The Model-Theoretic Ordinal Analysis of Theories of Predicative Strength

Jeremy Avigad and Richard Sommer

October 14, 1997

Abstract

We use model-theoretic methods described in [3] to obtain ordinal analyses of a number of theories of first- and second-order arithmetic, whose proof-theoretic ordinals are less than or equal to Γ_0.

1 Introduction

In [3] we introduced a model-theoretic approach to ordinal analysis as an interesting alternative to cut elimination. Here we extend these methods to the analysis of stronger theories of first- and second-order arithmetic which are nonetheless predicatively justifiable.

When used in this sense, the word "predicative" refers to a foundational stance under which one is willing to accept the set of natural numbers as a completed totality, but not the set of all subsets of the natural numbers. In this spirit, predicative theories bar definitions that require quantification over the full power set of $\mathbb{N}$, depicting instead a universe of sets of numbers that is constructed "from the bottom up." Work of Feferman and Schütte has established that the ordinal Γ_0 is the least upper bound to the strength of such theories (see for example, [4]). More recently a number of theories that are not *prima facie* justifiable on predicative grounds have been shown to have, in fact, predicative strength, in the sense of proving the same arithmetic statements as their predicatively justifiable counterparts. The analysis of such theories is our present concern.

This paper is best read as a sequel to [3]. In Section 2 we recap basic definitions from that source and introduce some new notation. In Section 3 we present a lemma, due to the second author, that allows one to build transfinite jump hierarchies, yielding ordinal analyses of the theories $(\Pi_1^0\text{-}CA)_{\prec\omega^\alpha}$. In Sections 4 and 5 we analyze the theories ACA and $\Sigma_1^1\text{-}AC$ respectively, and in the two remaining sections we

1

treat the theories $\widehat{ID}_n$, $\widehat{ID}_{<\omega}$, ATR_0, and ATR. (Stricly speaking, ATR goes beyond the bounds of predicativity, since its ordinal is Γ_{ε_0}; but we have included it here because its analysis is not much more difficult than that of ATR_0.)

2   Preliminaries

Suppose we have fixed an initial segment of the countable ordinals, and assigned cofinal sequences $\lambda[0], \lambda[1], \lambda[2], \ldots$ to limit ordinals λ. The notion of an α-large set of ordinals is defined inductively, as follows:

- Every set is 0-large.

- A set A is $(\beta + 1)$-large if it is nonempty and $A - \{\min(A)\}$ is β-large.

- If λ is a limit, then a set A is λ-large if it is nonempty and $A - \{\min(A)\}$ is $\lambda[\min(A)]$-large.

Fix a nonstandard model of arithmetic $\mathcal{M}$. The approach to ordinal analysis described in [3] involves starting with an appropriately large interval $[a, b]$ with nonstandard endpoints in $\mathcal{M}$, and using it to build a model $\mathcal{N}$ of the theory in question. The constructions proceed by extracting from $[a, b]$ a nonstandardly-large subset A, and possibly other $\mathcal{M}$-finite sets S, with various combinatorial properties. The first-order part of $\mathcal{N}$ is then taken to be any limit I of points in A (that is, any initial segment of $\mathcal{M}$ with no greatest element, in which points of A occur cofinally), and elements of the second-order universe of $\mathcal{N}$ are obtained by taking intersections of the sets S with I (these intersections are denoted S^I). The trick is to design the combinatorial properties of A and S so that, "in the limit," S^I will have desired properties in $\mathcal{N}$. In practice we often blur the distinction between an $\mathcal{M}$-finite set S and its potentially unbounded counterpart in I, and drop the superscript from S^I.

For example, suppose we want to guarantee that, in the limit, S will be the Turing jump of T; that is

$$I \models S = T'$$

where $T' = \{x \mid \mathrm{Tr}_{\Sigma_1^0}(x, T)\}$ and $\mathrm{Tr}_{\Sigma_1^0}(x, Z) \equiv \exists y \; \Theta(x, y, Z)$ is a complete Σ_1^0 truth predicate relative to Z. Our goal is to define a finitary combinatorial notion "S approximates the Turing jump of T in A," written

$$A \approxeq S = T',$$

which will guarantee that $S = T'$ holds in I. The motivation behind the definition is that in order determine the jump of a set Z it is

sufficient to have bounds on where to find witnesses y to the formula $\exists y \, \Theta(x, y, Z)$. We introduce the notation $j^{a,b}$ by

$$j^{a,b}(Z) = \{e < a \mid \exists y \leq b \, \Theta(e, y, Z)\},$$

and we take

$$j^{a,\infty}(Z) = \{e < a \mid \exists y \, \Theta(e, y, Z)\}.$$

Clearly for all natural numbers a there is a value a_Z such that

$$j^{a,a_Z}(Z) = j^{a,\infty}(Z),$$

and any integer greater than a_Z will also satisfy this equation. Also note that

$$Z' = \bigcup_a j^{a,\infty}(Z).$$

Our definition of $A \approx\!\!\!\!\!\approx S = T'$, where $A = \{a_0, \ldots, a_k\}$, is motivated by the desire to have a_k behave like ∞ and the mapping $a_i \mapsto a_{i+1}$ satisfy the properties of a mapping $a \mapsto a_Z$ described above.

In the next definition we use the notation S^a to denote the set $\{x \in S \mid x < a\}$; note that this agrees with the definition of S^I if we identify a with the set of natural numbers less than a.

Definition 2.1 *Let $A = \{a_0, a_1, \ldots, a_k\}$, S, and T be finite sets. Say that S approximates the Turing jump of T in A, written $A \approx\!\!\!\!\!\approx S = T'$, if the following hold:*

i. for every $i < k$, $j^{a_i, a_{i+1}}(T) = j^{a_i, a_k}(T)$, and

ii. $k \geq 1$ implies $S^{a_{k-1}} = j^{a_{k-1}, a_k}(T)$.

Although there is, for each T, a unique set S satisfying $S = T'$, note that sets S such that $A \approx\!\!\!\!\!\approx S = T'$ are not uniquely determined, since the definition does not say anything about what numbers greater than a_{k-1} are in S.

The following lemma states the fundamental property of the "approximates the jump" relation.

Lemma 2.2 *Let A, S, and T be finite sets in $\mathcal{M}$ such that*

$$\mathcal{M} \models (A \approx\!\!\!\!\!\approx S = T').$$

Then for any limit I of A,

$$I \models S = T'.$$

The proof of this lemma is straightforward and can be found in [3]. Recall that a set $A = \{a_0, a_1, \ldots, a_k\}$ is *spread out* if for all $i < k - 3$, $2^{a_i} < a_{i+1}$. In [3] it is also shown that if A is spread out and there is a set S such that $A \approx\!\!\!\!\!\approx S = \emptyset'$, then I will be a model of $I\Sigma_1$.

The next lemma lists some basic properties of "approximates the jump."

3

Lemma 2.3   *1. For any S, T, and a,*

$$\emptyset \not\approx S = T' \ \text{ and } \ \{a\} \not\approx S = T'.$$

2. *If $B \subset A$ and $A \not\approx S = T'$ then $B \not\approx S = T'$.*

3. *Suppose B is obtained from A by replacing the minimum element of A by a smaller number (i.e., $B = (A - \{\min A\}) \cup \{\min B\}$ and $\min B \leq \min A$), then $A \not\approx S = T'$ implies $B \not\approx S = T'$.*

4. *Suppose $A = \{a_0, \ldots, a_k\}$ and $k \geq 1$. Then*

$$A \not\approx S = T'$$

if and only if

$$S^{a_{k-1}} = j^{a_{k-1}, a_k}(T), A - \{a_0\} \not\approx S = T' \ \text{ and } \ j^{a_0, a_1}(T) = j^{a_0, a_k}(T).$$

For the analysis of predicative theories we need a sufficiently strong notation system. The one that follows is based on Veblen's sequence of ordinal functions φ_α, each of which enumerates the fixed points of its predecessors (for more information see [3, 4]).

Definition 2.4 *Our set of* ordinal notations *is defined inductively, as follows:*

- 0 *is an ordinal notation.*

- *If $\alpha_1, \alpha_2, \ldots, \alpha_k$ are ordinal notations other than 0, then so is*

$$\alpha_1 + \alpha_2 + \ldots + \alpha_k.$$

- *If α and β are ordinal notations, so is $\varphi(\alpha, \beta)$.*

- *If α is an ordinal notation, so is Γ_α.*

Notations of the form $\alpha + 1$ (that is, $\alpha + \varphi(0,0)$) are called successor *notations. A notation that is neither 0 nor a successor notation is called a* limit *notation.*

The symbol $\preceq$ denotes the usual order relation for notations of this form. When we refer to notations such as 1, ω^α, ω_n, ε_α, γ_n, and so on, these are to be taken as abbreviations for their usual representations with 0, $+$, and φ. In particular, we use $\alpha \cdot n$ to denote the term

$$\alpha + \alpha + \ldots + \alpha$$

in which there are n terms in the sum, ω^α to denote $\varphi(0, \alpha)$, and $\varphi_\alpha^n(\beta)$ to denote the n-fold iteration

$$\varphi(\alpha, \varphi(\alpha, \ldots \varphi(\alpha, \beta))).$$

4

Our treatment of ordinal addition violates unique readability, since, for example, the term $\alpha + \beta + \gamma$ can be interpreted by associating to the left or to the right. As it turns out, blurring this distinction is convenient, and one can check that the definitions and proofs below are insensitive to the way such a term is parsed.

Definition 2.5 *Sequences are assigned to limit notations as follows. (Here λ always denotes a limit ordinal.)*

1. $(\alpha + \beta)[n] =_{\text{def}} \alpha + (\beta[n])$.

2. $\omega^{\alpha+1}[n] =_{\text{def}} \omega^{\alpha} \cdot (n+2)$

3. $\varphi(\alpha, \lambda)[n] =_{\text{def}} \varphi(\alpha, \lambda[n] + 1)$.

4. $\varphi(\alpha + 1, 0)[n] =_{\text{def}} \varphi_{\alpha}^{n+2}(1)$.

5. $\varphi(\alpha + 1, \beta + 1)[n] =_{\text{def}} \varphi_{\alpha}^{n+2}(\varphi(\alpha + 1, \beta) + 1)$.

6. $\varphi(\lambda, 0)[n] =_{\text{def}} \varphi(\lambda[n], \lambda[n])$.

7. $\varphi(\lambda, \beta + 1)[n] =_{\text{def}} \varphi(\lambda[n] + 1, \varphi(\lambda, \beta) + 1)$.

8. $\Gamma_0[n] = \gamma_{n+1}$

9. $\Gamma_{\alpha+1}[n] = \gamma_{n+1}^{\Gamma_{\alpha}+1}$.

10. $\Gamma_{\lambda}[n] = \Gamma_{\lambda[n]+1}$.

We have chosen these particular limit sequences to facilitate our constructions, though they differ from the "standard" assignments only slightly.

Note that different notations can denote the same ordinal, as is the case with ε_0 and ω^{ε_0}. Further note that equivalent notations need not have equivalent limit sequences; for example, $\varepsilon_0[n] = \omega_{n+2}$, but $\omega^{\varepsilon_0}[n] = \omega^{\omega_{n+2}+1}$. We assume that to each notation α there has been assigned a canonical *normal form* $\overline{\alpha}$, satisfying the following:

Lemma 2.6 *For any notations α and β, we have*

1. *If $\alpha \equiv \beta$ (that is, $\alpha \preceq \beta$ and $\beta \preceq \alpha$) then $\overline{\alpha} = \overline{\beta}$.*

2. *$\alpha \equiv \overline{\alpha}$.*

We will also assume, for simplicity, that $\overline{\alpha + 1} = \overline{\alpha} + 1$.

Because our constructions take place in a model of arithmetic, we need to assume that notations α have been coded as numbers $\ulcorner \alpha \urcorner$ in a reasonable way. The requirements in [3] were very minimal; here, because the models we construct contain jump hierarchies that are again indexed by ordinals, we need to assume that the following two lemmas are satisfied.

Lemma 2.7 *If $\alpha = \alpha_1 + \cdots + \alpha_k$ then $\ulcorner \alpha \urcorner > \ulcorner \alpha_i \urcorner$ for $i = 1, 2, \ldots, k$, and $\ulcorner \alpha \urcorner > k$.*

5

Lemma 2.8 *There is a (standard) number k such that for every notation α and natural number x,*

$$\ulcorner\alpha[x]\urcorner \leq (\ulcorner\alpha\urcorner)^{k\cdot x},$$

and

$$\ulcorner\omega^\alpha\urcorner \leq (\ulcorner\alpha\urcorner)^k.$$

Choosing a coding that satisfies these is not difficult. In fact, under a reasonable coding scheme the first condition of the second lemma will follow from the fact that the "length" of $\alpha[x]$ is less than x times the length of α. Furthermore, the bounds stated in this lemma are not essential: any bound that is elementary in x and $\ulcorner\alpha\urcorner$ will do.

Lemma 2.9 below is a corollary of Lemma 2.7, by a straightforward induction on codes of notations. Note that statements (1) and (2) are logically equivalent.

Lemma 2.9

1. *If λ is a limit notation and $\lambda[n] \preceq \gamma \prec \lambda$ then $\ulcorner\gamma\urcorner \geq n$.*

2. *If λ is a limit notation, $\gamma \prec \lambda$, and $\ulcorner\gamma\urcorner < n$ then $\gamma \prec \lambda[n]$.*

3   Approximating transfinite jump hierarchies

In this section we use appropriately large intervals to build approximations to jump hierarchies indexed by ordinals notations. In this context, it is traditional to use only notations in normal form. To simplify notation we adopt the convention that whenever an ordinal is used as an index to such a hierarchy, it is implicitly "cast" to normal form. In other words, H_α is to be interpreted as $H_{\overline{\alpha}}$, and we define $H_{\prec\alpha}$ to be the disjoint union $\bigoplus_{\gamma\prec\alpha} H_{\overline{\gamma}}$.

Definition 3.1 *The set H is an α-level jump hierarchy, written $\mathcal{H}_\alpha(H)$, if the following conditions hold:*

1. Successor conditions: *if $\gamma \prec \alpha$ then*

$$H_{\gamma+1} = (H_{\preceq\gamma})'.$$

2. Limit conditions: *if $\lambda \preceq \alpha$ is a limit notation then*

$$H_\lambda = H_{\prec\lambda}.$$

If $H_0 = S$ then H is an α-level jump hierarchy from S*, which is written $\mathcal{H}_\alpha^S(H)$.*

Note that the definition does not specify what H_i is when i is not a notation in normal form, or when i is a notation in normal form that denotes an ordinal greater than α, so H is not uniquely determined.

Definition 3.2 *Suppose $A = \{a_0, \dots, a_k\}$. The set H approximates an α-level jump hierarchy in A, written $A \approx \mathcal{H}_\alpha(H)$, if the following conditions hold:*

1. *Successor conditions: if $\gamma \prec \alpha$ and $\ulcorner \gamma \urcorner < a_i$, where $i \leq k$, then*

$$\{a_i, a_{i+1}, \dots, a_k\} \approx H_{\gamma+1} = (H_{\preceq \gamma})'.$$

2. *Limit conditions: if $\lambda \preceq \alpha$ is a limit notation and $\ulcorner \lambda \urcorner < a_k$, then*

$$H_\lambda = H_{\prec \lambda}.$$

If $H_0 = S$ then H approximates an α-level jump hierarchy from S in A, which is written $A \approx \mathcal{H}_\alpha^S(H)$.

The following lemma asserts the fundamental property of the relation $A \approx \mathcal{H}_\alpha(H)$.

Lemma 3.3 *Suppose*

$$M \models (A \approx \mathcal{H}_\alpha(H)),$$

A is spread out, and I is any limit of A. Then

$$I \models \mathcal{H}_\alpha(H).$$

Proof. Assume $A \approx \mathcal{H}_\alpha(H)$ and suppose I is a limit of A. If $\ulcorner \gamma \urcorner \in I$ then for some $i \leq k$, $\ulcorner \gamma \urcorner < a_i \in I$. Thus if $\gamma \prec \alpha$, by the successor conditions of Definition 3.2 we have that $\{a_i, \dots, a_k\} \approx H_{\gamma+1} = (H_{\preceq \gamma})'$. Since $a_i \in I$ we have that I is a limit of $\{a_i, \dots, a_k\}$, and Lemma 2.2 implies

$$I \models H_{\gamma+1} = (H_{\preceq \gamma})'.$$

If $\gamma \preceq \alpha$ is a limit notation in I then $\ulcorner \gamma \urcorner < a_k$, and so

$$I \models H_\gamma = H_{\prec \gamma}$$

follows from the limit conditions of Definition 3.2. $\qquad \square$

The following definition will be useful in proving Lemma 3.6 below. A lemma listing basic properties of approximate jump hierarchies follows the definition.

Definition 3.4 *We say H agrees with J up to α, and write $H \equiv_\alpha J$, if $H_{\preceq \alpha} = J_{\preceq \alpha}$.*

7

Lemma 3.5

1. $\equiv_\alpha$ is an equivalence relation.

2. If $\beta \preceq \alpha$ and $H \equiv_\alpha J$ then $H \equiv_\beta J$.

3. For any H and α, $\emptyset \not\approx \mathcal{H}_\alpha(H)$.

4. For any A and H, $A \not\approx \mathcal{H}_0(H)$.

5. If $B \subseteq A$ and $A \not\approx \mathcal{H}_\alpha(H)$ then $B \not\approx \mathcal{H}_\alpha(H)$.

6. If $\beta \preceq \alpha$ and $A \not\approx \mathcal{H}_\alpha(H)$ then $A \not\approx \mathcal{H}_\beta(H)$.

7. If $H \equiv_\alpha J$ and $A \not\approx \mathcal{H}_\alpha(H)$ then $A \not\approx \mathcal{H}_\alpha(J)$.

8. Suppose B is obtained from A by replacing A's minimum element by a smaller integer (i.e., $B = (A - \{\min A\}) \cup \{\min B\}$ where $\min B \leq \min A$), then $A \not\approx \mathcal{H}_\alpha(H)$ implies $B \not\approx \mathcal{H}_\alpha(H)$.

The proofs of the above are straightforward from the definitions. To augment our notation, we add to the assertion $A \not\approx \mathcal{H}_\beta(H)$ the information that A is α-large by writing

$$A \stackrel{\alpha}{\not\approx} \mathcal{H}_\beta(H).$$

The following lemma can be thought of as a model-theoretic counterpart to the predicative cut-elimination lemma (see, for example, [8]).

Lemma 3.6 *Suppose there is a set H such that $A \stackrel{\varphi(\rho,\alpha)}{\not\approx} \mathcal{H}_\beta(H)$. Then there is a spread out $B \subset A$ and a set $J \equiv_\beta H$ such that $B \stackrel{\alpha}{\not\approx} \mathcal{H}_{\beta+\omega^\rho}(J)$.*

Proof.

Suppose $A \stackrel{\varphi(\rho,\alpha)}{\not\approx} \mathcal{H}_\beta(H)$. We will prove the lemma by transfinite induction on $\varphi(\rho, \alpha)$. In the event $\alpha = 0$ or α is a limit the result is easy. If $\alpha = 0$ then, by Lemma 3.5.3, we get the result by taking $B = \emptyset$ and $J = H$.

If α is a limit then $A - \{a_0\}$ is $\varphi(\rho, \alpha[a_0] + 1)$-large, where $a_0 = \min A$. Applying the induction hypothesis, we get $\hat{B}$ and J such that $\hat{B} \subset A - \{a_0\}$, $J \equiv_\beta H$, and

$$\hat{B} \stackrel{\alpha[a_0]+1}{\not\approx} \mathcal{H}_{\beta+\omega^\rho}(J).$$

Taking $B = (\hat{B} - \{\min \hat{B}\}) \cup \{a_0\}$, we have that $B \subset A$, B is α-large and, by Lemma 3.5.8, $B \stackrel{\alpha}{\not\approx} \mathcal{H}_{\beta+\omega^\rho}(J)$. To see that B is spread out it is enough to note that, for any set X, if X is spread out and Y is obtained from X by replacing $\min X$ with something smaller, then Y is spread out.

We will handle three separate cases for $\alpha = \alpha_0 + 1$, according to whether ρ is 0, a successor, or a limit. In each of these cases we will appeal to the induction hypothesis to get sets $\hat{B}$ and $\hat{J}$ such that

$$\hat{B} \subset A - \{a_0\}, \tag{1}$$

$$\hat{J} \equiv_\beta H, \tag{2}$$

and

$$\hat{B} \overset{\alpha_0}{\approx} \mathcal{H}_{\beta+\omega^\rho}(\hat{J}). \tag{3}$$

The set $\hat{B}$ is one element short of being α-large, and in each case we add a_0 to $\hat{B}$ to get B. We suppose

$$B = \{b_0, b_1, \ldots b_k\},$$

and so $b_0 = a_0$ and k is the cardinality of $\hat{B}$. The trick is to pick $\hat{B}$ such that $B = \hat{B} \cup \{b_0\}$ has the right properties; this will go differently in the different cases, but there are points of similarity that we mention now.

Our main focus is to establish

$$B \overset{}{\approx} \mathcal{H}_{\beta+\omega^\rho}(J), \tag{4}$$

For this we need to show that the successor and limit conditions of Definition 3.2 hold. In part, we will use (3), but we also use

$$\text{if } \rho \neq 0 \text{ then } B \overset{\alpha}{\approx} \mathcal{H}_{\beta+\omega^\rho[b_0]}(\hat{J}), \tag{5}$$

and

$$\text{if } \rho = 0 \text{ and } k \geq 1 \text{ then } j^{b_0,b_1}(J_{\preceq\beta}) = j^{b_0,b_k}(J_{\preceq\beta}). \tag{6}$$

Before showing how to get $\hat{B}$ and $\hat{J}$ satisfying (1), (2), (3), (5), and (6), we will use these conditions to get (4). For this we need to say how J is defined.

Define J by

$$\text{if } \rho \neq 0 \text{ set } J_{\prec\beta+\omega^\rho} =_{\text{def}} \hat{J}_{\prec\beta+\omega^\rho} \text{ and } J_{\beta+\omega^\rho} =_{\text{def}} J_{\prec\beta+\omega^\rho}, \tag{7}$$

and

$$\text{if } \rho = 0 \text{ set } J_{\preceq\beta} =_{\text{def}} \hat{J}_{\preceq\beta}, \text{ and if } k \geq 1 \text{ set } J_{\beta+1} =_{\text{def}} j^{b_{k-1},b_k}(J_{\preceq\beta}). \tag{8}$$

To show (4) from the above, we need to pay special attention to the $k = 0$ case; this corresponds to $\hat{B} = \emptyset$. First we show the limit

conditions for (4). If $\hat{B} \neq \emptyset$ then using Lemma 3.5, (1), (2), (3) and the definition of J given by (7) and (8) we get the limit conditions for (4), but if $\hat{B} = \emptyset$ then the limit conditions for (3) are trivially satisfied, and so (3) doesn't help for establishing the limit conditions for (4). Nonetheless, when $\rho \neq 0$, we can use (5) to get the limit conditions for $\lambda \prec \beta + \omega^\rho[b_0]$, and if $\beta + \omega^\rho$ is a limit, which it is when $\rho \neq 0$, the conditions $\lambda \prec \beta + \omega^\rho$ and $\ulcorner\lambda\urcorner < b_k$ imply, by Lemma 2.9, that $\lambda \prec \beta + \omega^\rho[b_k]$; since $k = 0$ this means $\lambda \prec \beta + \omega^\rho[b_0]$, and we have the desired limit conditions in the case $\rho \neq 0$. If $\rho = 0$ then, for λ a limit, $\lambda \preceq \beta + \omega^\rho$ if and only if $\lambda \preceq \beta$, so the limit conditions for (4) follow from (1), (2), (8) and the fact $A \approx\mathcal{H}_\beta(H)$ (using Lemma 3.5.5 and 3.5.7).

The successor conditions for (4) in the cases $\gamma \prec \beta + \omega^\rho$ and $\ulcorner\gamma\urcorner < b_i$, when $i > 0$, follow from (3); in the case $\gamma \prec \beta + \omega^\rho$, $\ulcorner\gamma\urcorner < b_0$, and $\rho \neq 0$, we use Lemma 2.9 to get $\gamma \preceq \beta + \omega^\rho[b_0]$, and then the result follows from (5). If $\rho = 0$ then $\omega^\rho = 1$ and $\gamma \prec \beta + \omega^\rho$ if and only if either $\gamma \prec \beta$ or $\gamma = \beta$. The successor conditions for $\gamma \prec \beta$, in the event $\rho = 0$, follow from (1), (2), (8) and the assumption $A \approx\mathcal{H}_\beta(H)$ (using Lemma 3.5.5 and 3.5.7). In order to take care of the case $\gamma = \beta$, first note that if $k = 0$ then this case follows from Lemma 2.3.1. If $k \geq 1$, the desired successor condition follows from Lemma 2.3.4, (6), and the second part of (8).

The fact $B \subset A$ follows from $\hat{B} \subset A$, $b_0 = \min A$ and $B =_{\text{def}} \hat{B} \cup \{b_0\}$. The fact $J \equiv_\beta H$ follows from the definition of J and (2). A bit of an argument is needed in order to show that B is spread out; we will take care of that as we go through the cases for proving (3), (5), and (6).

Case $\rho = 0$ and $\alpha = \alpha_0 + 1$: Notice that this is the main case of Lemma 2.2 which was proved as Lemma 8.3 in [3]; since the argument is presented there in detail (as well as in several cited references), we will be brief here. Note that $A - \{a_0\}$ is $\omega^{\alpha_0}(a_0 + 2)$-large. By a property for partitioning $\omega^\delta l$-large sets (see [3, Lemma 5.5]), there is an increasing partition of $A - \{a_0\}$ into $a_0 + 2$ many ω^{α_0}-large sets P_0, P_1, $\dots$, P_{a_0+1}. By the pigeon-hole principle we can select $j \geq 1$ such that for all $e < a_0$,

$$(\mu y < \max A)\Theta(e, y, H_{\preceq\beta}) \notin (\min P_j, \max P_j]. \qquad (9)$$

By the induction hypothesis there is a set $\hat{B} \subset P_j$ and a set $\hat{J} \equiv_\beta H$ such that

$$\hat{B} \overset{\alpha_0}{\approx} \mathcal{H}_{\beta+1}(\hat{J});$$

i.e., such that (3) holds.

Assume $k \geq 1$. Since $\min P_j \leq b_1 \leq b_k \leq \max P_j$, (9) implies

$$j^{b_0,b_1}(\hat{J}_{\preceq\gamma}) = j^{b_0,b_k}(\hat{J}_{\preceq\gamma});$$

10

i.e., (6) holds.

To see that B is spread out, note that, by the induction hypothesis, $\hat{B}$ is spread out, so we need only verify that if $k > 3$ then $2^{b_0} < b_1$. Since we selected $j \geq 1$, the ω^{α_0}-large set P_0 sits between b_0 and b_1. If $\alpha_0 < 3$ then (removing elements from $\hat{B}$ if necessary) we can get by with $k \leq 2$, so assume $\alpha_0 \geq 3$, then by a property of ω^δ-large sets when $\delta \geq 3$ (see [3, Lemma 5.6]), $2^{\min P_0} < \max P_0$, and so $2^{b_0} < b_1$, as desired.

Case $\rho = \rho_0 + 1$ and $\alpha = \alpha_0 + 1$: Apply the induction hypothesis $a_0 + 2$ times starting with the $\varphi_{\rho_0}^{a_0+2}(\varphi(\rho, \alpha_0) + 1)$-large set $A - \{a_0\}$ in place of A resulting in sets B^* and J^* such that $J^* \equiv_\beta H$ and

$$B^* \overset{\varphi(\rho,\alpha_0)+1}{\approx} \mathcal{H}_{\beta+\omega^{\rho_0}(a_0+2)}(J^*) \tag{10}$$

Then apply the induction hypothesis again with $B^* - \{\min B^*\}$ in place of A and J^* in place of H, to get

$$\hat{B} \subset B^* - \{\min B^*\} \tag{11}$$

and

$$\hat{J} \equiv_{\beta+\omega^{\rho_0}(a_0+2)} J^* \tag{12}$$

such that

$$\hat{B} \overset{\alpha_0}{\approx} \mathcal{H}_{\beta+\omega^\rho}(\hat{J});$$

i.e., (3) holds. Using the fact that $\beta + \omega^\rho[b_0] = \beta + \omega^{\rho_0} \cdot (b_0 + 2)$, (10) and (12) imply

$$B^* \approx \mathcal{H}_{\beta+\omega^\rho[b_0]}(\hat{J}). \tag{13}$$

Using Lemma 3.5.5 and (11) we have

$$\hat{B} \cup \{\min B^*\} \overset{\alpha}{\approx} \mathcal{H}_{\beta+\omega^\rho[b_0]}(\hat{J}). \tag{14}$$

So, by Lemma 3.5.8 we have

$$B \overset{\alpha}{\approx} \mathcal{H}_{\beta+\omega^\rho[b_0]}(\hat{J});$$

i.e., (5) holds.

To see that B is spread out, use $b_0 < \min B^*$, (11), and the fact B^* is spread out (which is given by the induction hypothesis).

Case ρ is a limit and $\alpha = \alpha_0 + 1$: Apply the induction hypothesis with the $\varphi(\rho[b_0] + 1, \varphi(\rho, \alpha_0) + 1)$-large set $A - \{a_0\}$ in place of A resulting in sets B^* and J^* such that

$$B^* \overset{\varphi(\rho,\alpha_0)+1}{\approx} \mathcal{H}_{\beta+\omega^{\rho[b_0]+1}}(J^*) \tag{15}$$

11

Apply the induction hypothesis again with the $\varphi(\rho, \alpha_0)$-large set $B^* - \{\min B^*\}$ in place of A and J^* in place of H to get $\hat{B}$ and $\hat{J}$ such that

$$\hat{B} \overset{\alpha_0}{\approx} \mathcal{H}_{\beta + \omega^\rho}(\hat{J}).$$

So (3) holds. By Lemma 3.5.5 and (15),

$$\hat{B} \cup \{\min B^*\} \approx \mathcal{H}_{\beta + \omega^{\rho[b_0]+1}}(\hat{J}). \tag{16}$$

Using Lemma 3.5.8, we then have

$$B \approx \mathcal{H}_{\beta + \omega^{\rho[b_0]+1}}(\hat{J}). \tag{17}$$

Since ρ is a limit, $\beta + \omega^\rho[b_0] = \beta + \omega^{\rho[b_0]+1}$, and so

$$B \overset{\alpha}{\approx} \mathcal{H}_{\beta + \omega^\rho[b_0]}(\hat{J}) \tag{18}$$

i.e., (5) holds.

To see that B is spread out, use $b_0 < \min B^*$, $B - \{b_0\} \subset B^* - \{\min B^*\}$, and the fact B^* is spread out (which is given by the induction hypothesis). $\qquad\square$

Relativizing the construction in Lemma 3.6 and setting $\beta = 0$ yields

Lemma 3.7 *Suppose C and T are sets and C is $\varphi(\rho, \alpha)$-large. Then there are sets A and H such that A is an α-large subset of C and spread out, and $A \approx \mathcal{H}_{\omega^\rho}^T(H)$.*

The following lemma asserts that the various levels of a transfinite jump hierarchy code the truth of arithmetic formulas involving previous levels. To state it, we expand the language of arithmetic to include constants that denote sets of the form H_γ, so that if ψ is a formula with parameters $H_{\gamma_1}, \ldots, H_{\gamma_k}$, the code $\ulcorner \psi \urcorner$ can refer to these parameters.

Lemma 3.8 *There is an $I\Delta_0 + (\exp)$-definable function*

$$\mathrm{TruthCode}(\ulcorner \psi \urcorner, x)$$

with the following property: whenever

$$\mathcal{N} \models I\Delta_0 + (\exp) + \mathcal{H}_\alpha(H),$$

ψ is a $\Sigma_m^0(H_{\gamma_1}, \ldots, H_{\gamma_k})$ formula, and

$$\alpha \succeq \beta \succeq (\sup\{\gamma_1, \ldots, \gamma_k\} + m),$$

then the equivalence

$$\psi \leftrightarrow \mathrm{TruthCode}(\ulcorner \psi \urcorner, \beta) \in H_\beta$$

holds in $\mathcal{N}$.

The proof of this lemma is routine.

The constructions in this section enable us to build models of theories that assert the existence of jump hierarchies. If α is a limit notation in normal form, $(\Pi_1^0\text{-}CA)_{\prec\alpha}$ is a theory in the language of second-order arithmetic that consists of the basic defining axioms for successor, plus, and times, arithmetic comprehension with parameters, an induction axiom for sets of natural numbers, and axioms

$$\exists X\ \mathcal{H}_\gamma(X)$$

for every (standard) notation $\gamma \prec \alpha$ (see [7]).

Theorem 3.9 *Suppose a and b are nonstandard elements of $\mathcal{M}$ such that*

$$\mathcal{M} \models [a,b]\ is\ \varphi(\alpha,0)\text{-}large,$$

where α is a limit notation. Then there are a cut $a < I < b$ and a finite set S in $\mathcal{M}$ such that

$$\mathcal{N} = \langle I, \{S_i^I \mid i \in I\}\rangle$$

is a model of $(\Pi_1^0\text{-}CA)_{\prec\omega^\alpha}$. (In particular, if α is an ε-number, this is equivalent to $(\Pi_1^0\text{-}CA)_{\prec\alpha}$.)

Proof. Since $[a,b]$ is $\varphi(\alpha,0)$-large, $[a+1,b]$ is $\varphi(\alpha[a+1],\alpha[a+1])$-large. By Lemma 3.7 there are sets A and H such $A \overset{\alpha[a+1]}{\approx} \mathcal{H}_{\omega^{\alpha[a+1]}}(H)$. One can verify that since α is a limit notation and A is $\alpha[a+1]$- large, A is also ω-large. (In the "counting down" procedure it is impossible to pass from an ordinal greater than ω to an ordinal less than ω, without hitting ω first.) As a result, the set A has at least $a+1$ elements. Let

$$A' =_{\mathrm{def}} \{a_0, a_1, \ldots, a_a\}$$

denote the first $a+1$ elements of A, let I be any limit of A', and let

$$J =_{\mathrm{def}} \{j \mid a_j \in I\},$$

so that J is a limit of $\{0, 1, \ldots, a\}$.

Let $\mathcal{N}$ be the model defined by

$$\mathcal{N} =_{\mathrm{def}} \langle I, \{(H_{\omega^{\alpha[j]}}^{[k]})^I \mid k \in I, j \in J\}\rangle$$

We claim that $\mathcal{N}$ models $(\Pi_1^0\text{-}CA)_{\prec\omega^\alpha}$. Verifying that arithmetic comprehension holds is not difficult, using Lemma 3.8. And since every standard $\gamma \prec \omega^\alpha$ is less than $\omega^{\alpha[n]}$ for some standard n, and H_γ is then equal to $H_{\omega^{\alpha[n]}}^{[k]}$ for some appropriate k,

$$\mathcal{N} \models \exists X\ \mathcal{H}_\gamma(X),$$

as desired. As in [3, Theorem 9.4] the second-order universe of $\mathcal{N}$ can be coded into a single set S. □

4 Constructing a model of ACA

The theory ACA is a subsystem of second-order arithmetic consisting of quantifier-free defining equations for successor, plus, and times, an axiom schema (ACA) of comprehension for arithmetic formulas with numeric and set parameters, and full second-order induction. Since the arithmetically definable sets form a natural interpretation of the second-order variables of this theory, it isn't surprising that an ω-level jump hierarchy can be used to build a model. Indeed, one can verify that if H is such a hierarchy in a model

$$\mathcal{N} = \langle \mathcal{K}, \{H\} \rangle$$

of the weak base theory $I\Delta_0 + (\exp)$, then

$$\mathcal{N}' =_{\text{def}} \langle \mathcal{K}, \{H_i^{[k]} \mid i, k \in \mathcal{K}\} \rangle$$

will satisfy (ACA). But now the second-order universe of $\mathcal{N}'$ is definable in $\mathcal{N}$, so, for example, given any arithmetic formula $\theta(X, Y)$ we have that

$$\mathcal{N}' \models \exists X \ \forall Y \ \theta(X, Y)$$

if and only if

$$\mathcal{N} \models \exists x, i \ \forall y, j \ \theta(H_i^{[x]}, H_j^{[y]}).$$

In short, second-order quantification in $\mathcal{N}'$ reduces to first-order quantification relative to the parameter H in $\mathcal{N}$. This suggests that to build a model of ACA we only need to build a model in which first-order induction holds relative to an ω-level jump hierarchy H. But Lemma 3.7 makes this easy: starting from a suitably large interval we can obtain finite sets H and A such that A is ε_0-large and H approximates an ω-level jump hierarchy in A. Then we can use the techniques of [3] to thin A down so that first-order induction will hold relative to the parameter H in any limit of A.

Theorem 4.1 *Suppose $\mathcal{M}$ is a model of true arithmetic, and a and b are nonstandard elements of $\mathcal{M}$ such that*

$$\mathcal{M} \models [a, b] \ is \ \varepsilon_{\varepsilon_0}\text{- large.}$$

Then there are a cut $a < I < b$ and a finite set S coded in $\mathcal{M}$ such that

$$\langle I, \{S_j^I \mid j \in I\} \rangle$$

is a model of ACA.

Proof. Recall that $\varepsilon_{\varepsilon_0}$ abbreviates $\varphi(1, \varepsilon_0)$. Applying Lemma 3.7 we obtain a sets A and a set H such that $A \overset{\varepsilon_0}{\approx} \mathcal{H}_\omega(H)$. As in [3] we can

14

thin A to an ω-large set A' (and build another jump hierarchy from H), so that first-order induction is guaranteed to hold relative to H in any limit I. Now if we take

$$\mathcal{N} =_{\text{def}} \langle I, \{H_i^{[k]} \mid i, k \in I\}\rangle$$

the previous discussion shows that $\mathcal{N}$ will be a model of ACA. As usual, we can code the second-order universe of $\mathcal{N}$ into a single set S. □

5 Constructing a model of Σ_1^1-AC

We would like to extend the construction of the previous section to model the theory Σ_1^1-AC, which adds to ACA the Σ_1^1 axiom of choice, (Σ_1^1-AC). It is important to realize, however, that the simple approach to building a model of Σ_1^1-AC_0 described in [3] falls short. That construction relied on the fact that the assertion

$$\forall x \ \exists Y \ \varphi(x, Y) \tag{19}$$

in the final model implied

$$I \models \forall x \ \exists y \ \varphi(x, H_e^{[y]}) \tag{20}$$

for a single parameter e. Furthermore— and this is crucial— such a parameter e could be found in the model, that is, in the cut I. This fact allowed us to express the truth of (20) in any limit I with a formula in $\mathcal{M}$, and then find the least such e.

In the construction in the previous section, however, we needed to index the jump hierarchy by all the elements of I, so that the fact that (19) holds in the final model does *not* necessarily imply that there is an e in I satisfying (20). On the other hand, if e is beyond I then the formula (20) is not coded below any a_i in I, and hence the truth of (20) in I is not expressible in $\mathcal{M}$. As a result, the construction for Σ_1^1-AC_0 falls apart here.

The solution is to use the jump lemma to build a much larger hierarchy, allowing us to refer to more sets from within I. We will show that if we have an α-level hierarchy for a suitable α, we can find an initial segment K of α such that whenever (19) holds of the hierarchy along K, then (20) holds when e is replaced by any notation $\gamma \succ K$ in I; and K has the further property that whenever such a formula holds for every γ above K, it holds for some γ *in* K as well.

In previous constructions we've proceeded by building a single set S that codes the sets of a second-order universe $\{S_x\}$. We would like to make this process more explicit now. If φ is a second-order formula,

15

we define φ *relativized to S* to be the translation of φ whereby second-order quantifiers are taken to range over the collection of sets coded by S. For example, if φ is the second-order formula

$$\forall X_1 \, \exists X_2 \, \forall X_3 \, \theta(X_1, X_2, X_3),$$

where θ is arithmetic, then φ^S is the formula

$$\forall x_1 \, \exists x_2 \, \forall x_3 \, \theta(S_{x_1}, S_{x_2}, S_{x_3}).$$

Using relativization we can express the fact that a S codes the second-order part of a model of a particular theory.

Definition 5.1 *Let S be a set of natural numbers. Say that S is an ω-model of Σ_1^1-AC if (ACA) and $(\Sigma_1^1$-AC$)$ hold relativized to S.*

Lemma 5.2 *Suppose*

$$\langle \mathcal{K}, \{S\} \rangle$$

is a model of Peano Arithmetic in which induction holds relative to the parameter S, and

$$\mathcal{K} \models S \text{ is an } \omega\text{-model of } \Sigma_1^1\text{-}AC.$$

Then

$$\mathcal{N} =_{\mathrm{def}} \langle \mathcal{K}, \{S_i \mid i \in K\} \rangle$$

is a model of Σ_1^1-AC.

Proof. Straightforward. Second-order induction in $\mathcal{N}$ follows from first-order induction in $\langle \mathcal{K}, \{S\} \rangle$. $\qquad\square$

Having reduced our task to that of constructing an ω-model of Σ_1^1-AC, we now come to the main lemma in this section. We pause to note that the proof was inspired by similar "pseudohierarchy" construc-tions in [7, 6] (see also [10, 11]). The constructions in [7, 6] however, rely on deep proof-theoretic results and Gödel's second incompleteness theorem, whereas, in contrast, the constructions here are more direct.[1]

In clause (2) of the following lemma, as in Lemma 3.8, we allow arithmetic formulas φ to include parameters of the form H_γ. Intuitively it states that anything (coded low enough) that happens in the jump hierarchy at stage β_i already happens before stage α_i.

---

[1]We would like to point out that our methods also allow us to build a model of the Σ_1^1 axiom scheme of dependent choice, $(\Sigma_1^1$-$DC)$, by building a jump hierarchy H, using the methods of [12] to guarantee that transfinite induction holds relative to H, and then employing the techniques of [7, 11]. In this case, however, we do not know of a more direct construction.

Lemma 5.3 *Suppose C is a $\varphi(\alpha, \alpha)$-large set, and ω^α is coded below $2^{\min(C)^2}$. Then there is a set*

$$A = \{a_0, a_1, \dots, a_k\} \subseteq C$$

such that A is α-large and spread out, a set H such that $A \approx \mathcal{H}_{\omega^\alpha}(H)$, and sequences of ordinal notations

$$0 = \alpha_0 \prec \alpha_1 \prec \alpha_2 \prec \dots \prec \alpha_k \prec \beta_k \prec \dots \prec \beta_1 \prec \beta_0 = \omega^{\alpha[a_0]}$$

such that for each $0 < i < k$ the following hold:

1. *If $i < k - 4$, the codes for α_i and β_i are less than $2^{a_{i+1}^2}$.*

2. *If $\varphi(X)$ is an arithmetic formula coded below a_{i-1}, and*

$$\mathrm{TruthCode}(\varphi(H_{\beta_i}), \beta_0) \in H_{\beta_0},$$

   then for some $\gamma \preceq \alpha_i$ that is coded below a_{i+1} we have

$$\mathrm{TruthCode}(\varphi(H_\gamma), \beta_0) \in H_{\beta_0}.$$

Proof. Since C is $\varphi(\alpha, \alpha)$-large, by Lemma 3.7 we can find sets A and H, such that

$$A = \{a_0, a_1, \dots, a_k\}$$

is α-large and spread out, and $A \approx \mathcal{H}_{\omega^\alpha}(H)$. As a result, we only need to construct the sequences of α_i and β_i satisfying the conclusion of the lemma.

Setting $\alpha_0 = 0$ and $\beta_0 = \omega^{\alpha[a_0]}$, we will carry out the construction in k steps, where at each stage $i > 0$ we construct α_i and β_i so that

$$\beta_i = \alpha_i + \omega^{\alpha[a_0, \dots, a_i]} \tag{21}$$

and

$$\alpha_{i-1} \prec \alpha_i \prec \beta_i \prec \beta_{i-1}, \tag{22}$$

and clauses (1) and (2) of the lemma are satisfied.

Suppose we've constructed $\alpha_0, \alpha_1, \dots, \alpha_i$ and $\beta_0, \beta_1, \dots, \beta_i$. At stage $i + 1$, there are two possibilities: either

$$\alpha[a_0, a_1, \dots, a_i] = \delta + 1$$

for some δ or

$$\alpha[a_0, a_1, \dots, a_i] = \lambda,$$

where λ is a limit.

In the first case we have

$$\beta_i = \alpha_i + \omega^{\delta+1}.$$

17

For each $j \leq a_i + 3$, define

$$\eta_j = \alpha_i + \omega^\delta \cdot j,$$

so that

$$\alpha_i = \eta_0 \prec \eta_1 \prec \ldots \eta_{a_i+3} \prec \beta_i$$

and for each $j < a_i + 2$ we have

$$
\begin{aligned}
\eta_{j+1} &= \eta_j + \omega^\delta \\
&= \eta_j + \omega^{\alpha[a_0, a_1, \ldots, a_{i+1}]}.
\end{aligned}
$$

Since there are $a_i + 1$ many values η_j, for $j > 0$, the pigeonhole principle implies that we can find an $l > 0$ so that

$$l \neq \mu j \leq a_i + 2 \ (\mathrm{TruthCode}(\ulcorner \varphi(H_{\eta_{j+1}}) \urcorner, \beta_0) \in H_{\beta_0})$$

for any of the at most a_i formulas φ coded below a_i. Set

$$\alpha_{i+1} =_{\mathrm{def}} \eta_l$$

and

$$\beta_{i+1} =_{\mathrm{def}} \eta_{l+1}.$$

Then if

$$\mathrm{TruthCode}(\ulcorner \varphi(H_{\beta_{i+1}}) \urcorner, \beta_0) \in H_{\beta_0}$$

for some φ coded below a_i, we have that

$$\mathrm{TruthCode}(\ulcorner \varphi(H_{\eta_j}) \urcorner, \beta_0) \in H_{\beta_0}$$

for some $j \leq l$. Using Lemmas 2.8 and 2.7 and the induction hypothesis one can verify that the codes of α_{i+1} and β_{i+1} are less than $2^{a_{i+1}^2}$. So clauses (1) and (2) of the lemma are satisfied, and the induction hypotheses (21) and (22) are also maintained.

The case where $\alpha[a_0, a_1, \ldots, a_i] = \lambda$ is almost identical; in this case, take

$$\eta_j =_{\mathrm{def}} \alpha_i + \omega^{\lambda[a_{i+1}]} \cdot j$$

and proceed as before. $\qquad\square$

The following definition is slightly awkward since it defines a property of a set A by giving conditions on $A - \{\min(A)\}$. This concession pays off later on, in that it simplifies that statement of Lemma 5.7. The definition should not be taken to imply that the conditions of Lemma 5.3, expressible in the language of first-order arithmetic, provide the *only* way of ensuring that S will be an ω-model of Σ_1^1-AC in any limit I of a nonstandard A. What is important is that they provide *one* way of doing so, as is evidenced by the lemmas that follow. Notice that we use the usual coding trick of [3] to guarantee that only certain sets are "seen" by the limit I.

Definition 5.4 *Let $A = \{a_0, a_1, \ldots, a_k\}$ and S be finite sets. Say that S approximates an ω-model of Σ_1^1-AC in A if there is a set H and sequences $\langle \alpha_i \rangle$ and $\langle \beta_i \rangle$ satisfying the conclusion of Lemma 5.3 with A replaced by $A - \{min(A)\}$; and, letting*

$$T_{\langle a_i, \gamma, d \rangle} = H_\gamma^{[d]}$$

we have that

$$S = \bigoplus_{a_i \in A, \gamma \prec \alpha_i, d < a_k} T_{\langle a_i, \gamma, d \rangle}.$$

The idea behind this definition is as follows. Suppose S approximates an ω-model of Σ_1^1-AC in $A = \{a_0, a_1, \ldots, a_k\}$, and let I be any limit of A. Then I determines a limit of the sequence

$$\alpha_0, \alpha_1, \ldots, \alpha_k,$$

namely the set of α_i corresponding to some a_i in I. If we let K denote the set of ordinal notations in I below some such α_i, then K denotes a set of notations in I with no greatest element, and in fact, an initial segment with the favorable properties discussed in the opening paragraphs of this section. The set S^I then codes the sets that are definable in some level of H corresponding to a notation in K.

The next two lemmas show that Definition 5.4 has the desired properties. The first was inspired by [7].

Lemma 5.5 *Suppose $\langle \mathcal{N}, \{H\} \rangle$ is a model of $I\Delta_0 + (\exp)$, such that $\mathcal{N} \models \mathcal{H}_\beta(H)$. Let K be an initial segment of the notations less than β in $\mathcal{N}$ such that K has no greatest element, and define*

$$\mathcal{S} =_{\mathrm{def}} \{H_\gamma^{[d]} \mid \gamma \in K, d \in \mathcal{N}\}.$$

Suppose further that whenever $\varphi(X)$ is an arithmetic formula with set parameters from $\mathcal{S}$ and $\varphi(H_\delta)$ holds for every δ above K, then it holds for some δ in K as well. Then

$$\mathcal{K} =_{\mathrm{def}} \langle \mathcal{N}, \mathcal{S} \rangle$$

is a model of (ACA) and $(\Sigma_1^1$-$AC)$.

Proof. That (ACA) holds is implied by the fact that K has no greatest element, as follows. Let $\varphi(x, \vec{P})$ be a Σ_l^0 formula with set parameters $\vec{P}$ in $\mathcal{S}$. Each parameter P is of the form $H_\gamma^{[d]}$ for some $\gamma \in K$, so we can pick γ' to be the largest of these. Since K has no greatest element, it is closed under successors, so

$$\delta =_{\mathrm{def}} \gamma' + l$$

is also in K. By Lemma 3.8 we have that

$$\mathcal{N} \models \forall x \; (\varphi(x) \leftrightarrow \text{TruthCode}(\ulcorner \varphi(x) \urcorner, \delta) \in H_\delta).$$

But then

$$\mathcal{N} \models \forall x \; (\varphi(x) \leftrightarrow x \in H_\delta^{[\varphi]}),$$

and this latter set is an element of $\mathcal{S}$.

To see that $\mathcal{K}$ is a model of $(\Sigma_1^1\text{-}AC)$, suppose that

$$\mathcal{K} \models \forall x \; \exists Y \; \varphi(x, Y) \tag{23}$$

for some arithmetic formula φ with parameters in $\mathcal{S}$. The fact that (23) holds implies that for every x there is a γ in K such that

$$\mathcal{N} \models \exists y \; \varphi(x, H_\gamma^{[y]}).$$

Since every level of the hierarchy codes all the ones that come before it, we have that

$$\mathcal{N} \models \forall x \; \exists y \; \varphi(x, H_\delta^{[y]}) \tag{24}$$

for *any* δ above K. By the hypothesis of the lemma there is some δ in K satisfying (24) as well. But the set H_δ is an element of $\mathcal{S}$. Now, using arithmetic comprehension in $\mathcal{K}$, define Y so that for every x

$$Y_x = H_\delta^{[f_x]},$$

where f_x is the least natural number such that $\varphi(x, H_\delta^{[f_x]})$ holds. Then we have

$$\forall x \; \varphi(x, Y_x),$$

witnessing the conclusion of $(\Sigma_1^1\text{-}AC)$. $\qquad\qquad\square$

Lemma 5.6 *Suppose S and A are sets in $\mathcal{M}$ such that*

$$\mathcal{M} \models S \text{ approximates an } \omega\text{-model of } \Sigma_1^1\text{-}AC \text{ in } A.$$

If I is any limit of A, then

$$I \models S^I \text{ is an } \omega\text{-model of } \Sigma_1^1\text{-}AC.$$

Proof. Suppose S, A, and I are in the statement of the lemma. Since I is a limit of A it is a limit of $A - \{\min(A)\}$, so there are a set H and sequences $\langle \alpha_i \rangle$ and $\langle \beta_i \rangle$ satisfying the conclusion of Lemma 5.3. Define

$$\mathcal{S} = \{S_i^I \mid i \in I\}.$$

20

and let
$$K =_{\mathrm{def}} \{\gamma \in I \mid \exists i\, (a_i \in I \wedge \gamma \prec \alpha_i)\}.$$

It is not difficult to verify that

$$\mathcal{S} = \{(H_\gamma^{[d]})^I \mid \gamma \in K, d \in I\}.$$

so we only need to verify that I, H^I, and K satisfy the hypotheses of Lemma 5.5.

The fact that K has no greatest element in I follows from the fact that if $\gamma \in K$, then for some i such that $a_i \in I$, we have that $\gamma \prec \alpha_i$. Since I is a limit, a_{i+1} is in I as well, and hence $\alpha_i \prec \alpha_{i+1}$ is also in K.

Suppose that $\varphi(\mathcal{S})$ is an arithmetic formula with set parameters from $\mathcal{S}$, and $\varphi(H_\delta)$ holds in I for every δ above K. Note that as long as δ is sufficiently below β_0, we will have that

$$\mathrm{TruthCode}(\ulcorner\varphi(H_\delta)\urcorner, \beta_0) \in H_{\beta_0}. \qquad (25)$$

Find a_i so that in addition $\ulcorner\varphi(X)\urcorner$ is less than a_i. Then we have that (25) holds for $\delta = \beta_{i+1}$, and so clause (2) of Lemma 5.3 implies that it also holds for some $\delta \in K$ as well. By Lemma 3.8

$$I \models \varphi(H_\delta)$$

for this δ. $\qquad\square$

From Lemma 5.3 we now obtain the following

Lemma 5.7 *Suppose C is $\varphi(\lambda, 0)$-large and λ is coded below $\min(C)$. Then there are sets A and S, such that A is λ-large and spread out, and S approximates an ω-model of Σ_1^1-AC in A.*

Proof. Set
$$a_0 =_{\mathrm{def}} \min(C).$$

Since C is $\varphi(\lambda, 0)$-large, $C - \{a_0\}$ is $\varphi(\lambda[a_0], \lambda[a_0])$-large, and one can use Lemma 2.8 to show that $\ulcorner\omega^{\lambda[a_0]}\urcorner$ is coded below $2^{a_0^2}$. Apply Lemma 5.3 to obtain a $\lambda[a_0]$-large set A', a set H, and sequences $\langle \alpha_i \rangle$ and $\langle \beta_i \rangle$. Set
$$A =_{\mathrm{def}} A' \cup \{a_0\}$$

and define S as in Definition 5.4. $\qquad\square$

Putting it all together yields the following

Theorem 5.8 *Suppose a and b are nonstandard elements of $\mathcal{M}$ such that*

$$\mathcal{M} \models [a,b] \text{ is } \varphi(\varepsilon_0, 0)\text{-large.}$$

Then there are a cut $a < I < b$ and a finite set S coded in $\mathcal{M}$ such that

$$\langle I, \{S_j^I \mid j \in I\}\rangle$$

is a model of Σ_1^1-AC.

Proof. By Lemma 5.7 there are an ε_0-large $A \subset [a,b]$ and a set S such that S approximates an ω-model of Σ_1^1-AC in A. As in the proof of Theorem 4.1 we can thin A down to an ω-large set A' such that first-order induction relative to the parameters S holds in any limit I of A'. Lemmas 5.6 and 5.2 then yield a model of $(\Sigma_1^1$-$AC)$ and (ACA) in which full second-order induction holds. $\qquad\square$

6   Approximating finite Σ_1^1-AC hierarchies

Using our relativized version of the transfinite jump lemma, we can iterate the construction in the previous section, to obtain nested ω-models of Σ_1^1-AC. This gives rise to the following sequence of definitions and lemmas.

Definition 6.1 *Say S is an ω-model of Σ_1^1-AC containing T if S is an ω-model of Σ_1^1-AC and for some i, $S_i = T$.*

Definition 6.2 *Say that S approximates an ω-model of Σ_1^1-AC containing T in A if S is as in Definition 5.4, except that H approximates a jump hierarchy from T in A.*

Of course, if S approximates an ω-model of Σ_1^1-AC containing T in A, where S, T, and A are all coded in some nonstandard model of true arithmetic, then S^I is an ω-model of Σ_1^1-AC containing T^I in any limit I of A. The lemmas in the previous section can be relativized to yield the following

Lemma 6.3 *Suppose C is an $\varphi(\lambda, 0)$-large set, and T is any finite set in $\mathcal{M}$. Then there are a set S and a λ-large $A \subseteq C$ such that S approximates an ω-model of Σ_1^1-AC containing T in A.*

Analogous to the notion of a c-level jump hierarchy, we introduce the notion of a c-level hierarchy of nested models of Σ_1^1-AC.

Definition 6.4 *Say H is a c-level nested Σ_1^1-AC hierarchy if for each i, $0 < i \leq c$, H_i is an ω-model of Σ_1^1-AC containing $H_{<i}$.*

Definition 6.5 *Say H approximates a c-level nested Σ_1^1-AC hierarchy in A if for each i, $0 < i \leq c$, H_i approximates an ω-model of Σ_1^1-AC containing $H_{<i}$ in A.*

If H_0 is some set T in Definition 6.4 (6.5), we will say that H is (approximates) a c-level nested Σ_1^1-AC hierarchy *from* T. Once again, it is not hard to show that if H approximates a c-level nested Σ_1^1-AC hierarchy from T in A, and $\min(A)$ is nonstandard, then H^I *is* a c-level nested Σ_1^1-AC hierarchy from T^I in any limit I of A.

As in the case of finite jump hierarchies, we can approximate finite nested Σ_1^1-AC hierarchies if we start from a suitably large interval.

Lemma 6.6 *Suppose C is γ_c^α-large, and T is any finite set. Then there are an α-large set A and a set H such that H approximates a c-level nested Σ_1^1-AC hierarchy from T in A.*

Proof. Recall that $\gamma_0^\alpha = \alpha$ and $\gamma_{i+1}^\alpha = \varphi(\gamma_i^\alpha, 0)$. The proof is just a simple iteration of Lemma 5.7. $\qquad\square$

Finally, as in Section 3, we can extend these constructions to the transfinite. Though we only state the results for ω iterations, we note that by an appropriate extension to our notation systems they can easily be generalized.

Definition 6.7 *Say H is an ω-level nested Σ_1^1-AC hierarchy if for every $i > 0$, H_i is an ω-model of Σ_1^1-AC containing $H_{<i}$.*

Definition 6.8 *Suppose $A = \{a_0, a_1, \ldots, a_k\}$. Say H approximates an ω-level nested Σ_1^1-AC hierarchy in A if, whenever $0 < i < a_{j-1}$, H_i approximates an ω-model of Σ_1^1-AC containing $H_{<i}$ in $\{a_{j+1}, a_{j+2}, \ldots, a_k\}$.*

Lemma 6.9 *Suppose C is Γ_α-large. Then there are a set H and an α-large set $A \subseteq C$ such that H approximates an ω-level nested Σ_1^1-AC hierarchy in A.*

Proof. The proof is analogous to that of Lemma 3.6. $\qquad\square$

7 Constructing models of $\widehat{ID}_n$, ATR_0, and ATR

As it turns out, nested Σ_1^1-AC hierarchies are useful in constructing models of the theories $\widehat{ID}_n$, ATR_0, and ATR, which we now address.

Let $\varphi(x, Y)$ be an arithmetic formula, in which the sole set parameter Y occurs positively (that is, in the scope of an even number of negation symbols, assuming that φ is written using the connectives $\exists$,

$\forall$, $\wedge$, $\vee$, and $\neg$). We can think of such an arithmetic formula as a "positive arithmetic operator" since it defines the monotone function

$$\Gamma_\varphi : P(\omega) \to P(\omega)$$

given by

$$\Gamma_\varphi(A) = \{x | \varphi(x, A)\}.$$

(The monotonicity means that for any sets A and B, $A \supset B$ implies $\Gamma_\varphi(A) \supset \Gamma_\varphi(B)$.) Classically such operators are known to have fixed-points: for example, defining

- $\Gamma_\varphi^0 =_{\text{def}} \emptyset$
- $\Gamma_\varphi^{\alpha+1} =_{\text{def}} \Gamma_\varphi(\Gamma_\varphi^\alpha)$
- $\Gamma_\varphi^\lambda =_{\text{def}} \bigcup_{\gamma \prec \lambda} \Gamma_\varphi^\gamma$, for limit ordinals λ,

the monotonically increasing sequence Γ_φ^α must stabilize at some countable stage δ. At this point, we will have $\Gamma_\varphi(\Gamma_\varphi^\delta) = \Gamma_\varphi^\delta$, so that Γ_φ^δ defines a fixed-point of Γ_φ, and, in fact, the least fixed-point (in the sense that it is contained in any other fixed-point).

The theory $\widehat{ID}_1$ is a first-order theory in the language of Peano Arithmetic, with an additional predicate P_φ for each positive arithmetic operator $\varphi(x, X)$. $\widehat{ID}_1$ then extends the axioms of PA with axioms

$$\forall x (P_\varphi(x) \leftrightarrow \varphi(x, P_\varphi))$$

that assert that each P_φ is a fixed-point (though not necessarily least) of the operator Γ_φ. Similarly, each theory $\widehat{ID}_{n+1}$ adds new constants for positive arithmetic formulas in the language of $\widehat{ID}_n$ and the corresponding fixed point axioms, and $\widehat{ID}_{<\omega}$ is the union of the $\widehat{ID}_n$'s. See [5, 2] more more information on $\widehat{ID}_{<\omega}$, and [1] for more information on inductive definitions in general.

The connection between the theories $\widehat{ID}_n$ and nested Σ_1^1-AC hierarchies is given by the following two lemmas.

Lemma 7.1 *The theory $\widehat{ID}_1$ can be interpreted in Σ_1^1-AC.*

Proof (sketch). Let $\varphi(x, Y)$ be any arithmetic (or even Σ_1^1) formula in which the set parameter Y occurs positively. Aczel has observed that by using a universal Σ_1^1 truth predicate and diagonalizing as in the proof of Gödel's fixed-point theorem, there is a Σ_1^1 formula $\psi(z)$ that defines a fixed point of φ. In particular, for this ψ, Σ_1^1-AC proves

$$\forall x \ (\psi(x) \leftrightarrow \varphi(x, \{z \mid \psi(z)\})).$$

(The axiom scheme (Σ_1^1-AC) is required to bring set quantifiers to the front of appropriate formulas.) One can then interpret the fixed-point constants of $\widehat{ID}_1$ using such ψ, so that induction in $\widehat{ID}_1$ is reduced to second- order induction in (Σ_1^1-AC). See [5] for more details. $\square$

Lemma 7.2 *The theory $\widehat{ID}_n$ can be interpreted in ACA_0 together with the extra axiom "there is an n-level nested Σ_1^1-AC hierarchy."*

Proof (sketch). One inductively shows that each fixed-point constant from the language of $\widehat{ID}_i$ can be interpreted by a set that is arithmetic in the ith level of the nested Σ_1^1-AC hierarchy. $\qquad\square$

Combined with Lemma 6.6 this yields

Theorem 7.3 *Suppose a and b are nonstandard elements of $\mathcal{M}$ such that*

$$\mathcal{M} \models [a, b] \text{ is } \gamma_n\text{-large.}$$

Then there are a cut $a < I < b$ and a finite set T coded in $\mathcal{M}$ such that

$$\langle I, \dots, T^I_{\ulcorner\varphi\urcorner}, \dots\rangle$$

is a model of $\widehat{ID}_n$, where the sets $T^I_{\ulcorner\varphi\urcorner}$ interpret the fixed-point constants.

Proof. Recall that $\gamma_1 =_{\text{def}} \gamma_1^{\varepsilon_0}$. By Lemma 6.6 we can find a set S' and an ε_0-large set A' such that S' approximates an n-level nested Σ_1^1-AC hierarchy in A'. As in the proof of Theorem 4.1 we can thin A' down to an ω-large set A, and obtain a set S which will code the universe of a model of (ACA) containing S'. We can read off interpretations in S for the fixed-point constants of $\widehat{ID}_n$ from the proof of Lemma 7.2. $\square$

If $[a, b]$ is Γ_0-large then $[a + 1, b]$ is γ_a-large, and we can use the same construction to obtain an a-level nested Σ_1^1-AC hierarchy. If a is nonstandard, we can interpret all the "standard" fixed-point constants of $\widehat{ID}_{<\omega}$. This yields

Theorem 7.4 *Suppose a and b are nonstandard elements of $\mathcal{M}$ such that*

$$\mathcal{M} \models [a, b] \text{ is } \Gamma_0\text{-large.}$$

Then there are a cut $a < I < b$ and a finite set T coded in $\mathcal{M}$ such that

$$\langle I, \dots, T^I_{\ulcorner\varphi\urcorner}, \dots\rangle$$

is a model of $\widehat{ID}_\omega$.

The theories ATR_0 and ATR extend ACA_0 and ACA respectively, by adding an a schema (ATR) which allows definitions by arithmetic transfinite recursion along any well ordering:

$$WO(\prec) \to \exists Y\, \forall b, x\, (x \in Y_b \leftrightarrow \varphi(x, Y_{\prec b})) \qquad (ATR)$$

where φ ranges over arithmetic formulas, possibly involving set parameters. Here $WO(\prec)$ represents the Π_1^1 assertion that the set $\prec$ codes a well-ordering, that is, every set X contains a $\prec$-least element. Intuitively, (ATR) asserts that given any well-ordering, we can build a hierarchy Y such that each level b is obtained from an arithmetic comprehension over all the levels that have preceded it. For more information about ATR_0, see [11, 9, 10, 2].

The connection between arithmetic transfinite recursion and nested Σ_1^1-AC hierarchies is given by the following

Lemma 7.5 *Over ACA_0, the scheme (ATR) is equivalent to the assertion "for every X, there is an ω-model of Σ_1^1-AC containing X."*

Proof. We sketch the right-to-left direction, which is the only direction we need below. From within ACA_0, suppose every set X is contained in an ω-model of Σ_1^1-AC and $\prec$ is a well-ordering. Letting $\varphi(x, Y)$ be any arithmetic formula, we need to show that there is a transfinite hierarchy defined by φ along $\prec$. By coding all the set parameters of φ into a single set, we can find an ω-model S of Σ_1^1-AC that contains these parameters. We claim that for every c there is a (unique) set W in S that codes the hierarchy up to c, i.e. W satisfies

$$\forall b \prec c \; \forall x \; (x \in W_b \leftrightarrow \varphi(x, W_{\prec b})), \tag{26}$$

and $W_b = \emptyset$ if b is not a predecessor of c. Notice that the claim is an arithmetic assertion in the parameter S. Suppose there is a c for which the claim is false, i.e. for this c there is no W in S satisfying (26). By (ACA) and the assumption that $\prec$ is a well-ordering, we can find the least such c. Then for every $d \prec c$, there is a hierarchy up to d in S. But using $(\Sigma_1^1$-$AC)$ in S we can combine all these hierarchies into a single set, and then using arithmetic comprehension in S we can turn this into a hierarchy up to c, contrary to our assumption.

We've shown that for every c there is a hierarchy up to c in S. Again, using $(\Sigma_1^1$-$AC)$ and (ACA) in S, we can find a hierarchy defined for all the elements of $\prec$, completing the proof.

The left-to-right direction of the lemma can be found in [11, 10]. $\square$

Lemma 7.5, combined with Lemmas 6.6 and 6.9, yield the last two theorems in this paper.

Theorem 7.6 *Suppose a and b are nonstandard elements of $\mathcal{M}$ such that*

$$\mathcal{M} \models [a, b] \text{ is } \Gamma_0\text{-large.}$$

Then there are a cut $a < I < b$ and a finite set S coded in $\mathcal{M}$ such that

$$\langle I, \{S_j^I \mid j \in I\}\rangle$$

26

is a model of ATR_0.

Proof. Since $[a, b]$ is Γ_0-large, $[a+1, b]$ is γ_{a+1}-large. As in the proof of Theorem 7.3 we obtain an ε_0-large set A' and a set T that approximates an $(a+1)$-level nested Σ_1^1-AC hierarchy in A', and thin A' down to an ω-large set A. Suppose I is any limit of A and J is any limit of the set $\{0, 1, \dots, a+1\}$. Since each T_{j+1}^I is an ω-model of Σ_1^1-AC containing T_j^I, it is not difficult to verify that

$$\mathcal{K} =_{\text{def}} \langle I, \{(T_j^I)_i \mid i \in I, j \in J\}\rangle$$

will satisfy (ACA) as well as the assertion "for every X there is an ω-model of Σ_1^1-AC containing X." Hence, by the preceeding lemma, $\mathcal{K}$ will be a model of ATR_0. By the usual trick we can find a single set S such that

$$\langle I, \{S_i \mid i \in I\}\rangle$$

is of this form. $\qquad\square$

Theorem 7.7 *Suppose a and b are nonstandard elements of $\mathcal{M}$ such that*

$$\mathcal{M} \models [a, b] \text{ is } \Gamma_{\varepsilon_0}\text{- large}.$$

Then there are a cut $a < I < b$ and a finite set S coded in $\mathcal{M}$ such that

$$\langle I, \{S_j^I \mid j \in I\}\rangle$$

is a model of ATR.

Proof. The proof bears the same relationship to the construction of a model of ACA in Theorem 4.1 as the previous proof bore to the construction of a model of ACA_0. Which is to say, first we use Lemma 6.9 to construct a set S that approximates an ω-level nested Σ_1^1-AC-hierarchy in an ε_0-large set $A \subseteq [a, b]$. Then we thin A down to an ω-large set A' and guarantee that induction will hold relative to A in any limit. Taking I to be a limit of A and

$$\mathcal{N} =_{\text{def}} \langle I, \{S_i^I \mid i \in I\}\rangle,$$

$\mathcal{N}$ will satisfy full second-order induction as well as (ACA) and (ATR). $\square$

References

[1] Aczel, Peter, "An introduction to inductive definitions," in J. Barwise, ed. *The Handbook of Mathematical Logic*, North-Holland, 1977, pages 739–782.

[2] Avigad, Jeremy, "On the relationship between ATR_0 and $\widehat{ID}_{<\omega}$," *Journal of Symbolic Logic*, 61:768–779, 1996.

[3] Avigad, Jeremy, and Richard Sommer, "A model-theoretic approach to ordinal analysis," *Bulletin of Symbolic Logic*, 3:17–52, 1997.

[4] Feferman, Solomon, "Systems of predicative analysis," *Journal of Symbolic Logic*, 29:1–30, 1964.

[5] Feferman, Solomon, "Iterated inductive fixed-point theories: application to Hancock's conjecture", in G. Metakides ed. *Patras Logic Symposium*, North-Holland, 1982.

[6] Friedman, Harvey, Kenneth McAloon, and Stephen Simpson, "A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis," in G. Metakides ed. *Patras Logic Symposium*, North-Holland, 1982, pages 197–230.

[7] Friedman, Harvey, "Iterated inductive definitions and $\Sigma_2^1\text{-}AC$," in A. Kino et al. eds., *Intuitionism and Proof Theory*, North-Holland, 1970, pages 435–442.

[8] Pohlers, Wolfram, *Proof Theory: An Introduction*, *Springer Verlag Lecture Notes in Mathematics 1407*, 1989.

[9] Simpson, Stephen G., "Subsystems of Z_2 and Reverse Mathematics," appendix to Gaisi Takeuti, *Proof Theory* (second edition), North-Holland, 1987.

[10] Simpson, Stephen G., "On the strength of König's duality theorem of countable bipartite graphs," Journal of Symbolic Logic, 59:113-123, 1994.

[11] Simpson, Stephen G., *Subsystems of Second Order Arithmetic*, preprint.

[12] Sommer, Richard, "Transfinite induction within Peano Arithmetic," *Annals of Pure and Applied Logic*, 76:231–289, 1995.