ELSEVIER Annals of Pure and Applied Logic 76 (1995) 231–289

ANNALS OF PURE AND APPLIED LOGIC

Transfinite induction within Peano arithmetic

Richard Sommer*

Department of Mathematics, Stanford University, Stanford, CA 94305-2125, USA

Received 10 August 1992; revised 10 May 1995; communicated by D. Van Dalen

Abstract

The relative strengths of first-order theories axiomatized by transfinite induction, for ordinals less-than ε_0, and formulas restricted in quantifier complexity, is determined. This is done, in part, by describing the provably recursive functions of such theories. Upper bounds for the provably recursive functions are obtained using model-theoretic techniques. A variety of additional results that come as an application of such techniques are mentioned.

1. Introduction

The main aim of this paper is to present results regarding subsystems of arithmetic axiomatized by transfinite induction.[1] Also, this paper can be viewed as an introduction to techniques that are capable of providing a variety of results in the study of subsystems of arithmetic and analysis, extending beyond what is proved here.

Well-known work of Gentzen [6] implies that *PA* proves transfinite induction up to ordinals less than ε_0, but not up to ε_0 itself. If $\mathscr{C}$ is a class of formulas and α is an ordinal less than ε_0, then the axiom schema of transfinite induction up to α for $\mathscr{C}$-formulas, $TI(\alpha, \mathscr{C})$, is given by

$$\forall\beta((\forall\gamma < \beta)\varphi(\gamma) \rightarrow \varphi(\beta)) \rightarrow (\forall\gamma < \alpha)\varphi(\gamma): \quad \varphi \in \mathscr{C}, \tag{1}$$

where α is the Gödel number of an ordinal less than ε_0, β and γ are variables ranging over elements of the universe that are Gödel numbers of ordinals less than ε_0, and $<$ denotes the less-than relation for ordinals ($<$ will be a defined predicate symbol in all

* E-mail: sommer@math.stanford.edu.

[1] Most of what appears in this paper is taken from the author's Ph.D. Thesis [19], directed by R.M. Solovay.

SSDI 0168-0072(95)00029-1

of our theories). We will include a weak base theory in $TI(\alpha,\mathscr{C})$. Using the work of Gentzen:

$$PA \vdash TI(\alpha,\mathscr{C}) \Leftrightarrow \alpha < \varepsilon_0, \tag{2}$$

where $\mathscr{C}$ is the set of arithmetic formulas. Our main goal is to measure the strength of the theories $TI(\alpha,\mathscr{C})$, where $\alpha < \varepsilon_0$ and $\mathscr{C}$ is a subset of the arithmetic formulas. The techniques used for this are primarily model-theoretic, in contrast with the methods applied by Gentzen, as well as techniques used by others in this area, which for the most part are proof-theoretic methods based on a version of cut-elimination.

Sections 2 and 3 present some preliminary material; Section 3 reports results stating that a very weak fragment of arithmetic, namely $I\Delta_0$, is a sufficient base theory for developing a significant amount of "basic ordinal arithmetic". $I\Delta_0$ is Peano arithmetic with the induction axioms restricted to bounded arithmetic formulas – it is much weaker than PRA, even unable to prove that exponentiation is a total function. Section 2 develops some basic coding theory in $I\Delta_0$. In Section 3, notions used in the definition of $TI(\alpha,\mathscr{C})$ are shown to make sense in $I\Delta_0$; e.g., $I\Delta_0$ is shown to be a base theory sufficient to interpret the notions of "ordinal", "ordinal less than" and prove many basic facts about these notions. Furthermore, $I\Delta_0$ is capable of defining many ordinal functions (addition, fundamental sequences, etc.), as well as prove many basic facts about these functions. One familiar with the capability of PRA to carry out such ordinal arithmetic may want to skip Sections 2 and 3, replacing $I\Delta_0$ with PRA in the later sections. Most proofs are omitted in Sections 2 and 3–the reader is referred to the literature.

In Section 4 we prove the relative strengths, over $I\Delta_0$, of the theories $TI(\alpha,\mathscr{C})$ for all $\alpha < \varepsilon_0$, and $\mathscr{C} = \Sigma_n$, Π_n, and Δ_0, for all $n \geqslant 1$. The results of this section will partially rely on results of Section 5, where the techniques are primarily model-theoretic and are more interesting than the proofs in Section 4. The relative strengths of these theories for $\mathscr{C} = \Pi_n$ with $n \geqslant 1$, and over the theory PRA, were originally proven proof-theoretically in [17]; the work here was obtained without knowledge of [17].

Some of the relative strengths, of these theories, are determined by characterizing the provably recursive functions of these theories in terms of the Wainer (equivalently, the extended Ackerman or fast-growing) hierarchy; this is done in Section 5. In particular, we will show that for all $n \geqslant 1$ and all $\alpha < \varepsilon_0$:

$$TI(\omega^{\omega^{\alpha}},\Pi_n) \vdash \forall x \exists y (F_\beta(x) = y) \Leftrightarrow \beta < \omega_n^{\alpha+1}, \tag{3}$$

where F_β is the βth function in the Wainer hierarchy, and ω_n^α is defined by: $\omega_0^\alpha \stackrel{\text{def}}{=} \alpha$, and $\omega_{n+1}^\alpha \stackrel{\text{def}}{=} \omega^{\omega_n^\alpha}$. In the process of proving this result we show that there is a Δ_0-formula, in the free variables α, x, and y, expressing $F_\alpha(x) = y$, for which many properties of the functions in this hierarchy are provable in $I\Delta_0$.

Together with results from Section 4, (3) yields a precise characterization of the functions in the Wainer hierarchy that are provably total in the theories $TI(\beta,\mathscr{C})$ for all $\beta < \varepsilon_0$ and $\mathscr{C} = \Sigma_n$, Π_n, or Δ_0. In Section 6 this characterization is used to classify the $TI(\beta, \mathscr{C})$-provably recursive functions. This is obtained by first proving the conservation result that for $n \geqslant 1$ and $\alpha < \varepsilon_0$,

$$\Pi_2(TI(\omega^{\omega^\alpha}, \Pi_n)) = \Pi_2(I\Delta_0 + \{\text{“}F_\gamma \text{ is total”}: \gamma < \omega_n^{\alpha+1}\}), \tag{4}$$

where $\Pi_n(T) \overset{\text{def}}{=}$ the Π_n-consequences of T.

Additionally in Section 6 it is claimed that the model-theoretic proofs of Section 5 can be carried out in the theory $I\Delta_0 + \text{Exp}$ (where Exp is the formula in the language of arithmetic expressing that the exponential function is total), yielding the following result:

$$I\Delta_0 \vdash 1\text{-Con}(TI(\omega^{\omega^\alpha}, \Pi_n)) \leftrightarrow \forall x \exists y (F_{\omega_n^{\alpha+1}}(x) = y). \tag{5}$$

Here n-Con(T) is the formalization of “the union of T and the set of all true Π_n-sentences is consistent”.

Also in Section 6 a hierarchy of theories generated by “iterated reflection” is considered, i.e. let $S_0 = T_0$, where T_0 is some base theory; $S_{\alpha+1} = S_\alpha + 1\text{-CON}(S_\alpha)$; and $S_\lambda = \bigcup_{\alpha<\lambda} S_\alpha$ for λ a limit ordinal. We will show a correspondence between this type of iteration and the type of iteration used to generate the F_β hierarchy. The study of such hierarchies began with the work of Feferman in [3]; Ulf Schmerl studied this hierarchy for $T_0 = PRA$ in [17].

Furthermore, Section 6 states various generalizations of the above results including: extending to arbitrary recursive ordinals beyond ε_0, considering axiom schemata for all arithmetic formulas (rather than restricting the number of quantifiers), extending to arbitrary recursive well-orderings (of a given order type), and extending to second-order theories. Many of the results in Section 6 are stated without proof, referring to future papers.

2. Preliminaries

2.1. Δ_0 and $I\Delta_0$

Let $\mathscr{L}(X)$ denote the first-order language with nonlogical symbols from X, and if $L = \mathscr{L}(X)$, and Y is a set of nonlogical symbols, then $L(Y)$ will denote the language obtained by adjoining the symbols in Y to L. L_0 will denote the usual language of arithmetic i.e. the language $\mathscr{L}(\{0, S, +, \cdot, <\})$.

A *Δ_0-formula* is any formula built up from the atomic formulas of L_0 using $\rightarrow$, $\leftrightarrow$, $\wedge$, $\vee$, $\neg$, $(\exists x \leqslant \tau)$, and $(\forall \tau \leqslant \tau)$ (where $(\exists x \leqslant \tau)\varphi$ and $(\forall \tau \leqslant \tau)\varphi$ are defined, for τ an L_0-term not containing x, by $\exists x (x \leqslant \tau \wedge \varphi)$ and $\forall x (x \leqslant \tau \rightarrow \varphi)$ respectively). For $n \geqslant 1$, a *Σ_n-formula* is a formula of the form $\exists \vec{x_1} \forall \vec{x_2} \exists \vec{x_3} \ldots Q\vec{x_n}\, \varphi(\vec{x_1}, \ldots, \vec{x_n}, \vec{y})$, where $\varphi \in \Delta_0$, $Q \in \{\forall, \exists\}$, and the notation $\vec{z}$ is used to denote a string of variables, say $z_1, \ldots, z_m$, for some $m \in \omega$. A *Π_n-formula* is defined similarly except its first quantifier is $\forall$ rather than $\exists$.

Let P^- be the set of axioms for a discretely ordered commutative semi-ring. Then P^- is a version of the "usual" set of basic axioms of arithmetic that is commonly used together with an axiom schema to axiomatize theories such as Peano arithmetic (PA). Let $I\mathscr{C}$ denote the theory axiomatized by P^- together with the axiom schema:

$$\forall\vec{z}(\varphi(0,\vec{z}) \wedge \forall y(\varphi(y,\vec{z}) \to \varphi(y+1,\vec{z})) \to \forall x\,\varphi(x,\vec{z})): \quad \varphi \in \mathscr{C}.$$

PA is then IL_0; $I\Sigma_n$, $I\Pi_n$, and $I\Delta_0$ are also commonly studied theories of this form. We note that in these cases, i.e. for $\mathscr{C} = L_0$, Σ_n, Π_n, or Δ_0, $I\mathscr{C}$ is equivalent to the theory P^- together with the strong induction axioms:

$$\forall\vec{z}(\forall x((\forall y < x)\varphi(y,\vec{z}) \to \varphi(x,\vec{z})) \to \forall x\,\varphi(x,\vec{z})): \quad \varphi \in \mathscr{C}.$$

Let "exp" be an expression denoting a symbol for the exponential function. $I\Delta_0(\exp)$ is defined similarly to $I\Delta_0$, allowing "exp" in the language when defining the bounded formulas for the induction axioms, and including the recursive defining equations for the exponential function in P^-. $I\Delta_0(\exp)$ is a conservative extension of $I\Delta_0+$ Exp, where Exp is an L_0-formula expressing that the exponential function is total.

2.2. *Sequences and coding*

In this section we introduce a Gödel numbering of sequences that has nice properties in $I\Delta_0$. We use the terms "Gödel number" and "code" interchangeably. We use ω to denote the set $\{0,1,2,\ldots\}$. For any set A, we use A^* to denote $\bigcup_{n\in\omega} A^n$, where A^n denotes the set of all sequences of length n. That is $A^n \stackrel{\text{def}}{=} \{(a_1,\ldots,a_n)\colon \text{for } i = 1,\ldots,n,\ a_i \in A\}$. We will use $*$ to denote concatenation of elements of A^*, i.e.,

$$(a_1,\ldots,a_n) * (b_1,\ldots,b_m) \stackrel{\text{def}}{=} (a_1,\ldots,a_n,b_1,\ldots,b_m).$$

As long as we are in a situation where we are not considering sequences of sequences, there will be no ambiguity if we use $a_0 * (a_1,\ldots,a_n)$ to denote $(a_0) * (a_1,\ldots,a_n)$, and if we denote the sequence $(a_1,\ldots,a_n)$ by $a_1 * \cdots * a_n$. If we have enough symbols on hand so that each element of A can be denoted by a single symbol then we will leave off the $*$'s. For example, for $1*1*2*1 \in \{1,2\}^*$ we may write 1121; however, when $1*1*2*1$ is regarded as an element of ω^*, and we are using strings of digits not separated by $*$'s to denote numerals in base ten, we will always include the $*$'s to avoid ambiguity. The *length* of $(a_1,\ldots,a_n)$ is n. We use Λ to denote the *empty* sequence, the unique sequence of length 0. For any sequences s and t we say s is an *initial segment* of t, or t is an *end extension* of s, if for some sequence u, $t = s*u$, and we write $s \subseteq_{\mathrm{e}} t$. And we say s is a *part of* t if for some sequences u and v, $t = u*s*v$; we write $s \subseteq_{\mathrm{p}} t$. We use the convention that if $j < i$, then $a_i * \cdots * a_j = \Lambda$.

Sequences will be coded using the following version of n-ary notation for the natural numbers. For $n \geqslant 2$, let the *Smullyan base n representation* (cf. [18]) of an integer x, denoted $\bar{x}^n$, be given by

$$\bar{x}^n = a_0 * a_1 * \ldots * a_k,$$

where $a_0, \ldots, a_k$ are the unique natural numbers such that for all $i \leqslant k$, $1 \leqslant a_i \leqslant n$, and

$$x = a_0 \cdot n^k + a_1 \cdot n^{k-1} + \cdots + a_{k-1} \cdot n + a_k.$$

For all $n \geqslant 2$, $\overline{}^n$ is a bijection from ω to $\{1, \ldots, n\}^*$; so for any $n \geqslant 2$, each $x \in \omega$ may be construed as an element of $\{1, \ldots, n\}^*$. We use $x *_n y$ to denote the natural number z such that $\overline{z}^n = \overline{x}^n * \overline{y}^n$, and we use $|x|_n$ to denote the length of $\overline{x}^n$. If $\overline{x}^n = a_0 * \cdots * a_k$ then $|x|_n = k + 1$ and we let

$$(x)_{n,i} \stackrel{\text{def}}{=} \begin{cases} a_i & \text{if } i \leqslant k; \\ 0 & \text{otherwise.} \end{cases}$$

Using the fact that there is a Δ_0-formula for the exponential function with nice properties provable in $I\Delta_0$ (cf., e.g., [7]), it can be shown that for each $n \geqslant 2$, the functions $x \mapsto |x|_n$, $*_n$, and $(x, i) \mapsto (x)_{n,i}$ are Δ_0-definable. We write $x \subsetneq_n y$ if $\overline{x}^n \subsetneq_e \overline{y}^n$. Since for each $n \geqslant 2$, $x \subsetneq_n y \leftrightarrow (\exists z \leqslant y)(x *_n z = y)$, it is clear that the relation $\subsetneq_n$ is Δ_0-definable. We state some properties of these functions and relations that hold in $I\Delta_0$.

Proposition 2.1. *For all $n \geqslant 2$, the following hold in $I\Delta_0$:*

(i) *"$|\ |_n$, $*_n$ and $(x, i) \mapsto (x)_{n,i}$ are total functions".*

(ii) *"$*_n$ is associative".*

(iii) *The Δ_0-definable relations are closed under the quantifiers $\forall x \subsetneq_n \tau$ and $\exists x \subsetneq_n \tau$, where τ is any L_0-term.*

(iv) $n^{|x|_n} - 1 \leqslant x \cdot (n - 1)$, *hence* $2^{|x|_2} \leqslant x + 1$.

(v) $x \cdot (n - 1) \leqslant n \cdot (n^{|x|_n} - 1)$; *also* $x + 1 < n^{|x|_n + 1}$.

(vi) $|x + 1|_n \leqslant |x|_n + 1$.

(vii) $x \leqslant y \rightarrow |x|_n \leqslant |y|_n$.

(viii) $|x|_n - 1 < \log_n(x) < |x|_n + 1$.

Next we show how elements of ω^* are coded. $\langle s_0, \ldots, s_k \rangle$ is used to denote the code of the sequence $(s_0, \ldots, s_k)$. The Smullyan base 3 notation for the code of $(s_0, \ldots, s_k)$, is obtained by replacing each s_i in $(s_0, \ldots, s_k)$ by $3 * \overline{s_i}^2$, i.e. a "3" followed by the Smullyan base 2 notation for s_i. This is given by

$$\overline{\langle \Lambda \rangle}^3 \stackrel{\text{def}}{=} \Lambda, \text{ and } \overline{\langle s_0, \ldots, s_k \rangle}^3 \stackrel{\text{def}}{=} 3 * \overline{s_0}^2 * 3 * \overline{s_1}^2 * \cdots * 3 * \overline{s_k}^2.$$

If $s = \langle s_0 * \cdots * s_k \rangle$ then let $\text{len}(s) \stackrel{\text{def}}{=} k + 1$, and

$$(s)_i \stackrel{\text{def}}{=} \begin{cases} s_i & \text{if } i \leqslant k; \\ 0 & \text{otherwise.} \end{cases}$$

We note that $\langle\ \rangle$ is a bijection from ω^* to a Δ_0-definable subset of ω; we let Sq denote this subset.

To obtain Δ_0-definitions for some of the above functions, note that there is a Δ_0-formula $v(x, y, i, k)$ expressing "there are exactly y occurrences of the digit i in the Smullyan base k representation of x", that has a number of nice properties provable in $I\Delta_0$. For an example of such a formula, ν, see Lemma 2.2 of [25]; in this example "Smullyan base k representation" is replaced by "k-ary expansion of".

Using the formula $\nu(x,y,i,k)$ above, it is easy to see that the relations $\subsetneq_{\mathrm{e}}$ and $\subsetneq_{\mathrm{p}}$, as well as the functions $x \mapsto \langle x \rangle$, len, $*$, and $(x,i) \mapsto (x)_i$ are Δ_0-definable. We will use obvious conventions such as writing $(x)_i < (y)_j$ in place of

$$\mathrm{Sq}(x) \wedge \mathrm{Sq}(y) \wedge \mathrm{len}(x) > i \wedge \mathrm{len}(y) > j \wedge \exists u, v((x)_i = u \wedge (y)_j = v \wedge u < v).$$

The following is a variation of well-known results:

Proposition 2.2. (i) $\langle\ \rangle$ *is a bijection from* ω^* *to* Sq.

(ii) *For all* $k \in \omega$, $I\Delta_0 \vdash \forall x_1, \ldots, x_k \exists z(z = \langle x_0, \ldots, x_k \rangle)$.

(iii) $I\Delta_0 \vdash$ "len, $*$, *and* $(x,i) \mapsto (x)_i$ *are total functions*".

(iv) *The* Δ_0*-definable relations are closed under the quantifiers* $\exists x \subsetneq_{\mathrm{e}} \tau$, $\forall x \subsetneq_{\mathrm{e}} \tau$, $\exists x \subsetneq_{\mathrm{p}} \tau$, *and* $\forall x \subsetneq_{\mathrm{p}} \tau$.

3. Ordinal arithmetic

3.1. Representability

Definition 3.1. Suppose we are given the following:

(a) A set of ordinals U and sets R and F of relations and functions on the ordinals in U. We require that $U \in R$.

(b) A set B of basic properties of the ordinal functions and relations in $R \cup F$, expressed in the language $\mathscr{L}(R \cup F)$.

(c) A theory, T, in the language $L_0 \overset{\mathrm{def}}{=} \mathscr{L}(\{0, S, +, \cdot, \leqslant\})$, for proving the basic properties, B, about representations of $R \cup F$.

(d) A subset C of L_0, to serve as a complexity class for representing the elements of $R \cup F$.

(e) A subset D of C of the form $\{\theta_s\colon\ s \in R \cup F\}$, which will be used for defining representations of the elements of $R \cup F$.

We say that D, *C-represents* (U,R,F,B) *in* T if there is a Gödel numbering $\alpha \mapsto \ulcorner\alpha\urcorner$ of U, i.e. an injection, $\ulcorner\ \urcorner$, from U into ω, such that

(i) For all $r \in R$, $\{(\ulcorner\vec{\alpha}\urcorner)\colon r\vec{\alpha}\} = \{(\vec{n})\colon \theta_r(\vec{n})\}$.

(ii) For all $f \in F$, $\{(\ulcorner\vec{\alpha}, \beta\urcorner)\colon f(\vec{\alpha}) = \beta\} = \{(\vec{n}, m)\colon \theta_f(\vec{n}, m)\}$.

(iii) For all $f \in F$, $T \vdash \forall \vec{x} \exists! y \theta_f(\vec{x}, y)$.

(iv) $T + D' \vdash B^U$, where B^U is the set of sentences obtained from the sentences in B by restricting all quantifiers to U, and

$$D' \overset{\mathrm{def}}{=} \{(\theta_r(\vec{x}) \leftrightarrow r\vec{x})\colon\ r \in R\} \cup \{(\theta_f(\vec{x}, y) \leftrightarrow f(\vec{x}) = y)\colon\ f \in F\}.$$

If for a given U, R, F, B, C, and T such a D exists we say (U,R,F,B) *is* (C,T)*-representable.*

[20] and [21] show that for an arbitrary recursive ordinal U, and sets, R and F, of common relations and functions on ordinals, and a set, B, that includes many basic properties of these relations and functions, (U,R,F,B) is $(\Delta_0, I\Delta_0)$-representable (this is also done in [19], but there all results pertain only up to the ordinal ε_0). The generality of [20, 21] ought to allow the extension of the main results of [19] to arbitrary recursive ordinals, but the author has not checked this in every detail (this is discussed further in Section 6.6). For the most part, in this paper, the restriction to ordinals less than or equal to ε_0 will remain in place.

We will state, but not prove, results showing that a significant amount of ordinal arithmetic can be carried out in $I\Delta_0$; the reader is referred to [19–21]. Alternatively, the reader who is familiar with carrying out basic ordinal arithmetic in *PRA*, and who is willing to settle for weaker results, may replace $I\Delta_0$ everywhere below with *PRA*.

3.2. Cantor normal form

In order to describe the Gödel numbering of ordinals as well to provide definitions of certain ordinal functions and relations we recall the Cantor normal form theorem. We also note that codings of the ordinals less than ε_0 are easily obtained using this theorem.

Proposition 3.2 (Cantor). *For α an ordinal such that $\alpha > 0$, α has a unique representation*:

$$\alpha = \omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k,$$

where $0 < k, n_1, \ldots, n_k \in \omega$, and $\alpha_1, \ldots, \alpha_k$ are ordinals such that $\alpha_1 > \cdots > \alpha_k$.

We refer the reader to [1] for a proof of this Proposition; this will also serve as a reference for some of the "well-known facts about ordinals" that we sight, without proof, below.

Definition 3.3. We call the representation, $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$, given in the above proposition, the *Cantor normal form*, or the *normal form* of α. We will let 0 be, by definition, the normal form of 0. So that statements such as "suppose α has normal form $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$ ..." make sense even if $\alpha = 0$, we use the convention that if $k = 0$ then "$\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$" denotes "0".

Definition 3.4. For α an ordinal and n a natural number, let ω_n^α be defined inductively by $\omega_0^\alpha \overset{\text{def}}{=} \alpha$, and $\omega_{n+1}^\alpha \overset{\text{def}}{=} \omega^{\omega_n^\alpha}$. Also let, $\omega_n \overset{\text{def}}{=} \omega_n^1$, and $\varepsilon_0 \overset{\text{def}}{=} \sup_{n\in\omega} \omega_n$.

We note that ε_0 is the least ordinal greater than ω closed under $+$ and the function $(\alpha, \beta) \mapsto \alpha^\beta$, and ε_0 is the least α such that $\alpha = \omega^\alpha$. That is, ε_0 is the first *epsilon number*.

We will follow [19] and use the following Gödel numbering of ordinals. If $\alpha < \varepsilon_0$ and the Cantor normal form of α is $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$, then the Smullyan base 4 representation of $\ulcorner\alpha\urcorner$ is given by

$$\overline{\ulcorner\alpha\urcorner}^4 = 4 * \overline{\ulcorner\alpha_1\urcorner}^4 * 3 * \overline{n_1}^2 * \cdots 4 * \overline{\ulcorner\alpha_k\urcorner}^4 * 3 * \overline{n_k}^2.$$

In [19], the mapping $\alpha \mapsto \ulcorner\alpha\urcorner$ on ε_0 is shown to be a bijection onto a Δ_0-definable subset of ω.

The choice of coding of the ordinals less than ε_0 described above allows for the handling of the Cantor normal form of ordinals in $I\Delta_0$. That is, there are Δ_0-definable, $I\Delta_0$-provably total, functions that, on argument $\ulcorner\alpha\urcorner$, give as values the codes of the terms of the Cantor normal form of α, etc. More precisely, there are Δ_0-definable, $I\Delta_0$-provably total, functions $\#t$, t, e, and c such that if the Cantor normal form of α is $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$, then for $1 \leqslant i \leqslant j \leqslant k$,

(i) $\#t(\ulcorner\alpha\urcorner) = k$,
(ii) $t(\ulcorner\alpha\urcorner, i, j) = \ulcorner\omega^{\alpha_i} \cdot n_i + \cdots + \omega^{\alpha_j} \cdot n_j\urcorner$,
(iii) $t(\ulcorner\alpha\urcorner, i) = t(\ulcorner\alpha\urcorner, i, i) = \ulcorner\omega^{\alpha_i} \cdot n_i\urcorner$,
(iv) $e(\ulcorner\alpha\urcorner, i) = \ulcorner\alpha_i\urcorner$, and
(v) $c(\ulcorner\alpha\urcorner, i) = n_i$.

$\#t$, t, e and c stand for "number of terms", "terms", "exponent", and "coefficient" respectively. Using these functions, statements of the form "suppose the Cantor normal form of α is ..." have natural translations into the language of arithmetic.

3.3. *Some ordinal functions and relations*

Below we define some ordinal functions and relations that will be used later on.

Definition 3.5. For $\alpha < \varepsilon_0$ with Cantor normal form $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$, the following functions are defined recursively,

(i) $h(0) = 0$, and
(ii) For $\alpha > 0$, $h(\alpha) = h(\alpha_1) + 1$.
(iii) $c_{\max}(0) = 0$, and
(iv) For $\alpha > 0$, $c_{\max}(\alpha) = \max(\{c(\alpha_i)\colon\ 1 \leqslant i \leqslant k\} \cup \{n_i\colon\ 1 \leqslant i \leqslant k\})$.

$h(\alpha)$ is called the *height* of α, and $c_{\max}(\alpha)$ is the *maximum coefficient* of α.

Definition 3.6. If the normal forms of α and β are $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$ and $\omega^{\beta_1} \cdot m_1 + \cdots + \omega^{\beta_l} \cdot m_l$, respectively, we say α *meshes with* β, and write $\alpha \gg \beta$, if(def) either $\alpha_k \geqslant \beta_1$, $\alpha = 0$, or $\beta = 0$.

The notation $\{\alpha\}(n)$, for α a limit ordinal and $n \in \omega$, is often used to denote the nth ordinal in a fundamental sequence for α (i.e., an increasing sequence cofinal in α). We note that the definitions vary slightly in the literature. The definition will be given using the fact, following from the Cantor normal form Theorem, that if $0 < \alpha < \varepsilon_0$, then there are unique $\beta, \gamma < \varepsilon_0$ and $0 < m \in \omega$ such that

$$\alpha = \beta + \omega^\gamma \cdot m \text{ and } \beta \gg \omega^{\gamma+1}. \tag{6}$$

We define the fundamental sequence function on all ordinals (rather than only on limit ordinals).

Definition 3.7. Suppose α, β, γ, and m satisfy (6). The nth *ordinal in the fundamental sequence of* α, denoted $\{\alpha\}(n)$, is given by the following:

$$\{\alpha\}(n) = \beta + \omega^\gamma \cdot (m-1) + \begin{cases} 0 & \text{if } \gamma = 0; \\ \omega^\delta \cdot (n+1) & \text{if } \gamma = \delta + 1; \\ \omega^{\{\gamma\}(n)} & \text{if } \gamma \text{ is a limit ordinal.} \end{cases}$$

In [19] it is shown that this fundamental sequence function (on codes of ordinals) has a Δ_0-definable graph, but its growth rate is super polynomial (consider the ordinal α given by

$$\alpha = \underbrace{(\omega^{(\omega^{(\omega^{\cdots(\omega^2 \cdot 2)} \cdot 2)} \cdot 2)} \cdot 2)}_{n \ \omega\text{'s}},$$

then $\{\alpha\}(0)$ is on the order of $\ulcorner\alpha\urcorner^{\log \ulcorner\alpha\urcorner}$), so the fundamental sequence function is not $I\Delta_0$-provably total. Also, in [19], the following upper bound is given,

$$\ulcorner\{\alpha\}(n)\urcorner \leqslant 4 \cdot (3\ulcorner\alpha\urcorner + 1)^{|\ulcorner\alpha\urcorner|_4} \cdot (n+2)^2. \tag{7}$$

Using this bound the fundamental sequence function can be shown to be $I\Delta_0 + \Omega_1$-provably total, where Ω_1 denotes the formula $\forall x \exists y (y = x^{|x|_2})$. In [21], fundamental sequences are given that are $I\Delta_0$-provably total and that have a number of nice properties provable in $I\Delta_0$; furthermore, this is done for arbitrary recursive ordinals. The author has not fully investigated whether such fundamental sequences could be used to obtain all of the results of this paper.

3.4. Notational conventions

We state the following notational conventions:

(i) Lower case Greek letters $\alpha, \beta, \gamma, \ldots$ will be used to range over codes of ordinals less than ε_0; so, for instance, $\varphi(\alpha)$ will abbreviate the formula $\theta_{\varepsilon_0}(x) \to \varphi(x)$, where x is a free variable not occurring in φ, and θ_{ε_0} is the formula used define the set of codes of ordinals less than ε_0. $\forall \alpha \varphi(\alpha)$ then abbreviates $\forall x(\theta_{\varepsilon_0}(x) \to \varphi(x))$, etc.

(ii) We will be using the symbols $<$, $+$, and $\cdot$ for two different relations or functions, i.e. the one on integers and the one on ordinals, we must take care to do this only if the meaning is clear from the context. For example, in "$\alpha < \beta$", "$<$" denotes the less-than relation on ordinals, whereas "$\ulcorner\alpha\urcorner < \ulcorner\beta\urcorner$" denotes "the Gödel number of α is less than the Gödel number of β".

(iii) We will use lower case Roman letters x, y, z, etc. to range over integers, but since every integer corresponds to a finite ordinal, we may wish to use these symbols to denote such ordinals; e.g. $\alpha + x$ will be used in place of $\alpha + \omega^0 \cdot x$. Occasionally this convention may cause confusion so we introduce the following functions (both Δ_0-definable and $I\Delta_0$-provably total):

$$\mathrm{ord}(x) \overset{\mathrm{def}}{=} \ulcorner\omega^0 \cdot x\urcorner; \text{ and}$$

$$\mathrm{num}(y) \overset{\mathrm{def}}{=} \begin{cases} z \text{ if } y = \ulcorner\omega^0 \cdot z\urcorner & \text{for some } z \in \omega; \\ 0 & \text{otherwise.} \end{cases}$$

We will only abide by such conventions if the meaning is absolutely clear from the context.

3.5. Basic ordinal arithmetic

Now we are ready to list the elements of the sets R, F, and B from Section 3.1. Since we are taking $T = I\Delta_0$, and neither multiplication, exponentiation, nor the fundamental sequence function are provably total in $I\Delta_0$, we do not list these functions in F (multiplication and the fundamental sequence function are $I\Delta_0 + \Omega_1$-provably total); however, all three have Δ_0-graphs, and since we are taking $C = \Delta_0$, we include the graphs of these functions in R (The functions $\alpha \mapsto \omega^\alpha$ and $\alpha \mapsto \alpha^\omega$ are $I\Delta_0$-provably total, and hence are included in F. The graph of $(\alpha, \beta) \mapsto \alpha^\beta$ is included in R, totality of this function is equivalent to totality of exponentiation of the natural numbers.) A number of properties of these functions are provable in the case that the functions converge; for example, if $(\alpha \cdot \beta) \cdot \gamma$ and $\alpha \cdot (\beta \cdot \gamma)$ both exist then they are equal. In order to list such facts in B, while keeping the statement of these facts as brief as possible, we use suggestive notation such as,

$$(\alpha \cdot \beta) \cdot \gamma \simeq \alpha \cdot (\beta \cdot \gamma)$$

that imply the convergence condition. In particular, $\simeq$, $\prec$, and $\preccurlyeq$ are used to state that if everything on either side of the symbol exists then the relation ($=$, $<$, or $\leqslant$) holds. (*Note:* this definition is different than the usual one from recursion theory).

The notation $\sup_{\alpha<\beta} \tau(\alpha) = \gamma$ is used below, where τ is a term involving ordinal functions, to denote:

$$(\forall \alpha < \beta)(\tau(\alpha) \leqslant \gamma) \wedge (\forall \delta < \gamma)(\exists \alpha < \beta)(\delta \leqslant \tau(\alpha)).$$

The elements of R include:

$$<, \leqslant, \mathrm{SUC}, \mathrm{LIM}, \gg, \{\langle \alpha, \beta, \alpha \cdot \beta \rangle : \alpha, \beta < \varepsilon_0\},$$
$$\{\langle \alpha, \beta, \alpha^\beta \rangle : \alpha < \varepsilon_0, n \geqslant 0\}, \{\langle \alpha, \{\alpha\}(n) \rangle : \alpha < \varepsilon_0, n \geqslant 0\}\}.$$

The elements of F include:

$$0,\ +,\ \alpha \mapsto \omega^\alpha,\ \alpha \mapsto \alpha^\omega,\ \text{num},\ \text{ord},\ h,\ c_{\max},\ \#t,\ t,\ c,\ e.$$

The elements of B include:
Definitions of SUC and LIM:
O_1. $\mathrm{SUC}(\alpha) \leftrightarrow \exists\beta(\alpha = \beta + 1)$,
O_2. $\mathrm{LIM}(\alpha) \leftrightarrow \alpha \neq 0 \wedge \neg\mathrm{SUC}(\alpha)$.

(Strong) defining formula for $<$:
O_3. $\alpha < \beta \leftrightarrow \exists!\gamma(\gamma \neq 0 \wedge \alpha + \gamma = \beta)$.

Defining formulas for addition:
O_4. $\alpha + 0 = \alpha$,
O_5. $\alpha + (\beta + 1) = (\alpha + \beta) + 1$,
O_6. $\mathrm{LIM}(\lambda) \rightarrow (\alpha + \lambda = \sup_{\delta<\lambda}(\alpha + \delta))$.

Defining formulas for multiplication:
O_7. $\alpha \cdot 0 = 0$,
O_8. $\alpha \cdot (\beta + 1) \simeq (\alpha \cdot \beta) + \alpha$,
O_9. $\mathrm{LIM}(\lambda) \rightarrow (\alpha \cdot \lambda \simeq \sup_{\delta<\lambda}(\alpha \cdot \delta))$.

Defining formulas for exponentiation:
O_{10}. $\alpha^0 = 1$,
O_{11}. $\alpha^{\beta+1} \simeq \alpha^\beta \cdot \alpha$,
O_{12}. $\mathrm{LIM}(\lambda) \rightarrow (\alpha^\lambda \simeq \sup_{\delta<\lambda} \alpha^\delta)$.

Defining equations of the fundamental sequence function:
O_{13}. $\{0\}(n) = 0$,
O_{14}. $\{\beta + 1\}(n) = \beta$,
O_{15}. $\beta \gg \omega^{\gamma+1} \rightarrow \{\beta + \omega^{\gamma+1}\}(n) \simeq \beta + \omega^\gamma \cdot (n+1)$,
O_{16}. $\mathrm{LIM}(\lambda) \wedge \beta \gg \omega^\lambda \rightarrow \{\beta + \omega^\lambda\}(n) \simeq \beta + \omega^{\{\lambda\}(n)}$.

Trichotomy:
O_{17}. Exactly one of the following holds: $\alpha = 0$, $\mathrm{SUC}(\alpha)$, or $\mathrm{LIM}(\alpha)$.

$<$ is a linear order with least element and successor function:
O_{18}. Exactly one of the following holds: $\alpha < \beta$, $\beta < \alpha$, or $\alpha = \beta$.
O_{19}. $\alpha < \beta \wedge \beta < \gamma \rightarrow \alpha < \gamma$,
O_{20}. $0 \leqslant \alpha$,
O_{21}. $\alpha < \alpha + 1 \wedge (\beta \leqslant \alpha \vee \alpha + 1 \leqslant \beta)$.

Monotonicity of ordinal functions:
O_{22}. $\alpha < \beta \rightarrow (\gamma + \alpha < \gamma + \beta) \wedge (\alpha + \gamma \leqslant \beta + \gamma)$,
O_{23}. $\alpha < \beta \wedge \gamma \neq 0 \rightarrow (\gamma \cdot \alpha \prec \gamma \cdot \beta) \wedge (\alpha \cdot \gamma \preccurlyeq \beta \cdot \gamma)$,
O_{24}. $\alpha < \beta \wedge \gamma \neq 0 \rightarrow (\gamma^\alpha \prec \gamma^\beta) \wedge (\alpha^\gamma \preccurlyeq \beta^\gamma)$.

Associative laws:

O_{25}. $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$,

O_{26}. $(\alpha\cdot\beta)\cdot\gamma\simeq\alpha\cdot(\beta\cdot\gamma)$.

Left distributive law:

O_{27}. $\alpha\cdot(\beta+\gamma)\simeq\alpha\cdot\beta+\alpha\cdot\gamma$.

Properties of exponentiation:

O_{28}. $\gamma^{\alpha}\cdot\gamma^{\beta}\simeq\gamma^{\alpha+\beta}$,

O_{29}. $\alpha<\beta\rightarrow\omega^{\alpha}+\omega^{\beta}=\omega^{\beta}$,

O_{30}. $(\gamma^{\alpha})^{\beta}\simeq\gamma^{\alpha\cdot\beta}$.

Properties of fundamental sequences:

O_{31}. $\mathrm{LIM}(\lambda)\rightarrow\sup_{n\in\omega}\{\lambda\}(n)\simeq\lambda$,

O_{32}. $\mathrm{LIM}(\lambda)\wedge m<n\rightarrow\{\lambda\}(m)\prec\{\lambda\}(n)$.

Properties of ord and num:

O_{33}. $\alpha<\omega\rightarrow\exists x(\alpha=\mathrm{ord}(x))$,

O_{34}. $(\exists\alpha<\omega)(x=\mathrm{num}(\alpha))$,

O_{35}. $x<y\leftrightarrow\mathrm{ord}(x)<\mathrm{ord}(y)$,

O_{36}. $\alpha,\beta<\omega\rightarrow(\alpha<\beta\leftrightarrow\mathrm{num}(\alpha)<\mathrm{num}(\beta))$.

Properties of the Cantor normal form: Suppose the normal form of α is $\omega^{\alpha_1}\cdot n_1+\cdots+\omega^{\alpha_k}\cdot n_k$, then

O_{37}. $k=0\;\Leftrightarrow\;\alpha=0$,

O_{38}. $(\forall i<k)(n_{i+1}>0)$,

O_{39}. $\forall i,j(1\leqslant i<j\leqslant k\rightarrow\alpha_i<\alpha_j)$,

O_{40}. $(\forall i<k)(\alpha_{i+1}<\alpha)$.

O_{41}. Definition of less-than in terms of the Cantor normal form.

O_{42}. Definition of addition in terms of the Cantor normal form.

O_{43}. Definition of multiplication in terms of the Cantor normal form.

O_{44}. Definition of exponentiation in terms of the Cantor normal form.

O_{45}. Definition of the fundamental sequence function in terms of the Cantor normal form.

Precise statements for O_{41}–O_{45} are omitted, since the corresponding formulas are cumbersome – they are motivated by considering several cases.

4. Transfinite induction

For $\alpha<\varepsilon_0$, let $TI(\alpha,\varphi)$, *transfinite induction up to* α *for* φ, be the universal closure of:

$$\forall\beta((\forall\gamma<\beta)\varphi(\gamma)\rightarrow\varphi(\beta))\rightarrow(\forall\gamma<\alpha)\varphi(\gamma).$$

We abbreviate "$\forall\beta((\forall\gamma < \beta)\varphi(\gamma) \to \varphi(\beta))$" by $\mathrm{Prog}(\varphi,\gamma)$, and say φ *is progresive with respect to* γ; if the free variable γ of φ is understood we may simply write $\mathrm{Prog}(\varphi)$. Then $TI(\alpha,\varphi)$ is (the universal closure of) $\mathrm{Prog}(\varphi) \to (\forall\gamma < \alpha)\varphi(\gamma)$. Let $TI(\alpha,\mathscr{C})$, *transfinite induction up to* α *for* $\mathscr{C}$*-formulas*, be the theory $I\Delta_0 + \{TI(\alpha,\varphi)\colon \varphi \in \mathscr{C}\}$. We let $TI(<\alpha,\mathscr{C})$ denote $\bigcup_{\beta<\alpha} TI(\beta,\mathscr{C})$. We will be concerned with the relative strengths of the theories $TI(\alpha,\mathscr{C})$ and $TI(<\alpha,\mathscr{C})$ for all values of $\alpha < \varepsilon_0$, and with $\mathscr{C} = \Pi_n$, Σ_n, or Δ_0 where $n \geqslant 1$. The relative strengths can be summed up as follows (*note:* since $TI(\alpha,\mathscr{C})$ and $TI(<\alpha,\mathscr{C})$ always include $I\Delta_0$ as a base theory, we are comparing the relative strengths of the axioms of transfinite induction over $I\Delta_0$):

Theorem 4.1. *$TI(\alpha,\Pi_n)$ vs. $TI(\beta,\Pi_m)$.*

(a) *For $0 < m \leqslant n$ and $\omega \leqslant \alpha < \varepsilon_0$:*

$$TI(\alpha,\Pi_n) \vdash TI(\beta,\Pi_m) \Leftrightarrow \beta < \omega_{n-m}^{\alpha^{\omega}}.$$

(b) *For $0 < n < m$ and $0 \leqslant \alpha < \varepsilon_0$:*

$$TI(\alpha,\Pi_n) \vdash TI(\beta,\Pi_m) \Leftrightarrow \beta < \omega.$$

(c) *For $m,n \in \omega$ and $\alpha,\beta < \omega$:*

$$TI(\alpha,\Pi_n) \equiv TI(\beta,\Pi_m) \equiv I\Delta_0.$$

$TI(\alpha,\Pi_n)$ vs. $TI(\beta,\Sigma_m)$:

(d) *For $\alpha \geqslant \omega$ and $n \in \omega$:*

$$TI(\omega\cdot\alpha,\Sigma_n) \equiv TI(\alpha,\Pi_{n+1}).$$

(e) *For $0 < \alpha < \omega$ and $n \in \omega$:*

$$TI(\omega\cdot\alpha,\Sigma_n) \equiv TI(\omega,\Pi_n).$$

(f) *For $\alpha,\beta < \omega$ and $n,m \in \omega$:*

$$TI(\alpha,\Sigma_n) \equiv TI(\beta,\Pi_m) \equiv I\Delta_0.$$

$TI(\beta,\Pi_n)$ vs. $TI(<\alpha,\Pi_n)$:

(g) *For $\omega \leqslant \alpha < \varepsilon_0$ and $n \geqslant 1$:*

$$TI(<\alpha,\Pi_n) \vdash TI(\beta,\Pi_n) \Leftrightarrow \begin{cases} \beta < \alpha & \text{if } \alpha = \omega^{\omega^{\gamma}} \text{ for some } \gamma. \\ \beta < \alpha^{\omega} & \text{otherwise.} \end{cases}$$

(h) *For $\alpha,\beta < \omega$ and $n,m \in \omega$:*

$$TI(<\alpha,\Pi_n) \equiv TI(<\beta,\Pi_n) \equiv I\Delta_0.$$

Parts (a)–(c) of the theorem give the relative strengths of theories of the form $TI(\alpha,\Pi_n)$; parts (d)–(f) show that each theory of the form $TI(\beta,\Sigma_m)$ is equivalent to one of the form $TI(\alpha,\Pi_n)$, and parts (g) and (h) allow us to evaluate the strength of a theory of the form $TI(<\beta,\Pi_m)$, by considering theories of the form $TI(\alpha,\Pi_n)$. Thus Theorem 4.1 gives a complete picture of the relative strengths of $TI(\alpha,\mathscr{C})$ and $TI(<\alpha,\mathscr{C})$ for $0 \leqslant \alpha < \varepsilon_0$, and $\mathscr{C} = \Sigma_n$, Π_n, or Δ_0 where $n \geqslant 1$. The reader may find the discussion in Section 4.3 useful for clarifying this further. We note that (a), (b) and (g) were also proven by Ulf Schmerl in [17].

This theorem will be proven, for the most part, in this section, although some parts will rely on results not included in this paper and results given in the following section. In particular, we will give complete proofs of Theorem 4.1(c) – 4.1(h) and the $\Leftarrow$-directions of Theorem 4.1(a) and 4.1(b) in this section. The $\Rightarrow$-direction of Theorem 4.1(a) will be proven by classifying the provably total recursive functions of these theories; we do this in Section 5, where the proofs will tend to be more sophisticated and generally more interesting than the proofs in this section. The $\Rightarrow$-direction of Theorem 4.1(b) will rely on well-known results not proven in this paper.

4.1. Basic lemmas

Lemma 4.2. *For $\mathscr{C} = \Sigma_n$, Π_n, or Δ_0:*
(i) $\alpha > \beta \Rightarrow TI(\alpha,\mathscr{C}) \vdash TI(\beta,\mathscr{C})$.
(ii) $I\mathscr{C} \equiv TI(\omega,\mathscr{C})$.
(iii) *For all* $\alpha < \omega$: $I\Delta_0 \equiv TI(\alpha,\mathscr{C})$.

Proof. (i) is clear using the fact that $I\Delta_0$, and hence $TI(\alpha,\mathscr{C})$, proves the transitivity of ordinal $<$ (cf. O_{19} of Section 3). For the $\vdash$-direction of ii, note that $\varphi(\alpha) \in \mathscr{C}$ implies $\varphi(\mathrm{ord}(x)) \in \mathscr{C}$; also by O_{19}, O_{33}, and O_{35}, we have

$$I\Delta_0 \vdash \mathrm{Prog}(\varphi) \to \forall y((\forall x < y)\varphi(\mathrm{ord}(x)) \to \varphi(\mathrm{ord}(y))).$$

Hence $I\mathscr{C} \vdash \mathrm{Prog}(\varphi) \to \forall y \varphi(\mathrm{ord}(y))$ which, using O_{33} again, gives

$$I\mathscr{C} \vdash \mathrm{Prog}(\varphi) \to (\forall \alpha < \omega)\varphi(\alpha),$$

which is what we needed. The $\dashv$-direction of (ii) is similar except we use O_{34}, O_{36}, and the fact $\varphi(x) \in \mathscr{C}$ implies $(\varphi(\mathrm{num}(\alpha)) \vee \alpha \geqslant \omega) \in \mathscr{C}$.

The $\dashv$-direction of (iii) is immediate since, by definition, $TI(\alpha,\mathscr{C})$ always includes $I\Delta_0$. The proof of the $\vdash$-direction of (iii) uses meta-induction on α. Clearly, $I\Delta_0 \vdash TI(0,\varphi)$ for any φ. Now suppose $I\Delta_0 \vdash TI(c,\varphi)$, i.e. $I\Delta_0 \vdash \mathrm{Prog}(\varphi) \to (\forall\gamma < c)\varphi(\gamma)$. Since $I\Delta_0 \vdash \mathrm{Prog}(\varphi) \wedge (\forall\gamma < c)\varphi(\gamma) \to \varphi(c)$, we have, using O_{21}, $I\Delta_0 \vdash \mathrm{Prog}(\varphi(\gamma)) \to (\forall\gamma < c+1)\varphi(\gamma)$, i.e. $I\Delta_0 \vdash TI(c+1,\varphi)$. By induction on c we have, for all $c \in \omega$ and all formulas φ, $I\Delta_0 \vdash TI(c,\varphi)$. $\square$

Lemma 4.3. *For all $n, c \in \omega$, where $n \geqslant 1$: $TI(\alpha,\Pi_n) \vdash TI(\alpha^c,\Pi_n)$.*

Proof. If $\alpha < \omega$ then $\alpha^c < \omega$, so using (iii) from the above lemma we have the result. Hence we may assume that $\alpha \geqslant \omega$. So, by (i) and (ii) of the previous lemma, $TI(\alpha,\Pi_n) \vdash I\Sigma_1$; thus ordinal multiplication is a total function and all of the properties of ordinal arithmetic listed in Section 3 hold. If the normal form of α is $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$ then using $\alpha_1 > 0$, O_{41}, and O_{44}, as well as other facts about ordinals provable in $I\Delta_0$, we have

$$\omega^{\alpha_1 \cdot c} \leqslant \alpha^c \leqslant \omega^{\alpha_1 \cdot (c+1)},$$

for any $c \in \omega$. So by Lemma 4.2(i), we need only consider α of the form ω^β; since we are assuming $\alpha \geqslant \omega$, we may assume $\beta > 0$. We will use meta-induction on c to prove the lemma. The $c = 0$ case is clear, so we need, noting that $(\omega^\beta)^c = \omega^{\beta \cdot c}$,

$$TI(\omega^{\beta \cdot c}, \Pi_n) \vdash TI(\omega^{\beta \cdot (c+1)}, \Pi_n).$$

Suppose $\varphi(\gamma) \in \Pi_n$, and let

$$\psi(\delta) \stackrel{\text{def}}{=} (\forall \gamma < \omega^{\beta \cdot c} \cdot \delta)\varphi(\gamma).$$

Note that $\psi \in \Pi_n$. Working in $TI(\omega^{\beta \cdot c}, \Pi_n)$ we will show $\mathrm{Prog}(\varphi, \gamma) \to \mathrm{Prog}(\psi, \delta)$. Suppose $\mathrm{Prog}(\varphi, \gamma)$ and suppose $(\forall \delta' < \delta)\psi(\delta')$, we will show $\psi(\delta)$ by cases.

Case 1: $\psi(\delta)$ is clear for $\delta = 0$.

Case 2: If $\delta = \delta_0 + 1$ then, from $(\forall \delta' < \delta)\psi(\delta')$ and O_{21}, we have, $(\forall \delta' \leqslant \delta_0)\psi(\delta')$, so using the definition of ψ,

$$(\forall \gamma < \omega^{\beta \cdot c} \cdot \delta_0)\varphi(\gamma). \tag{8}$$

Using (8) and $\mathrm{Prog}(\varphi, \gamma)$ we have

$$\forall \alpha((\forall \alpha' < \alpha)\varphi(\omega^{\beta \cdot c} \cdot \delta_0 + \alpha') \to \varphi(\omega^{\beta \cdot c} \cdot \delta_0 + \alpha)).$$

That is, $\mathrm{Prog}(\varphi(\omega^{\beta \cdot c} \cdot \delta_0 + \alpha'), \alpha')$, so we have, by $TI(\omega^{\beta \cdot c}, \Pi_n)$,

$$(\forall \alpha < \omega^{\beta \cdot c})\varphi(\omega^{\beta \cdot c} \cdot \delta_0 + \alpha). \tag{9}$$

(8) and (9) yield,

$$(\forall \gamma < \omega^{\beta \cdot c} \cdot \delta_0 + \omega^{\beta \cdot c})\varphi(\gamma),$$

So by O_8, $(\forall \gamma < \omega^{\beta \cdot c} \cdot \delta)\varphi(\gamma)$, and hence $\psi(\delta)$.

Case 3: If δ is a limit ordinal then $(\forall \gamma < \omega^{\beta \cdot c} \cdot \delta)\varphi(\gamma)$ follows from O_9 and

$$(\forall \delta' < \delta)(\forall \gamma < \omega^{\beta \cdot c} \cdot \delta')\varphi(\gamma).$$

The three cases give us $\mathrm{Prog}(\varphi) \to \mathrm{Prog}(\psi)$. Using $TI(\omega^{\beta \cdot c}, \psi)$, from $\mathrm{Prog}(\varphi)$, we conclude $(\forall \delta < \omega^\beta)\psi(\delta)$, i.e., $(\forall \delta < \omega^\beta)(\forall \gamma < \omega^{\beta \cdot c} \cdot \delta)\varphi(\gamma)$. So by O_9, $(\forall \gamma < \omega^{\beta \cdot c} \cdot \omega^\beta)\varphi(\gamma)$, and hence by O_8 and O_{28}, $(\forall \gamma < \omega^{\beta \cdot (c+1)})\varphi(\gamma)$. By induction on c we have the lemma. □

We note that the above argument cannot be carried out if we replace c by a free variable x; for this we would, at one point in the proof, assume $TI(\omega^{\beta \cdot x}, \Pi_n)$. With

x free this axiom schema implies $TI(\omega^{\beta\cdot\omega}, \Pi_n)$ which, as will be shown in the next section, cannot be obtained from $TI(\omega^\beta, \Pi_n)$. Next we present a result of Gentzen.

Lemma 4.4. *For all $n \geqslant 1$ and $\alpha < \varepsilon_0$,*

$$I\Sigma_n + TI(\alpha, \Pi_{n+1}) \vdash TI(\omega^\alpha, \Pi_n).$$

Proof. Suppose $\varphi(\delta) \in \Pi_n$ and let

$$\psi(\beta) \stackrel{\text{def}}{=} \forall\gamma((\forall\delta < \gamma)\varphi(\delta) \to (\forall\delta < \gamma + \omega^\beta)\varphi(\delta)).$$

Note $\psi \in \Pi_{n+1}$. We will show

Claim. $I\Sigma_n \vdash \text{Prog}(\varphi) \to \text{Prog}(\psi)$.

To see that the claim implies the lemma note that from the claim we can conclude that

$$I\Sigma_n + TI(\alpha, \Pi_{n+1}) \vdash \text{Prog}(\varphi) \to (\forall\beta < \alpha)\psi(\beta).$$

Hence,

$$I\Sigma_n + TI(\alpha, \Pi_{n+1}) \vdash \text{Prog}(\varphi) \to \psi(\alpha).$$

Setting $\gamma = 0$ in ψ, we get

$$I\Sigma_n + TI(\alpha, \Pi_{n+1}) \vdash \text{Prog}(\varphi) \to (\forall\delta < \omega^\alpha)\varphi(\delta),$$

which is what we needed for the lemma; hence we need only prove the claim.

We will work in $I\Sigma_n$. Assume $\text{Prog}(\varphi)$ and $(\forall\beta' < \beta)\psi(\beta')$; we want to conclude $\psi(\beta)$. By cases we have

Case 1: $\beta = 0$. $\psi(0) = \forall\gamma((\forall\delta < \gamma)\varphi(\delta) \to (\forall\delta < \gamma + 1)\varphi(\delta))$, which is immediate from $\text{Prog}(\varphi)$ and O_{21}.

Case 2: $\beta = \beta' + 1$. In $\psi(\beta') = \forall\gamma((\forall\delta < \gamma)\varphi(\delta) \to (\forall\delta < \gamma + \omega^{\beta'})\varphi(\delta))$, we may replace γ with $\gamma' + \omega^{\beta'} \cdot x$, for any γ', to get

$$\forall\gamma'((\forall\delta < (\gamma' + \omega^{\beta'} \cdot x))\varphi(\delta) \to (\forall\delta < (\gamma' + \omega^{\beta'} \cdot x + \omega^{\beta'}))\varphi(\delta)).$$

That is, using O_8, $\forall\gamma'(\theta(x) \to \theta(x+1))$, where

$$\theta(x) \stackrel{\text{def}}{=} (\forall\delta < \gamma' + \omega^{\beta'} \cdot x)\varphi(\delta).$$

Noting that $\theta \in \Pi_n$ and that $I\Sigma_n \equiv I\Pi_n$, we may use $I\Sigma_n$ and properties of the function ord, to get $\forall\gamma'(\theta(0) \to \forall x\theta(x))$, i.e., using O_4, O_7, and O_8,

$$\forall\gamma'((\forall\delta < \gamma')\varphi(\delta) \to \forall x(\forall\delta < \gamma' + \omega^{\beta'} \cdot x)\varphi(\delta)).$$

And hence, using O_6 and O_9,

$$\forall\gamma'((\forall\delta < \gamma')\varphi(\delta) \to (\forall\delta < \gamma' + \omega^{\beta'} \cdot \omega)\varphi(\delta)),$$

i.e., by O_{11},

$$\forall\gamma'((\forall\delta < \gamma')\varphi(\delta) \to (\forall\delta < \gamma' + \omega^{\beta'+1})\varphi(\delta)),$$

which is $\psi(\beta)$.

Case 3: β is a limit ordinal. By O_{12}, we have $(\forall\beta' < \beta)\psi(\beta') \to \psi(\beta)$, whenever β is a limit ordinal. This completes the proof of Case 3, and hence the proof of the claim, and hence the proof of the lemma.

The converse of Lemma 4.4 is false, i.e. by the $\Rightarrow$-direction of 4.1(b) and Lemma 4.2, for any α such that $\omega \leqslant \alpha < \varepsilon_0$,

$$I\Sigma_n + TI(\omega^\alpha, \Pi_n) \nvdash TI(\alpha, \Pi_{n+1}).$$

We continue by proving another lemma.

Lemma 4.5. *For all* $n \in \omega$,

$$TI(\omega \cdot \alpha, \Sigma_n) \vdash TI(\alpha, \Pi_{n+1}).$$

Note that since the $n = 0$ *case is included, we have* $TI(\omega \cdot \alpha, \Delta_0) \vdash TI(\alpha, \Pi_1)$.

Proof. Using Lemma 4.2, this lemma is clear for $\alpha < \omega$; so suppose $\alpha \geqslant \omega$. Suppose $\varphi(\gamma) = \forall x \psi(x, \gamma)$ and $\psi \in \Sigma_n$. The key idea of the proof is that for each $\gamma < \alpha$ and each $x \in \omega$ we can code γ and x by the ordinal $\omega \cdot \gamma + x$ which is less than $\omega \cdot \alpha$. We will work in $TI(\omega \cdot \alpha, \Sigma_n)$ to obtain $TI(\alpha, \varphi)$. We may obtain, in $I\Delta_0$, that each ordinal $\gamma < \varepsilon_0$ can be expressed in the form

$$\gamma = \gamma_0 + \omega^{m_k} \cdot n_k + \cdots + \omega^{m_1} \cdot n_1 + n_0, \tag{10}$$

where $\gamma_0 \gg \omega^\omega$ and $\omega > m_k > \cdots > m_1 > 0$, and $n_1, \ldots, n_k \neq 0$ (we do allow $n_0 = 0$). By "$k = 0$" we mean $\omega^{m_k} \cdot n_k + \cdots + \omega^{m_1} \cdot n_1$ is not included. We define $g_0 : \varepsilon_0 \to \omega$ and $g_1 : \varepsilon_0 \to \varepsilon_0$ by

$$g_0(\gamma) \stackrel{\text{def}}{=} n_0$$

and

$$g_1(\gamma) \stackrel{\text{def}}{=} \gamma_0 + \omega^{m_k - 1} \cdot n_k + \cdots + \omega^{m_1 - 1} \cdot n_1,$$

where γ has the form described in (10). Using the representability of Cantor normal form functions, it is clear that both g_0 and g_1 are Δ_0-definable and $I\Delta_0$-provably total. Next we show,

Claim 1. $I\Delta_0 \vdash \forall\gamma\forall x(g_0(\omega \cdot \gamma + x) = x \wedge g_1(\omega \cdot \gamma + x) = \gamma)$.

For the proof of Claim 1, note that if the normal form of δ is $\omega^{\delta_1} \cdot d_1 + \cdots + \omega^{\delta_l} \cdot d_l$ then by the definition of $\alpha \mapsto \omega \cdot \alpha$, and the fact this function is $I\Delta_0$-provably total,

$$\omega \cdot \delta = \omega^{1+\delta_1} \cdot d_1 + \cdots + \omega^{1+\delta_l} \cdot d_l.$$

For $\delta_i \geqslant \omega$, by O_{29}, we have $1 + \delta_i = \delta_i$; hence

$$\omega \cdot \gamma + x = \gamma_0 + \omega^{1+m_k} \cdot n_k + \cdots + \omega^{1+m_1} \cdot n_1 + \omega \cdot n_0 + x.$$

This clearly gives $g_0(\omega \cdot \gamma + x) = \gamma$ and $g_1(\omega \cdot \gamma + x) = x$. Carrying the proof out in $I\Delta_0$ is routine using basic ordinal arithmetic provable in $I\Delta_0$. We now make a second claim.

Claim 2. $I\Delta_0 \vdash \alpha < \beta \rightarrow \forall x(\omega \cdot \alpha + x < \omega \cdot \beta)$.

Suppose $\alpha < \beta$. If $\beta = \gamma + 1$ then, by O_{21}, $\alpha \leqslant \gamma$ so, so by O_{23}, $\omega \cdot \alpha \leqslant \omega \cdot \gamma$, and hence, by O_{22}, $\omega \cdot \alpha + x < \omega \cdot \alpha + \omega \leqslant \omega \cdot \gamma + \omega = \omega \cdot (\gamma + 1) = \omega \cdot \beta$, the next to last equality following from O_8. If β is a limit ordinal then, by O_{17} and O_{21}, $\alpha + 1 < \beta$ so, by O_{23}, $\omega \cdot (\alpha + 1) \leqslant \omega \cdot \beta$; hence, by O_{22} and O_8, $\omega \cdot \alpha + x < \omega \cdot \beta$. This completes the proof of Claim 2.

Let $\theta(\gamma) \stackrel{\text{def}}{=} \psi(g_0(\gamma), g_1(\gamma))$. Note $\theta \in \Sigma_n$. We make the following claim.

Claim 3. $I\Delta_0 \vdash \mathrm{Prog}(\varphi) \rightarrow \mathrm{Prog}(\theta)$.

For the proof of this claim suppose $\mathrm{Prog}(\varphi)$ and $(\forall \gamma < \beta)\theta(\gamma)$, we will derive, in $I\Delta_0$, $\theta(\beta)$. By Claim 2, if $\gamma < g_1(\beta)$ then $\omega \cdot \gamma + x < \omega \cdot g_1(\beta)$ for all x. Noting $\omega \cdot g_1(\beta) \leqslant \beta$, we have, by Claim 2, $\omega \cdot \gamma + x < \beta$, so

$$(\forall \gamma < g_1(\beta))\forall x \theta(\omega \cdot \gamma + x).$$

Hence, by the definition of θ,

$$(\forall \gamma < g_1(\beta))\forall x \psi(g_0(\omega \cdot \gamma + x), g_1(\omega \cdot \gamma + x)).$$

By Claim 1, $(\forall \gamma < g_1(\beta))\forall x \psi(x, \gamma)$, i.e., $(\forall \gamma < g_1(\beta))\varphi(\gamma)$. Using $\mathrm{Prog}(\varphi)$, we get $\varphi(g_1(\beta))$, i.e. $\forall x \psi(x, g_1(\beta))$, and hence $\psi(g_0(\beta), g_1(\beta))$. This is just $\theta(\beta)$, which is what we needed to complete the proof of the claim.

To complete the proof of Lemma 4.5, we will work in $TI(\omega \cdot \alpha, \Sigma_n)$. We assume $\mathrm{Prog}(\varphi)$ and we will conclude that $(\forall \gamma < \alpha)\varphi(\alpha)$; this will give us $TI(\alpha, \varphi)$ and hence the lemma. We use Claim 3 and $TI(\omega \cdot \alpha, \theta)$ to get $(\forall \gamma < \omega \cdot \alpha)\theta(\gamma)$. By Claim 2 and O_{23}, we have $(\forall \gamma < \alpha)\forall x \theta(\omega \cdot \gamma + x)$. Thus by Claim 1 and the definition of θ, we have $(\forall \gamma < \alpha)\forall x \psi(x, \gamma)$, and so $(\forall \gamma < \alpha)\varphi(\gamma)$. □

4.2. The Proof of Theorem 4.1

The $\Leftarrow$-direction of 4.1(a) can be stated as: for $n \geqslant m > 0$ and $\omega \leqslant \alpha < \varepsilon_0$,

$$TI(\alpha, \Pi_n) \vdash TI(\omega^{\beta}_{n-m}, \Pi_m) \quad \text{all } \beta < \alpha^{\omega}.$$

Suppose $\beta < \alpha^{\omega}$ then, using the fact $\alpha \geqslant \omega$, one has that $\beta < \alpha^{c}$ for some $c \in \omega$. By Lemma 4.3 and Lemma 4.2(i), we have

$$TI(\alpha, \Pi_n) \vdash TI(\beta, \Pi_n).$$

By Lemma 4.2(i), (ii) and the fact $I\Sigma_n \equiv I\Pi_n$ for all $n \in \omega$, we have $TI(\alpha, \Pi_n) \vdash I\Sigma_n$; since $n \geqslant 1$, we may apply Lemma 4.4, $n - m$ times, to get

$$TI(\alpha, \Pi_n) \vdash TI(\omega^{\beta}_{n-m}, \Pi_m).$$

Hence we have the $\Leftarrow$-direction of Theorem 4.1(a).

Theorem 4.1(c), (f), (h) and the $\Leftarrow$-direction of 4.1(b) are immediate from Lemma 4.2(iii). The $\vdash$-direction of Theorem 4.1(d) follows directly from Lemma 4.5. For the $\dashv$-direction first use Lemma 4.2(i), Lemma 4.3 (with $c = 2$), and the fact $\alpha \geqslant \omega$, to get $TI(\alpha, \Pi_{n+1}) \vdash TI(\omega \cdot \alpha, \Pi_{n+1})$. Then noting $\Sigma_n \subseteq \Pi_{n+1}$ we are done.

For Theorem 4.1(e) note that $I\Sigma_n \equiv I\Pi_n$ all $n \in \omega$, so it is enough to show that for all $c \in \omega - \{0\}$,

$$TI(\omega \cdot c, \Sigma_n) \equiv TI(\omega, \Sigma_n).$$

The $\vdash$-direction follows from Lemma 4.2(i). For the $\dashv$-direction we use meta-induction on c. The $c = 1$ case is immediate. Working in $TI(\omega \cdot c, \Sigma_n)$, suppose $\varphi \in \Sigma_n$ and Prog(φ); we wish to derive $(\forall \gamma < \omega \cdot (c+1))\varphi(\gamma)$. Using $TI(\omega \cdot c, \varphi)$ we have that Prog(φ) implies $(\forall \gamma < \omega \cdot c)\varphi(\gamma)$, so if we let $\psi(\alpha) \stackrel{\text{def}}{=} \varphi(\omega \cdot c + \alpha)$, it will be enough to show that Prog(φ) implies $(\forall \alpha < \omega)\psi(\alpha)$. Note $\psi \in \Sigma_n$. Clearly Prog(φ) together with $(\forall \gamma < \omega \cdot c)\varphi(\gamma)$ implies Prog(ψ), so Prog(φ) alone implies Prog(ψ). By $TI(\omega, \psi)$, we then have Prog(φ) $\rightarrow (\forall \gamma < \omega)\psi(\gamma)$, which is all we needed to complete the proof of Theorem 4.1(e).

To prove Theorem 4.1(g) we first argue that, for $n \geqslant 1$ and $\alpha \geqslant \omega$, $TI(\alpha, \Pi_n)$ can be axiomatized by a single sentence; first note in this case $TI(\alpha, \Pi_n) \vdash I\Pi_n \vdash I\Delta_0$ and, for $n \geqslant 1$, $I\Pi_n$ is finitely axiomatizable. Also using a Π_n-truth definition for Π_n-formulas, the axioms of transfinite induction can be expressed by a single sentence. Next we claim, that for $n > 0$ and $\alpha \geqslant \omega$,

$$TI(<\alpha, \Pi_n) \vdash TI(\beta, \Pi_n) \Leftrightarrow \beta < \sup_{\delta < \alpha} \delta^{\omega}. \tag{11}$$

For the $\Rightarrow$-direction, we will show $TI(<\alpha, \Pi_n) \nvdash TI(\sup_{\delta<\alpha} \delta^{\omega}, \Pi_n)$. The result then follows from Lemma 4.2(i). Now since $TI(\sup_{\delta<\alpha} \delta^{\omega}, \Pi_n)$ is finitely axiomatizable, using the definition of $TI(<\alpha, \Pi_n)$, we have $TI(<\alpha, \Pi_n) \vdash TI(\sup_{\delta<\alpha} \delta^{\omega}, \Pi_n)$ only if for some $\beta < \alpha$, $TI(\beta, \Pi_n) \vdash TI(\sup_{\delta<\alpha} \delta^{\omega}, \Pi_n)$. If this happens, by Lemma 4.2(i), $TI(\beta, \Pi_n) \vdash TI(\beta^{\omega}, \Pi_n)$, which fails, if $\beta \geqslant \omega$, by Theorem 4.1(a) with $m = n$, and fails by Lemma 4.2 and the fact $I\Delta_0 \nvdash I\Pi_n$ if $\beta < \omega$.

For the $\Leftarrow$-direction, suppose $\beta < \sup_{\delta<\alpha} \delta^{\omega}$, then for some $\delta < \alpha$ we have $\beta < \delta^{\omega}$. By definition, $TI(<\alpha, \Pi_n) \vdash TI(\delta, \Pi_n)$. If $\delta \geqslant \omega$, we get $TI(\delta, \Pi_n) \vdash TI(\beta, \Pi_n)$ from Theorem 4.1(a) with $n = m$. If $\delta < \omega$, the same result follows from Lemma 4.2(iii). Thus $TI(<\alpha, \Pi_n) \vdash TI(\beta, \Pi_n)$, completing the $\Leftarrow$-direction. this completes the proof of (11).

Now suppose $\delta < \alpha$ and the normal forms of δ and α are $\omega^{\delta_1} \cdot m_1 + \cdots + \omega^{\delta_l} \cdot m_l$ and $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$, respectively, and the normal forms of δ_1 and α_1 are $\omega^{\delta_{1,1}} \cdot m_{1,1} + \cdots + \omega^{\delta_{1,l_1}} \cdot m_{1,l_1}$ and $\omega^{\alpha_{1,1}} \cdot n_{1,1} + \ldots + \omega^{\alpha_{1,k_1}} \cdot n_{1,k_1}$ respectively. Then

$\delta^\omega = \omega^{\omega^{\delta_{1,1}+1}}$. If $\alpha = \omega^{\omega^{\alpha_{1,1}}}$ then we must have $\delta_1 < \alpha_1$, and $\delta_{1,1} < \alpha_{1,1}$. Thus by O_{21}, $\delta_{1,1} + 1 \leqslant \alpha_{1,1}$, and hence, $\delta < \delta^\omega \leqslant \alpha$. This yields $\sup_{\delta<\alpha} \delta^\omega = \alpha$, for this case. If α is not of the form $\omega^{\omega^{\cdot\cdot}}$ then $\alpha > \omega^{\omega^{\alpha_{1,1}}}$ and so

$$\alpha^\omega = \omega^{\omega^{\alpha_{1,1}+1}} = (\omega^{\omega^{\alpha_{1,1}}})^\omega \leqslant \sup_{\delta<\alpha} \delta^\omega \leqslant \alpha^\omega.$$

That is $\sup_{\delta<\alpha} \delta^\omega = \alpha^\omega$, and this completes the proof of Theorem 4.1(g).

The only parts of Theorem 4.1 left to prove are the $\Rightarrow$-directions of (a) and (b). As already mentioned, for the $\Rightarrow$-direction of a we use a result of Section 5. In particular, we show that there is a formula φ such that for any

$$\beta \geqslant \omega_{n-m}^{\alpha^\omega},\ TI(\beta, \Pi_m) \vdash \varphi, \tag{12}$$

but,

$$TI(\alpha, \Pi_n) \nvdash \varphi.$$

By Lemma 4.2(i), and Lemma 4.4, rather than show (12), it is enough to show that

$$TI(\alpha^\omega, \Pi_n) \vdash \varphi.$$

If the normal form of α is $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$ and the normal form of α_1 is $\omega^{\alpha_{1,1}} \cdot n_{1,1} + \ldots + \omega^{\alpha_{1,k_1}} \cdot n_{1,k_1}$ then it is enough to show that

$$TI(\omega^{\omega^{\alpha_{1,1}+1}}, \Pi_n) \vdash \varphi \quad \text{and} \quad TI(\omega^{\omega^{\alpha_{1,1}}}, \Pi_n) \nvdash \varphi.$$

That such a φ exists is a direct consequence of Theorem 5.2.

For the $\Rightarrow$-direction of 4.1(b), we use a well-known result, a proof of which is given in [19,22]. In particular, we use the fact that for any Π_{n+2}-sentence, φ,

$$I\Sigma_{n+1} \vdash \varphi \rightarrow \mathrm{Con}(\ulcorner\varphi\urcorner).$$

We have already noted that for $\alpha \geqslant \omega$ and $n > 0$, $TI(\alpha, \Pi_n)$ can be axiomatized by a single sentence. We note that the sentence can be the conjunction of a single sentence axiomatization of the transfinite induction axioms, which has complexity Π_{n+2}, and a single sentence axiomatization of $I\Sigma_n$, which also has complexity Π_{n+2}. Thus $TI(\alpha, \Pi_n)$ can be axiomatized by a single Π_{n+2}-sentence. Thus,

$$I\Sigma_{n+1} \vdash TI(\alpha, \Pi_n) \rightarrow \mathrm{Con}(\ulcorner TI(\alpha, \Pi_n)\urcorner).$$

So by Gödel's second incompleteness theorem,

$$TI(\alpha, \Pi_n) \nvdash I\Sigma_{n+1}.$$

By Lemma 4.2 (i) and (ii), for all

$$\beta \geqslant \omega,\ TI(\alpha, \Pi_n) \nvdash TI(\beta, \Pi_{n+1}).$$

Since $m > n$ implies $\Pi_{n+1} \subseteq \Pi_m$, we have the $\Rightarrow$-direction of Theorem 4.1(b).

We have completed the proof of Theorem 4.1. Before proceeding to Section 5, we will add to our discussion of transfinite induction.

4.3. Some comments about Theorem 4.1

Suppose $\omega \leqslant \alpha < \varepsilon_0$ and the normal form of α is $\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k$ and the normal form of α_1 is $\omega^{\alpha_{1,1}} \cdot n_{1,1} + \ldots + \omega^{\alpha_{1,k_1}} \cdot n_{1,k_1}$ Then Theorem 4.1(a), for $m = n$, can be stated as

$$TI(\alpha, \Pi_n) \vdash TI(\beta, \Pi_n) \Leftrightarrow \beta < \omega^{\omega^{\alpha_{1,1}+1}}.$$

In other words $TI(\omega^{\omega^\gamma}, \Pi_n)$ is just as strong as $TI(\delta, \Pi_n)$ for any δ such that $\omega^{\omega^\gamma} \leqslant \delta < \omega^{\omega^{\gamma+1}}$, and $TI(\omega^{\omega^{\gamma+1}}, \Pi_n)$ is strictly stronger. That is we have the hierarchy:

$$TI(0, \Pi_n), TI(\omega, \Pi_n), TI(\omega^\omega, \Pi_n), TI(\omega^{\omega^2}, \Pi_n), \ldots,$$
$$TI(< \omega^{\omega^\omega}, \Pi_n), TI(\omega^{\omega^\omega}, \Pi_n), TI(\omega^{\omega^{\omega+1}}, \Pi_n), \ldots,$$

where each theory is strictly weaker than the following one but of the same strength as the theories that would fall in between. So by Theorem 4.1, every theory of the form $TI(\alpha, \Pi_n)$ or $TI(\alpha, \Sigma_n)$ is equivalent either to $I\Delta_0$ or to a theory of the form $TI(\omega^{\omega^\gamma}, \Pi_n)$, for some $\gamma < \varepsilon_0$, and some $n \geqslant 1$. Also theories of the form $TI(< \alpha, \Pi_n)$ and $TI(< \alpha, \Sigma_n)$ are equivalent to either $I\Delta_0$ or to a theory of the form $TI(< \omega^{\omega^\gamma}, \Pi_n)$, for some $\gamma < \varepsilon_0$ and some $n \geqslant 1$. For this reason we will restrict our attention only to theories of the form $TI(\omega^{\omega^\gamma}, \Pi_n)$, where $\gamma < \varepsilon_0$ and $n \geqslant 1$.

4.4. Least ordinal number principles

Often ordinary induction is easier to work with as a least number principle. Let $L\mathscr{C}$ denote the theory axiomatized by P^- together with the axiom schema:

$$\forall \vec{z}(\exists x\, \varphi(x, \vec{z}) \rightarrow \exists x(\varphi(x, \vec{z}) \wedge (\forall y < x)\neg\varphi(y, \vec{z}))): \quad \varphi \in \mathscr{C}.$$

It is well-known that $L\Pi_n \equiv L\Sigma_n \equiv I\Sigma_n \equiv I\Pi_n$ (cf. [14]). In fact, $L\Sigma_n$ can be obtained from $I\Pi_n$ by taking contrapositives of the induction axioms; $L\Pi_n$ can be obtained in a similar manner from $I\Sigma_n$. Although Theorem 4.1 implies that $TI(\alpha, \Sigma_n)$ is strictly stronger than $TI(\alpha, \Pi_n)$ for all $\alpha \geqslant \omega^2$, we can show a parallel with ordinary induction in the least number principle versus induction sense.

To see this let $LN(\alpha, \varphi)$, *the least number principle up to* α *for* φ, denote the universal closure of

$$(\exists \gamma < \alpha)\varphi(\gamma) \rightarrow \exists \beta(\varphi(\beta) \wedge (\forall \gamma < \beta)\neg\varphi(\gamma)).$$

Let $LN(\alpha, \mathscr{C})$, *the least number principle up to* α *for* $\mathscr{C}$*-formulas*, be the theory $I\Delta_0 + \{LN(\alpha, \varphi)\colon \varphi \in \mathscr{C}\}$. Then $LN(\alpha, \varphi)$ can be obtained from $TI(\alpha, \neg\varphi)$ by taking contrapositives; this yields:

Theorem 4.6. *For all* $n \in \omega$ *and* $\alpha < \varepsilon_0$,

(i) $LN(\alpha, \Sigma_n) \equiv TI(\alpha, \Pi_n)$, *and*

(ii) $LN(\alpha, \Pi_n) \equiv TI(\alpha, \Sigma_n)$.

4.5. Ordinal induction as a least sequence principle

Each ordinal less than ω^c can be identified with a sequence of natural numbers of length c. This can be seen by noting every $\alpha < \omega^c$ has the form

$$\alpha = \omega^{c-1} \cdot n_1 + \omega^{c-2} \cdot n_2 + \cdots + \omega \cdot n_{c-1} + n_c,$$

where $n_1, \dots, n_c \in \omega$. So we may identify α with the sequence $(n_1, \dots, n_c)$.

Now if we are given that $(\exists \alpha < \omega^c)\varphi(\alpha)$, in order to find $\mu\alpha\varphi(\alpha)$, the least α such that $\varphi(\alpha)$, we may first find

$$n_1 = \mu x(\exists x_2, \dots, x_c \varphi(\omega^{c-1} \cdot x + \omega^{c-2} \cdot x_2 + \cdots + x_c)).$$

Second we find

$$n_2 = \mu x(\exists x_3, \dots, x_c \varphi(\omega^{c-1} \cdot n_1 + \omega^{c-2} \cdot x + \omega^{c-3} \cdot x_3 + \cdots + x_c)),$$

etc., until the sequence $(n_1, \dots, n_c)$ is obtained. Clearly,

$$\omega^{c-1} \cdot n_1 + \cdots + \omega \cdot n_{c-1} + n_c = \mu\alpha\varphi(\alpha).$$

If $\varphi \in \Sigma_n$ then we may accomplish this using c applications of $L\Sigma_n$. We conclude that $L\Sigma_n \vdash LN(\omega^c, \Sigma_n)$. We already had this from Theorems 4.1 and 4.6; however, this technique will be useful in the next section where viewing ordinal induction as a least sequence principle is a key idea in some of the proofs. We note that the proof of Lemma 4.3 could have been accomplished for general α, as it was above for $\alpha = \omega$, by coding ordinals less than α^c by sequences of length c of ordinals less than α. We will not do this but the key ideas for such a coding of ordinals is described in Section 5.6.1.

One may attempt to prove $LN(\alpha^\omega, \Sigma_n)$ from $LN(\alpha, \Sigma_n)$, using the above technique, by first finding the least c such that $(\exists \beta < \alpha^c)\varphi(\beta)$, and then proceed as above. The problem is, of course, that c may be nonstandard, and so we may not be able to apply $LN(\alpha, \Sigma_n)$ c times. In the following section we will show that in fact:

$$LN(\alpha, \Sigma_n) \nvdash LN(\alpha^\omega, \Sigma_n).$$

5. Transfinite induction and the fast-growing hierarchy

For L an extension of L_0, and T an L-theory containing P^-, we say f *is T-provably recursive* if (def) for some $\varphi \in \Sigma_1$,

(i) $f(\vec{n}) = m \Leftrightarrow \omega \models \varphi(\vec{n}, m)$, and

(ii) $T \vdash \forall \vec{x} \exists! y \varphi(\vec{x}, y)$

Using a version of the Kleene T-predicate it is clear that the T-provably recursive functions are exactly the functions $\vec{x} \mapsto \{e\}(\langle \vec{x} \rangle)$ such that T proves for each $\vec{x}$ there is a computation of program e on input $\langle \vec{x} \rangle$.

Given f and φ as in (i), we write "$f(\vec{x}) = y$" to denote $\varphi(\vec{x}, y)$, and borrowing notation from recursion theory, we write $f(\vec{x})\downarrow$ and $f(\vec{x})\uparrow$ to denote $\exists y\varphi(\vec{x}, y)$ and $\neg\exists y\varphi(\vec{x}, y)$, respectively. Additionally, we use $(f\downarrow)$ to denote $\forall\vec{x}(f(\vec{x})\downarrow)$. Suppose f_1 and f_2 are recursive functions and $\varphi_1, \varphi_2 \in \Sigma_1$ are such that $f_i(\vec{n}) = m \Leftrightarrow \omega \models \varphi_i(\vec{n}, m)$, for $i = 1, 2$. We write $f_1(\vec{x}) \simeq f_2(\vec{x})$ to denote $\forall y(\varphi_1(\vec{x}, y) \leftrightarrow \varphi_2(\vec{x}, y))$, and $f_1 \simeq f_2$ to denote $\forall\vec{x}(f_1(\vec{x}) \simeq f_2(\vec{x}))$. (*Note:* this definition of $\simeq$ differs from the definition used in Section 3.)

5.1. The fast-growing hierarchy

Classifying the provably recursive functions of a theory T is a natural way of partially describing T. For instance, as is well known, the provably recursive functions of $I\Sigma_1$ are exactly the primitive recursive functions. Results of Paris in [10] together with results of Ketonen and Solovay in [8] yield a characterization of the provably recursive functions of $I\Sigma_n$, for each $n \geqslant 1$, in terms of the *Wainer hierarchy* [24] (also known as the *fast growing hierarchy* and the *extended Ackermann hierarchy*). The provably recursive functions of the $I\Sigma_n$'s were first established, independently, by Mints and Takeuti using proof-theoretic methods (cf. [9]). The definitions of the functions of the Wainer hierarchy vary slightly in the literature (the differences partially due to different definitions of the fundamental sequence function); we use the following:

(i) $F_0(x) \stackrel{\text{def}}{=} x + 1$,

(ii) $F_{\alpha+1}(x) \stackrel{\text{def}}{=} F_\alpha^{x+1}(x)$,

(iii) $F_\lambda(x) \stackrel{\text{def}}{=} F_{\{\lambda\}(x)}(x)$, for λ a limit ordinal,

where $F^y(x)$ is F iterated y times on x, i.e,

$$F^y(x) \stackrel{\text{def}}{=} \underbrace{F(F(\cdots(F}_{y\text{ times}}(x))\cdots)).$$

This hierarchy is a very natural one, obtained at successor stages by iterating the previous function, and at limit stages by diagonalization. In the following section we will be specific about how "$F_\alpha(x) = y$" is formalized; in particular, we will introduce a Δ_0-formula to denote $F_\alpha(x) = y$ (for α, x, and y free variables) such that (i)–(iii) above are provable in $I\Delta_0$ with "$=$" replaced by "$\simeq$". For such a formalization the results of Paris, Ketonen, and Solovay mentioned above imply:

Proposition 5.1. *For all $n \geqslant 1$,*

$$I\Sigma_n \vdash F_\alpha\downarrow \ \Leftrightarrow\ \alpha < \omega_n.$$

So, for example, $I\Sigma_1 \vdash$ "F_c is total" for all $c \in \omega$, but $I\Sigma_1 \nvdash$ "F_ω is total". We note that F_ω is essentially Ackermann's function.

In this section we generalize the above to:

Theorem 5.2. *For all $n \geqslant 1$ and all $\alpha < \varepsilon_0$,*

$$TI(\omega^{\omega^\alpha}, \Pi_n) \vdash F_\beta \downarrow \;\Leftrightarrow\; \beta < \omega_n^{\alpha+1}.$$

Proposition 5.1 is just Theorem 5.2 for $\alpha = 0$. We note that, as explained in Section 4.3, it is enough to only consider $I\Delta_0$ and $TI(\omega^{\omega^\alpha}, \Pi_n)$ for all $\alpha < \varepsilon_0$ and $n \geqslant 1$, in order to describe the situation for $TI(<\gamma, \mathscr{C})$ and $TI(\gamma, \mathscr{C})$ for all $\gamma < \varepsilon_0$ and $\mathscr{C} = \Delta_0$, Σ_n, and Π_n for all $n \geqslant 1$. Thus using Theorem 5.2 and the well-known fact that $I\Delta_0 \vdash F_\alpha \downarrow$ if and only if $\alpha = 0$ or $\alpha = 1$, we have a precise description of the functions in the fast growing hierarchy that are provably total in the theories $TI(\alpha, \mathscr{C})$ and $TI(<\alpha, \mathscr{C})$ for all $\alpha < \varepsilon_0$ and all $\mathscr{C} = \Delta_0$, Σ_n, or Π_n, for $n \geqslant 1$. In Section 6.3 we give a precise characterization of all of the provably recursive functions of these theories. Before proceeding with the proof of Theorem 5.2 we will, in the following section, show the existence of a Δ_0-formula expressing $F_\alpha(x) = y$.

We remark that in the upcoming sections, where the construction of models of arithmetic take place, we will suppose that we are working in a metatheory sufficiently strong to handle these constructions; certainly ZFC will do. However, in Section 6.4, finitistic interpretations of such model-theoretic techniques are discussed.

5.2. A Δ_0-formula denoting $F_\alpha(x) = y$

The computation of $F_\alpha(n)$ can easily be described as a sequence of manipulations of expressions of the form

$$F_{\alpha_1}^{n_1}(F_{\alpha_2}^{n_2}(\cdots(F_{\alpha_k}^{n_k}(n))\cdots)), \tag{13}$$

where $n_1, \ldots, n_k \in \omega - \{0\}$ and $\alpha_1 > \cdots > \alpha_k \geqslant 0$. We are abusing notation by writing n and n_i when we really mean expressions such as the Smullyan base 2 representations of n and n_i, i.e. $\overline{n}^2$ and $\overline{n_i}^2$; similarly we should use such expressions in place of the α_i's. A single *computation step* is described as follows. If $n_k > 1$ replace n_k by $n_k - 1$, and if $n_k = 1$ erase $F_{\alpha_k}^{n_k}$ along with its parentheses; also: if $\alpha_k = 0$ replace n by $n+1$; if $\alpha_k = \gamma+1$ replace n by $F_\gamma^{n+1}(n)$, and if α_k is a limit ordinal replace n by $F_{\{\alpha_k\}(n)}^1(n)$. Continue as long as some F_γ^i's remain. If we begin with the expression "$F_\alpha^1(n)$" then the value of the expression we end up with is the value of $F_\alpha(n)$.

We may have our expressions meet the requirement $\alpha_k > 0$, if we compute $F_0^m(n)$ all at once by writing the expression for $m+n$ rather than "$F_0^m(n)$" whenever a computation step requires us to write "$F_0^m(n)$". Doing this allows us to assign ordinals to expressions of the form (13) in a natural way such that a computation step has a very nice description as a function on ordinals. In particular, if we assign to (13) the ordinal

$$\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k + n,$$

(where $\alpha_k > 0$) then a computation step is given by the function $C : \varepsilon_0 \to \varepsilon_0$ defined by

$$C(\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k + n) \stackrel{\text{def}}{=} \{\omega^{\alpha_1} \cdot n_1 + \cdots + \omega^{\alpha_k} \cdot n_k\}(n) + n.$$

So if $\alpha = \beta+n$, where $\beta \gg \omega$, then $C(\alpha) \stackrel{\text{def}}{=} \{\beta\}(n)+n$. By reviewing the definition of $\{\ \}$ it is clear that the expression corresponding to the ordinal $C(\alpha)$ is the expression obtained by performing one computation step on the expression corresponding to α.

The number of steps needed to compute $F_\alpha(x)$ using the computation function described above can be close to $F_\alpha(x)$, and hence an exponential in $F_\alpha(x)$ can be required to code the computation of $F_\alpha(x)$. By using a modified computation function the number of steps needed may be reduced sufficiently to allow us to code the computation by a number smaller than a polynomial in $F_\alpha(x)$, and hence enabling us to produce a Δ_0-formula expressing $F_\alpha(x) = y$. We reduced the number of steps in the computation when we chose to compute $F_0^m(n)$ "all at once". A further reduction is made by computing $F_1^m(n)$ all at once. Noting $F_1^m(n) = 2^m(n+1) - 1$, we define for $\beta \gg \omega^2$,

$$C_1(\beta + \omega \cdot m + n) \stackrel{\text{def}}{=} \{\beta\}(2^m(n+1) - 1) + 2^m(n+1) - 1.$$

Using a Δ_0-formula expressing $2^x = y$ (cf. e.g. [7] or the appendix of [4]), the graph of C_1 can clearly be represented by a Δ_0-formula.

We code sequences of ordinals as follows. We denote the Gödel number of the sequence $\alpha_0, \ldots, \alpha_{k-1}$ by $\langle \alpha_0, \ldots, \alpha_{k-1} \rangle_o$. We let the Smullyan base 5 representation of $\langle \alpha_0, \ldots, \alpha_{k-1} \rangle_o$ be given by

$$5 * \overline{\ulcorner\alpha\urcorner_0}^4 * \cdots * 5 * \overline{\ulcorner\alpha\urcorner_{k-1}}^4,$$

where $\overline{\ulcorner\alpha\urcorner}^4$ denotes the Smullyan base 4 representation of $\ulcorner\alpha\urcorner$ (cf. Section 3). We let $\mathrm{len}_o(\langle \alpha_0, \ldots, \alpha_{k-1} \rangle_o) \stackrel{\text{def}}{=} k$, and

$$(\langle \alpha_0, \ldots, \alpha_{k-1} \rangle_o)_{o,i} \stackrel{\text{def}}{=} \begin{cases} \alpha_i & \text{if } i < k; \\ 0 & \text{otherwise.} \end{cases}$$

In a manner similar to that used in Section 2.2 for sequences in ω^*, we can obtain Δ_0-definitions for len_o and $(s,i) \mapsto (s)_{o,i}$, as well as other functions for working with sequences of ordinals, that have nice properties provable in $I\Delta_0$. To make notation less cumbersome we will drop the subscript "o" and use the same notation for elements of $(\varepsilon_0)^*$ as for elements of ω^*. Since we are using different symbols to denote ordinals than to denote natural numbers, no ambiguity will arise from the fact that we are using the same notation for sequences of ordinals as for sequences of natural numbers. We emphasize that our conventions are such that $\langle \alpha_0, \ldots, \alpha_k \rangle$ and $\langle \ulcorner\alpha_0\urcorner, \ldots, \ulcorner\alpha_k\urcorner \rangle$ denote different numbers; similarly $\langle \mathrm{ord}(n_0), \ldots, \mathrm{ord}(n_k) \rangle$ and $\langle n_0, \ldots, n_k \rangle$ denote different numbers; in both cases, for the former, we use the definition of sequences of ordinals less than ε_0; and for the latter, we use our definition of sequences from the natural numbers.

We use the phrase "s is a sequence of ordinals" to denote a formula expressing "$s = 0$ or the Smullyan base 5 representation of s is of the form $5 * x_i * \cdots * 5 * x_{k-1}$, where for $i < k$, x_i is the Smullyan base 4 representation of an ordinal less than ε_0", and we say "$s = \langle \alpha_0, \ldots, \alpha_{k-1} \rangle$ is a sequence of ordinals" if s is a sequence of ordinals of length k, and for all $i < k$, $(s)_{\mathrm{o,i}} = \alpha_i$. We use the phrase *s is a computation sequence* to denote the formula:

$$\text{“}\langle\alpha_0,\ldots,\alpha_k\rangle \text{ is a sequence of ordinals”} \\ \wedge(\forall j < k)(\alpha_{j+1} = C_1(\alpha_j)). \tag{14}$$

We use the phrase $\langle\alpha_0,\ldots,\alpha_k\rangle$ *is a computation of* $F^i_\alpha(x)$ to denote the formula:

$$\text{“}\langle\alpha_0,\ldots,\alpha_k\rangle \text{ is a computation sequence”} \\ \wedge(\alpha_0 = \omega^\alpha \cdot i + x) \wedge (C_1(\alpha_k) < \omega). \tag{15}$$

Actually, $\langle\alpha_0,\ldots,\alpha_k\rangle$ may code all but the last computation step in the computation of $F^i_\alpha(x)$; this allows for a further reduction in the size of $\langle\alpha_0,\ldots,\alpha_k\rangle$. The formula (15) is easily seen to be Δ_0.

As a notational convenience, for $s = \langle\alpha_0,\ldots,\alpha_k\rangle$, we let,

$$U(s) \stackrel{\text{def}}{=} \mathrm{num}(C_1(\alpha_k)).$$

By earlier conventions, if $C_1(\alpha_k) \geqslant \omega$ then $U(s) = 0$. For $a = 135$ and $b = 5^{2421}$ we will use the following formula to denote $F^i_\alpha(x) = y$ (we note that a and b can be chosen much smaller):

$$((x > 0 \vee i = 0) \\ \wedge(\exists s \leqslant y^a + b)\ (\text{“}s \text{ is a computation of } F^i_\alpha(x)\text{”} \wedge U(s) = y)) \\ \vee((x = 0 \wedge i > 0) \\ \wedge(\exists s \leqslant y^a + b)\ (\text{“}s \text{ is a computation of } F^{i-1}_\alpha(1)\text{”} \wedge U(s) = y)). \tag{16}$$

We will often write “$F_\alpha(x)$” in place of “$F^1_\alpha(x)$”. The formula (16) is easily seen to be equivalent to a Δ_0-formula, and we have:

Theorem 5.3. *Let* $F^i_\alpha(x) = y$ *denote the formula given in* (16); *then* $I\Delta_0$ *proves the following*:

(i) $(F_\alpha(x) = y_1 \wedge F_\alpha(x) = y_2) \rightarrow y_1 = y_2$.
(ii) $F_0(x) = x + 1$.
(iii) $F_{\alpha+1}(x) \simeq F^{x+1}_\alpha(x)$.
(iv) $\mathrm{LIM}(\lambda) \wedge (\{\lambda\}(x)\downarrow) \rightarrow (F_\lambda(x) \simeq F_{\{\lambda\}(x)}(x))$.
(v) $(F^0_\gamma(x) = x) \wedge (F^{i+1}_\gamma(x) \simeq F^i_\gamma(F_\gamma(x)))$.

By $F^{i+1}_\gamma(x) \simeq F^i_\gamma(F_\gamma(x))$ *we mean* $\forall y(F^{i+1}_\gamma(x) = y \leftrightarrow \exists z(F_\gamma(x) = z \wedge F^i_\gamma(z) = y)$.

In order to free this paper of a large number of technical details we omit the proof of this theorem. This theorem is proved in Appendix A of [19]. We conclude this section by stating some properties of the functions F_α. Other than their provability in $I\Delta_0$ these properties are well-known; proofs of these properties are included in Appendix A of [19].

Proposition 5.4. *The following are provable in $I\Delta_0$:*
(i) *For $\alpha \geqslant 2$, $(F_\alpha(x)\downarrow)$ implies that $x^2 < F_\alpha(x)$,*
(ii) *If $i > 0$ and $(F^i_\alpha(x)\downarrow)$ then $x < F^i_\alpha(x)$),*
(iii) *Suppose $n < \omega$, $n < \alpha < \varepsilon_0$, $n \leqslant x$, and $F_\alpha(x) = y$, then $(\exists z \leqslant y)(F_n(x) = z)$.*
(iv) *$\alpha > \beta \to ((F_\alpha\downarrow) \to (F_\beta\downarrow))$, and*
(v) *If exponentiation is total, $x \geqslant w$, and $F_\alpha(x) = y$, then $(\exists z \leqslant y)(F_\alpha(w) = z)$.*

5.3. The $\Leftarrow$-direction of Theorem 5.2

Since $(F_\alpha\downarrow)$ is a Π_2-formula, there is an easy proof of the $\Leftarrow$-direction of Theorem 5.2 for $n \geqslant 2$. To handle this direction for all $n \geqslant 1$, we prove:

Lemma 5.5. *For all $\alpha > 0$,*

$$TI(\omega^\alpha, \Pi_1) \vdash F_\alpha\downarrow.$$

Proof. We let $C^i_1(\gamma) < \delta$ denote the formula:

$$\exists s(\text{“}s = \langle \alpha_0, \ldots, \alpha_{i-1}\rangle \text{ is a computation sequence”} \wedge (\alpha_0 = \gamma) \wedge (C_1(\alpha_{i-1}) < \delta)).$$

So $\exists i(C^i_1(\gamma) < \delta)$ is a Σ_1-formula. Also, noting the construction of our Δ_0-formula for the F_α's, it is easy to get

$$I\Delta_0 \vdash (\forall x > 0)((\exists i)(C^i_1(\omega^\alpha + x) < \omega) \to F_\alpha(x)\downarrow).$$

Noting $\langle \omega^0 \cdot 1\rangle < 5^{5421} = b$ is a computation of $F^0_\gamma(1)$ for any γ, and considering (16), we obtain $I\Delta_0 \vdash (F_\gamma(0)\downarrow)$; hence, it is enough to show that

$$TI(\omega^\alpha, \Pi_1) \vdash (\forall x > 0)\exists i(C^i_1(\omega^\alpha + x) < \omega).$$

We now work in $TI(\omega^\alpha, \Pi_1)$. By Lemma 4.2, we have $I\Pi_1$, and hence $C_1\downarrow$ (since $\{\ \}\downarrow$ and $\exp\downarrow$). Using results of Section 2, it is easy to show that $(\forall \gamma \geqslant \omega)(C_1(\gamma) < \gamma)$, so $\forall x(C_1(\omega^\alpha + x) < \omega^\alpha + x)$, and hence,

$$\forall x(\exists \gamma < \omega^\alpha + \omega)(\exists i)(C^i_1(\omega^\alpha + x) < \gamma). \tag{17}$$

By the $\Leftarrow$-direction of Theorem 4.1(a) with $m = n = 1$, and by the fact that if $\alpha > 0$ then $\omega^\alpha + \omega < (\omega^\alpha)^\omega$, we have $TI(\omega^\alpha + \omega, \Pi_1)$. So by Theorem 4.6 we have $LN(\omega^\alpha + \omega, \Sigma_1)$. By (17), there is a least γ, say γ_0, such that $\exists i(C^i_1(\omega^\alpha + x) < \gamma)$. Let i_0 be such that $C^{i_0}_1(\omega^\alpha + x) < \gamma_0$, and let $\gamma_1 \stackrel{\text{def}}{=} C^{i_0}_1(\omega^\alpha + x)$. If $\gamma_1 \geqslant \omega$ then $C_1(\gamma_1) < \gamma_1$, implying $C^{i_0+1}_1(\omega^\alpha + x) < \gamma_1 < \gamma_0$, contradicting the leastness of γ_0. Thus $\gamma_1 < \omega$, so $C^{i_0}_1(\omega^\alpha + x) < \omega$. We have established $\forall x \exists i(C^i_1(\omega^\alpha + x) < \omega)$, which gives us the lemma. $\square$

Using the $\Leftarrow$-direction of Theorem 4.1(a) we have $TI(\omega^{\omega^\alpha}, \Pi_n) \vdash TI(\beta, \Pi_1)$, all $\beta < \omega^{\omega_n^{\alpha+1}}$, so using the previous lemma,

$$TI(\omega^{\omega^\alpha}, \Pi_n) \vdash F_\beta\downarrow \text{ all } \beta < \omega_n^{\alpha+1},$$

which is the $\Leftarrow$-direction of Theorem 5.2. The remainder of Section 5 is devoted to proving the $\Rightarrow$-direction.

5.4. The $I\Sigma_1$-provably recursive functions

We have already noted that the provably recursive functions of $I\Sigma_1$ are exactly the primitive recursive ones. This was originally proven proof-theoretically. The simple model theoretic proofs of Kirby and Paris in [13] lead to the same result. Also Paris made significant progress in the area of models of arithmetic by developing this technique, for example in [10]. The same proof techniques will be applied the proof of the $\Rightarrow$-direction of Theorem 5.2; in this section we demonstrate the technique for the $I\Sigma_1$ case. This example will serve as the prototype for future applications. Recall, for these model theoretic constructions, we are working in a sufficiently strong theory such as ZFC.

Before proceeding we mention some notational conventions. For L_0-models I and M we say *I is an initial segment of M* or *M is an end extension of I*, and we write $I \subsetneq_e M$, if $I \subseteq M$, and for all $a \in I$ if $M \models b < a$ then $b \in I$. We use ω to denote the standard part of any model of $I\Delta_0$, so for any $M \models I\Delta_0$, we have $\omega \subsetneq_e M$. Also "$a \in M - \omega$" means a is nonstandard. If $I \subsetneq_e M$ then we will write $I < b$ to mean $b \in M - I$. Sometimes we write $a < I$ to indicate $a \in I$. We use $[a,b]$, $[a,b)$, $(a,b]$, and (a,b) to denote $\{x\colon\ M \models a \leqslant x \leqslant b\}$, $\{x\colon\ M \models a \leqslant x < b\}$, $\{x\colon\ M \models a < x \leqslant b\}$, and $\{x\colon\ M \models a < x < b\}$, respectively. If $a \in M$ often we will identify a with $\{x\colon M \models x < a\}$; for example if $\{a_n\colon n \in \omega\} \subseteq M$, then,

$$\bigcup_{n\in\omega} a_n \stackrel{\text{def}}{=} \{x\colon \text{ for some } n \in \omega,\ M \models x < a_n\}.$$

If for all $m \in \omega$, $I \stackrel{\text{def}}{=} \bigcup_{n\in\omega} a_n \neq a_m$ we may use the terminology *I is a limit point of M*.

The key lemma used to classify the provably recursive functions of $I\Sigma_1$ is the following. (Recall, $I\Delta_0(\exp)$ is the natural extension of $I\Delta_0$ obtained by including exponentiation.)

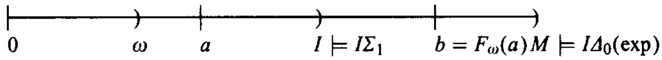
Lemma 5.6. *Suppose $M \models I\Delta_0(\exp) + \{F_\omega(a) = b\}$ where $a,b \in M - \omega$, then there is an $I \subsetneq_e M$ such that $a < I < b$ and $I \models I\Sigma_1$.*

We note that the existence of such an M is easy to show. By the absoluteness of Δ_0-formulas with respect to initial segments of models of $I\Delta_0$, we have

$$I \not\models \exists y(F_\omega(a) = y),$$

so $I\Sigma_1 \ \not\vdash F_\omega\downarrow$; since $I \models I\Sigma_1$, using Proposition 5.4(iv), we have the $\Rightarrow$-direction of Theorem 5.2 for $\alpha = 0$ and $n = 1$. We note that this result would be obtained even if M is a nonstandard model of PA or of $Th(\omega)$. The fact that M need only be a model of a theory such as $I\Delta_0(\exp) + \{F_\omega(a) = b\} + \{a > n\colon\ n \in \omega\}$ is used to formalize this construction in weak theories, as is done in Chapter 5 of [19] and in [22].

The lemma may be represented by the following picture.



Proof of Lemma 5.6. Suppose $M \models I\Delta_0(\exp) + \{F_\omega(a) = b\} + \{a > n\colon\ n \in \omega\}$, and suppose $\{\psi_i\colon\ i \in \omega\}$ is an infinitely repetitive enumeration of all Δ_0-formulas. We will define $\{a_n\colon\ n \in \omega\}$ and $\{b_n\colon\ n \in \omega\}$ with the following properties:

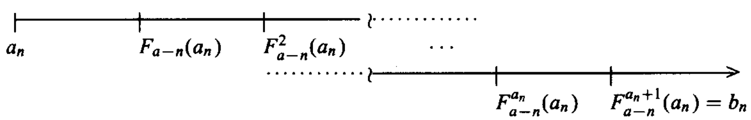
$(1)_n$ $n \neq 0 \Rightarrow (a_{n-1})^2 < a_n$.

$(2)_n$ $a_0 < a_1 < \cdots < a_n \leqslant b_n \leqslant b_{n-1} \leqslant \cdots \leqslant b_0$.

$(3)_n$ If $n \neq 0$ and $\langle \vec{y} \rangle < (a_{n-1} - 1)$ then $\mu w \psi_{n-1}(w, \vec{y}) \notin (a_n, b_n]$.

$(4)_n$ $b_n = F_{a-n+1}(a_n)$.

Note that $\mu w \psi_n(w, \vec{y})$, i.e. the least witness to the formula $\exists w \psi(w, \vec{y})$ if there is such a w or 0 otherwise, exists since $M \models I\Delta_0$. We proceed inductively (in the meta-theory). If we let $a_0 \overset{\text{def}}{=} a$, and $b_0 \overset{\text{def}}{=} b$, then clearly $(1)_0$–$(3)_0$ all hold. Since $F_\omega(a) = F_{a+1}(a)$, $(4)_0$ holds as well. Now suppose we have a_i and b_i for $i = 0, \ldots, n$ such that $(1)_i$–$(4)_i$ all hold. Note that, using the fact $b_n = F_{a-n}^{a_n+1}(a_n)$, we can express the interval $[a_n, b_n]$ by the following picture:



Since there are $a_n + 1$ subintervals of the form $(F_{a-n}^j(a_n), F_{a-n}^{j+1}(a_n)]$, and the set $\{\mu w \psi_n(w, \vec{y})\colon\ \langle \vec{y} \rangle < (a_n - 1)\}$ has at most $a_n - 1$ elements, we know that there is an interval of the form $(F_{a-n}^j(a_n), F_{a-n}^{j+1}(a_n)]$ disjoint from $\{\mu w \psi_n(w, \vec{y})\colon\ \langle \vec{y} \rangle < (a_n - 1)\}$, and such that $j \neq 0$; we note that this argument uses the Δ_0-pigeon-hole principle, which is known to be provable in $I\Delta_0(\exp)$ (cf. [15]). Let $a_{n+1} \overset{\text{def}}{=} F_{a-n}^j(a_n)$ and $b_{n+1} \overset{\text{def}}{=} F_{a-n}^{j+1}(a_n)$ for such a j. $(2)_{n+1}$, $(3)_{n+1}$, and $(4)_{n+1}$ are all immediate. For $(1)_{n+1}$ we use Proposition 5.4(i) and the fact that $a - n > 2$ to get $(a_n)^2 < F_{a-n}(a_n)$. So by Proposition 5.4(ii) and the fact that $j \neq 0$, we have $(a_n)^2 < F_{a-n}^j(a_n) = a_{n+1}$; $(1)_{n+1}$ follows.

Now let $I \overset{\text{def}}{=} \bigcup_{n \in \omega} a_n$, we claim $I \models I\Sigma_1$. By $(1)_n$ for $n \in \omega$, I is closed under $+$ and $\cdot$, so I is a Δ_0-elementary substructure of M. Since $I\Delta_0$ has a Π_1 set of axioms (the induction axioms of $I\Delta_0$ can be given in the form $\forall \vec{z} \forall w(\varphi(0, \vec{z}) \wedge (\forall w \leqslant w)(\varphi(x, \vec{z}) \to \varphi(x+1, \vec{z})) \to (\forall w \leqslant w)\varphi(x, \vec{z}))$ where φ is Δ_0), we have $I \models I\Delta_0$. So, using results discussed in Section 4.4 it will be enough to show that the least number principle holds in I for all Σ_1-formulas.

Suppose $c_0, c_1, \ldots, c_k \in I$, $\varphi \in \Delta_0$, and $I \models \exists w \varphi(w, c_0, \vec{c})$, where $\vec{c}$ denotes $c_1, \ldots, c_k$. We want to show that there is a least such c_0. Since $I \models I\Delta_0$ and since for all $k \in \omega$, $I\Delta_0 \vdash \forall x_0, \ldots, x_k \exists s (s = \langle x_0, \ldots, x_k \rangle)$ (cf. Proposition 2.2(i)), we have $\langle c_0, \ldots, c_k \rangle \in I$.

Hence for some $n \in \omega$, $\langle c_0, \ldots, c_k \rangle < (a_n - 1)$, and since $\{\psi_n \colon n \in \omega\}$ is infinitely repetitive, we may select n such that $\varphi = \psi_n$ as well.

Now since $I \models \exists w \psi_n(w, c_0, \vec{c})$ and $I < b_{n+1}$ we have

$$M \models (\exists w \leqslant b_{n+1}) \psi_n(w, c_0, \vec{c}).$$

Using Δ_0-induction in M, let $\widehat{c_0}$ be the least such c_0, then $\langle \widehat{c_0}, \vec{c} \rangle \leqslant \langle c_0, \vec{c} \rangle < (a_n - 1)$ also. By $(3)_{n+1}$ we have

$$M \models (\exists w \leqslant a_{n+1}) \psi_n(w, \widehat{c_0}, \vec{c}),$$

so $I \models \exists w \psi_n(w, \widehat{c_0}, \vec{c})$. $\widehat{c_0}$ is also least in I, since if $c < \widehat{c_0}$ and $I \models \exists w \psi_n(w, c, \vec{c})$ then $M \models (\exists w \leqslant b_{n+1}) \psi_n(w, c, \vec{c})$ contradicting the leastness of $\widehat{c_0}$ in M. $\square$

The goal of the remainder of Section 5 is to prove a version of Lemma 5.6 with $I\Sigma_1$ replaced by $TI(\omega^{\omega^\alpha}, \Pi_n)$ and $F_\omega \downarrow$ replaced by $F_{\omega_n^{\alpha+1}} \downarrow$. The above proof will be generalized in two ways. For one we will need to apply the above technique for n unbounded quantifiers rather than just one; also we will need to deal with ordinals less than ω^{ω^α} for an arbitrary $\alpha < \varepsilon_0$, rather than for the particular case where $\alpha = 0$. The machinery for generalizing to n unbounded quantifiers is developed in the next section, and then, in the following section, we deal with arbitrary ordinals.

5.5. Generalizing to n unbounded quantifiers

The technique used to generalize to more than one unbounded quantifier is a variation of techniques used by Paris in [10]. To see how this is done, first consider the result of carrying out the construction of the a_n's and b_n's (for the proof of Lemma 5.6) inside M. How this can be accomplished is described in the next subsection. Now we note that inside M the construction would necessarily terminate once $a - n + 1$ reaches 0, i.e. once $n = a + 1$ (consider $(4)_n$). In this section we show that if the interval $[a, b]$ satisfies a certain largeness condition then the construction, with the property $(4)_n$ of the proof of Lemma 5.6 replaced by a different property, can be carried out much further. So we would end up with sequences $a_0, \ldots, a_r$ and $b_0, \ldots, b_r$ where r is a nonstandard number usually much larger than a_0.

We note that for any limit point I of the a_n's, any $\vec{c} \in I$, and any $\varphi \in \Delta_0$,

$$I \models \exists w \varphi(w, \vec{c}) \Leftrightarrow M \models (\exists w \leqslant a_r) \varphi(w, \vec{c}). \tag{18}$$

The $\Rightarrow$-direction follows from $I < a_r$ and $I \prec_{\Delta_0} M$ (since I is closed under $+$ and $\cdot$). The $\Leftarrow$-direction uses the fact that if $w \leqslant a_r$ then $w \leqslant b_i$ for all $i \leqslant r$; so if $\langle \vec{c} \rangle < (a_n - 1)$, $\psi_n = \varphi$, and $M \models (\exists w \leqslant a_r) \psi_n(w, \vec{c})$ then $M \models (\exists w \leqslant b_{n+1}) \psi_n(w, \vec{c})$, and so by $(3)_{n+1}$, $M \models (\exists w \leqslant a_{n+1}) \psi_n(w, \vec{c})$. If $\vec{c} \in I$ then there is such an n where $a_n \in I$; since I is a limit point $a_{n+1} \in I$ as well. Since $I \prec_{\Delta_0} M$ we have $I \models \exists w \psi_n(w, \vec{c})$. Note that by taking the negation of both sides of the $\Leftrightarrow$ we can replace the "$\exists$" in (18) by "$\forall$".

Now if the set $\{a_0, \ldots, a_r\}$ is sufficiently large we can handle a second quantifier by performing a second construction, selecting intervals in the set $\{a_0, \ldots, a_r\}$ rather than the set $[a,b]$, and by working with formulas of the form $\exists w(Qx \leqslant a_r)\psi_n$, where $Q \in \{\forall, \exists\}$, rather than $\exists w \psi_n$. We will get sequences $a'_1, \ldots, a'_s$ and $b'_1, \ldots, b'_s$ with properties similar to $(1)_n$–$(4)_n$ of Lemma 5.6. Then we will show that for any limit point of the a'_n's, for $Q \in \{\forall, \exists\}$, and for any $\varphi \in \Delta_0$,

$$I \models \exists w Q x \varphi(w,x,\vec{c}) \;\Leftrightarrow\; M \models (\exists w \leqslant a'_s)(Qx \leqslant a_r)\varphi(w,x,\vec{c}).$$

As in (18) we may replace "$\exists$" with "$\forall$". If the interval $[a,b]$ is sufficiently large this process may be carried out for n quantifiers. We will show that replacing γ by ω^γ in the subscript of $F_\gamma(a) = b$ yields a sufficient largeness condition to accomplish this task for one additional quantifier.

5.5.1. *Carrying out the construction inside M*

To carry out the construction inside M it is enough to have a truth definition for Δ_0-formulas to which M can apply the least number principle (i.e. for which M has induction). Since we are only concerned with standard formulas with parameters below b, and since $M \models I\Delta_0$, it is enough to get a formula $\Theta_{\Delta_0} \in \Delta_0$ such that for all $\varphi \in \Delta_0$,

$$M \models (\forall b \leqslant b)(\varphi(\vec{c}) \leftrightarrow \Theta_{\Delta_0}(\ulcorner\varphi\urcorner, \langle\vec{c}\rangle)).$$

From [11] we know that if for some $e \in M$ and all $n \in \omega$ we have $e > 2^{b^n}$, then there is such a formula Θ_{Δ_0} with e as a parameter. So there is a $\Delta_0(\exp)$-truth definition for Δ_0-formulas; we let Θ_{Δ_0} denote such a formula.

We note that by using the formula Θ_{Δ_0}, we need not consider the enumeration $\{\psi_i \colon i \in \omega\}$ in the construction of the a_n's and b_n's. Instead we work only with the formula Θ_{Δ_0}. For example, we consider $\mu w \Theta_{\Delta_0}(\ulcorner\psi_i\urcorner, \langle w, \vec{y}\rangle)$ rather than $\mu w \psi_i(w, \vec{y})$.

5.5.2. *α-Large sets*

Along the lines of [8] we extend the domain of the fundamental sequence function to include finite sets and sequences of natural numbers. In this section we work primarily in $I\Delta_0(\exp)$.

Definition 5.7. For $a_0, \ldots, a_k \in \omega$ let $\{\alpha\}(a_0, \ldots, a_k)$ be given by

(a) $\{\alpha\}(\Lambda) \stackrel{\text{def}}{=} \alpha$, and

(b) $\{\alpha\}(a_0, \ldots, a_k) \stackrel{\text{def}}{=} \{\{\alpha\}(a_0, \ldots, a_{k-1})\}(a_k)$.

More precisely we let "$\{\alpha\}(a_0, \ldots, a_k) = \beta$" denote the formula,

$$\begin{aligned}&\exists s(\text{“}s = \langle \alpha_0, \ldots, \alpha_{k+1}\rangle \text{ is a sequence of ordinals”}\\ &\quad \wedge(\alpha_0 = \alpha) \wedge (\forall i \leqslant k)(\alpha_{i+1} = \{\alpha_i\}(a_i)) \wedge (\alpha_k = \beta)).\end{aligned} \tag{19}$$

We say $(a_0, \ldots, a_k)$ *is α-large* if(def) $\{\alpha\}(a_0, \ldots, a_k) = 0$, and $(a_0, \ldots, a_k)$ *is exactly α-large* if(def) $(a_0, \ldots, a_k)$ is α-large and $(a_0, \ldots, a_{k-1})$ is not α-large. We say $A = \{a_0, \ldots, a_k\}$ *in increasing order* if(def) $a_0 < \cdots < a_k$. We code finite sets as increasing sequences, i.e. we identify the set $A = \{a_0, \ldots, a_k\}$ (in increasing order) with the sequence $(a_0, \ldots, a_k)$. For $A = \{a_0, \ldots, a_k\}$ in increasing order, let

(c) $\{\alpha\}(A) \stackrel{\text{def}}{=} \{\alpha\}(a_0, \ldots, a_k)$.

Furthermore let,

(d) $G_\alpha(x) \overset{\text{def}}{=} \mu y([x,y]$ is α-large).

More precisely we let "$G_\alpha(x) = y$" denote the formula,

$$\exists s(\text{"}s = \langle \alpha_0, \ldots, \alpha_{k+1}\rangle \text{ is a sequence of ordinals"} \wedge (\alpha_0 = \alpha)$$
$$\wedge(\forall i \leqslant k)(\alpha_{i+1} = \{\alpha_i\}(x+i)) \wedge (\alpha_k \neq 0) \wedge (\alpha_{k+1} = 0) \wedge (y = x+k)). \quad (20)$$

Remark. According to (7),

$$\{\ulcorner\alpha\urcorner\}(n) \leqslant 4 \cdot (3^{\ulcorner\alpha\urcorner} + 1)^{|\ulcorner\alpha\urcorner|_4} \cdot (n+2)^2.$$

So clearly, s in (19) can be bounded by a term using the exponential function exp; similarly s in (20) can be bounded by a term in exp. Thus there are Δ_0(exp)-formulas expressing "$\{\alpha\}(t) = \beta$", "t is α-large", and "$G_\alpha(x) = y$", such that their defining formulas given above are provable in $I\Delta_0$(exp). Rather than seek a more efficient coding of these computations we will work over the base theory $I\Delta_0$(exp).

In light of the definition of F_α in terms of the computation function C_1, it is not surprising that the function F_α is closely related to the function G_{ω^α}. In fact we have:

Proposition 5.8. *The following are provable in* $I\Delta_0$(exp) :

(i) $\forall x((F_\alpha(x+1)\downarrow) \rightarrow (G_{\omega^\alpha}(x)\downarrow) \wedge G_{\omega^\alpha}(x) \leqslant F_\alpha(x+1))$,

(ii) $\forall x((G_{\omega^\alpha}(x)\downarrow) \rightarrow (F_\alpha(x)\downarrow) \wedge F_\alpha(x) \leqslant G_{\omega^\alpha}(x)+1)$.

Other than provability in $I\Delta_0$(exp) this proposition is well-known (cf. [8]). The proofs of these, over $I\Delta_0$(exp), are included in Appendix B of [19], and are omitted here.

The condition that $[a,b]$ is $\omega_{n+1}^{\alpha+1}$-large is a sufficient largeness condition to carry out the two types of generalizations described at the end of Section 5.4. We will show that for each of the n ω's represented by the n in the subscript $n+1$, we can handle a single unbounded quantifier as previously described. By Proposition 5.8(i) the condition $F_{\omega_n^{\alpha+1}}(a) = b$ will provide us with a sufficient largeness condition.

We next state some of the key properties of α-large sets, but first we introduce some terminology. We say $A_0, \ldots, A_k$ *is an increasing partition of A* if (def):

(a) $A = A_0 \cup \cdots \cup A_k$ and

(b) for all $i < k$, max $A_i <$ min A_{i+1}.

The following, over a stronger base theory, are versions of well-known results (cf. [8, 10]); for the base theory $I\Delta_0$(exp), the proof, omitted here, is included in Appendix B of [19].

Proposition 5.9. *The following are provable in* $I\Delta_0$(exp) :

(i) *If $X = \{x_0, \ldots, x_c\}$ in increasing order is $\omega^\alpha + 1$-large where $\alpha > 0$, then there are sets $X_1, X_2, \ldots, X_{x_0+1}$ such that $\{x_0, x_1\}, X_1, \ldots, X_{x_0+1}$ is an increasing partition of X, and for $i = 1, \ldots, x_0+1$, X_i is $\omega^{\{\alpha\}(x_0)}$-large.*

(ii) *Suppose* $X = \{x_0, \ldots, x_r\}$ *and* $Y = \{y_0, \ldots, y_s\}$ *both in increasing order,* $s \geqslant r$, *and* $x_i \geqslant y_i$ *all* $i = 0, \ldots, r$. *Then:* X *is* α*-large* $\Rightarrow$ Y *is* α*-large.*

(iii) *Suppose* $X \subseteq Y$ *and* X *is* α*-large, then* Y *is* α*-large.*

(iv) s *is exactly* $\omega^\alpha \cdot a$*-large if and only if* $s = s_1 * \cdots * s_a$ *where for* $i = 1, \ldots, a$, s_i *is exactly* ω^α*-large.*

The proof of the next lemma imitates the construction of the a_n's carried out in the proof of Lemma 5.6.

Lemma 5.10. *Suppose* $\theta(u,v,w) \in \Delta_0$. *The following is provable in* $I\Delta_0(\exp)$. *If* $X = \{x_0, \ldots, x_c\}$ *in increasing order is* $\omega^\gamma + 1$*-large where* $0 < \gamma < \varepsilon_0$, *then there are sets* $Y = \{y_0, \ldots, y_r\} \subseteq X$ *and* $Z = \{z_0, \ldots, z_r\} \subseteq X$ *such that* Y *is exactly* γ*-large, and for* $i \leqslant r$,

$(1)_i$ *If* $0 < i \leqslant r - k + 1$ *and* $k \leqslant y_{i-1}$ *then* $F_k(y_{i-1})\downarrow$ *and* $F_k(y_{i-1}) \leqslant y_i$.

$(2)_i$ $x_0 = y_0 < y_1 < \cdots < y_i \leqslant z_i \leqslant z_{i-1} \leqslant \cdots \leqslant z_0 = x_c$.

$(3)_i$ *If* $i \neq 0$ *and* $p < (y_{i-1} - 1)$ *then* $\mu w\theta(p,v,w) \notin (y_i, z_i]$.

$(4)_i$ $[y_i, z_i] \cap X$ *is* $\omega^{\{\gamma\}(y_0, y_1, \ldots, y_{i-1})} + 1$*-large.*

Proof. We remark that the free variable v in θ is a parameter fixed throughout the construction of Y and Z. We prove the lemma with $(1)_i$ replaced by

$(1')_i$ *If* $i \neq 0$ *then* $F_{\{\gamma\}(y_0, \ldots, y_{i-1})}(y_{i-1})\downarrow$ *and* $F_{\{\gamma\}(y_0, \ldots, y_{i-1})}(y_{i-1}) \leqslant y_i$.

Since Y is exactly γ-large, if $i \leqslant r-k+1$ then $\{\gamma\}(y_0, \ldots, y_{i-1}) \geqslant k$, so by Proposition 5.4(iii), if $y_{i-1} \geqslant k$, we have $F_k(y_{i-1}) \leqslant F_{\{\gamma\}(y_0, \ldots, y_{i-1})}(y_{i-1})$, hence $(1')_i \Rightarrow (1)_i$.

Let $\psi(r,X,Y,Z) \stackrel{\text{def}}{=} (\forall i \leqslant r)((1')_i \wedge (2)_i \wedge (3)_i \wedge (4)_i)$; note that ψ is a $\Delta_0(\exp)$-formula. We will show:

(i) $(\exists Y, Z \leqslant X)\psi(0,X,Y,Z)$, and

(ii) $(\exists Y, Z \leqslant X)(\psi(j,X,Y,Z) \wedge \{\gamma\}(Y) \neq 0) \rightarrow (\exists Y, Z \leqslant X)\psi(j+1,X,Y,Z)$.

Clearly $\exists Y, Z\psi(c+1,X,Y,Z)$ fails (e.g. consider $(2)_{c+1}$). So using $\Delta_0(\exp)$-induction, (i) and (ii), there is an $r \leqslant c$ such that $\exists Y, Z(\psi(r,X,Y,Z) \wedge$ "Y is exactly γ-large"). Thus if we prove (i) and (ii) we will have the lemma.

For (i), simply let $Y = \{x_0\}$ and $Z = \{x_c\}$, then $(1')_0$, $(2)_0$, $(3)_0$, and $(4)_0$ are all clear. For (ii) suppose we have $Y = \{y_0, \ldots, y_j\}$ and $Z = \{z_0, \ldots, z_j\}$ such that $(1')_i$, $(2)_i$–$(4)_i$ all hold for $i = 0, \ldots, j$ and such that $\{\gamma\}(Y) \neq 0$. Since $Y \subset X$ there is a $i' \leqslant c$ such that $y_j = x_{i'}$. Applying Lemma 5.9 (i) and $(4)_j$, there are sets $W_1, \ldots, W_{y_j+1}$ such that $\{x_{i'}, x_{i'+1}\}, W_1, \ldots, W_{y_j+1}$ is an increasing partition of $[y_j, z_j] \cap X$, and for $l = 1, \ldots, y_{j+1}$, W_l is $\omega^{\{\gamma\}(y_0, \ldots, y_j)}$-large. Let l be such that $1 \leqslant l \leqslant y_j$ and $(\max W_l, \max W_{l+1}] \cap \{\mu w\theta(p,v,w)\colon\ p < (y_j - 1)\} \neq \emptyset$. Such an l exists since there are y_j intervals of the form $(\max W_l, \max W_{l+1}]$ with $1 \leqslant l \leqslant y_j$ and the set $\{\mu w\theta(p,v,w)\colon\ p < (y_j - 1)\}$ has at most $y_j - 1$ elements. For this we use the Δ_0-pigeon-hole principle; noting that θ is a Δ_0-formula and we can code the set

$\{\max W_l\colon\ 1 \leqslant l \leqslant y_j + 1\}$. Let $y_{j+1} \stackrel{\text{def}}{=} \max W_l$ and $z_{j+1} \stackrel{\text{def}}{=} \max W_{l+1}$. $(1')_i$, $(2)_i$–$(4)_i$ all $i \leqslant j$, $(2)_{j+1}$ and $(3)_{j+1}$ are all clear. For $(4)_{j+1}$, we note that

$$\{\max W_l\} \cup W_l = [y_{j+1}, z_{j+1}] \cap X$$

is $\omega^{\{\gamma\}(y_0,\dots,y_j)} + 1$-large. For $(1')_{j+1}$ note that $W_1 \subset (y_j, \max W_1] \subseteq (y_j, y_{j+1}]$, and W_1 is $\omega^{\{\gamma\}(y_0,\dots,y_j)}$-large, so by Proposition 5.9(iii),

$$(y_j, y_{j+1}] \text{ is } \omega^{\{\gamma\}(y_0,\dots,y_j)}\text{-large.}$$

So, $G_{\omega^{\{\gamma\}(y_0,\dots,y_j)}}(y_j + 1)\downarrow$, and

$$G_{\omega^{\{\gamma\}(y_0,\dots,y_j)}}(y_j + 1) \leqslant y_{j+1}. \tag{21}$$

Using Proposition 5.9(ii), $[y_j, G_{\omega^{\{\gamma\}(y_0,\dots,y_j)}}(y_j + 1) - 1]$ is $\omega^{\{\gamma\}(y_0,\dots,y_j)}$-large, so,

$$G_{\omega^{\{\gamma\}(y_0,\dots,y_j)}}(y_j) + 1 \leqslant G_{\omega^{\{\gamma\}(y_0,\dots,y_j)}}(y_j + 1). \tag{22}$$

By Proposition 5.8(ii),

$$F_{\{\gamma\}(y_0,\dots,y_j)}(y_j) \leqslant G_{\omega^{\{\gamma\}(y_0,\dots,y_j)}}(y_j) + 1. \tag{23}$$

By (21)–(23), we have $(1')_{j+1}$, completing the proof of (ii) and the proof of the lemma. □

5.5.3. Generalizing to n unbounded quantifiers

For generalizing to n unbounded quantifiers we use the following notation. If $\psi(\vec{y}) = Q_n x_n \dots Q_1 x_1 \varphi(\vec{x}; \vec{y})$, where $\varphi \in \Delta_0$ and $Q_i \in \{\forall, \exists\}$ for all $i = 1, \dots, n$, let

$$\psi^*(\vec{y}; \vec{z}) \stackrel{\text{def}}{=} (Q_n x_n \leqslant z_n) \dots (Q_1 x_1 \leqslant z_1) \varphi(\vec{x}; \vec{y}),$$

where $z_1, \dots, z_n$ do not appear in φ.

The following lemma will allow us to generalize the proof of Section 5.4 to n unbounded quantifiers; also it is tailored for dealing with arbitrary ordinals as described in the next two sections.

Lemma 5.11. (a) *For $\beta < \varepsilon_0$ and $n \geqslant 1$ the following is provable in $I\Delta_0(\exp)$. Suppose $a \geqslant n + 1$ and $F_{\omega^\beta_{n-1}}(a) = b$; then there is a set $A = \{a_0, \dots, a_r\}$ in increasing order, and $\vec{d} = d_1, \dots, d_{n-1}$ such that*

(i) *$a_0 = a - n - 1$, $a_r \leqslant b$, and if $0 < i \leqslant r - k + 1$ and $k \leqslant a_{i-1}$ then $F_k(a_{i-1})\downarrow$ and $F_k(a_{i-1}) \leqslant a_i$;*

(ii) *A is β-large;*

(iii) *and for $i = 1, \dots, r$, if $\varphi(\vec{v}, w) \in \Sigma_{n-1} \cup \Pi_{n-1}$ and $\langle \ulcorner \varphi^* \urcorner, \vec{p} \rangle < a_{i-1} - 1$ then,*

$$(\exists w \leqslant a_i)\varphi^*(\vec{p}, w; \vec{d}) \leftrightarrow (\exists w \leqslant a_r)\varphi^*(\vec{p}, w; \vec{d}).$$

We use a $\Delta_0(\exp)$-truth definition for Δ_0-formulas to express (iii).

(b) *Furthermore, if* $M \models I\Delta_0(\exp) + (F_{\omega^\beta_{n-1}}(a) = b) + \{a > m\colon m \in \omega\}$, *then for any limit point* I *of* A *(given in* a*), any* $\psi \in \Sigma_n \cup \Pi_n$, *and any* $\vec{p} \in I$*:*

$$I \models \psi(\vec{p}) \Leftrightarrow M \models \psi^*(\vec{p}; \vec{d}, a_r);$$

Proof. We will use meta-induction on n to prove (a) and (b) of the lemma simultaneously. For (a) we work in $I\Delta_0(\exp) + (F_{\omega^\beta_{n-1}}(a) = b)$. For (b) we let $M \models I\Delta_0(\exp) + (F_{\omega^\beta_{n-1}}(a) = b)$; we suppose $a \in M - \omega$, and we suppose I is a limit point of the set A given in (a) (we note that if $\beta \geqslant \omega$ then such an I exists). Let $\theta(u, v, w)$ be a formula such that, in $I\Delta_0(\exp)$, for any $\varphi \in \Delta_0$, and for any $\vec{p}$, q, and d,

$$\theta(\langle \ulcorner\varphi\urcorner, \vec{p}\rangle, \langle\vec{d}\rangle, q) \text{ expresses } \varphi(\vec{p}, q; \vec{d}).$$

Using the formula Θ_{Δ_0} of Section 5.5.1 we may assume that θ is a $\Delta_0(\exp)$-formula.

Base step: $n = 1$. By Proposition 5.8(i), $G_{\omega^\beta_n}(a-1) \leqslant b$, so $[a-1, b]$ is ω^β-large, hence $[a-2, b]$ is $\omega^\beta+1$-large. For $X \stackrel{\text{def}}{=} [a-2, b]$ and $\theta(u, v, w)$ as given above (using $v = \Lambda$ as a parameter in θ), let $A \stackrel{\text{def}}{=} Y$ and $B \stackrel{\text{def}}{=} Z$ be given by Lemma 5.10, and let $a_i \stackrel{\text{def}}{=} y_i$ and $b_i \stackrel{\text{def}}{=} z_i$, for $i \leqslant r$. Also let $\vec{d} \stackrel{\text{def}}{=} \Lambda$. (i) and (ii) are immediate from Lemma 5.10.

The $\rightarrow$-direction of (iii) follows from $a_i \leqslant a_r$. For the $\leftarrow$-direction, suppose $\varphi \in \Delta_0$ and $\langle \ulcorner\varphi^*\urcorner, \vec{p}\rangle < a_{i-1} - 1$ where $0 < i \leqslant r$, and $(\exists w \leqslant a_r)\varphi^*(\vec{p}, w)$. Then using the definition of θ and the fact $a_r \leqslant b_i$, we have $(\exists w \leqslant b_i)\theta(\langle\ulcorner\varphi^*\urcorner, \vec{p}\rangle, \Lambda, w)$. By Lemma 5.10(3)$_i$, we have $(\exists w \leqslant a_i)\theta(\langle\ulcorner\varphi^*\urcorner, \vec{p}\rangle, \Lambda, w)$. Thus $(\exists w \leqslant a_i)\varphi^*(\vec{p}, w)$, giving us (iii).

For (b), suppose $\psi \in \Sigma_1$, say $\psi = \exists w\varphi(\vec{v}, w)$, where $\varphi \in \Delta_0$. By (i) of (a), I is closed under $+$ and $\cdot$; hence, I is a Δ_0-elementary substructure of M. Since $\varphi \in \Delta_0$ we have $\varphi = \varphi^*$, so for all $\vec{p}, q \in I$,

$$I \models \varphi(\vec{p}, q) \Leftrightarrow M \models \varphi^*(\vec{p}, q),$$

Since $I < a_r$, we have $I \models \exists w\varphi(\vec{p}, w) \Rightarrow M \models (\exists w \leqslant a_r)\varphi(\vec{p}, w)$, i.e.,

$$I \models \psi(\vec{p}) \Rightarrow M \models \psi^*(\vec{p}; a_r).$$

On the other hand, suppose $M \models \psi^*(\vec{p}; a_r)$, where $\vec{p} \in I$, i.e. $M \models (\exists w \leqslant a_r)\varphi(\vec{p}, w)$. Since $\ulcorner\varphi^*\urcorner$ is standard and $I \models I\Delta_0$, we have $\langle\ulcorner\varphi^*\urcorner, \vec{p}\rangle \in I$; hence, since I is a limit point of A, $\langle\ulcorner\varphi^*\urcorner, \vec{p}\rangle < a_{i-1} - 1$ for some $a_{i-1} \in I$. By (iii) of (a), $M \models (\exists w \leqslant a_i)\varphi^*(\vec{p}, w)$. Since I is a limit point $a_i \in I$, so $I \models \exists w\varphi(\vec{p}, w)$. We have proven (b) for $\psi \in \Sigma_1$; for $\psi \in \Pi_1$ we take the negation of both sides of the $\Leftrightarrow$. Thus (b) holds.

Induction step: Suppose $a \geqslant n+2$ and $F_{\omega^\beta_n}(a) = b$. Let $A' = \{a'_0, \dots, a'_{r'}\}$ and $\vec{d} \stackrel{\text{def}}{=} d_1, \dots, d_{n-1}$ be given by the induction hypothesis, using $\beta' = \omega^\beta$ in place of β, and noting $F_{\omega^{\beta'}_{n-1}}(a) = b$. Note A' is ω^β-large. Let $X = \{a-n-2\} \cup A'$ (so X is $\omega^\beta+1$-large), and let $d_n \stackrel{\text{def}}{=} a'_{r'}$. Apply Lemma 5.10, with θ as above using $v = \langle d_1, \dots, d_n\rangle$

as a parameter, to obtain $A \stackrel{\text{def}}{=} Y$ and $B \stackrel{\text{def}}{=} Z$, and let $a_i \stackrel{\text{def}}{=} y_i$ and $b_i \stackrel{\text{def}}{=} z_i$ for $i \leqslant r$. Then (i) and (ii) are immediate.

The $\rightarrow$-direction of (iii) follows from $a_i \leqslant a_r$. For the $\leftarrow$-direction, suppose $\varphi \in \Sigma_n \cup \Pi_n$ and $\langle \ulcorner \varphi^* \urcorner, \vec{p} \rangle < a_{i-1} - 1$ where $0 < i \leqslant r$, and $(\exists w \leqslant a_r)\varphi^*(\vec{p}, w; \vec{d})$. Then using the definition of θ and the fact $a_r \leqslant b_i$, we have $(\exists w \leqslant b_i)\theta(\langle \ulcorner \varphi^* \urcorner, \vec{p} \rangle, \langle \vec{d} \rangle, w)$. By Lemma 5.10(3)$_i$ we have $(\exists w \leqslant a_i)\theta(\langle \ulcorner \varphi^* \urcorner, \vec{p} \rangle, \langle \vec{d} \rangle, w)$. Thus $(\exists w \leqslant a_i)\varphi^*(\vec{p}, w; \vec{d})$, giving us (iii).

For (b), suppose ψ is such that $\psi(\vec{u}) = \exists w \varphi(\vec{u}, w)$ for some $\varphi \in \Pi_n$, and suppose for some $\vec{p} \in I$, $M \models (\exists w \leqslant a_r)\varphi^*(\vec{p}, w; \vec{d})$. Since $\ulcorner \varphi \urcorner$ is standard and $\vec{p} \in I$ there is a j such that $0 < j < r$ and $\langle \ulcorner \varphi^* \urcorner, \vec{p} \rangle < a_{j-1} - 1 < I$. By (a)(iii), we have $M \models (\exists w \leqslant a_j)\varphi^*(\vec{p}, w; \vec{d})$. Since I is a limit point, $a_j \in I$, so we have, using (b) in the induction hypothesis, $I \models \exists w \varphi(\vec{p}, w)$. We have shown that

$$I \models \psi(\vec{p}) \Leftarrow M \models \psi^*(\vec{p}; \vec{d}, a_r).$$

For the $\Rightarrow$-direction suppose $I \models \psi(\vec{p})$, i.e. suppose $I \models \exists w \varphi(\vec{p}, w)$. Since $I < a_r$ we clearly have, using (b) in the induction hypothesis, $M \models (\exists w \leqslant a_r)\varphi^*(\vec{p}, w; \vec{d})$, or equivalently $M \models \psi^*(\vec{p}; \vec{d}, a_r)$, completing the proof, for all $\psi \in \Sigma_{n+1}$, of,

$$I \models \psi(\vec{p}) \Leftrightarrow M \models \psi^*(\vec{p}; \vec{d}, a_r).$$

To get this for $\psi \in \Pi_{n+1}$, simply take the negation of both sides of the $\Leftrightarrow$. This completes the proof of (b), and hence the proof of the induction step and the proof of the lemma. $\square$

5.5.4. Remarks on the proof of the $\Rightarrow$-direction of Theorem 5.2

We will apply Lemma 5.11 as follows. Suppose

$$M \models I\Delta_0(\exp) + (F_{\omega_n^{\alpha+1}}(a) = b) + \{a > m\colon\ m \in \omega\};$$

in our metatheory, we can easily show that such an M exists. Let A and $\vec{d}$ be given by Lemma 5.11 using $\beta = \omega^{\alpha+1}$. A is $\omega^{\alpha+1}$-large, so $A - \{a_0\}$ is $\omega^\alpha \cdot (a_0 + 1)$-large. By Proposition 5.9(iv) there is an increasing partition $A_1, \ldots, A_{a_0+1}$ of $A - \{a_0\}$ such that for $j = 1, \ldots, a_0 + 1$, A_j is ω^α-large. Let $i_1, \ldots, i_{a_0+1}$ be such that $a_{i_j} = \min A_j$, and let,

$$I \stackrel{\text{def}}{=} \bigcup_{j \in \omega} a_{i_j}. \tag{24}$$

Since $a < I < b = F_{\omega_n^{\alpha+1}}(a)$ we have $I \not\models F_{\omega_n^{\alpha+1}}\downarrow$. In Section 5.6.2 it is shown that $I \models TI(\omega^{\omega^\alpha}, \Pi_n)$. Before proceeding to that task we point out that for the $\alpha = 0$ case, where we only need $I \models I\Sigma_n$, the result is easily obtained. By (b) of Lemma 5.11:

$$I \models \psi(\vec{p}) \Leftrightarrow M \models \psi^*(\vec{p}; \vec{d}, a_r).$$

Hence the least number principle for Δ_0-formulas in M yields the least number principle for Σ_n-formulas in I.

5.6. *Dealing with arbitrary ordinals*

To see how we will deal with arbitrary ordinals consider the limit point I defined in (24). Lemma 5.11 implies that with parameters far enough below a_i the least witness, in I, to a formula, is below a_{i+1}. If $I \models \varphi(p,s)$ where we view s as a sequence of natural numbers, we may find the least e_1 such that $I \models \exists s' \varphi(p, \langle e_1 \rangle * s')$. A careful application of the lemma will show that if p is far enough below a_i, then $\langle p, e_1 \rangle$ will be far enough below a_{i+1} to repeat the argument to obtain the least e_2 such that $I \models \exists s' \varphi(p, \langle e_1, e_2 \rangle * s')$ and such that $\langle p, e_1, e_2 \rangle$ is far enough below a_{i+2}, and so on. If I is sufficiently long, i.e. if $a_{i+k} \in I$ for a sufficiently large value of k, then the construction may terminate yielding $e_1, \ldots, e_k \in I$ such that $I \models \varphi(p, \langle e_1, \ldots, e_k \rangle)$, and for all $i = 1, \ldots, k$,

$$I \models e_i = \mu e(\exists s' \varphi(p, \langle e_1, \ldots, e_{i-1}, e \rangle * s')).$$

$\widehat{s} \stackrel{\text{def}}{=} \langle e_1, \ldots, e_k \rangle$ will then be the $<_L$-least sequence such that $I \models \varphi(p, \widehat{s})$ (the lexicographic order, $<_L$, is defined in Definition 5.22). In Section 4.5 we showed that $TI(\omega^c, \Pi_n)$ is equivalent to the least sequence principle for sequences of length c on Σ_n-formulas. In this section we show that $TI(\omega^{\omega^\gamma}, \Pi_n)$ is equivalent to the least sequence principle for "exactly ω^γ-large sequences" on Σ_n-formulas. Using I as given in (24) we will show, for $\varphi \in \Sigma_n$, that the above construction can be carried out to find the $<_L$-least sequence s such that

$$I \models \varphi(p,s) \wedge \text{"}s \text{ is exactly } \omega^\alpha\text{-large"}.$$

This will yield $I \models TI(\omega^{\omega^\alpha}, \Pi_n)$.

5.6.1. *Ordinals less than ω^{ω^γ} coded by exactly ω^γ-large sequences*

Let $S_\alpha \stackrel{\text{def}}{=} \{s\colon\ s \text{ is an exactly } \alpha\text{-large sequence}\}$.

Theorem 5.12. *The following is provable in $I\Delta_0 + F_3 \downarrow$. For all $\gamma < \varepsilon_0$ there is a total function H_γ such that*

$$(S_{\omega^\gamma}, <_L) \underset{H_\gamma}{\cong} (\omega^{\omega^\gamma}, <).$$

That is $I\Delta_0 + F_3 \downarrow$ proves H_γ is an order preserving bijection from $(S_{\omega^\gamma}, <_L)$ to $(\omega^{\omega^\gamma}, <)$, where $<_L$ is the lexicographic order (defined in Definition 5.22) and $<$ is the ordinal less-than relation.

The proof of this theorem will follow several lemmas. First we give a lemma that will motivate our definition of H_γ. For each fixed α, this lemma describes a kind of normal form for the ordinals less than $\omega^{\omega^{\alpha+1}}$.

Lemma 5.13. *The following is provable in $I\Delta_0(\exp)$. For all $\alpha, \gamma < \varepsilon_0$, if $\omega^{\omega^\alpha} \leqslant \gamma < \omega^{\omega^{\alpha+1}}$ then γ can be expressed uniquely in the form*

$$\gamma = \omega^{\omega^\alpha \cdot k} \cdot (1 + \delta_k) + \omega^{\omega^\alpha \cdot (k-1)} \cdot \delta_{k-1} + \cdots + \omega^{\omega^\alpha \cdot 1} \cdot \delta_1 + \omega^{\omega^\alpha \cdot 0} \cdot \delta_0,$$

where $k \geqslant 1$ and $\delta_0, \ldots, \delta_k < \omega^{\omega^\alpha}$. In particular, if $\omega^{\omega^\alpha} \leqslant \gamma < \omega^{\omega^{\alpha+1}}$ then γ is uniquely identified with a sequence of ordinals $\delta_k, \delta_{k-1}, \ldots, \delta_0$, where $\delta_i < \omega^{\omega^\alpha}$. "$1 + \delta_k$" is used in place of "$\delta_k$" to ensure that the identification is bijective.

Proof. Suppose $\omega^{\gamma_1} \cdot n_1 + \cdots + \omega^{\gamma_l} \cdot n_l$ is the Cantor normal form of γ. Then $\omega^{\alpha+1} > \gamma_i$, for $i = 1, \ldots, l$. By O_{31} of Section 3 for each $i = 1, \ldots, l$ there is a k_i such that $\gamma_i < \{\omega^{\alpha+1}\}(k_i) = \omega^\alpha \cdot (k_i + 1)$. Using O_{32} and induction on k_i, we may assume k_i is least; i.e. we may assume that $\omega^\alpha \cdot k_i \leqslant \gamma_i < \omega^\alpha \cdot (k_i + 1)$. Using O_3, for each $i = 1, \ldots, l$ there is a unique $\gamma_i' < \omega^\alpha$ such that $\gamma_i = \omega^\alpha \cdot k_i + \gamma_i'$. Let $k \stackrel{\text{def}}{=} k_1$, and for each $j = -1, 0, 1, \ldots, k$ let,

$$i_j \stackrel{\text{def}}{=} \begin{cases} l+1 & \text{if } j = -1; \\ \text{the least } i \text{ such that } (k_i = j) & \text{if such an } i \text{ exists;} \\ i_{j-1} & \text{otherwise.} \end{cases}$$

By O_{27}, O_{28}, and induction,

$$\begin{aligned} \gamma = {} & \omega^{\omega^\alpha \cdot k} \cdot (\omega^{\gamma'_{i_k}} \cdot n_{i_k} + \cdots + \omega^{\gamma'_{i_{k-1}-1}} \cdot n_{i_{k-1}-1}) + \cdots \\ & + \omega^{\omega^\alpha \cdot 1} \cdot (\omega^{\gamma'_{i_1}} \cdot n_{i_1} + \cdots + \omega^{\gamma'_{i_0-1}} \cdot n_{i_0-1}) \\ & + \omega^{\omega^\alpha \cdot 0} \cdot (\omega^{\gamma'_{i_0}} \cdot n_{i_0} + \cdots + \omega^{\gamma'_l} \cdot n_l). \end{aligned}$$

For $j = 0, \ldots, k-1$, let $\delta_j \stackrel{\text{def}}{=} \omega^{\gamma'_{i_j}} \cdot n_{i_j} + \cdots + \omega^{\gamma'_{i_{j-1}-1}} \cdot n_{i_{j-1}-1}$, and let $\delta_k \stackrel{\text{def}}{=}$ the unique δ such that $1 + \delta = \omega^{\gamma'_{i_k}} \cdot n_{i_k} + \cdots + \omega^{\gamma'_{i_{k-1}-1}} \cdot n_{i_{k-1}-1}$ (the existence of such a δ is guaranteed by O_3). Clearly $\delta_j < \omega^{\omega^\alpha}$ for all $j = 0, \ldots, k$, and $\gamma = \omega^{\omega^\alpha \cdot k} \cdot (1 + \delta_k) + \omega^{\omega^\alpha \cdot (k-1)} \cdot \delta_{k-1} + \cdots + \omega^{\omega^\alpha} \cdot \delta_1 + \omega^{\omega^\alpha \cdot 0} \cdot \delta_0$. The uniqueness of k is clear since k is the least integer such that $\gamma < \omega^{\omega^\alpha \cdot (k+1)}$. Now suppose,

$$\begin{aligned} & \omega^{\omega^\alpha \cdot k} \cdot (1 + \delta_k) + \omega^{\omega^\alpha \cdot (k-1)} \cdot \delta_{k-1} + \cdots + \omega^{\omega^\alpha} \cdot \delta_1 + \delta_0 \\ & = \omega^{\omega^\alpha \cdot k} \cdot (1 + \delta'_k) + \omega^{\omega^\alpha \cdot (k-1)} \cdot \delta'_{k-1} + \cdots + \omega^{\omega^\alpha} \cdot \delta'_1 + \delta'_0. \end{aligned} \tag{25}$$

Let i be greatest such that $\delta_i \neq \delta'_i$, say $\delta_i < \delta'_i$. Using O_{23}, $\omega^{\omega^\alpha \cdot i} \cdot \delta_i < \omega^{\omega^\alpha \cdot i} \cdot \delta'_i$; so,

$$\begin{aligned} & \omega^{\omega^\alpha \cdot k} \cdot (1 + \delta_k) + \omega^{\omega^\alpha \cdot (k-1)} \cdot \delta_{k-1} + \cdots + \omega^{\omega^\alpha} \cdot \delta_1 + \delta_0 \\ & < \omega^{\omega^\alpha \cdot k} \cdot (1 + \delta'_k) + \omega^{\omega^\alpha \cdot (k-1)} \cdot \delta'_{k-1} + \cdots + \omega^{\omega^\alpha} \cdot \delta'_1 + \delta'_0 \end{aligned}$$

contradicting (25). Thus $\delta_i = \delta'_i$ all $i = 0, \ldots, k$; i.e. $\delta_0, \ldots, \delta_k$ are unique. □

We now define the H_α's of Theorem 5.12, first in the metatheory; then we sketch how this can be carried out in $I\Delta_0 + (F_3 \downarrow)$.

Definition 5.14. For $\alpha < \varepsilon_0$, let H_α, defined on $\{\gamma\colon\ \gamma < \omega^{\omega^\alpha}\}$, be given by

(a) $H_0(\gamma) = \langle \mathrm{num}(\gamma) \rangle$, all $\gamma < \omega$.

$$\text{(b)}\qquad H_{\beta+1}(\gamma) = \begin{cases} \langle 0 \rangle * H_\beta(\gamma) & \text{if } \gamma < \omega^{\omega^\beta} \\ \langle k \rangle * H_\beta(\delta_k) * \cdots * H_\beta(\delta_0) & \text{if } \omega^{\omega^\beta} \leqslant \gamma < \omega^{\omega^{\beta+1}}, \\ & \text{where } k, \delta_0, \ldots, \delta_k \text{ are} \\ & \text{given by Lemma 5.13.} \end{cases}$$

(c) For λ a limit ordinal,

$$H_\lambda(\gamma) = \begin{cases} \langle 0 \rangle * H_{\{\lambda\}(0)}(\gamma) & \text{if } \gamma < \omega^{\omega^{\{\lambda\}(0)}} \\ \langle k \rangle * H_{\{\lambda\}(k)}(\delta) & \text{if } \omega^{\omega^{\{\lambda\}(k-1)}} \leqslant \gamma < \omega^{\omega^{\{\lambda\}(k)}}, \\ & \text{where } \delta, \text{ given by } \mathrm{O}_3, \text{ is} \\ & \text{such that } \gamma = \omega^{\omega^{\{\lambda\}(k-1)}} + \delta. \end{cases}$$

We set $H_\alpha(\gamma) \overset{\text{def}}{=} 0$ if $\gamma \geqslant \omega^{\omega^\alpha}$.

We note that in every case of the definition, if $\gamma < \omega^{\omega^\alpha}$, then $H_\alpha(\gamma)$ is of the form $\langle k \rangle * s$, where k is least such that $\gamma < \{\omega^{\omega^\alpha}\}(k)$, and $s = \Lambda$ or $s = H_{\{\alpha\}(k)}(\gamma_0) * \cdots * H_{\{\alpha\}(k)}(\gamma_{l-1})$ for some $\gamma_0, \ldots, \gamma_{l-1} < \omega^{\omega^{\{\alpha\}(k)}}$.

Viewing H as a binary function on ordinals, H is given by transfinite recursion (noting that $\{\alpha\}(k) < \alpha$ and $\gamma_i < \gamma$ for $i = 0, \ldots, l-1$). Below we sketch how the transfinite recursion used here is of a form that can be carried out in $I\Delta_0 + (F_3\downarrow)$. Further we will show that there is a form of transfinite induction provable in $I\Delta_0 + (F_3\downarrow)$ that can be used to prove properties of H_α. We note that this is a variation of techniques applied in [10, 8]. Only a sketch of the key ideas used to prove this result are given; this is worked out in detail in [19]. The following finite set of ordinals will be useful for this purpose as well as in some later applications.

Definition 5.15. Let $\mathrm{O}(h,c) \overset{\text{def}}{=} \{\alpha\colon\ h(\alpha) \leqslant h \text{ and } c_{\max}(\alpha) \leqslant c\}$. Compare Section 3.3 for the definitions of the functions h and $c_{\max}$. We note that $\gamma \in \mathrm{O}(h,c)$ is equivalent to $h(\gamma) \leqslant h \wedge c_{\max}(\gamma) \leqslant c$, and hence is a Δ_0-formula.

Lemmas below imply that for $\alpha \in \mathrm{O}(h, c+1)$ and $\gamma \in \mathrm{O}(h,c)$, H_γ is definable using "recursion in $\mathrm{O}(h,c+1)$"; in particular, Lemma 5.16 shows that $\{\alpha\}(k) \in \mathrm{O}(h,c+1)$ and $\gamma_i \in \mathrm{O}(h,c)$ for $i = 0, \ldots, l-1$ (where k and the γ_i's are from the definition of H_α). Further, according to Lemma 5.17, $\mathrm{O}(h,c)$ has cardinality bounded by a $\Delta_0(F_3)$-term; so there is a $\Delta_0(F_3)$-enumeration of finite sets of ordinals (the $\mathrm{O}(h,c)$'s) such that whenever α and γ are in the enumeration and $H_{\alpha'}(\gamma')$ is involved in the definition of $H_\alpha(\gamma)$, then α' and γ' appear earlier in the enumeration than α and γ. These facts can be used to show that there is a $\Delta_0(F_3)$-definition of H. Next the lemmas just mentioned are given.

Lemma 5.16. ($I\Delta_0 + \forall x \exists y (x^{\log x} = y)$)

(1) $\mathrm{LIM}(\lambda), \{\lambda\}(k) \in \mathrm{O}(h,c) \Rightarrow k < c$.

(2) $\gamma \in \mathrm{O}(h,c)$, $(\{\alpha\}(k) \leqslant \gamma < \{\alpha\}(k+1)) \Rightarrow k < c$.
(3) $k < c$, $\gamma \in \mathrm{O}(h,c) \Rightarrow \{\gamma\}(k) \in \mathrm{O}(h,c)$.

The proof of the above lemma is omitted; it uses induction on h and is included in detail in [19]. Next, bounds on the cardinality of $\mathrm{O}(h,c)$ and on the size of the codes of elements in $\mathrm{O}(h,c)$ are given. Let

$$n(h,c) \stackrel{\text{def}}{=} \text{the cardinality of the set } \mathrm{O}(h,c), \tag{26}$$

and

$$l(h,c) \stackrel{\text{def}}{=} \max\{|\ulcorner\gamma\urcorner|_4 : \gamma \in \mathrm{O}(h,c)\}.$$

For all c, $\mathrm{O}(0,c) = \{0\}$, so we have $n(0,c) = 1$ and $l(h,c) = 0$. For all $\gamma \in \mathrm{O}(h+1,c)$ there are $\gamma_1,\ldots,\gamma_k \in \mathrm{O}(h,c)$ and $n_1,\ldots,n_k \leqslant c$ such that

$$\overline{\ulcorner\gamma\urcorner}^4 = 4 * \overline{\ulcorner\gamma\urcorner_1}^4 * 3 * \overline{n_1}^2 * \cdots * 4 * \overline{\ulcorner\gamma\urcorner_k}^4 * 3 * \overline{n_k}^2.$$

Thus
- $l(h+1,c) \leqslant (2 + l(h,c) + |c|_2) \cdot n(h,c)$, and
- $n(h+1,c) \leqslant (c+1)^{n(h,c)}$.

We may obtain, by induction on h, if $h > 0$,
- $n(h,c) \leqslant \underbrace{(c+1)^{(c+1)^{\cdot^{\cdot^{\cdot(c+1)}}}}}_{h\ (c+1)\text{'s}}$, and
- $l(h,c) \leqslant \underbrace{(3c)^{(3c)^{\cdot^{\cdot^{\cdot(3c)}}}}}_{h\ (3c)\text{'s}}$.

The first bound is clear. The second bound is much more than adequate; note that $l(h+1,c)$ is a little more than $l(h,c) \cdot n(h,c)$ which, as long as $c \geqslant 1$, is less than the bound for $l(h,c)$ cubed, and certainly $B^3 \leqslant (3c)^B$.

Next we claim that for $k \geqslant 1$,

$$\underbrace{m^{m^{\cdot^{\cdot^{\cdot m}}}}}_{k\ m\text{'s}} \leqslant F_3(m \cdot k). \tag{27}$$

This is easy to prove in $I\Delta_0 + F_3\downarrow$ (we note that the bound on the right is excessive). To obtain this it is useful to note that

$$\underbrace{2^{2^{\cdot^{\cdot^{\cdot 2^x}}}}}_{x\ 2\text{'s}} \leqslant F_3(x),$$

which is provable in $I\Delta_0 + F_3\downarrow$. We will not fill in the details of the proof of (27). Using this bound it is straightforward to prove:

Lemma 5.17. ($I\Delta_0 + F_3\downarrow$) *The functions n and l are both total and each is bounded by an $L_0(F_3)$-term.*

We now describe the restricted form of ordinal induction mentioned prior to Lemma 5.16, in particular induction on the ordinals in the finite set $\mathrm{O}(h,c)$ for fixed values of h and c. The idea for this comes from Paris in [10].

Proposition 5.18. *For all $\varphi \in \Delta_0$, the following is provable in $I\Delta_0 + F_3 \downarrow$:*

$$(\forall \delta \in \mathrm{O}(h,c))((\forall \gamma \in \mathrm{O}(h,c) \cap \delta)\varphi(\gamma) \to \varphi(\delta)) \to (\forall \gamma \in \mathrm{O}(h,c))\varphi(\gamma). \tag{28}$$

We use the phrase "induction on γ in $\mathrm{O}(h,c)$ for $\varphi(\gamma)$" to denote (28).

Proof. Using the totality of the functions l and n, given by (26), we can show, in $I\Delta_0 + (F_3 \downarrow)$, that for each h and c there is a sequence $\langle \gamma_1, \ldots, \gamma_{n(h,c)} \rangle$ such that $\mathrm{O}(h,c) = \{\gamma_i : 1 \leqslant i \leqslant n(h,c)\}$, and for $i = 2, \ldots, n(h,c)$, we have $\gamma_{i-1} < \gamma_i$. Note that

$$(\forall \delta \in \mathrm{O}(h,c))((\forall \gamma \in \mathrm{O}(h,c) \cap \delta)\varphi(\gamma) \to \varphi(\delta)) \tag{29}$$

is then equivalent to $(\forall i \leqslant n(h,c))((\forall j < i)\varphi(\gamma_j) \to \varphi(\gamma_i))$; so by ordinary induction, noting that for a fixed h and c the formula $\varphi(\gamma_j)$ is a Δ_0-formula, if (29) then $(\forall i \leqslant n(h,c))\varphi(\gamma_i)$, which is equivalent to $(\forall \gamma \in \mathrm{O}(h,c))\varphi(\gamma)$. This yields the proposition. □

In light of Lemmas 5.16 and 5.17 and the discussion preceding the statement of Lemma 5.16, the following is straightforward. The proof uses ideas similar to ideas in the proof of the previous proposition.

Proposition 5.19. *The function $(\alpha, \gamma) \mapsto H_\alpha(\gamma)$ is $I\Delta_0(F_3)$-provably recursive.*

Lemma 5.20. *The following is provable in $I\Delta_0 + F_3 \downarrow$. If $\gamma < \omega^{\omega^\alpha}$ then $H_\alpha(\gamma)$ is exactly ω^α-large.*

Proof. We will use induction on $\alpha \in \mathrm{O}(h,c+1)$ to show that for any $\gamma \in \mathrm{O}(h,c)$, $H_\alpha(\gamma)$ is exactly ω^α-large. For the $\alpha = 0$ case, $H_0(\gamma) = \langle \mathrm{num}(\gamma) \rangle$ is clearly exactly 1-large, i.e. exactly ω^0-large.

If $\alpha = \beta + 1$, then $H_\alpha(\gamma) = \langle k \rangle * H_\beta(\delta_k) * \cdots * H_\beta(\delta_0)$ for some $k \in \omega$ and $\delta_0, \ldots, \delta_k < \omega^{\omega^\beta}$ where $\delta_0, \ldots, \delta_k \in \mathrm{O}(h,c)$. Applying the induction hypothesis we have $H_\beta(\delta_0), \ldots, H_\beta(\delta_k)$ are each exactly ω^β-large; applying Proposition 5.9(iv), we have $H_\beta(\delta_0) * \cdots * H_\beta(\delta_k)$ is exactly $\omega^\beta \cdot (k+1)$-large, and hence $H_{\beta+1}(\gamma)$ is exactly $\omega^{\beta+1}$-large.

If $\alpha = \lambda$ is a limit ordinal then $H_\alpha(\gamma) = \langle k \rangle * H_{\{\lambda\}(k)}(\delta)$ for some $k \in \omega$ and $\delta < \omega^{\omega^{\{\lambda\}(k)}}$ where $\delta \in \mathrm{O}(h,c)$. By Lemma 5.16 we have $k \leqslant c$ and $\{\lambda\}(k) \in \mathrm{O}(h,c+1)$, so by the induction hypothesis $H_{\{\lambda\}(k)}(\delta)$ is exactly $\omega^{\{\lambda\}(k)}$-large, and thus $H_\alpha(\gamma)$ is exactly ω^α-large. □

Lemma 5.21. $I\Delta_0 + F_3 \downarrow \vdash H_\alpha$ *is a bijection from $\{\gamma : \gamma < \omega^{\omega^\alpha}\}$ to S_{ω^α}.*

Proof. Applying Proposition 5.19 and Lemma 5.20, it is enough to show that H_α is one-to-one and onto. We will use induction on α in $\mathrm{O}(h,c+1)$ to show that for all $l \leqslant n(h,c)$, and all $n_1,\dots,n_l \leqslant c$, if $(n_1,\dots,n_l)$ is exactly ω^α-large, then there is a unique $\gamma \in \mathrm{O}(h+2,c+l) \cap \omega^{\omega^\alpha}$ such that $H_\alpha(\gamma) = \langle n_1,\dots,n_l\rangle$. In proving $H_\alpha(\gamma)$ is unique, restricting γ to $\mathrm{O}(h+2,c+l)$ is no problem since if $H_\alpha(\gamma_1) = H_\alpha(\gamma_2)$, we can pick h and c such that $\alpha \in \mathrm{O}(h,c+1)$ and $\gamma_1,\gamma_2 \in \mathrm{O}(h+2,c)$. Also we note that, as proven earlier, l is bounded by an $L_0(F_3)$-term in h and c (since $n(h,c)$ is), and so $\langle n_1,\dots,n_l\rangle$ is bounded by an $L_0(F_3)$-term in h and c. Also, as noted earlier, a code for the set $\mathrm{O}(h,c)$ is bounded by an $L_0(F_3)$-term in h and c. Thus we are doing induction on a $\Delta_0(F_3)$-formula which is allowed in $I\Delta_0 + F_3\downarrow$.

We divide into three cases according to whether α is 0, a successor, or a limit ordinal. If $\alpha = 0$ then we must have $l = 1$. Clearly $H_\alpha(\gamma) = \langle n_1\rangle$ if and only if $\gamma = \mathrm{ord}(n_1)$. Also $\gamma \in \mathrm{O}(1,n_1) \subseteq \mathrm{O}(h+2,c+1) \cap \omega^{\omega^0}$. This completes this case.

If $\alpha = \beta + 1$, we note that $n_2,\dots,n_l$ is exactly $\omega^\beta \cdot (n_1+1)$-large; by Proposition 5.9(iv) there are unique integers $i_0,\dots,i_{n_1+1}$ such that $1 = i_0 < \cdots < i_{n_1+1} = l$ and for $j \leqslant n_1$, $(n_{i_j+1},\dots,n_{i_{j+1}})$ is exactly ω^β-large. By the induction hypothesis there are unique $\delta_0,\dots,\delta_{n_1} < \omega^{\omega^\beta}$ such that $H_\beta(\delta_j) = \langle n_{i_j+1},\dots,n_{i_{j+1}}\rangle$, for $j = 0,\dots,n_1$. Let $k = n_1$, and let,

$$\gamma \stackrel{\text{def}}{=} \begin{cases} \delta_0 & \text{if } n_1 = 0;\\ \omega^{\omega^\beta\cdot k}\cdot(1+\delta_0) + \omega^{\omega^\beta\cdot(k-1)}\cdot\delta_1 + \cdots & \\ +\omega^{\omega^\beta\cdot 1}\cdot\delta_{k-1} + \omega^{\omega^\beta\cdot 0}\cdot\delta_k & \text{otherwise.}\end{cases}$$

Clearly $H_\beta(\gamma) = \langle n_1,\dots,n_l\rangle$. Also $\gamma < \omega^{\omega^{\beta+1}}$, so we need to show that $\gamma \in \mathrm{O}(h+2,c+l)$. Since $\alpha \in \mathrm{O}(h,c+1)$ we have $\beta \in \mathrm{O}(h,c+1)$. By the induction hypothesis $\delta_j \in \mathrm{O}(h+2,c+i_{j+1}-i_j) \subseteq \mathrm{O}(h+2,c+l-1)$, for $j = 0,\dots,n_1$. From these facts it is clear that $\gamma \in \mathrm{O}(h+2,c+l)$. Uniqueness follows from the uniqueness of the i_j's and of the δ_j's.

If $\alpha = \lambda$ is a limit ordinal then $n_2,\dots,n_l$ is exactly $\omega^{\{\lambda\}(n_1)}$-large. Since $n_1 \leqslant c$, by Lemma 5.16, we have $\{\lambda\}(n_1) \in \mathrm{O}(h,c+1)$, so by the induction hypothesis there is a unique $\delta \in \omega^{\omega^{\{\lambda\}(n_1)}} \cap \mathrm{O}(h+2,c+l-1)$ such that $H_{\{\lambda\}(n_1)}(\delta) = \langle n_2,\dots,n_l\rangle$. Let

$$\gamma \stackrel{\text{def}}{=} \begin{cases} \delta & \text{if } n_1 = 0,\\ \omega^{\omega^{\{\lambda\}(n_1-1)}} + \delta & \text{otherwise.}\end{cases}$$

By O_{31}, $\gamma < \omega^{\omega^{\{\lambda\}(n_1)}} < \omega^{\omega^\lambda}$ and if $n_1 \neq 0$ then $\omega^{\omega^{\{\lambda\}(n_1-1)}} \leqslant \gamma$. So clearly $H_\lambda(\gamma) = \langle n_1,\dots,n_l\rangle$. Since $\delta \in \mathrm{O}(h+2,c+l-1)$ and $\omega^{\omega^{\{\lambda\}(n_1-1)}} \in \mathrm{O}(h+2,c+1)$, we clearly have $\gamma \in \mathrm{O}(h+2,c+l)$. By O_3 and the induction hypothesis, γ is unique. $\square$

Definition 5.22. For $a_1,\dots,a_k,b_1,\dots,b_l \in \omega$, let $a_1 * \cdots * a_k <_L b_1 * \cdots * b_l$ if (def) either:

(a) for some $i < \min\{k,l\}$: $a_j = b_j$ (all $j = 1,\dots,i$), and $a_{i+1} < b_{i+1}$; or,

(b) $k < l$ and $a_1 * \cdots * a_k \subsetneq_{\mathrm{e}} b_1 * \cdots * b_l$,

where $<$ agrees with the usual less than on natural numbers.

Lemma 5.23. *$I\Delta_0 \vdash H_\alpha : (\{\gamma\colon \gamma < \omega^{\omega^\alpha}\}, <) \to (S_{\omega^\alpha}, <_L)$ is order preserving.*

Proof. We will show by induction on α in $\mathrm{O}(h, c+1)$ that if $\gamma_1, \gamma_2 \in \mathrm{O}(h,c)$, $\gamma_1 < \gamma_2 < \omega^{\omega^\alpha}$, $H_\alpha(\gamma_1) = \langle n_1, \dots n_{k_1}\rangle$, and $H_\alpha(\gamma_2) = \langle m_1, \dots, m_{k_2}\rangle$, then for some $i \leqslant k_1$, $n_1 = m_1, \dots, n_{i-1} = m_{i-1}$ and $n_i < m_i$. This is clearly a $\Delta_0(F_3)$-induction. If $\alpha = 0$ then $\mathrm{num}(\gamma_1) < \mathrm{num}(\gamma_2)$ so clearly $n_1 < m_1$ and we are done.

If $\alpha = \beta + 1$ then for some $\delta_0, \dots, \delta_{n_1}, \delta'_0, \dots, \delta'_{m_1}$, we have $H_\alpha(\gamma_1) = \langle n_1\rangle * H_\beta(\delta_0) * \cdots * H_\beta(\delta_{n_1})$ and $H_\alpha(\gamma_2) = \langle m_1\rangle * H_\beta(\delta'_0) * \cdots * H_\beta(\delta'_{m_1})$. If $n_1 < m_1$ the result is clear; otherwise, since $\gamma_1 < \gamma_2$, there is an $i < n_1$ such that $\delta_0 = \delta'_0, \dots, \delta_{i-1} = \delta'_{i-1}$ and $\delta_i < \delta'_i$, in which case the result follows from the induction hypothesis.

If $\alpha = \lambda$ is a limit ordinal then $H_\lambda(\gamma_1) = \langle n_1\rangle * H_{\{\lambda\}(n_1)}(\delta)$ for some $\delta < \omega^{\omega^{\{\lambda\}(n_1)}}$, and $H_\lambda(\gamma_2) = \langle m_1\rangle * H_{\{\lambda\}(m_1)}(\delta')$, for some $\delta' < \omega^{\omega^{\{\lambda\}(m_1)}}$. If $n_1 < m_1$ then we are done. Otherwise, since $\gamma_1 < \gamma_2$, we have $n_1 = m_1$ and so, by O_{22}, $\delta < \delta'$. Also by Lemma 5.16, $\{\lambda\}(m_1) \in \mathrm{O}(h,c)$, and hence by the induction hypothesis, we are done. $\square$

Proof of Theorem 5.12. The theorem follows directly from Lemmas 5.21 and 5.23. $\square$

5.6.2. The $\Rightarrow$-direction of Theorem 5.2

Motivated by the results of the last section, we use results of Section 5.5 to prove the following lemma. Recall that our model-theoretic constructions take place in a strong metatheory such as ZFC. The statement of the lemma is followed by an explanation of some of the intuition behind the lemma.

Lemma 5.24. (a) *For $\alpha < \varepsilon_0$ and $n \geqslant 1$, there is a polynomial (i.e. an L_0-term) P and an $m \in \omega$, such that the following is provable in $I\Delta_0(\exp)$. Suppose $F_{\omega_{n-1}^{(\omega^\alpha \cdot t)}}(a) = b$ and $a \geqslant m + n + 1$; then there is a set $C = \{c_1, \dots, c_t\}$, in increasing order, and $d_1, \dots, d_n$, such that,*

(i) *$a - n - 1 \leqslant c_1$, $c_t \leqslant b$, and if $0 < i \leqslant t - k + 1$ and $k \leqslant c_{i-1}$ then $F_k(c_{i-1})\downarrow$ and $F_k(c_{i-1}) \leqslant c_i$, and*

(ii) *supposing j, $\vec{p}$, and $\theta(\vec{u}, v, w)$, are such that*

$$1 \leqslant j \leqslant t-2,\ \ \theta \in \Pi_{n-1},\ \ \text{and } P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle) \leqslant c_j; \tag{30}$$

and supposing that for some s there is a w such that,

$$P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s * \langle w\rangle) \leqslant d_n, \tag{31}$$

s is exactly ω^α-large, and

$$\theta^*(\vec{p}, s, w; d_1, \dots, d_{n-1}),$$

then the $<_L$-least s such that for some w (31) holds exists, is less than c_{j+2}, and is witnessed by a w that is also less than c_{j+2}.

We use a $\Delta_0(\exp)$*-truth definition for* Δ_0*-formulas to express* (ii).

(b) *Furthermore, if* $M \models I\Delta_0(\exp) + (F_{\omega_{n-1}^{\omega^\alpha \cdot t}}(a) = b) + \{a > l:\ l \in \omega\}$, *then for any limit point* I *of* C *(given in* a*), any* $\psi \in \Sigma_n \cup \Pi_n$, *and any* $\vec{p} \in I$:

$$I \models \psi(\vec{p}) \ \Leftrightarrow\ M \models \psi^*(\vec{p}; d_1, \ldots, d_n),$$

where $d_1, \ldots, d_n$ *are given in clause* (a) *of the lemma.*

The polynomial P is used to express "the parameters, $\vec{p}$, are sufficiently far below c_j" (in (30)) and "$\vec{p}$, s, and w are sufficiently far below d_n" (in (31)) where $\vec{c}$ is a sequence that will be used to define a limit point, I, that is a model of $TI(\omega^{\omega^\alpha}, \Pi_n)$ (similar to the role of $\vec{a}$ in Lemma 5.6) and $\vec{d}$ is used for bounding quantifiers in Π_n formulas (to show that Π_n-truth in I can be tested by a Δ_0-formula in M). Usually a will be taken to be nonstandard; however, there are finitistic counterparts of the lemma where the standard lower bound $m + n + 1$ plays an important role (cf. Section 6.4).

Proof of Lemma 5.24. Using $\beta = \omega^\alpha \cdot t$, let $A = \{a_0, \ldots, a_r\}$, and $\vec{d} = d_1, \ldots, d_{n-1}$, be given by Lemma 5.11; let $d_n \overset{\text{def}}{=} a_r$. Then A is $\omega^\alpha \cdot t$-large, and hence by Proposition 5.9(iv) there is an increasing partition $A_1, \ldots, A_t$ of A such that for $j = 1, \ldots, t$, A_j is ω^α-large. Let $c_j \overset{\text{def}}{=} \min A_j$, for $j = 1, \ldots, t$. Then (a)(i) is clear. Also any limit point of C is a limit point of A as well, so (b) follows from Lemma 5.11(b).

We have left to prove (a)(ii). We suppose P is a polynomial such that,

$$P(x) \geqslant \langle x, x \rangle + 2.$$

In fact, using Proposition 2.1, it is clear that we can take $P(x) = (x+2)^k$ for some standard k; thus P is standard. Furthermore, x is easily recovered from $P(x)$. Let ψ be a formula such that

$$\psi(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s', P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s * \langle w \rangle))$$

expresses:

$$\text{"}\theta \in \Pi_{n-1}\text{"} \wedge \theta(\vec{p}, s, w) \wedge \text{"}s \text{ is exactly } \omega^\alpha\text{-large"} \wedge s' \subsetneq_{\mathrm{e}} s.$$

Using a $\Delta_0(\exp)$-truth definition for Δ_0-formulas, it is not difficult to show that, over $I\Delta_0(\exp)$, if $n = 1$ then ψ is equivalent to a $\Delta_0(\exp)$-formula, and if $n > 1$, then ψ is equivalent to a Π_{n-1}-formula. Furthermore, if $n > 1$ we can have the $n-1$ unbounded quantifiers of ψ correspond to the $n-1$ unbounded quantifiers of θ. That is let

$$\psi_0(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s', P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s * \langle w \rangle), \vec{x}),$$

be a $\Delta_0(\exp)$-formula expressing "$\theta(\vec{p}, s, w)$ is equivalent to

$$\forall x_1 \exists x_2 \ldots Q x_{n-1} \theta_0(\vec{p}, s, w, \vec{x}),$$

where $\theta_0 \in \Delta_0$, $\theta_0(\vec{p}, s, w, \vec{x})$ is true, s is exactly ω^α-large, and $s' \subsetneq_{\mathrm{e}} s$". This is just the conjunction of $\Delta_0(\exp)$-formulas. Then let,

$$\psi \overset{\text{def}}{=} \forall x_1 \exists x_2 \ldots Q x_{n-1} \psi_0.$$

Such a ψ has the property,

$$\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s', P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s * \langle w\rangle); \vec{d}) \to \theta^*(\vec{p}, s, w; \vec{d}).$$

We note that $\ulcorner\psi^*\urcorner$ is standard. We suppose we have $m \stackrel{\text{def}}{=} P(\ulcorner\psi^*\urcorner)$, so m is standard.

Suppose j, θ, $\vec{p}$, s, and w satisfy (30) and (31). Let i_0, i_1, and i_2 be such that $c_j = a_{i_0}$, $c_{j+1} = a_{i_1}$, and $c_{j+2} = a_{i_2}$. The result will follow from the following claim.

Claim. *Let $k \stackrel{\text{def}}{=} i_2 - i_1$; then for some $l \leqslant k$ there is an exactly ω^α-large sequence $e_1, \ldots, e_l$ such that for all $i \leqslant l$,*

$(1)_i$ $i \neq 0 \Rightarrow e_i = \mu e(\exists x \leqslant d_n)\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_{i-1}, e\rangle, x; \vec{d})$,

$(2)_i$ $\langle \ulcorner\psi^*\urcorner, \langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i\rangle\rangle < a_{i_1+i-1} - 1$,

$(3)_i$ $(\exists x \leqslant a_{i_1+i})\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i\rangle, x; \vec{d})$.

Proof of Claim. Let $\chi(\langle e_1, \ldots, e_i\rangle) \stackrel{\text{def}}{=} (\forall j \leqslant i)((1)_j \wedge (2)_j \wedge (3)_j)$. We will show,

(1) $\chi(\Lambda)$, and

(2) for $i < k$, if $\chi(\langle e_1, \ldots, e_i\rangle)$ and "$e_1, \ldots, e_i$ is not ω^α-large" then there is an e_{i+1} such that $\chi(\langle e_1, \ldots, e_{i+1}\rangle)$.

Before proving (1) and (2) we will show that they imply the claim. Noting that

$$(\exists \langle e_1, \ldots, e_i\rangle \leqslant a_{i_1+i-1})\chi(\langle e_1, \ldots, e_i\rangle)$$

is a $\Delta_0(\exp)$-formula, we use (1), (2), and $\Delta_0(\exp)$-induction on i to get that either:

(3) For some $l \leqslant k$ and some $e_1, \ldots, e_l$, $\chi(\langle e_1, \ldots, e_l\rangle)$ and $e_1, \ldots, e_l$ is exactly ω^α-large, or

(4) for some $e_1, \ldots, e_k$, $\chi(\langle e_1, \ldots, e_k\rangle)$ and $e_1, \ldots, e_k$ is not ω^α-large.

We note that (3) is just the claim; we will show that (4) fails. Suppose (4) holds; then by $(2)_i$, $e_i \leqslant a_{i_1+i-1}$ for all $i = 1, \ldots, k$. Since $A_{j+1} = \{a_{i_1}, a_{i_1+1}, \ldots, a_{i_1+k-1}\}$ is ω^α-large, we may apply Proposition 5.9(ii) to get that $e_1, \ldots, e_k$ is ω^α-large, a contradiction; hence (4) fails.

Now we prove (1) and (2). For (1), note that $(1)_0$ is immediate. For $(2)_0$, note that $i_0 < i_1$, so $c_j = a_{i_0} \leqslant a_{i_1-1}$. By (30), and the definition of P,

$$\langle\langle \ulcorner\theta\urcorner, \vec{p}\rangle, \langle \ulcorner\theta\urcorner, \vec{p}\rangle\rangle + 2 \leqslant P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle) \leqslant c_j;$$

also,

$$\langle \ulcorner\psi^*\urcorner, \ulcorner\psi^*\urcorner\rangle + 2 \leqslant P(\ulcorner\psi^*\urcorner) = m \leqslant a - n - 1 \leqslant c_1 \leqslant c_j. \tag{32}$$

So we have,

$$\langle \psi^*, \langle \ulcorner\theta\urcorner, \vec{p}\rangle\rangle < a_{i_i-1} - 1,$$

and hence $(2)_0$ holds. By (31), the definition of ψ, and by the fact $d_n \stackrel{\text{def}}{=} a_r$, we have

$$(\exists x \leqslant a_r)\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}\rangle, x; \vec{d}).$$

By $(2)_0$ and Lemma 5.11(a)(iii) we have $(\exists x \leqslant a_{i_1})\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}\rangle, x; \vec{d})$, i.e. $(3)_0$ holds.

For (2) let $e_1, \ldots, e_i$ be such that $i < k$, for all $j \leqslant i$, $(1)_j$, $(2)_j$, and $(3)_j$ hold, and $e_1, \ldots, e_i$ is not ω^α-large. By $(3)_i$ and the definition of ψ, for some w, s', and x,

$$\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i\rangle, x; \vec{d}) \wedge \tag{33}$$

$$x = P(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i\rangle * s' * \langle w\rangle) \leqslant a_{i_1+i}. \tag{34}$$

Since $e_1, \ldots, e_i$ is not ω^α-large we have $s' \neq \Lambda$, so for some e,

$$\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i, e\rangle, x; \vec{d}).$$

Noting that $a_{i_1+i} \leqslant d_n$ we have by $\Delta_0(\exp)$-induction on

$$(\exists x \leqslant d_n)\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i, e\rangle, x; \vec{d})$$

that e_{i+1} given by $(1)_{i+1}$ exists. Since $e_{i+1} \leqslant e$, and since $*$ and P are increasing functions we have

$$\begin{aligned} P(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_{i+1}\rangle) &\leqslant P(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i, e\rangle) \\ &\leqslant P(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_i\rangle * s' * \langle w\rangle). \end{aligned}$$

Using (34) and the fact $\langle z, z\rangle + 2 \leqslant P(z)$, we have

$$\langle \langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_{i+1}\rangle, \langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_{i+1}\rangle\rangle + 2 \leqslant a_{i_1+i}.$$

This together with (32) yields,

$$\langle \psi^*, \langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_{i+1}\rangle\rangle < a_{i_1+i} - 1;$$

i.e. $(2)_{i+1}$ holds. By $(1)_{i+1}$, $(2)_{i+1}$, the fact $d_n = a_r$, and Lemma 5.11(a)(iii), we have

$$(\exists x \leqslant a_{i_1+i+1})\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e_1, \ldots, e_{i+1}\rangle, x; \vec{d}),$$

i.e. $(3)_{i+1}$ holds. $\square$

For the proof of (a)(ii) of the lemma, let $\hat{s} \stackrel{\text{def}}{=} \langle e_1, \ldots, e_l\rangle$ be given by the claim. Using $(2)_l$ of the claim, $\hat{s} < a_{i_1+l-1}$, since $l \leqslant k$, we have $\hat{s} < a_{i_1+k-1} < a_{i_2} = c_{j+2}$. By $(3)_l$, for some $\hat{w}$ and x,

$$\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * \hat{s}, x; \vec{d}) \wedge x = P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * \hat{s} * \langle \hat{w}\rangle) \leqslant a_{i_1+l}.$$

From this, the fact $a_{i_1+l} \leqslant c_{j+2}$, and the definition of ψ, we conclude,

$$(\hat{w} < c_{j+2}) \wedge \theta^*(\vec{p}, \hat{s}, \hat{w}; \vec{d}).$$

All that is left to show is that $\hat{s}$ is the $<_L$-least s such that there is a w satisfying (31). Suppose $s' = \langle e'_1, \ldots, e'_{l'}\rangle$ and w' satisfy (31), and $s' <_L \hat{s}$. Since $\hat{s}$ and s' are both exactly ω^α-large we have $s' \subseteq_e \hat{s} \Rightarrow s' = \hat{s}$. Thus for some $i \leqslant l, l'$ we have $e'_i < e_i$; let i_0 be the least such i. Then by $(1)_{i_0}$,

$$e'_{i_0} < \mu e(\exists x \leqslant d_n)\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e'_1, \ldots, e'_{i_0-1}, e\rangle, x; \vec{d}). \tag{35}$$

Letting $x' \stackrel{\text{def}}{=} P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s' * \langle w'\rangle)$, we have

$$\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e'_1, \ldots, e'_{i_0-1}, e'_{i_0}\rangle, x'; \vec{d});$$

this clearly implies that

$$(\exists x \leqslant d_n)\psi^*(\langle \ulcorner\theta\urcorner, \vec{p}, e'_1, \ldots, e'_{i_0-1}, e'_{i_0}\rangle, x; \vec{d}).$$

This contradicts (35), which implies there is no such s', completing the proof of the lemma. $\square$

The following theorem implies that there is a model of $TI(\omega^{\omega^\alpha}, \Pi_n)$ which is not a model of $F_{\omega_n^{\alpha+1}}\downarrow$; this together with Proposition 5.4(iv), yields the $\Rightarrow$-direction of Theorem 5.2. Together with the results of Section 5.3 we will have completed the proof of Theorem 5.2.

Theorem 5.25. *Suppose $M \models I\Delta_0(\exp) + (F_{\omega_{n-1}^{\omega^\alpha \cdot c}}(a) = b) + \{a, c > m : m \in \omega\}$, then there is an $I \subsetneq_e M$ such that $a < I < b$ and $I \models TI(\omega^{\omega^\alpha}, \Pi_n)$.*

It is straightforward, in a strong metatheory, to show that there is such an M with $c = a + 1$; in this case $I \models TI(\omega^{\omega^\alpha}, \Pi_n) + \neg\exists y(F_{\omega_n^{\alpha+1}}(a) = y)$.

Proof. Let C and $d_1, \ldots, d_n$ be given by Lemma 5.24 (using $t = c$). Since c is nonstandard, there is a limit point of C; let I be such a limit point. We need to show that $I \models TI(\omega^{\omega^\alpha}, \Pi_n)$. By Theorem 4.6, it is enough to show that $I \models LN(\omega^{\omega^\alpha}, \Sigma_n)$. Using Lemma 5.24(a)(i), we have $I \models I\Delta_0 + F_3\downarrow$. Applying Theorem 5.12, it is then enough to show that for all $\theta \in \Pi_{n-1}$, $\vec{p} \in I$, if

$$I \models \exists w\theta(\vec{p}, s, w) \wedge \text{“}s \text{ is exactly } \omega^\alpha\text{-large”}, \tag{36}$$

then there is a $<_L$-least such sequence s.

Let $w, s \in I$ be such that

$$I \models \theta(\vec{p}, s, w) \wedge \text{ “}s \text{ is exactly } \omega^\alpha\text{- large”}.$$

Since θ is standard, $\ulcorner\theta\urcorner \in I$, and hence

$$P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle), P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s * \langle w\rangle) \in I.$$

So for some j, $P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle) \leqslant c_j \in I$. Thus (30) is satisfied. By Lemma 5.24(a)(i), $I < d_n$, so $P(\langle \ulcorner\theta\urcorner, \vec{p}\rangle * s * \langle w\rangle) < d_n$. By Lemma 5.24(b),

$$M \models \theta^*(\vec{p}, s, w; d_1, \ldots, d_{n-1}) \wedge \text{“}s \text{ is exactly } \omega^\alpha\text{-large”}.$$

So (31) holds in M.

Let $\hat{s}$ be the $<_L$-least s such that for some w (31) holds in M. The existence of such an $\hat{s}$ is assured by Lemma 5.24(a)(ii), and $\hat{s} < c_{j+2}$; furthermore, there is a witness $\hat{w}$

such that $\hat{w} < c_{j+2}$. Since I is a limit point $c_{j+2} < I$, and hence $\hat{w}, \hat{s} \in I$. By Lemma 5.24(b),

$$I \models \theta(\vec{p}, \hat{s}, \hat{w}) \wedge \text{“}\hat{s} \text{ is exaclty } \omega^{\alpha}\text{-large”}.$$

Using the argument of the previous paragraph, any s satisfying (36) with $s <_L \hat{s}$ would clearly contradict the $<_L$-leastness of $\hat{s}$ in M. This completes the proof of the theorem. □

5.7. *Indicator functions*

Theorem 5.25 can be stated in the language of Paris and Kirby using the definition of indicator function (cf. [13] or [10]) as follows.

Corollary 5.26. *For $M \models I\Delta_0(\exp)$ the function f defined by,*

$$f(a, b) = \max\{c : F_{\omega_{n-1}^{\omega^{\alpha} \cdot c}}(a) \leqslant b\},$$

is an indicator for initial segments of M that are models of $TI(\omega^{\omega^{\alpha}}, \Pi_n)$.

6. Further results

6.1. *Relativizing the Proof*

Significant generalizations of Theorem 5.2, as well as generalizations of other results given below, may be obtained by relativizing the proofs in the following way. First replace the fast-growing hierarchy given in Section 5 with a relativized version that replaces the initial function, F_0, with a different initial function (as in [2]); in particular, for a relativization with respect to Σ_n, where $n \geqslant 1$, take, for F_0, a “Σ_n-complete function”, e.g., the function given by

$$F_{n,0}(x) = \langle y_0, y_1, \ldots, y_x \rangle, \tag{37}$$

where y_z is the least w such that $\Theta_{\Pi_{n-1}}((z)_0, \langle w, (z)_1 \rangle)$, if such a w exists, and is 0 otherwise ($\Theta_{\Pi_{n-1}}(x, y)$ expresses “x is (the Gödel number of) a Π_{n-1}-formula that is satisfied by the parameters (coded by) y). Note that the statement “$(F_{n,0}\downarrow)$” is equivalent, over $I\Delta_0$, to the statement of the Σ_n-least number principle; i.e., for $n \in \omega$, $I\Delta_0 + (F_{n,0}\downarrow) \equiv I\Sigma_n$. $F_{n,\alpha}$ for α a successor or limit ordinal is defined as in Section 5.1; in particular,

$$F_{n,\beta+1}(x) \overset{\text{def}}{=} F_{n,\beta}^{x+1}(x),$$

and

$$F_{n,\lambda}(x) \overset{\text{def}}{=} F_{n,\{\lambda\}(x)}(x), \quad \text{if } \lambda \in \text{LIM}.$$

Also let,

$$F_{0,\alpha} \stackrel{\text{def}}{=} F_\alpha.$$

For $n \geqslant 1$, any initial segment, I, of a model, M, of $I\Sigma_n$, that is closed under $F_{n,0}$ will be a Σ_n-elementary substructure of M. In this way, n quantifiers are handled "for free". Straightforward relativizations of the proof of Theorem 5.2 yield the following generalization.

Proposition 6.1. *For all $n \geqslant 1$, $m \geqslant 0$, and $\alpha < \varepsilon_0$,*

$$TI(\omega^{\omega^\alpha}, \Pi_n) \vdash (F_{m,\beta} \downarrow) \Leftrightarrow \beta < \omega_{n-m}^{\alpha+1},$$

where if $m > n$ then $\omega_{n-m}^{\alpha+1} \stackrel{\text{def}}{=} 0$.

6.2. Conservation results

For $n \in \omega$, and T an L_0-theory, let

$$\Pi_n(T) \stackrel{\text{def}}{=} \{\varphi\colon\ \varphi \text{ is a } \Pi_n\text{-sentence and } T \vdash \varphi\}.$$

Next we prove,

Theorem 6.2. *For all $n \geqslant 1$ and $\alpha < \varepsilon_0$,*

$$\Pi_2(TI(\omega^{\omega^\alpha}, \Pi_n)) = \Pi_2(I\Delta_0 + \{F_\beta \downarrow\colon \beta < \omega_n^{\alpha+1}\}).$$

Proof. Theorem 5.2 yields the $\subseteq$-direction. To obtain the $\supseteq$-direction, suppose $\psi \in \Delta_0$, and suppose for all $m \in \omega$,

$$I\Delta_0 + (F_{\omega_{n-1}^{\omega^\alpha \cdot m}} \downarrow) \nvdash \forall x \exists y \varphi(x, y).$$

Using a straightforward compactness argument we may obtain a model M such that, $a, c \in M - \omega$, and

$$M \models I\Delta_0(\exp) + (F_{\omega_{n-1}^{\omega^\alpha \cdot c}}(a) = b) + (\exists x \leqslant a)(\forall y)\neg\varphi(x, y).$$

To see this simply note that $I\Delta_0(\exp) + (F_{\omega_{n-1}^{\omega^\alpha \cdot c}}(a) = b) + (\exists x \leqslant a)(\forall y)\neg\varphi(x, y) + \{c \geqslant m\colon m \in \omega\} + \{a \geqslant l\colon l \in \omega\}$ is finitely satisfiable. By Theorem 5.25 there is an $I \subsetneq_e M$ such that $a < I < b$, and $I \models TI(\omega^{\omega^\alpha}, \Pi_n)$. By the absoluteness of Δ_0-formulas,

$$I \models \forall y \neg\varphi(x, y),$$

for some $x \leqslant a$; i.e., $I \models \neg\forall x \exists y \varphi(x, y)$. This implies that $TI(\omega^{\omega^\alpha}, \Pi_n) \nvdash \forall x \exists y \varphi(x, y)$, which yields the theorem. $\square$

Theorem 6.2 implies that the $TI(\omega^{\omega^\alpha}, \Pi_n)$-provably recursive functions are exactly the $(I\Delta_0 + \{(F_\gamma \downarrow)\colon \gamma < \omega_n^{\alpha+1}\})$-provably recursive ones. This allows for a classification of the $TI(\omega^{\omega^{\alpha}}, \Pi_n)$-provably recursive functions, which is done in Section 6.3. Before moving on to that task we state the generalization of Theorem 6.2 obtained by imitating its proof, but making use of Proposition 6.1 instead of Theorem 5.2.

Proposition 6.3. *For all $n \geqslant 1$, $n \geqslant m \geqslant 0$, and $\alpha < \varepsilon_0$,*

$$\Pi_{m+2}(TI(\omega^{\omega^{\alpha}}, \Pi_n)) = \Pi_{m+2}(I\Delta_0 + \{F_{m,\beta}\downarrow : \beta < \omega_{n-m}^{\alpha+1}\}).$$

6.3. The provably recursive functions of transfinite induction

We now show a correspondence between the $I\Delta_0 + (F_\gamma\downarrow)$-provably recursive functions and functions in other familiar hierarchies. Applying Theorem 6.2 we will have a correspondence between the $TI(\omega^{\omega^{\alpha}}, \Pi_n)$-provably recursive functions and the same hierarchies. Using Theorem 4.1, as well as Lemma 4.2, we will have a classification of the $TI(\alpha, \mathscr{C})$-provably recursive functions, for all $\alpha < \varepsilon_0$ and $\mathscr{C} = \Sigma_n$, Π_n, or Δ_0, in terms of these hierarchies with the inclusion of, as the smallest set in each hierarchy, the $I\Delta_0$-provably recursive functions.

Proposition 6.4. *For $\gamma \geqslant 2$, the following are equivalent:*

(i) *f is $I\Delta_0 + (F_\gamma\downarrow)$-provably recursive.*

(ii) *There exist $e, n \in \omega$ such that for all $\vec{x} \in \omega$, $f(\vec{x}) = \{e\}(\vec{x})$, and there is a computation of $\{e\}(\vec{x})$ that is coded by a number less than $F_\gamma^n(\max \vec{x})$.*

(iii) *f is elementary recursive in F_γ.*

(iv) *f is in the γth class of the Grzegorczyk hierarchy (as defined in* [16]).

For a proof of the equivalence of (iii) and (iv) the reader is referred to [16]; the reader is also referred to [16] for the definition of the Grzegorczyk hierarchy and for relationships between this hierarchy and other natural hierarchies. Below we sketch a proof of the equivalence of (i)–(iii). First note that the notation $\{e\}(\vec{x})$ in (ii) is the usual one from recursion theory, defined in terms of a Δ_0-version of Kleene's T-predicate, i.e. a Δ_0-formula, $\mathbf{T}(e,x,z)$, expressing "z is a computation, on a universal Turing machine, of program e on input x". $\{e\}(x) = y$ is the formula,

$$\exists z(\mathbf{T}(e,x,z) \wedge U(z) = y),$$

where U is a function expressing "$U(z)$ is the output of the program z". For (iii) we use the following definition (cf. [16]).

Definition 6.5. For $\mathscr{F}$ a set of functions, $E(\mathscr{F})$, *the elementary closure of $\mathscr{F}$*, is the smallest class of functions including all functions in $\mathscr{F}$ as well as the zero, successor, exponential, and projection functions, that is closed under composition and limited recursion, where we say f is obtained from g and h with bound b by limited recursion if,

$$f(\vec{x}, 0) = g(\vec{x}),$$

and

$$f(\vec{x}, y+1) = h(\vec{x}, y, f(\vec{x}, y))$$

and for all $\vec{x}, y$,

$$f(\vec{x}, y) \leqslant b(\vec{x}, y).$$

So if $g, h, b \in E(\mathscr{F})$ then f, as given above, is in $E(\mathscr{F})$. f is *elementary recursive in* $\mathscr{F}$ if(def) $f \in E(\mathscr{F})$, and f is *elementary recursive* if (def) $f \in E(\emptyset)$.

Proof. (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii).

(i) $\Rightarrow$ (ii). Suppose f is $I\Delta_0 + (F_\gamma\downarrow)$-provably recursive; i.e., suppose for some Σ_1-formula φ,

$$I\Delta_0 + (F_\gamma \downarrow) \vdash \forall \vec{x} \exists! y \varphi(\vec{x}, y), \tag{38}$$

and for all $\vec{x}, y \in \omega$,

$$\omega \models \varphi(\vec{x}, y) \;\Leftrightarrow\; f(\vec{x}) = y.$$

It is well-known that, over $I\Delta_0(\exp)$, there is an $e \in \omega$ such that $\varphi(\vec{x}, y)$ can be replaced by $\exists z(\mathbf{T}(e, \langle\vec{x}\rangle, z) \wedge U(z) = y)$. Suppose, for a contradiction, that for all $n \in \omega$ there is a $\vec{x} \in \omega$ such that every computation of $\{e\}(\vec{x})$ is coded by a number greater than $F_\gamma^n(\max \vec{x})$. Using a compactness argument we can show that there is a model M such that for some $a \in M - \omega$ and some $\langle\vec{b}\rangle \in M$,

$$M \models I\Delta_0 + \forall z(\mathbf{T}(e, \langle\vec{b}\rangle, z) \rightarrow z \geqslant F_\gamma^a(\max \vec{b})).$$

Since $\omega \models \forall \vec{x} \exists z \mathbf{T}(e, \langle\vec{x}\rangle, z)$, we have that $\max \vec{b}$ is nonstandard. Letting

$$I \overset{\text{def}}{=} \bigcup_{i \in \omega} F_\gamma^i(\max \vec{b}),$$

it is straightforward to show that $I \models I\Delta_0 + (F_\gamma \downarrow)$, but $I \not\models \exists y(\{e\}(\vec{b}) = y)$, contradicting (38). This completes the proof of (i) $\Rightarrow$(ii).

(ii) $\Rightarrow$ (iii). Suppose for all $\vec{x} \in \omega$ there is a computation of $f(\vec{x})$ coded by a number less than $F_\gamma^n(\max \vec{x})$. Then clearly all primitive recursions and all applications of the μ-operator used to define f can be replaced by limited recursions and applications of a bounded μ-operator involving the function F_γ^n. Since F_γ^n is elementary recursive in F_γ we need only show that $E(\mathscr{F})$ is closed under bounded μ-operations; i.e. it is enough to show,

Claim. *If $f, g \in E(\mathscr{F})$ then,*

$$h(\vec{x}) \overset{\text{def}}{=} \begin{cases} (\mu y \leqslant g(\vec{x}))(f(\vec{x}, y) = 0) & \textit{if such a } y \textit{ exists;} \\ 0 & \textit{otherwise.} \end{cases}$$

is in $E(\mathscr{F})$.

To prove the claim first note that addition, multiplication, and modified subtraction are all elementary recursive, where modified subtraction is given by

$$x \dot{-} y \overset{\text{def}}{=} \begin{cases} x-y & \text{if } x \geqslant y; \\ 0 & \text{otherwise.} \end{cases}$$

Modified subtraction can be defined using limited recursion to successively define the functions N, P, and $\dot{-}$, where N is defined to be 0 if $x = 0$ and 1 otherwise, and P is the predecessor function. Then h is obtained from composition using the fact that, for $f \in E(\mathscr{F})$, the following function is in $E(\mathscr{F})$:

$$h_0(\vec{x},z) \overset{\text{def}}{=} \begin{cases} (\mu y \leqslant z)(f(\vec{x},y)=0) & \text{if such a } y \text{ exists;} \\ 0 & \text{otherwise.} \end{cases}$$

h_0 can be defined by limited recursion as follows:

$$h_0(\vec{x},0) \overset{\text{def}}{=} 0;$$

$$h_0(\vec{x},y+1) \overset{\text{def}}{=} (y+1)\cdot(1 \dot{-} h_0(\vec{x},y))\cdot(1 \dot{-} f(\vec{x},y+1))\cdot N(f(\vec{x},0)) + h_0(\vec{x},y).$$

(iii) $\Rightarrow$(i). We will prove, by induction on the elementary recursive definition of f, that if f is elementary recursive in F_γ then,

(1) f is $\Delta_0(F_\gamma)$-definable.

(2) f is bounded by an $L_0(F_\gamma)$-term.

(3) f is $I\Delta_0 + F_\gamma\downarrow$-provably recursive.

Clearly (1), (2), and (3) all hold if f is F_γ or the zero, successor, exponential (noting $\gamma \geqslant 2$ and $I\Delta_0 \vdash (F_2\downarrow) \to \exp$), or a projection function. Suppose $f = g \circ h$ where (1)–(3) hold for g and h. Suppose $\varphi(y,z)$ and $\psi(\vec{x},y)$ are $\Delta_0(F_\gamma)$-definitions of g and h, respectively, and σ and τ are $L_0(F_\gamma)$-terms that bound g and h, respectively. Clearly we may suppose σ and τ are increasing functions. Then,

$$f(\vec{x}) = y \Leftrightarrow (\exists y \leqslant \tau(\vec{x}))(\psi(\vec{x},y) \wedge \varphi(y,z)),$$

so (1) holds; also f is bounded by $\sigma \circ \tau$, so (2) holds. Clearly, since g and h are assumed to be $I\Delta_0(F_\gamma)$-provably total, (3) holds.

Now suppose f is obtained from g, h, and b by limited recursion, and (1), (2) and (3) hold for g, h, and b. Suppose $\varphi(\vec{x},y)$ and $\psi(\vec{x},y,z)$ are $\Delta_0(F_\gamma)$-definitions of g and h, respectively, and b is bounded by the $L_0(F_\gamma)$-term τ. Then,

$$\begin{aligned} F(\vec{x},y) = z \Leftrightarrow \exists s(\ & s \in \mathrm{Sq} \wedge \mathrm{len}(s) = y+1 \wedge (s)_0 = g(\vec{x}) \\ & \wedge (\forall i < y)((s)_{i+1} = h(\vec{x},i,(s)_i) \wedge (s)_y = z). \end{aligned}$$

Since g and h are $\Delta_0(F_\gamma)$-definable, to obtain (1), we need only show s is bounded by an $L_0(F_\gamma)$-term. Since $(s)_i$ is bounded by $\tau(\vec{x},i)$, and since we may assume τ is increasing, s is bounded by an exponential in y and $\tau(\vec{x},y)$, so s is bounded by an

$L_0(F_\gamma)$-term. (1) and (2) both follow. Using the fact that g and h are both $I\Delta_0+(F_\gamma\downarrow)$-provably recursive, it is straightforward, applying induction, to obtain f is $I\Delta_0+(F_\gamma\downarrow)$-provably recursive.

We now have that the functions elementary recursive in F_γ are $I\Delta_0+(F_\gamma\downarrow)$-provably recursive. This completes the proof of (iii) $\Rightarrow$ (i), and hence the proof of the proposition. $\square$

Corollary 6.6. *For $n \geqslant 1$ and $\alpha < \varepsilon_0$, f is $TI(\omega^{\omega^\alpha}, \Pi_n)$-provably recursive if and only if f is in the* γth *class of the Grzegorczyk hierarchy for some $\gamma < \omega_n^{\alpha+1}$.*

Using Theorem 4.1, as well as Lemma 4.2, we have a classification of the $TI(\alpha,\mathscr{C})$-provably recursive functions, for all $\alpha < \varepsilon_0$ and $\mathscr{C} = \Sigma_n$, Π_n, or Δ_0, in terms of the Grzegorczyk hierarchy (if we include the $I\Delta_0$-provably recursive functions as the lowest level in the hierarchy). Similar classifications of "higher order" (nonrecursive) provably total functions of the theories $TI(\omega^{\omega^\alpha}, \Pi_n)$ may be obtained using Proposition 6.3 in place of Theorem 6.2; we will not state any results of this form.

6.4. Consistency results

Model-theoretic techniques, although often viewed as being more elegant and more direct than proof-theoretic techniques, are often cited as having the disadvantage of being nonfinitistic. Finitistic methods, in the study of subsystems of arithmetic, are recognized as being more fruitful in their results and more sound on philosophical grounds. In this section, we claim that the model-theoretic techniques of Section 5 have finitistic counterparts, and that they yield the same sort of results as the proof-theoretic methods. We refer to [19, 22] for the proofs, some of which will be sketched here.

There are many examples of how one can obtain stronger results by formalizing a proof in a weak system. Although, in arithmetic, this technique is most frequently applied to proof-theoretic methods, there are examples of this technique being applied to model-theoretic proofs as well. For example in [12], the authors show that, over PA, the Paris–Harrington principle, $\forall x \exists y (y \underset{*}{\to} (x+1)^x_x)$, is equivalent to the 1-consistency of PA, by showing that their model-theoretic proof is formalizable in PA. To obtain similar results, over a weaker base theory, namely $I\Sigma_1$, Paris, in [10], shows that the type of proof used in Section 5 can be modified to a proof formalizable in $I\Sigma_1$; he obtains,

$$I\Sigma_1 \vdash \forall x \exists y([x,y] \text{ is } \omega_{n+1} - \text{large}) \to 1\text{-Con}(I\Sigma_n). \tag{39}$$

n-Con is defined as follows:

Definition 6.7. For a theory, T, let n-Con(T) denote the schema,

$$\psi(x) \to \mathrm{Con}_T(\ulcorner \psi(\bar{x}) \urcorner): \ \psi \in \Pi_n \text{ has only } x \text{ free,}$$

where $\mathrm{Con}_T(\ulcorner\theta\urcorner)$ is a natural Π_n-formula expressing "there are no proofs of '$\theta \to 0 = 1$' from the axioms of T", and $\bar{x}$ is the closed term with value x given by

$$(\cdots((\overline{a_0}\cdot\overline{2}+\overline{a_1})\cdot\overline{2}+\overline{a_2})\cdot\overline{2}+\cdots+\overline{a_{k-1}})\cdot\overline{2}+\overline{a_k},$$

where $a_0 * \cdots * a_k$ is the Smullyan base 2 representation of x, $\overline{1} \stackrel{\text{def}}{=} S(0)$, and $\overline{2} \stackrel{\text{def}}{=} S(S(0))$. (*Note*: for fixed ψ, $\ulcorner\psi(\bar{x})\urcorner$ is polynomial in x.)

The definition of n-Con given here is suited for $I\Delta_0$; there are "formalized" versions that use truth definitions such as Θ_{Π_n} which are provably equivalent to the definition given above over $I\Delta_0(\exp)$. It is well-known that n-Con, over a sufficiently strong base theory, is provably equivalent to the schema, Rfn_{Σ_n}, expressing "if '$\psi(\bar{x})$' is provable then '$\psi(x)$' is true", where ψ ranges over Σ_n-formulas.

In the proof of (39), a modified version of the original proof is formalized and Herbrand's theorem is applied; whereas in [12] the proof uses the fact that PA has a version of the completeness theorem, so in that case there is no need to modify the original proof (we note, however, that there appears to be a hidden use of proof theory in carrying out the proof of Claim 2.11 of [12]).

It is natural to raise the question, "can proof-theoretic techniques be eliminated completely from proofs of results of this form?" Admittedly, the notion of "proof-theoretic techniques" is not very precise. Another, but just as vague, version of this question is "can the proof of Theorem 5.25 of Section 5.6.2 be 'directly' formalized in a weak subsystem?" We claim that the answer to these questions is yes, if we replace "Con(T)" by an equivalent notion, denoted "H-Con(T)", which stands for "Herbrand consistency" and is closer to the statement "T has a model" than to the statement "there is no proof of a contradiction from T"; in this way, without proof-theoretic techniques, the following proposition can be obtained.

Proposition 6.8. *For all $n \geqslant 1$ and all $\alpha < \varepsilon_0$,*

$$I\Delta_0 \vdash 1\text{-Con}(TI(\omega^{\omega^\alpha}, \Pi_n)) \leftrightarrow (F_{\omega_n^{\alpha+1}}\downarrow).$$

For the $\alpha = 0$ case our result differs from (39), mainly by the improvement of replacing $I\Sigma_1$ by $I\Delta_0$, and by including the $\rightarrow$-direction. A brief sketch of the key ideas of the proof of Proposition 6.8 is given below; it is proved in detail in [19] and is planned to be included in a paper that deals more generally with methods of formalizing this type of model-theoretic argument [22].

H-Con(T) states, roughly, "there are arbitrarily 'large' finite approximations of a model of T". In Section 5, we showed how, from a model of $T_0 \stackrel{\text{def}}{=} I\Delta_0(\exp) + \{a > m\colon m \in \omega\} + F_{\omega_n^{\alpha+1}}(a) = b$, we can construct a model of $T_1 \stackrel{\text{def}}{=} TI(\omega^{\omega^\alpha}, \Pi_n)$. These arguments can be viewed as showing how, from finite approximations of a model of T_0, i.e. from H-Con(T_0), we can construct finite approximations of a model of T_1, i.e. we can obtain H-Con(T_1). In fact this construction can take place in $I\Delta_0(\exp)$. Furthermore, H-Con(T_0) can be proven from $I\Delta_0 + (F_{\omega_n^{\alpha+1}}\downarrow)$. Proofs in [19] obtain

H-Con(T_1) directly from $I\Delta_0+(F_{\omega_n^{\alpha+1}}\downarrow)$, thus skipping the intermediate step of showing H-Con(T_0).

Using Proposition 6.1 in place of Theorem 5.2 in the proof of Proposition 6.8, the following generalization may be obtained.

Proposition 6.9. *For all $n \geqslant 1$, $n \geqslant m \geqslant 0$, and $\alpha < \varepsilon_0$,*

$$I\Delta_0 \vdash (m+1)\text{-Con}(TI(\omega^{\omega^\alpha}, \Pi_n)) \leftrightarrow (F_{m,\omega_{n-m}^{\alpha+1}}\downarrow).$$

6.5. A hierarchy generated by 1-Con

Ulf Schmerl in [17] follows Feferman in [3] by considering a hierarchy of theories defined as follows. Given S_0, let

(1) $S_{\alpha+1} \stackrel{\text{def}}{=} S_\alpha + 1\text{-CON}(S_\alpha)$.

(2) $S_\lambda \stackrel{\text{def}}{=} \bigcup_{\alpha<\lambda} S_\alpha$, if λ is a limit ordinal.

Here 1-CON is the formalized version of 1-Con described following Definition 6.7. We do not present an arithmetic definition of the S_α's. We note Schmerl, in [17], uses $S_0 =$ PRA. Now consider,

$$T_\alpha \stackrel{\text{def}}{=} \begin{cases} I\Delta_0 + (F_{\alpha+2}\downarrow) & \text{if } \alpha < \omega; \\ I\Delta_0 + \{(F_\gamma\downarrow):\ \gamma < \alpha\} & \text{if } \omega \leqslant \alpha < \varepsilon_0. \end{cases}$$

Then for any limit ordinal λ,

$$T_\lambda \equiv \bigcup_{\alpha<\lambda} T_\lambda,$$

and, noting that, for all $\alpha < \varepsilon_0$, $T_{\alpha+1} \equiv I\Delta_0 + (F_{3+\alpha}\downarrow)$, the following proposition implies:

$$T_{\alpha+1} \equiv T_\alpha + 1\text{-CON}(T_\alpha).$$

So the T_α's satisfy the defining equations of the S_α's. We remark that this does not imply an equivalence of these theories, using $S_0 = I\Delta_0 + (F_2\downarrow)$; such an equivalence would depend on the formal definition of the S_α's. We do not attempt this here.

Proposition 6.10. $I\Delta_0 \vdash \forall\alpha((F_{3+\alpha}\downarrow) \leftrightarrow 1\text{-CON}(T_\alpha))$.

The proof of this proposition, included in [19], is not given here; results of this form will be included in [22]. We remark that it is easy to show that PRA is a conservative extension of T_ω. Thus we conjecture that the results of [17] may be extended by replacing PRA with $I\Delta_0 + (F_2\downarrow)$, or, equivalently, with $I\Delta_0(\exp)$.

6.6. Generalizing to larger ordinals

Although the author has not checked this in every detail, he believes the main results of the last two sections can be generalized to larger ordinals. In [20, 21] it was shown that a great deal of basic ordinal arithmetic can be carried out in $I\Delta_0$ for arbitrary recursive epsilon numbers (α is an epsilon number if α is a solution to $\alpha = \omega^\alpha$). This suggests that the work here can be generalized by replacing ε_0 with an arbitrary recursive epsilon number. Based on [20, 21], there are representations of each recursive ordinal, as well as representations of many functions and relations on such ordinals, so that basic properties of the functions and relations are provable in $I\Delta_0$. For a fixed representation of this kind, where ε_γ is an epsilon number, Theorems 4.1 and 5.2, as well as Propositions 6.1 and 6.3, hold with ε_0 replaced by ε_γ. The author hopes to include such generalizations in a later paper.

6.7. Generalizing to schemata for L_0-formulas

Here we consider theories of the form $TI(\alpha, L_0)$, for $\alpha \geqslant \omega$. By Lemma 4.4, we clearly have

$$TI(\alpha, L_0) \vdash TI(\omega_n^{\alpha+1}, L_0),$$

for all $n \in \omega$; this proves the $\Leftarrow$-direction of

Proposition 6.11. *Suppose $\alpha \geqslant \omega$ and ε_γ is the least epsilon number strictly bigger than α, then*

$$TI(\alpha, L_0) \vdash TI(\beta, L_0) \ \Leftrightarrow \ \beta < \varepsilon_\gamma.$$

The $\Rightarrow$-direction can be obtained by a straightforward generalization of the proof of Theorem 5.2. Mainly what needs to be done is the definition of the fundamental sequence function needs to be extended to epsilon numbers so that a few basic properties hold. For succesor epsilon numbers the following will work:

$$\{\varepsilon_{\gamma+1}\}(n) = \omega_{n+1}^{\varepsilon_\gamma+1}.$$

For limit epsilon numbers it is convenient, and natural, to have the fundamental sequence function go through epsilon numbers. The author hopes to include this generalization in a later paper.

6.8. Generalizing to arbitrary well-orderings

For $e \in \omega$, let the binary relation $\prec_e$ be given by

$$n \prec_e m \overset{\text{def}}{\leftrightarrow} \{e\}(\langle n, m\rangle) = 0, \tag{40}$$

where $\{e\}$ is the eth partial recursive function. If $\{e\}$ is total and $\prec_e$ is a well-ordering of ω, then let $|e|$ denote the order type of $\prec_e$. Let X be a unary predicate symbol, and let $WF(e, X)$ express "if X is nonempty then X has a $\prec_e$-least element". Let $PA(X)$ be the result of replacing L_0 by $L_0(X)$ in the definition of PA. Gentzen, in [5], proved.

Proposition 6.12. *For all* $e \in \omega$,

$$PA(X) \vdash WF(e,X) \Rightarrow |e| < \varepsilon_0.$$

Although the arguments in Section 5 made heavy use of special properties of the particular representation of ε_0, and functions and relations on ε_0, the argument can be modified to obtain results of the form given in the above proposition, that are independent of such representations. The following can be proved using techniques based on the model-theoretic arguments of Section 5.

Proposition 6.13. $TI(\omega^{\omega^\alpha}, \Pi_n(X)) \vdash WF(e,X) \Rightarrow |e| < \omega_{n+1}^{\alpha+1}$.

Of course, the axiomatization of $TI(\omega^{\omega^\alpha}, \Pi_n(X))$ is dependent on a particular representation of a fixed recursive ordinal (that should be at least as big as the least epsilon number greater than α); in particular, basic properties of the representation must be provable for the theorem to hold. However, the following corollary shows no dependence on ordinal representations:

Corollary 6.14. *For all* $e, n \in \omega$, $I\Sigma_n(X) \vdash WF(e,X) \Rightarrow |e| < \omega_{n+1}$.

The proof of Proposition 6.13, omitted here, but to be included in [23], relies on the results of Section 5. We will sketch the key ideas of the proof. Note that in the proof of Theorem 5.2 a model of $TI(\omega^{\omega^\alpha}, \Pi_n)$ is obtained that has a descending sequence in $\omega_{n+1}^{\alpha+1}$ for our fixed *special* representation of ordinals; this descending sequence is given by the ordinals obtained in an "attempted computation" of $F_{\omega_n^{\alpha+1}}(a)$, where a is the nonstandard number used in the definition of I. For the above proposition we would like to get a set X that is a descending sequence in some *arbitrary* representation of the ordinal $\omega_{n+1}^{\alpha+1}$. Now, if $\mathcal{M}$ is a model of true second-order arithmetic, then the second-order part of $\mathcal{M}$ has a bijection, b, between the arbitrary representation of $\omega_{n+1}^{\alpha+1}$ and the special one. Now suppose $\mathcal{M}$ has a nonstandard first-order universe M. We can carry out the construction of the initial segment I as we did in the proof of Theorem 5.2, but this time around we use a fast-growing hierarchy that replaces the successor function as the initial function (F_0) with a function that majorizes the bijection, b, and its inverse. In this way I will be closed under the bijection, b, and hence, from a descending sequence in the special representation of $\omega_{n+1}^{\alpha+1}$ we can obtain a descending sequence in the arbitrary representation. Furthermore, if the construction is carried out in the right way, I will be a model of $TI(\omega^{\omega^\alpha}, \Pi_n(X))$, where X is satisfied by the set of elements in the descending sequence through the arbitrary representation of the ordinal. Such a construction yields the proposition.

Generalizing as in Section 6.7, we have Proposition 6.12 as well as:

Proposition 6.15. *Suppose* $\alpha \geqslant \omega$ *and* ε_γ *is the least epsilon number such that* $\alpha < \varepsilon_\gamma$, *then*

$$TI(\alpha, L_0(X)) \vdash WF(e,X) \Rightarrow |e| < \varepsilon_\gamma.$$

Once again, we are assuming a special representation of ε_γ for use in the axiomatization of $TI(\alpha, L_0(X))$; the corollary, Proposition 6.12, does not depend on special representations (except in the proof).

6.9. *Applications to second-order theories*

In light of the previous section it is natural to wonder if the model-theoretic techniques of Section 5 can be applied to classify the provably recursive functions of second-order theories. The answer is yes; i.e. the author has worked this out for some second-order theories. Results of this sort will be included in [23]. Furthermore, the author is hopeful that such techniques will be fruitful in establishing proof-theoretic ordinals for second-order theories whose proof-theoretic strength has not yet been measured.

References

[1] H. Enderton, Elements of Set Theory (Academic Press, New York, 1977).

[2] M.V.H. Fairtlough and S.S. Wainer, Ordinal complexity of recursive definitions, Inform. Comput. (1992).

[3] S. Feferman, Transfinite recursive progressions of axiomatic theories, J. Symbolic Logic 27 (1962) 259–316.

[4] H. Gaifman and C. Dimitracopoulos, Fragments of Peano arithmetic and the MDRP theorem, in: Logic and Algorithmic, Monographie No. 30 of de L'Enseignement Mathématique, Université de Genève (1982) 187–206.

[5] G. Gentzen, Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der Wahrheitsbegriff in der formalisierten Sprachen, Studia Philosophica 1 (1936) 261–4405.

[6] G. Gentzen, Die Widerspruchsfreiheit der reinen Zahlentheorie, Math. Ann. 122 (1936) 493–565.

[7] P. Hájek and P. Pudlák, Metamathematics of First Order Arithmetic (Springer, Berlin, 1993).

[8] J. Ketonen and R. Solovay, Rapidly growing ramsey functions, Ann. Math. 113 (1981) 267–314.

[9] G.E. Mints, Exact estimates of the provability of transfinite induction in the initial segments of arithmetic, Soviet Math. 1 (1971) 85–91.

[10] J.B. Paris, A hierarchy of cuts in models of arithmetic, in: Lecture Notes in Mathematics, Vol. 834 (Springer, Berlin, 1980) 312–337.

[11] J.B. Paris and C. Dimitracopoulos, Truth definitions for Δ_0-formulas, in: Logic and Algorithmic, Monographie No. 30 of de L'Enseignement Mathématique, Université de Genève (1982) 317–329.

[12] J.B. Paris and L. Harrington, in: Handbook of Mathematical Logic (North-Holland, Amsterdam, 1977) 1132–1142.

[13] J.B. Paris and L.A.S. Kirby, Initial segments of models of Peano's axioms, in: Lecture Notes in Mathematics, Vol. 619 (Springer, Berlin, 1977) 211–226.

[14] J.B. Paris and L.A.S. Kirby, Σ_n-Collection Schemas in Arithmetic (North-Holland, Amsterdam, 1978) 199–209.

[15] J.B. Paris and A. Wilkie, Counting problems in bounded arithmetic, in: Methods in Mathematical Logic, Lecture Notes in Mathematics, Vol. 1130 (Springer, Berlin, 1985) 317–340.

[16] H.E. Rose, Subrecursion: Functions and Hierarchies (Clarendon Press, Oxford, 1984).

[17] U.R. Schmerl, A fine structure generated by reflection formulas over primitive recursive arithemetic, in: Logic Colloquium 78 (North-Holland, Amsterdam, 1979) 335–350.

[18] R. Smullyan, Theory of formal systems, Ann. Math. Studies 47 (1961).

[19] R. Sommer, Transfinite induction and hierarchies generated by Transfinite Recursion within Peano arithmetic, Ph.D. Thesis, U. C. Berkeley, 1990.

[20] R. Sommer, Ordinals in bounded arithmetic, in: P. Clote and J. Krajíček, eds., Arithmetic, Proof Theory, and Complexity (Oxford University Press, Oxford, 1992).

[21] R. Sommer, Feasible aspects of fast growing functions, in preparation (1995).
[22] R. Sommer, Formalizing model-theoretic ordinal analysis, in preparation (1995).
[23] R. Sommer, Model-theoretic ordinal analysis, in preparation (1995).
[24] S.S. Wainer, A classification of the ordinal recursive functions, Archiv für mathematische Logik 13 (1970) 136–153.
[25] A. Woods, Bounded arithmetic formulas and turing machines of constant alternation, in: Logic Colloquium 1984 (North-Holland, Amsterdam, 1986) 355–377.