

Proof-theoretic strengths of the well-ordering principles

Toshiyasu Arai *

Graduate School of Mathematical Sciences, The University of Tokyo,
3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, JAPAN
tosarai@ms.u-tokyo.ac.jp

May 22, 2019

Abstract

In this note the proof-theoretic ordinal of the well-ordering principle for the normal functions \mathbf{g} on ordinals is shown to be equal to the least fixed point of \mathbf{g} . Moreover corrections to the previous paper [2] are made.

1 Introduction

In this note we are concerned with a proof-theoretic strength of a Π_2^1 -statement $\text{WOP}(\mathbf{g})$ saying that ‘for any well-ordering X , $\mathbf{g}(X)$ is a well-ordering’, where $\mathbf{g} : \mathcal{P}(\mathbb{N}) \rightarrow \mathcal{P}(\mathbb{N})$ is a computable functional on sets X of natural numbers. $\langle n, m \rangle$ denotes the elementary recursive pairing function $\langle n, m \rangle = \frac{(n+m)(n+m+1)}{2} + m$ on \mathbb{N} .

Definition 1.1 $X \subset \mathbb{N}$ defines a binary relation $<_X := \{(n, m) : \langle n, m \rangle \in X\}$.

$$\begin{aligned} \text{Prg}[<_X, Y] &: \Leftrightarrow \forall m (\forall n <_X m Y(n) \rightarrow Y(m)) \\ \text{TI}[<_X, Y] &: \Leftrightarrow \text{Prg}[<_X, Y] \rightarrow \forall n Y(n) \\ \text{TI}[<_X] &: \Leftrightarrow \forall Y \text{TI}[<_X, Y] \\ \text{WO}(X) &: \Leftrightarrow \text{LO}(X) \wedge \text{TI}[<_X] \end{aligned}$$

where $\text{LO}(X)$ denotes a Π_1^0 -formula stating that $<_X$ is a linear ordering.

For a functional $\mathbf{g} : \mathcal{P}(\mathbb{N}) \rightarrow \mathcal{P}(\mathbb{N})$,

$$\text{WOP}(\mathbf{g}) : \Leftrightarrow \forall X (\text{WO}(X) \rightarrow \text{WO}(\mathbf{g}(X)))$$

The theorem due to J.-Y. Girard is a base for further results on the strengths of the well-ordering principles $\text{WOP}(\mathbf{g})$. For second order arithmetics RCA_0 , ACA_0 , etc. see [8]. For a set $X \subset \mathbb{N}$, ω^X denotes an ordering on \mathbb{N} canonically defined such that its order type is ω^α when $<_X$ is a well-ordering of type α .

*I’d like to thank A. Freund for pointing out a flaw in [2]

Theorem 1.2 (Girard[3])

Over RCA_0 , ACA_0 is equivalent to $\text{WOP}(\lambda X.\omega^X)$.

In the following theorem ACA_0^+ denotes an extension of ACA_0 by the axiom of the existence of the ω -th jump of a given set.

Theorem 1.3 (Marcone and Montalbán[4]) Over RCA_0 , ACA_0^+ is equivalent to $\text{WOP}(\lambda X.\varepsilon_X)$.

Theorem 1.3 is proved in [4] computability theoretically. M. Rathjen noticed that the principle $\text{WOP}(\mathbf{g})$ is tied to the existence of *countable coded ω -models*.

Definition 1.4 A *countable coed ω -model* of a second-order arithmetic T is a set $Q \subset \mathbb{N}$ such that $M(Q) \models T$, where $M(Q) = \langle \mathbb{N}, \{(Q)_n\}_{n \in \mathbb{N}}, +, \cdot, 0, 1, < \rangle$ with $(Q)_n = \{m \in \mathbb{N} : \langle n, m \rangle \in Q\}$.

Let $X \in_\omega Y :\Leftrightarrow (\exists n[X = (Y)_n])$ and $X =_\omega Y :\Leftrightarrow (\forall Z(Z \in_\omega X \leftrightarrow Z \in_\omega Y))$.

It is not hard to see that over ACA_0 , the existence of the ω -th jump is equivalent to the fact that there exists an arbitrarily large countable coded ω -model of ACA_0 , cf. [1]. The fact means that there is a countable coded ω -model Q of ACA_0 containing a given set X , i.e., $X = (Q)_0$. From this characterization, Afshari and Rathjen[1] gives a purely proof-theoretic proof of Theorem 1.3. Their proof is based on Schütte's method of complete proof search in ω -logic, cf. [7].

In [4], a further equivalence is established for the binary Veblen function. In M. Rathjen, et.al.[1, 6, 5] and [2] the well-ordering principles are investigated proof-theoretically. Note that in Theorem 1.2 the proof-theoretic ordinal $|\text{ACA}_0| = |\text{WOP}(\lambda X.\omega^X)| = \varepsilon_0$ is the least fixed point of the function $\lambda x.\omega^x$. Moreover the ordinal $|\text{ACA}_0^+| = |\text{WOP}(\lambda X.\varepsilon_X)|$ in [4, 1] is the least fixed point of the function $\lambda x.\varepsilon_x$, and $|\text{ATR}_0| = |\text{WOP}(\lambda X.\varphi X 0)| = \Gamma_0$ in [6] one of $\lambda x.\varphi_x(0)$. These results suggest a general result that the well-ordering principle for normal functions \mathbf{g} on ordinals is equal to the least fixed point of \mathbf{g} .

In this note we confirm this conjecture under a mild condition on normal function \mathbf{g} , cf. Definition 2.3 for the extendible term structures.

We assume that the strictly increasing function \mathbf{g} enjoys the following conditions. The computability of the functional \mathbf{g} and the linearity of $\mathbf{g}(X)$ for linear orderings X are assumed to be provable elementarily, and if X is a well-ordering of type α , then $\mathbf{g}(X)$ is also a well-ordering of type $\mathbf{g}(\alpha)$. Moreover $\mathbf{g}(X)$ is assumed to be a *term structure* over constants $\mathbf{g}(c)$ ($c \in X$), function constants $+$, ω , and possibly other function constants.

Theorem 1.5 Let $\mathbf{g}(X)$ be an extendible term structure, and $\mathbf{g}'(X)$ an exponential term structure for which (2) holds below.

Then the proof-theoretic ordinal of the second order arithmetic $\text{WOP}(\mathbf{g})$ over ACA_0 is equal to the least fixed point $\mathbf{g}'(0)$ of the \mathbf{g} -function, $|\text{ACA}_0 + \text{WOP}(\mathbf{g})| = \min\{\alpha : \mathbf{g}(\alpha) = \alpha\} = \min\{\alpha > 0 : \forall \beta < \alpha(\mathbf{g}(\beta) < \alpha)\}$.

On the other side the proof of the harder direction of Theorem 4 in [2] should be corrected as pointed out by A. Freund. The theorem is stated as follows.

Theorem 1.6 *Let $\mathbf{g}(X)$ be an extendible term structure, and $\mathbf{g}'(X)$ an exponential term structure for which (2) holds.*

Then the following two are mutually equivalent over ACA_0 :

1. $\text{WOP}(\mathbf{g}')$.
2. $(\text{WOP}(\mathbf{g}))^+ :\Leftrightarrow \forall X \exists \mathcal{Q} [X \in_\omega \mathcal{Q} \wedge M(\mathcal{Q}) \models \text{ACA}_0 + \text{WOP}(\mathbf{g})]$.

Let us mention the contents of the paper. In the next section 2, $\mathbf{g}(X)$ is defined as a term structure. Exponential term structures and extendible ones are defined. The easy direction in Theorem 1.5 is shown. In section 3 we prove Theorems 1.5 and 1.6, assuming an elimination theorem 3.5 of the well-ordering principle in infinitary sequent calculi. In section 4 we prove the elimination theorem 3.5.

2 Term structures

Let us reproduce definitions on term structure from [2].

The fact that \mathbf{g} sends linear orderings X to linear orderings $\mathbf{g}(X)$ should be provable in an elementary way. \mathbf{g} sends a binary relation $<_X$ on a set X to a binary relation $<_{\mathbf{g}(X)} = \mathbf{g}(<_X)$ on a set $\mathbf{g}(X)$. We further assume that $\mathbf{g}(X)$ is a Skolem hull, i.e., a term structure over constants 0 and $\mathbf{g}(c)$ ($c \in \{0\} \cup X$) with the least element 0 in the order $<_X$, the addition $+$, the exponentiation ω^x , and possibly other function constants in a list \mathcal{F} . When $\mathcal{F} = \emptyset$, let $\omega^\alpha := \mathbf{g}(\alpha)$. Otherwise we assume that $\lambda\xi.\omega^\xi$ is in the list \mathcal{F} .

Definition 2.1 1. $\mathbf{g}(X)$ is said to be a *computably linear* term structure if there are three $\Sigma_1^0(X)$ -formulas $\mathbf{g}(X)$, $<_{\mathbf{g}(X)}$, $=$ for which all of the following facts are provable in RCA_0 : let $\alpha, \beta, \gamma, \dots$ range over terms.

(a) (Computability) Each of $\mathbf{g}(X)$, $<_{\mathbf{g}(X)}$ and $=$ is $\Delta_1^0(X)$ -definable. $\mathbf{g}(X)$ is a computable set, and $<_{\mathbf{g}(X)}$ and $=$ are computable binary relations.

(b) (Congruence)
 $=$ is a congruence relation on the structure $\langle \mathbf{g}(X); <_{\mathbf{g}(X)}, f, \dots \rangle$.
Let us denote $\mathbf{g}(X)/=$ the quotient set.

In what follows assume that $<_X$ is a linear ordering on X .

(c) (Linearity) $<_{\mathbf{g}(X)}$ is a linear ordering on $\mathbf{g}(X)/=$ with the least element 0.

(d) (Increasing) \mathbf{g} is strictly increasing: $c <_X d \Rightarrow \mathbf{g}(c) <_{\mathbf{g}(X)} \mathbf{g}(d)$.

(e) (Continuity) \mathbf{g} is continuous: Let $\alpha <_{\mathbf{g}(X)} \mathbf{g}(c)$ for a limit $c \in X$ and $\alpha \in \mathbf{g}(X)$. Then there exists a $d <_X c$ such that $\alpha <_{\mathbf{g}(X)} \mathbf{g}(d)$.

2. A computably linear term structure $\mathbf{g}(X)$ is said to be *extendible* if it enjoys the following two conditions.

- (a) (Suborder) If $\langle X, <_X \rangle$ is a substructure of $\langle Y, <_Y \rangle$, then $\langle \mathbf{g}(X); =, <_{\mathbf{g}(X)}, f, \dots \rangle$ is a substructure of $\langle \mathbf{g}(Y); =, <_{\mathbf{g}(Y)}, f, \dots \rangle$.
- (b) (Indiscernible) $\langle \mathbf{g}(c) : c \in \{0\} \cup X \rangle$ is an indiscernible sequence for linear orderings $\langle \mathbf{g}(X), <_{\mathbf{g}(X)} \rangle$: Let $\alpha[0, \mathbf{g}(c_1), \dots, \mathbf{g}(c_n)], \beta[0, \mathbf{g}(c_1), \dots, \mathbf{g}(c_n)] \in \mathbf{g}(X)$ be terms such that constants occurring in them are among the list $0, \mathbf{g}(c_1), \dots, \mathbf{g}(c_n)$. Then for any increasing sequences $c_1 <_X \dots <_X c_n$ and $d_1 <_X \dots <_X d_n$, the following holds.

$$\begin{aligned} & \alpha[0, \mathbf{g}(c_1), \dots, \mathbf{g}(c_n)] <_{\mathbf{g}(X)} \beta[0, \mathbf{g}(c_1), \dots, \mathbf{g}(c_n)] \quad (1) \\ \Leftrightarrow & \alpha[0, \mathbf{g}(d_1), \dots, \mathbf{g}(d_n)] <_{\mathbf{g}(X)} \beta[0, \mathbf{g}(d_1), \dots, \mathbf{g}(d_n)] \end{aligned}$$

Proposition 2.2 Suppose $\mathbf{g}(X)$ is an extendible term structure. Then the following is provable in RCA_0 : *Let both X and Y be linear orderings.*

Let $f : \{0\} \cup X \rightarrow \{0\} \cup Y$ be an order preserving map, $n <_X m \Rightarrow f(n) <_Y f(m)$ ($n, m \in \{0\} \cup X$). Then there is an order preserving map $F : \mathbf{g}(X) \rightarrow \mathbf{g}(Y)$, $n <_{\mathbf{g}(X)} m \Rightarrow F(n) <_{\mathbf{g}(Y)} F(m)$, which extends f in the sense that $F(\mathbf{g}(n)) = \mathbf{g}(f(n))$.

Proof. This is seen from the indiscernibility (1), cf. [2]. □

Definition 2.3 Suppose that function symbols $+, \lambda\xi.\omega^\xi$ are in the list \mathcal{F} of function symbols for a computably linear term structure $\mathbf{g}(X)$. Let $1 := \omega^0$, and $2 := 1 + 1$, etc.

$\mathbf{g}(X)$ is said to be an *exponential* term structure (with respect to function symbols $+, \lambda\xi.\omega^\xi$) if all of the followings are provable in RCA_0 .

- 1. 0 is the least element in $<_{\mathbf{g}(X)}$, and $\alpha + 1$ is the successor of α .
- 2. $+$ and $\lambda\xi.\omega^\xi$ enjoy the following familiar conditions.
 - (a) $\alpha <_{\mathbf{g}(X)} \beta \rightarrow \omega^\alpha + \omega^\beta = \omega^\beta$.
 - (b) $\gamma + \lambda = \sup\{\gamma + \beta : \beta < \lambda\}$ when λ is a limit number, i.e., $\lambda \neq 0$ and $\forall \beta <_{\mathbf{g}(X)} \lambda(\beta + 1 <_{\mathbf{g}(X)} \lambda)$.
 - (c) $\beta_1 <_{\mathbf{g}(X)} \beta_2 \rightarrow \alpha + \beta_1 <_{\mathbf{g}(X)} \alpha + \beta_2$, and $\alpha_1 <_{\mathbf{g}(X)} \alpha_2 \rightarrow \alpha_1 + \beta \leq_{\mathbf{g}(X)} \alpha_2 + \beta$.
 - (d) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
 - (e) $\alpha <_{\mathbf{g}(X)} \beta \rightarrow \exists \gamma \leq_{\mathbf{g}(X)} \beta(\alpha + \gamma = \beta)$.
 - (f) Let $\alpha_n \leq_{\mathbf{g}(X)} \dots \leq_{\mathbf{g}(X)} \alpha_0$ and $\beta_m \leq_{\mathbf{g}(X)} \dots \leq_{\mathbf{g}(X)} \beta_0$. Then $\omega^{\alpha_0} + \dots + \omega^{\alpha_n} <_{\mathbf{g}(X)} \omega^{\beta_0} + \dots + \omega^{\beta_m}$ iff either $n < m$ and $\forall i \leq n(\alpha_i = \beta_i)$, or $\exists j \leq \min\{n, m\}[\alpha_j <_{\mathbf{g}(X)} \beta_j \wedge \forall i < j(\alpha_i = \beta_i)]$.

3. Each $f(\beta_1, \dots, \beta_n) \in \mathbf{g}(X)$ ($+ \neq f \in \mathcal{F}$) as well as $\mathbf{g}(c)$ ($c \in \{0\} \cup X$) is closed under $+$. In other words the terms $f(\beta_1, \dots, \beta_n)$ and $\mathbf{g}(c)$ denote additively closed ordinals (additive principal numbers) when $<_{\mathbf{g}(X)}$ is a well-ordering.

In what follows we assume that $\mathbf{g}(X)$ is an extendible term structure, and $\mathbf{g}'(X)$ is an exponential term structure. Constants in the term structure $\mathbf{g}'(X)$ are 0 and $\mathbf{g}'(c)$ for $c \in \{0\} \cup X$, and function symbols in $\mathcal{F} \cup \{0, +\} \cup \{\mathbf{g}\}$ with a unary function symbol \mathbf{g} . We are assuming that a function constant $\lambda\xi.\omega^\xi$ is in the list $\mathcal{F} \cup \{\mathbf{g}\}$. Furthermore assume that RCA_0 proves that

$$\begin{aligned} \beta_1, \dots, \beta_n <_{\mathbf{g}'(X)} \mathbf{g}'(c) &\rightarrow f(\beta_1, \dots, \beta_n) <_{\mathbf{g}'(X)} \mathbf{g}'(c) \ (f \in \mathcal{F} \cup \{+, \mathbf{g}\}) \\ \omega^{\mathbf{g}'(\beta)} &= \mathbf{g}(\mathbf{g}'(\beta)) = \mathbf{g}'(\beta) \\ \mathbf{g}'(0) &= \sup_n \mathbf{g}^n(0) \\ \mathbf{g}'(c+1) &= \sup_n \mathbf{g}^n(\mathbf{g}'(c) + 1) \ (c \in \{0\} \cup X) \end{aligned} \tag{2}$$

where \mathbf{g}^n denotes the n -th iterate of the function \mathbf{g} , and we are assuming in the last that the successor element $c+1$ of c in X exists. The last two in (2) hold for normal functions \mathbf{g} when $\mathbf{g}(0) > 0$.

Note that $\mathbf{g}'(c)$ is an epsilon number when $<_{\mathbf{g}'(X)}$ is a well-ordering since the exponential function is in $\mathcal{F} \cup \{\mathbf{g}\}$.

We show the easy direction in Theorem 1.5. Let $<$ be an order of type $\mathbf{g}'(0)$, which is defined from a family of structures $\mathbf{g}(X_n)$ where the order types of X_n is $\gamma_n + 1$ defined as follows. A series of ordinals $\{\gamma_n\}_n < \mathbf{g}'(0)$ is defined recursively by $\gamma_0 = 0$ and $\gamma_{n+1} = \mathbf{g}(\gamma_n)$. Then $\text{WOP}(\mathbf{g})$ yields inductively $\text{TI}[<_{\gamma_n}]$ for initial segments of type γ_n . Hence $|\text{WOP}(\mathbf{g})| \geq \mathbf{g}'(0) := \min\{\alpha > 0 : \forall \beta < \alpha (\mathbf{g}(\beta) < \alpha)\}$.

3 Proof schema

In this section we give a proof schema of Theorems 1.5 and 1.6, each of these is based on an elimination theorem 3.5 of the well-ordering principle in infinitary sequent calculi.

Formulas in our infinitary sequent calculi are generated from literals \top (truth), $\perp := \bar{\top}$ (absurdity), $P(n)$, $\bar{P}(n)$, $E_i(n)$, $\bar{E}_i(n)$, $X_i(n)$, $\bar{X}_i(n)$ ($i, n \in \mathbb{N}$) by applying infinitary disjunction $\bigvee_{n \in \mathbb{N}} A_n$, infinitary conjunction $\bigwedge_{n \in \mathbb{N}} A_n$ and second-order quantifications $\exists X, \forall X$. Binary disjunctions $A_0 \vee A_1$ are understood to be $\bigvee_n B_n$ with $B_0 \equiv A_0$ and $B_{1+n} \equiv A_1$, and similarly for binary conjunctions. A formula is said to be a well-formed formula, wff in short if there is no free occurrence of ‘bound variables’ X_i, \bar{X}_i in it. The negation A of a wff A is defined recursively by the de Morgan’s law and the elimination of double negations. Each wff is assumed to be a translation A^∞ of a formula A without

free first-order variables in the language of second-order arithmetic. The translation is defined recursively as follows. For an arithmetic literal L , $L^\infty \equiv \top$ if L is true in the standard model \mathbb{N} , $L^\infty \equiv \perp$ otherwise. For a closed terms t and $R \in \{P, \bar{P}, E_i, \bar{E}_i, X_i, \bar{X}_i : i \in \mathbb{N}\}$, $R(t)^\infty \equiv R(n)$ with the value n of the closed term t in \mathbb{N} . $(A_0 \vee A_1)^\infty \equiv (A_0^\infty \vee A_1^\infty)$, and similarly for conjunctions. $(\exists x A(x))^\infty \equiv \bigvee_n A(\bar{n})^\infty$ for the n -th numeral \bar{n} . $(\forall x A(x))^\infty$ is defined to be an infinitary conjunction similarly. $(\exists X A(X))^\infty \equiv (\exists X A(X)^\infty)$, and similarly for the second-order universal quantifiers. A formula is said to be a *first-order* if no second-order quantifier occurs in it, while it is *arithmetical* if it is the translation of a formula in the language of the first-order arithmetic. i.e., neither the predicate constant P nor second-order variable occurs in it.

Each first-order formula A defines a binary relation $n <_A m : \Leftrightarrow A(\langle n, m \rangle)$. The principle is formulated in the inference rule (WP) together with a rule for the progressiveness $\text{Prg}[\prec_A, E_A]$ of E_A with respect to \prec_A :

$$\frac{\{\Gamma, E_A(n) : n \in \mathbb{N}\} \quad \neg \text{TI}[\prec_{\mathbf{g}_A}], \Gamma}{\Gamma} (WP)$$

where E_A is a variable proper to the relation \prec_A , and does not occur in Γ . $n <_{\mathbf{g}_A} m : \Leftrightarrow \mathbf{g}(A)(\langle n, m \rangle)$.

Our proof proceeds as follows. Given cut-free derivations of $\Gamma, E_A(n)$ without the rule (WP) , suppose that we can obtain an embedding f from the relation \prec_A to an ordinal α such that $n <_A m \Rightarrow f(n) < f(m) < \alpha$. Then the embedding f can be extended to an embedding F from the relation $\prec_{\mathbf{g}_A}$ to an ordinal $\mathbf{g}(\alpha)$ by Proposition 2.2. The embedding F yields the transfinite induction $\text{TI}[\prec_{\mathbf{g}_A}]$ for the relation $\prec_{\mathbf{g}_A}$. Eliminating the false formula $\neg \text{TI}[\prec_{\mathbf{g}_A}]$, we obtain Γ .

However in order to extract such an embedding f from derivations, we have to fix a meaning of the relation \prec_A . In other words, we need to interpret the predicate constant P and free-variables E_i occurring in the formula A so that these denote sets of natural numbers. This motivates Definition 3.1 below.

Definition 3.1 Let $\mathcal{E} \subset \mathbb{N}$ be a family of sets $\mathcal{E}_i = \{n \in \mathbb{N} : \langle i, n \rangle \in \mathcal{E}\}$. Each variable E_i is understood to denote the set \mathcal{E}_i . Let

$$\text{Diag}(\mathcal{E}_i) = \{E_i(n) : n \in \mathcal{E}_i\} \cup \{\bar{E}_i(n) : n \notin \mathcal{E}_i\}.$$

The predicate P denotes a set $\mathcal{P} \subset \mathbb{N}$.

$$\text{Diag}(\mathcal{P}) = \{P(n) : n \in \mathcal{P}\} \cup \{\bar{P}(n) : n \notin \mathcal{P}\}.$$

$\text{Diag}(\mathcal{P}, \mathcal{E}) = \text{Diag}(\mathcal{P}) \cup \bigcup_{i \in \mathbb{N}} \text{Diag}(\mathcal{E}_i)$ is identified with the countable coded ω -model $\langle \mathbb{N}; \mathcal{P}, \mathcal{E}_i \rangle_{i \in \mathbb{N}}$, and $\text{Diag}(\mathcal{P}, \mathcal{E}) \models A : \Leftrightarrow \langle \mathbb{N}; \mathcal{P}, \mathcal{E} \rangle_{i \in \mathbb{N}} \models A$ for first-order formulas A . For Σ_1^1 -formulas $\exists X F(X)$ with first-order matrices F , define $\text{Diag}(\mathcal{P}, \mathcal{E}) \models \exists X F(X)$ iff there exists a first-order formula $A(x)$ in the language of arithmetic such that $\text{Diag}(\mathcal{P}, \mathcal{E}) \models F(A)$, where $F(A)$ denotes the result of replacing literals $X(n)$ [$\bar{X}(n)$] in $F(X)$ by $A(n)^\infty$ [by $\neg A(n)^\infty$], resp.

For a finite set Γ of first-order formulas, $\text{Var}(\Gamma)$ denotes the set of second-order variables E_i occurring in Γ . For a family \mathcal{E}^X of finite sets \mathcal{E}_i^X , let

$$\Delta(\mathcal{E}^X; \Gamma) := \{E_i(n) : n \in \mathcal{E}_i^X, E_i \in \text{Var}(\Gamma), i \in \mathbb{N}\}$$

Definition 3.2 Let $\mathcal{P} \subset \mathbb{N}$ be a set of natural numbers, and \mathcal{E} a family of sets $\mathcal{E}_i \subset \mathbb{N}$. We define two cut-free infinitary one-sided sequent calculi $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP)$, and $\text{Diag}(\mathcal{P}, \emptyset) + (\text{prg})^\emptyset$ as follows.

Let \mathcal{E}^X be a family of *finite* sets $\mathcal{E}_i^X \subset \mathbb{N}$ ($i \in \mathbb{N}$), β, α ordinals, and Γ a sequent, i.e., a finite set of formulas (in negation normal form). We define a derivability relation $\vdash_\alpha^\beta \Gamma$ in the calculus $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP)$, and one $\mathcal{E}^X \vdash^\beta \Gamma$ in $\text{Diag}(\mathcal{P}, \emptyset) + (\text{prg})^\emptyset$ as follows, where the depth of the derivation is bounded by β , and the depth of the *nested applications* of the inferences (WP) is bounded by α in the witnessed derivation in $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP)$.

Axioms or initial sequents:

1. For $L \equiv \top$ and $L \in \text{Diag}(\mathcal{P})$, both $\vdash_\alpha^\beta \Delta, L$ and $\mathcal{E}^X \vdash^\beta \Delta, L$ hold.
2. $\vdash_\alpha^\beta \Delta, \bar{L}, L$ for literals $L \in \{E_i(n) : i, n \in \mathbb{N}\}$.
3. $\mathcal{E}^X \vdash^\beta \Delta, \bar{L}$ for literals $L \in \{E_i(n) : i, n \in \mathbb{N}\}$.

Inference rules: The following inference rules $(\vee), (\wedge), (\text{Rep}), (\exists_{1st}^2), (\forall^2)$ are shared by two calculi. The left part \mathcal{E}^X of \vdash should be deleted for the calculus $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP)$, and the subscript α is irrelevant to the calculus $\text{Diag}(\mathcal{P}, \emptyset) + (\text{prg})^\emptyset$ in the following. Let $\gamma < \beta$

$$\frac{\mathcal{E}^X \vdash_\alpha^\gamma \Gamma, \bigvee_n A_n, A_i}{\mathcal{E}^X \vdash_\alpha^\beta \Gamma, \bigvee_n A_n} (\vee) \quad \frac{\{\mathcal{E}^X \vdash_\alpha^\gamma \Gamma, \bigwedge_n A_n, A_i : i \in \mathbb{N}\}}{\mathcal{E}^X \vdash_\alpha^\beta \Gamma, \bigwedge_n A_n} (\wedge) \quad \frac{\mathcal{E}^X \vdash_\alpha^\gamma \Gamma}{\mathcal{E}^X \vdash_\alpha^\beta \Gamma} (\text{Rep})$$

$$\frac{\mathcal{E}^X \vdash_\alpha^\gamma F(A), \exists X F(X), \Gamma}{\mathcal{E}^X \vdash_\alpha^\beta \exists X F(X), \Gamma} (\exists_{1st}^2) \quad \frac{\mathcal{E}^X \vdash_\alpha^\gamma \Gamma, \forall X F(X), F(E)}{\mathcal{E}^X \vdash_\alpha^\beta \Gamma, \forall X F(X)} (\forall^2)$$

where in (\exists_{1st}^2) , $A(x)$ is a first-order formula, and in (\forall^2) , E is an eigenvariable not occurring in $\Gamma \cup \{\forall X F(X)\}$.

A first-order formula A defines a binary relation $n <_A m :\Leftrightarrow A(\langle n, m \rangle)$. Let $n <_{\mathbf{g}_A} m :\Leftrightarrow \mathbf{g}(A)(\langle n, m \rangle)$. For each first-order formulas A , $\beta_0 < \beta$ and $\alpha_0 < \alpha$, we have the following:

$$\frac{\{\vdash_{\alpha_0}^{\beta_0} \Gamma, E_A(n) : n \in \mathbb{N}\} \quad \vdash_{\alpha_0}^{\beta_0} \neg \text{TI}[\leq_{\mathbf{g}_A}], \Gamma}{\vdash_\alpha^\beta \Gamma} (WP)$$

where the variable E_A with the Gödel number $i = \lceil A \rceil$ does not occur in Γ .

1. For each first-order formulas A , $\beta_0 < \beta$, we have the following. Let $<_A^*$ denote the transitive closure of the relation $<_A$ for a first-order formula A .

$$\frac{\{\vdash_{\alpha}^{\beta_0} \Gamma, E_A(m), n \not<_A^* m, E_A(n) : n \in \mathbb{N}\}}{\vdash_\alpha^\beta \Gamma, E_A(m)} (\text{prg})^\infty$$

where $E_A \equiv E_i$ with the Gödel number $i = \lceil A \rceil$ of the formula A , and $\text{Var}(A) \subset \text{Var}(\Gamma)$.

2. Let A be a first-order formula with the Gödel number $i = \lceil A \rceil$. $n <_A^{*,\emptyset} m$ denotes the transitive closure of the relation $n <_A^\emptyset m :\Leftrightarrow \text{Diag}(\mathcal{P}, \emptyset) \models A(\langle n, m \rangle)$. If

$$m \in \mathcal{E}_i^X \quad (3)$$

then the inference $(prg)^\emptyset$ can be applied for $\beta_0 < \beta$:

$$\frac{\{((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X \cup \{n\}) \vdash^{\beta_0} \Gamma, E_A(m), E_A(n) : n <_A^{*,\emptyset} m\}}{((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X) \vdash^\beta \Gamma, E_A(m)} (prg)^\emptyset$$

where we assume that the variable $E_A \equiv E_i$ does not occur in A , and $\text{Var}(A) \subset \text{Var}(\Gamma)$.

The rule $(prg)^\infty$ states the fact that the set E_A is progressive with respect to the relation $<_A^*$, i.e., $\text{Prg}[<_A^*, E_A]$. It is convenient for us in proving Theorem 3.6 in section 4 to have the weaker statement $\text{Prg}[<_A^*, E_A]$ instead of the stronger $\text{Prg}[<_A, E_A]$. (WP) together with (prg) yields the well-ordering principle for \mathbf{g} .

Definition 3.3 Let π be a derivation witnessing the fact $\text{Diag}(\mathcal{P}) + (prg)^\infty + (WP) \vdash_\alpha^\beta \Gamma_0$, and $T(\pi) \subset <^\omega \mathbb{N}$ the underlying tree of π . Let us assign recursively a family $\mathcal{E}^X(\sigma) = \mathcal{E}^X(\sigma; \pi)$ of *finite* sets $\mathcal{E}_i^X(\sigma)$ to each node $\sigma \in T(\pi)$ in a bottom-up way as follows. In the definition, $\mathcal{E}^X(\sigma) \vdash \Gamma$ designates that $\mathcal{E}^X(\sigma)$ is assigned to the node σ at which the sequent Γ is placed. In this case we write $\sigma : \Gamma$.

To the end-sequent, i.e., the bottom sequent $\emptyset : \Gamma_0$, assign the set $\mathcal{E}_i^X(\emptyset) = \{n : E_i(n) \in \Gamma_0\}$.

Suppose that finite sets $(\mathcal{E}_j^X(\sigma))_{j \neq i}$ are assigned to the lower sequent $\sigma : \Gamma$ of a rule (WP) for the relation $<_A$ with $i = \lceil A \rceil$. For the n -th left upper sequents $\sigma * (n) : \Gamma, E_A(n)$, assign the family $((\mathcal{E}_j^X(\sigma))_{j \neq i}, \{n\})$ with $\mathcal{E}_i^X(\sigma * (n)) = \{n\}$. For the right upper sequent $\sigma * (\omega) : \neg \text{TI}[<_{\mathbf{g}_A}], \Gamma$, assign the family $(\mathcal{E}_j^X(\sigma))_{j \neq i}$.

$$\frac{\{((\mathcal{E}_j^X(\sigma))_{j \neq i}, \{n\}) \vdash \Gamma, E_A(n) : n \in \mathbb{N} \quad (\mathcal{E}_j^X(\sigma))_{j \neq i} \vdash \neg \text{TI}[<_{\mathbf{g}_A}], \Gamma\}}{(\mathcal{E}_j^X(\sigma))_{j \neq i} \vdash \Gamma} (WP)$$

where the variable E_A with $i = \lceil A \rceil$ does not occur in Γ .

Next suppose that a family $((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X(\sigma))$ ($i = \lceil A \rceil$) is assigned to the lower sequent $\sigma : \Gamma, E_A(m)$ of the rule $(prg)^\infty$. For each number n , assign the family $((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X(\sigma) \cup \{n\})$ to the n -th upper sequent $\sigma * (n) : \Gamma, E_A(m), n \not<_A^* m, E_A(n)$ with $\mathcal{E}_i^X(\sigma * (n)) = \mathcal{E}_i^X(\sigma) \cup \{n\}$.

$$\frac{\{((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X \cup \{n\}) \vdash \Gamma, E_A(m), n \not<_A^* m, E_A(n) : n \in \mathbb{N}\}}{((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X(\sigma)) \vdash \Gamma, E_A(m)} (prg)^\infty$$

where $E_A \equiv E_i$ with the Gödel number $i = \lceil A \rceil$ of the formula A .

For rules other than (WP) , $(prg)^\infty$, the upper sequents receive the same family as the lower sequent receives. For example

$$\frac{\mathcal{E}^X(\sigma) \vdash F(A), \exists X F(X), \Gamma}{\mathcal{E}^X(\sigma) \vdash \exists X F(X), \Gamma} (\exists_{1st}^2)$$

where $\text{Var}(A) \subset \text{Var}(\Gamma, F)$.

A family $\mathcal{E}^X(\sigma)$ has been assigned to each node $\sigma \in T(\pi)$ in the tree of the derivation π showing the fact $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP) \vdash_\alpha^\beta \Gamma_0$.

Let us define an ordinal function $F(\beta, \alpha)$ for giving an upper bound in eliminating the well-ordering principle. For normal function $\mathbf{g}(\alpha)$ in Theorems 1.5 and 1.6, and ordinals β, α , let us define ordinals $F(\beta, \alpha)$ recursively on α as follows. $F(\beta, 0) = \omega^{1+\beta}$,

$$F(\beta, \alpha + 1) = F\left(\mathbf{g}(\omega^{2(F(\beta, \alpha) + \beta) + 1}) + 1 + \beta, \alpha\right) + \mathbf{g}(\omega^{2(F(\beta, \alpha) + \beta) + 1}) + 1 \quad (4)$$

and $F(\beta, \lambda) = \sup\{F(\beta, \alpha) + 1 : \alpha < \lambda\}$ for limit ordinals λ .

Proposition 3.4 1. $\gamma < \beta \Rightarrow F(\gamma, \alpha) \leq F(\beta, \alpha)$, and $\gamma < \alpha \Rightarrow F(\beta, \gamma) < F(\beta, \alpha)$.

2. $F(\beta, \omega(1 + \alpha)) = \mathbf{g}'(\alpha)$ for $\beta < \mathbf{g}'(\alpha)$.

3. If $\beta < \mathbf{g}'(\alpha)$ and $\gamma < \omega(1 + \alpha)$, then $F(\beta, \gamma) < \mathbf{g}'(\alpha)$.

Proof. 3.4.1. This follows from the fact that each of functions $\beta \mapsto \alpha + \beta$, $\beta \mapsto \omega^\beta$ and $\beta \mapsto \mathbf{g}(\beta)$ is strictly increasing.

3.4.3. This follows from the fact that $\mathbf{g}'(\alpha)$ is closed under $\lambda x. \omega^x$ and \mathbf{g} . \square

The following Elimination theorem 3.5 of the inference (WP) is a crux for us.

Theorem 3.5 (Elimination of (WP))

Suppose that for a finite set Φ of Σ_1^1 -formulas $\vdash_\alpha^\beta \Phi, \Gamma$ holds in the calculus $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP)$ for $\sigma : \Phi, \Gamma$ in a witnessing derivation π in which the condition (3) is enjoyed for each $(\text{prg})^\infty$. Moreover assume that $\Gamma \subset \Delta(\mathcal{E}^X(\sigma); \Phi, \Gamma)$, and $\text{Diag}(\mathcal{P}, \emptyset) \not\vdash B$ for any $B \in \Phi$.

Then $\mathcal{E}^X(\sigma) \vdash_0^{F(\beta, \alpha) + \beta} \Delta(\mathcal{E}^X(\sigma); \Phi, \Gamma)$ holds in the calculus $\text{Diag}(\mathcal{P}, \emptyset) + (\text{prg})^\emptyset + (WP)$.

In proving Theorem 3.5, a key is an extension, Theorem 3.6 below, of a result due to G. Takeuti[9, 10], cf. Theorem 5 in [2].

Theorem 3.6 The following is provable in $\text{ACA}_0 + \text{WO}(\alpha)$:

Let $n \prec m$ be a binary relation on \mathbb{N} , and $n \prec^* m$ the transitive closure of the relation $n \prec m$. $(\text{prg})_\prec^D$ denotes the following inference rule for a predicate E .

$$\frac{\{\mathcal{E}^X \cup \{n\} \vdash^{\beta_0} \Gamma, E(m), E(n) : n \prec^* m\}}{\mathcal{E}^X \vdash^\beta \Gamma, E(m)} (\text{prg})_\prec^D$$

where the condition (3), $m \in \mathcal{E}^X$, is enjoyed with a finite set \mathcal{E}^X . $\mathcal{E}^X \vdash^\beta \Gamma$ denotes the derivability relation in a calculus $\text{Diag}(\emptyset) + (\text{prg})_\prec^D$, in which $\text{Diag}(\emptyset) = \{\bar{E}(n) : n \in \mathbb{N}\}$.

Assume that there exists an ordinal α for which $\{n\} \vdash^\alpha E(n)$ holds for any natural number n .

Then there exist an embedding f such that $n \prec m \Rightarrow f(n) < f(m)$, $f(m) < \omega^{\alpha+1}$ for any $n, m \in \mathbb{N}$.

Proofs of Theorems 3.5 and 3.6 are postponed in section 4.

In what follows we work in ACA_0^+ . the set $\{\lceil A \rceil : \text{Diag}(\mathcal{P}, \mathcal{E}) \models A\}$ of the satisfaction relation $\text{Diag}(\mathcal{P}, \mathcal{E}) \models A$ for first-order formulas A is then computable from the ω -th jump of the set \mathcal{P} .

3.1 Proof of Theorem 1.5

First let us prove Theorem 1.5. In this subsection the predicate P plays no role, and $\text{Diag}(\mathcal{P})$ is omitted. Let us introduce a finitary calculus $\mathbf{G}_2 + (prg) + (WPL)$ obtained from a calculus \mathbf{G}_2 for the predicative second-order logic with inference rules (\exists_{1st}^2) and (\forall^2) by adding the following rules (VJ) , (prg) , (WPL) as follows. The following inference (VJ) for complete induction schema for first-order formulas A and the successor function $S(x)$ with an eigenvariable x .

$$\frac{\Gamma, A(0) \quad \neg A(x), \Gamma, A(S(x)) \quad \neg A(t), \Gamma}{\Gamma} (VJ)$$

For first-order formulas A and the eigenvariable x :

$$\frac{\Gamma, E_A(t), x \not\prec_A^* t, E_A(x)}{\Gamma, E_A(t)} (prg) \quad \frac{\Gamma, E_A(x) \quad \Gamma, \text{LO}(<_A) \quad \neg \text{TI}[<_{\mathbf{g}_A}], \Gamma}{\Gamma} (WPL)$$

where $E_A \equiv E_i$ with $i = \lceil A \rceil$, in (prg) , x is the eigenvariable not occurring freely in $\Gamma, E_A(t)$, and $\text{Var}(A) \subset \text{Var}(\Gamma)$. In (WPL) , the variable E_A does not occur in Γ (nor in A). The initial sequents are Γ, \bar{L}, L for literals L .

We can assume that in a finitary proof, a variable E occurs in an upper sequent of an inference, but not in the lower sequent only when the inference is a (\forall^2) , and the variable E is the eigenvariable of the inference. Moreover if the end-sequent contains no second-order free variable, then the variable E_A can be assumed not to occur in the lower sequent Γ of the rule (WPL) for the relation $<_A$.

The axiom of arithmetic comprehension is deduced from the inference rule (\exists_{1st}^2) , and the axiom $\text{WOP}(\mathbf{g})$ for the well-ordering principle of \mathbf{g} is deduced from the inference rules (prg) and (WPL) .

Assume that $\text{TI}[<]$ is provable from $\text{WOP}(\mathbf{g})$ in ACA_0 for an arithmetical relation $<$. Let Δ_0 denote a set of negations of axioms for first-order arithmetic except complete induction. By eliminating (cut) 's we obtain a proof of $\Delta_0, E_{<}(x)$ in $\mathbf{G}_2 + (prg) + (WPL)$, where $E_{<} \equiv E_i$ with $i = \lceil x_0 < x_1 \rceil$.

Let us embed the finitary calculus $\mathbf{G}_2 + (prg) + (WPL)$ to an intermediate infinitary calculus $(prg)^\infty + (WP) + (cut)_{1st}$, which is obtained from $(prg)^\infty + (WP)$ by adding the cut inference $(cut)_{1st}$ with a first-order cut formulas A :

$$\frac{\Gamma, \neg A^\infty \quad A^\infty, \Delta}{\Gamma, \Delta} (cut)_{1st}$$

The logical depth $\text{dg}(A) < \omega$ of first-order formulas A is defined recursively by $\text{dg}(L) = 0$ for literals L , $\text{dg}(A_0 \vee A_1) = \text{dg}(A_0 \wedge A_1) = \max\{\text{dg}(A_0), \text{dg}(A_1)\} + 1$, and $\text{dg}(\exists x A(x)), \text{dg}(\forall x A(x)) = \text{dg}(A(0)) + 1$. Then let $\text{dg}(A^\infty) := \text{dg}(A)$. Let $\Gamma(x, \dots)$ be a sequent possibly with free first-order variables x, \dots . Assuming $\mathbf{G}_2 + (\text{prg}) + (\text{WPL}) \vdash \Gamma(x, \dots)$, we see easily that there exist $d, p, k, m < \omega$ such that $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{d,p}^{\omega^{k+m}} \Gamma(n, \dots)^\infty$ holds for any natural numbers n, \dots , where the first subscript d indicates that the number of nested applications of the rule (WPL) is bounded by d , and the second p designates that any (first-order) cut-formula A^∞ occurring in the witnessing derivation has the logical depth $\text{dg}(A^\infty) < p$. Note that each variable E occurring in the induction formula A of a (VJ) can be assumed to occur also in the lower sequent Γ . We see that there exist $d, p < \omega$ such that $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{d,p}^{\omega^2} \Delta_0, E_{\prec}(n)$ holds for any natural number n . Eliminating the false arithmetic Δ_0 , we obtain $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{d,p}^{\omega^2} E_{\prec}(n)$.

Let $2_0(\beta) = \beta$ and $2_{p+1}(\beta) = 2^{2^p(\beta)}$ for $p < \omega$.

Proposition 3.7 *Suppose $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{d,p}^\beta \Gamma$. Then $(\text{prg})^\infty + (\text{WP}) \vdash_{2_p(d),0}^{2_p(\beta)} \Gamma$.*

Proof. Let A be one of formulas $\exists x B$, $B \vee C$, $\bar{E}_i(n)$ and arithmetic literals. We see by induction on α that if $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{d,p}^\beta \Gamma, \neg A^\infty$ and $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{e,p}^\alpha A^\infty, \Delta$ with $\text{dg}(A) \leq p$, then $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{d+e,p}^{\beta+\alpha} \Gamma, \Delta$.

From the fact we see the proposition by induction on $p < \omega$. \square

By Proposition 3.7 we obtain an ordinal $\beta < \varepsilon_0$ and $c < \omega$ for which $(\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_{c,0}^\beta E_{\prec}(n)$, i.e., $(\text{prg})^\infty + (\text{WP}) \vdash_c^\beta E_{\prec}(n)$ holds for any n .

In a witnessing derivation π of the fact $(\text{prg})^\infty + (\text{WP}) \vdash_c^\beta E_{\prec}(n)$, the condition (3) may be violated. Let us convert the derivation π to a derivation π^\top as follows.

Definition 3.8 For a formula A and a family $\mathcal{E} = (\mathcal{E}_i)_i$ of sets, $A(\mathcal{E})$ denotes the result of replacing each literal $E_i(m)$ by \top when $m \notin \mathcal{E}_i$. $A(\mathcal{E})$ is defined recursively as follows. Let $i = \lceil A \rceil$.

$$((E_A(m))(\mathcal{E}), (\bar{E}_A(m))(\mathcal{E})) \equiv \begin{cases} (\top, \perp) & \text{if } m \notin \mathcal{E}_i \\ (E_{A(\mathcal{E})}(m), \bar{E}_{A(\mathcal{E})}(m)) & \text{if } m \in \mathcal{E}_i \end{cases}$$

$L(\mathcal{E}) \equiv L$ when $L \in \{\top, \perp, X_i(n), \bar{X}_i(n) : i, n \in \mathbb{N}\}$. $(\bigvee_n A_n)(\mathcal{E}) \equiv \bigvee_n (A_n(\mathcal{E}))$ and $(\bigwedge_n A_n)(\mathcal{E}) \equiv \bigwedge_n (A_n(\mathcal{E}))$. $(\exists X F(X))(\mathcal{E}) \equiv \exists X (F(X)(\mathcal{E}))$ and $(\forall X F(X))(\mathcal{E}) \equiv \forall X (F(X)(\mathcal{E}))$. For a sequent Γ , let $\Gamma(\mathcal{E}) = \{A(\mathcal{E}) : A \in \Gamma\}$.

For each node $\sigma : \Gamma$ in the derivation π , let

$$A^\sigma := A(\mathcal{E}^X(\sigma)), \Gamma^\sigma := \{A^\sigma : A \in \Gamma\}$$

Proposition 3.9 *Let π be a derivation witnessing the fact $\{n\} \vdash_c^\beta E_{\prec}(n)$ in $(prg)^\infty + (WP)$. Then there exists a derivation π^\top witnessing the same fact in $(prg)^\infty + (WP)$ such that $\mathcal{E}_{A^\sigma}^X(\sigma; \pi^\top) = \mathcal{E}_A^X(\sigma; \pi)$ for each $\sigma \in T(\pi^\top) \subset T(\pi)$, and the condition (3) is enjoyed for each rule $(prg)^\infty$ occurring in π^\top .*

Proof. This is seen by induction on the tree order on the well-founded tree $T(\pi)$. Each axiom $\sigma : \Gamma, \bar{L}, L$ turns either to $\sigma : \Gamma^\sigma, \bar{L}, L$ or to $\sigma : \Gamma^\sigma, \perp, \top$.

Consider a rule $(prg)^\infty$ in π .

$$\frac{\{((\mathcal{E}_j^X(\sigma; \pi))_{j \neq i}, \mathcal{E}_i^X(\sigma; \pi) \cup \{n\}) \vdash \Gamma, E_A(m), n \not\prec_A^* m, E_A(n) : n \in \mathbb{N}\}}{((\mathcal{E}_j^X(\sigma; \pi))_{j \neq i}, \mathcal{E}_i^X(\sigma; \pi)) \vdash \Gamma, E_A(m)} (prg)^\infty$$

where $E_A \equiv E_i$ with the Gödel number $i = \lceil A \rceil$ of the formula A . If $m \in \mathcal{E}_A^X(\sigma; \pi)$, then $m \in \mathcal{E}_{A^\sigma}^X(\sigma; \pi^\top) = \mathcal{E}_A^X(\sigma; \pi)$, and

$$\frac{\{((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X(\sigma) \cup \{n\}) \vdash \Gamma^\sigma, E_{A^\sigma}(m), n \not\prec_{A^\sigma}^* m, E_{A^\sigma}(n) : n \in \mathbb{N}\}}{((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X(\sigma)) \vdash \Gamma^\sigma, E_{A^\sigma}(m)} (prg)^\infty$$

Otherwise $(E_A(m))^\sigma \equiv \top$. Γ^σ, \top is an axiom. Discard the upper part.

From the construction of π^\top we see easily that $\mathcal{E}_{A^\sigma}^X(\sigma; \pi^\top) = \mathcal{E}_A^X(\sigma; \pi)$ for each node $\sigma \in T(\pi^\top) \subset T(\pi)$, and the condition (3) is enjoyed for each rule $(prg)^\infty$ occurring in π^\top . \square

By Proposition 3.9 we obtain a derivation π^\top witnessing the fact $\{n\} \vdash_c^\beta E_{\prec}(n)$ in $(prg)^\infty + (WP)$ such that the condition (3) is enjoyed for each rule $(prg)^\infty$ occurring in π^\top .

We see from Theorem 3.5 that in the calculus $\text{Diag}(\emptyset) + (prg)^\emptyset$, $\{n\} \vdash_0^\alpha E_{\prec}(n)$ holds for any n , and the ordinal $\alpha = F(\beta, c) + \beta$, where $\{n\} = \mathcal{E}_i^X(\pi)$ with $i = \lceil x_0 \prec x_1 \rceil$ and $\Delta(\{n\}, \emptyset; E_{\prec}(n)) = \{E_{\prec}(n)\}$.

Theorem 3.6 yields an embedding f such that $n \prec m \Rightarrow f(n) < f(m) < \omega^{\alpha+1}$.

On the other hand we have $\omega^{\alpha+1} = \omega^{F(\beta, c) + \beta + 1} < g'(0)$ by Proposition 3.4.3 and $\beta < \varepsilon_0 \leq g'(0)$. Thus Theorem 1.5 is proved.

3.2 Corrections to [2]

The proof of the harder direction of Theorem 4 in [2] should be corrected as pointed out by A. Freund. In this subsection the predicate P will denote a given set of natural numbers. Let us augment another countable list Y_i, \bar{Y}_i ($i \in \mathbb{N}$) of second-order free variables. First-order formulas may contain these variables Y .

Assuming $\text{WOP}(g')$, we need to show the existence of a countable coded ω -model $(\mathcal{P}, (\mathcal{Q})_i)_{i < \omega}$ of $\text{ACA}_0 + \text{WOP}(g)$ for a given set $\mathcal{P} \subset \mathbb{N}$. In what follows argue in $\text{ACA}_0 + \text{WOP}(g')$. Since $\text{WOP}(g')$ implies $\text{WOP}(\lambda X. \varepsilon_X)$, which in turn yields ACA_0^+ by Theorem 1.3, we are working in $\text{ACA}_0^+ + \text{WOP}(g')$, and we can assume the existence of the ω -th jump of any sets.

Let us search a derivation of the contradiction \emptyset in the following infinitary calculus $\text{Diag}(\mathcal{P}) + (prg)^\infty + (WP) + (ACA)$, in which the variables E_i are not

interpreted. The calculus is obtained from the infinitary calculus $(prg)^\infty + (WP)$ by adding the following rule (ACA) for arithmetic comprehension axiom:

$$\frac{Y_j \neq A, \Gamma}{\Gamma} (ACA)$$

where A is a first-order formula, Y_j is the eigenvariable not occurring in $\Gamma \cup \{A\}$, and $Y_j \neq A :\Leftrightarrow (\neg \forall x [Y_j(x) \leftrightarrow A(x)])^\infty$. Note that variables E_i, Y_i are uninterpreted in the calculus.

A tree $\mathcal{T} \subset {}^{<\omega}\mathbb{N}$ is constructed recursively as follows. At each node σ , a sequent and a family $\mathcal{E}^X(\sigma)$ of finite sets are assigned. At the bottom \emptyset , we put the empty sequent, and $\mathcal{E}^X(\sigma) = \emptyset$. The assignment $\mathcal{E}^X(\sigma)$ is done similarly as in Definition 3.3.

Suppose that the tree \mathcal{T} has been constructed up to a node $\sigma \in {}^{<\omega}\mathbb{N}$. Let $\{A_i\}_i$ be an enumeration of all first-order formulas (abstracts).

Case 0. The length $lh(\sigma) = 3i$: Apply one of inferences $(\bigvee), (\bigwedge), (\exists_{1st}^2)$, and $(prg)^\infty$ if it is possible. Otherwise repeat, i.e., apply an inference (Rep) .

When (\exists_{1st}^2) is applied backwards, a first-order A_j is chosen so that j is the least such that A_j has not yet been tested for the major formula $\exists X F(X)$ of the (\exists_{1st}^2) , and $Var(A_j) \subset Var(\Gamma \cup \{F\}) \cup \{P\}$.

$$\frac{\Gamma, \exists X F(X), F(A_j)}{\Gamma, \exists X F(X)} (\exists_{1st}^2)$$

When $(prg)^\infty$ is applied backwards to a formula $E_A(m)$ with $i = \lceil A \rceil$, the condition (3), $m \in \mathcal{E}_i^X(\sigma)$, and $Var(A) \subset Var(\Gamma)$ have to be met. Otherwise repeat.

$$\frac{\{((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X(\sigma) \cup \{n\}) \vdash \Gamma, n \not\prec_A^* m, E_A(n) : n \in \mathbb{N}\}}{((\mathcal{E}_j^X(\sigma))_{j \neq i}, \mathcal{E}_i^X(\sigma)) \vdash \Gamma, E_A(m)} (prg)^\infty$$

Case 1. $lh(\sigma) = 3\langle i, n \rangle + 1$: Apply the inference (ACA) backwards with the first-order $A \equiv A_i$ and an eigenvariable Y_j if $Var(A) \subset Var(\Gamma) \cup \{P\}$.

$$\frac{Y_j \neq A, \Gamma}{\Gamma} (ACA)$$

Otherwise repeat

Case 2. $lh(\sigma) = 3i + 2$: Apply the inference (WP) backwards with the relation $<_{A_i}$.

If the tree \mathcal{T} is not well-founded, then let \mathcal{R} be an infinite path through \mathcal{T} . We see for any i, n that at most one of $Q(n)$ or $\bar{Q}(n)$ is on \mathcal{R} for $Q \in \{E_i, Y_i, P : i \in \mathbb{N}\}$, and $[(P(n)) \in \mathcal{R} \Rightarrow n \notin \mathcal{P}] \& [(\bar{P}(n)) \in \mathcal{R} \Rightarrow n \in \mathcal{P}]$ due to the axioms Γ, L with $L \in \text{Diag}(\mathcal{P})$. Let $(\mathcal{Q})_i$ be the set defined by $n \in (\mathcal{Q})_{2i} \Leftrightarrow (E_i(n)) \notin \mathcal{R}$ and $n \in (\mathcal{Q})_{2i+1} \Leftrightarrow (Y_i(n)) \notin \mathcal{R}$.

$(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}}$ is shown to be a countable coded ω -model of $\text{ACA}_0 + \text{WOP}(\mathbf{g})$ as follows. The search procedure is fair, i.e., each formula is eventually analyzed on every path. To ensure fairness, formulas in sequents Γ are assumed to stand in a queue. The head of the queue is analyzed in **Case 0**, and the analyzed formula moves to the end of the queue in the next stage. We see from the fairness that $\text{Diag}(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}} \not\models A$ first by induction on the number of occurrences of logical connectives in first-order formulas A on the path \mathcal{R} , and then for Π_1^1 -formulas $\text{TI}[\prec_A]$ and Σ_1^1 -formulas $\neg \text{TI}[\prec_{\mathbf{g}_A}]$. Moreover $\text{Diag}(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}} \models \text{ACA}_0$ since the inference rules (*ACA*) are analyzed for every A_i . Finally we show $\text{Diag}(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}} \models \text{WOP}(\mathbf{g})$. Assume that $\text{Diag}(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}} \models \text{WO}[\prec_A]$ for a first-order A . The path \mathcal{R} passes through an inference (*WP*) for the relation \prec_A . If \mathcal{R} passes through the rightmost upper sequent $\neg \text{TI}[\prec_{\mathbf{g}_A}]$, then $\text{Diag}(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}} \not\models \neg \text{TI}[\prec_{\mathbf{g}_A}]$, i.e., $\text{Diag}(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}} \models \text{TI}[\prec_{\mathbf{g}_A}]$, and we are done. Suppose that \mathcal{R} passes through an n_0 -th upper sequent $((\mathcal{E}_j^X(\sigma_0))_{j \neq i}, \mathcal{E}_i^X(\sigma_0)) \vdash \Gamma_0, E_A(n_0)$ and $E_A = E_i$ with $i = \lceil A \rceil$. Since the condition (3), $n_0 \in \mathcal{E}_i^X(\sigma_0) = \{n_0\}$ is met, the formula $E_A(n_0)$ is analyzed after a number of steps at a $(\text{prg})^\infty$, and \mathcal{R} passes through an n_1 -th branch $((\mathcal{E}_j^X(\sigma_1))_{j \neq i}, \mathcal{E}_i^X(\sigma_1)) \vdash \Gamma_1, n_1 \not\prec_A^* n_0, E_A(n_1), E_A(n_0)$. We obtain $\text{Diag}(\mathcal{P}, (\mathcal{Q})_i)_{i \in \mathbb{N}} \not\models n_1 \not\prec_A^* n_0$, i.e., $n_1 \prec_A^{*, \mathcal{P}, \mathcal{Q}} n_0$. Also $\{n_0, n_1\} \subset \mathcal{E}_i^X(\sigma_1)$. In this way we obtain an infinite descending chain $\dots \prec_A^{*, \mathcal{P}, \mathcal{Q}} n_2 \prec_A^{*, \mathcal{P}, \mathcal{Q}} n_1 \prec_A^{*, \mathcal{P}, \mathcal{Q}} n_0$ from \mathcal{R} , contradicting the assumption $\text{WO}[\prec_{\mathcal{P}, \mathcal{Q}}]$.

In what follows assume that the tree \mathcal{T} is well-founded. Let Λ denote the least epsilon number larger than the order type of the Kleene-Brouwer ordering \prec_{KB} on the well-founded tree \mathcal{T} . We have $\text{WO}(\mathbf{g}'(\Lambda))$ by $\text{WOP}(\mathbf{g}')$ and $\text{WO}(\Lambda)$.

For $b < \Lambda$ let us write $S + (\text{ACA}) \vdash_c^b \Gamma$ when there exists a derivation of Γ in $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (\text{WP}) + (\text{ACA})$ such that its depth is bounded by b , the depth of nested applications of the rules (*WP*) is bounded by c , and the condition (3) is enjoyed for each inference $(\text{prg})^\infty$ in the derivation, where a family $\mathcal{E}^X(\sigma)$ of finite sets is assigned to each node σ in the derivation tree as in Definition 3.3.

For the inference

$$\frac{Y_j \neq A, \Gamma}{\Gamma} (\text{ACA})$$

substitute A for the eigenvariable Y_j , and deduce the valid formula $A = A$ logically in a finite number of steps, and then a $(\text{cut})_{1st}$ yields the lower sequent Γ . Axioms $\Gamma, \bar{Y}_j(n), Y_j(n)$ turns to another valid sequent $\Gamma', \neg A(n), A(n)$. In (*WP*), if Y_j occurs in $B(Y_j)$, then the variable $E_{B(Y_j)} \equiv E_i$ with $i = \lceil B(Y_j) \rceil$ should be renamed to $E_{B(A)} \equiv E_k$ with $k = \lceil B(A) \rceil$.

$$\frac{\Gamma(Y_j), E_{B(Y_j)}(n) \quad \neg \text{TI}[\prec_{\mathbf{g}_{B(Y_j)}}], \Gamma(Y_j)}{\Gamma(Y_j)} \quad \rightsquigarrow \quad \frac{\Gamma(A), E_{B(A)}(n) \quad \neg \text{TI}[\prec_{\mathbf{g}_{B(A)}}], \Gamma(A)}{\Gamma(A)}$$

Thus we obtain $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (\text{WP}) + (\text{cut})_{1st} \vdash_b^{\omega+b} \emptyset$ from $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (\text{WP}) + (\text{ACA}) \vdash_b^b \emptyset$, and $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (\text{WP}) \vdash_{2_p(b)}^{2_p(\omega+b)} \emptyset$ for a $p < \omega$ as in Proposition 3.7.

Here the condition (3) is forced in the search.

Theorem 3.5 yields $\emptyset \vdash_0^\delta \emptyset$ with $\delta = F(2_p(\omega + b), 2_p(b)) + 2_p(\omega + b)$ and $\mathcal{E}^X(\emptyset) = \Delta(\emptyset; \emptyset) = \emptyset$. This means that in the ω -logic, there exists a cut-free derivation of \emptyset in depth $\delta < \mathbf{g}'(\Lambda)$, which is seen from Proposition 3.4.3 and $b < \Lambda$. We see by induction up to the ordinal $\mathbf{g}'(\Lambda)$ that this is not the case. Therefore the tree \mathcal{T} must not be well-founded. Thus our proof of Theorem 1.6 is completed.

4 Elimination of the inference for well-ordering principle

It remains to show Theorems 3.6 and 3.5.

(**Proof** of Theorem 3.6). We can assume that the transitive closure \prec^* of the relation $n \prec m$ is irreflexive. Namely there is no sequence (n_0, \dots, n_k) ($k \geq 1$) such that $n = n_0 = n_k$ and $\forall j < k (n_{j+1} \prec n_j)$. Suppose that there exists such a sequence. By the assumption we obtain $\{n_0\} \vdash^{\alpha_0} E(n_0)$ for $\alpha_0 = \alpha$. Any positive literal $E(n_0)$ is not an axiom in $\text{Diag}(\emptyset)$. We see from $n_0 \in \{n_0\}$ by induction on ordinals α_0 that there must be an inference $(prg)_{\prec}^D$ in the witnessed derivation, and we obtain $\{n_0, n_1\} \vdash^{\alpha_1} E(n_0), E(n_1)$ for an $\alpha_1 < \alpha_0$. Again $P(n_0), P(n_1)$ is not an axiom in $\text{Diag}(\emptyset)$. In this way we would obtain an infinite descending chain $\{\alpha_m\}_{m < \omega}$ of ordinals such that $\{n_j\}_{j < k} \vdash_0^{\alpha_m} \{P(n_j) : j < k\}$.

By recursion on m , we define a non-empty finite set $\mathcal{E}(m)$, and an ordinal $\beta(m) \leq \alpha$ for which the followings hold for $\Delta(\mathcal{E}(m)) := \{E(n) : n \in \mathcal{E}(m)\}$.

$$\begin{aligned} \mathcal{E}(m) &\subset \{n : m \preceq^* n \leq m\} \ \& \ \mathcal{E}(m) \vdash^{\beta(m)} \Delta(\mathcal{E}(m)) \\ \forall n < m (m \prec^* n &\rightarrow \beta(m) < \beta(n)) \end{aligned} \quad (5)$$

Case 1. $\neg \exists n < m (m \prec^* n)$: Let $\mathcal{E}(m) = \{m\}$ and $\beta(m) = \alpha$. Then the conditions in (5) are fulfilled with $\Delta(\mathcal{E}(m)) = \{E(m)\}$.

Case 2. Otherwise: Pick a $k < m$ such that $m \prec^* k$ and $\beta(k) = \min\{\beta(n) : n < m, m \prec^* n\}$.

Then let $\mathcal{E}(m) = \mathcal{E}(k) \cup \{m\}$. On the other hand we have $\mathcal{E}(k) \vdash^{\beta(k)} \Delta(\mathcal{E}(k))$. The sequent $\Delta(\mathcal{E}(k))$ is not an axiom in $\text{Diag}(\emptyset)$. Search the lowest inference $(prg)_{\prec}^D$ in the derivation showing the fact $\mathcal{E}(k) \vdash^{\beta(k)} \Delta(\mathcal{E}(k))$:

$$\frac{\{\mathcal{E}(k) \cup \{n\} \vdash^{\beta_0} \Delta(\mathcal{E}(k)), E(n) : n \prec^* k'\}}{\mathcal{E}(k) \vdash^{\beta'} \Delta(\mathcal{E}(k))} (prg)_{\prec}^D$$

where $\beta_0 < \beta' \leq \beta(k)$, there may be some (Rep) 's below the inference $(prg)_{\prec}^D$, and $E(k') \in \Delta(\mathcal{E}(k))$ is the main formula of the inference $(prg)_{\prec}^D$. We have $m \prec^* k \preceq^* k'$, and $m \prec^* k'$. Pick the m -th branch in the upper sequents. We obtain $\mathcal{E}(m) \vdash^{\beta(m)} \Delta(\mathcal{E}(m))$ for $\beta(m) := \beta_0 < \beta(k)$. The conditions in (5) are fulfilled.

Now define a function $f(m)$ as follows.

$$f(m) = \max\{\omega^{\beta(m_0)} \# \dots \# \omega^{\beta(m_k)} : m_0 \prec^* \dots \prec^* m_k = m, m_0, \dots, m_{k-1} < m\}$$

where $\#$ denotes the natural sum. Note that the set $\{(m_0, \dots, m_k) : m_0 \prec^* \dots \prec^* m_k = m, m_0, \dots, m_{k-1} < m\}$ is finite since \prec^* is irreflexive.

We show the function f is a desired embedding between \prec^* and $<$. Assume $m \prec^* n$, and let m_0, \dots, m_k be a sequence such that $f(m) = \omega^{\beta(m_0)} \# \dots \# \omega^{\beta(m_k)}$, with $m_0 \prec^* \dots \prec^* m_k = m$ and $m_0, \dots, m_{k-1} < m$. We obtain $m_i \prec^* n$ for any $i \leq k$. Let us partition the set $\{0, \dots, k\}$ into two sets $A = \{i \leq k : n < m_i\}$ and $B = \{i \leq k : m_i < n\}$. Note that $m_i \neq n$ since \prec^* is irreflexive.

By (5) we obtain $\beta(m_i) < \beta(n)$ for each $i \in A$, and hence $\#\{\omega^{\beta(m_i)} : i \in A\} < \omega^{\beta(n)}$, where $\#\{\alpha_1, \dots, \alpha_n\} = \alpha_1 \# \dots \# \alpha_n$.

On the other hand we have $\#\{\omega^{\beta(m_i)} : i \in B\} \leq \max\{\omega^{\beta(n_0)} \# \dots \# \omega^{\beta(n_\ell)} : n_0 \prec^* \dots \prec^* n_{\ell-1}, n_0, \dots, n_{\ell-1} < n\}$. Therefore we conclude $f(m) < f(n)$. \square

Lemma 4.1 *For each $j \leq \ell$, let $<_j$ be a first-order formula with $j = \lceil <_j \rceil$. Let $\mathcal{E}^X = (\mathcal{E}_j^X)_{j < \ell}$ be finite sets, and \mathcal{E}_ℓ^X a finite set. Let $\Gamma \subset \bigcup_{j < \ell} \Delta(\mathcal{E}_j^X)$ be a sequent and $\Gamma_\ell \subset \Delta(\mathcal{E}_\ell^X) = \{E_\ell(n) : n \in \mathcal{E}_\ell^X\}$ a sequent. In the calculus $\text{Diag}(\mathcal{P}, \emptyset) + (\text{prg})^\mathcal{E}$, assume that $(\mathcal{E}^X, \mathcal{E}_\ell^X) \vdash^\alpha \Gamma, \Gamma_\ell$ for $\mathcal{E}^X = (\mathcal{E}_j^X)_{j < \ell}$. Then either $\mathcal{E}^X \vdash^\alpha \Gamma$ holds, or $\mathcal{E}_\ell^X \vdash^\alpha \Gamma_\ell$ holds.*

Proof. We show the lemma by induction on α . Assume that $(\mathcal{E}^X, \mathcal{E}_\ell^X) \vdash^\alpha \Gamma$ does not hold. The set $\Gamma \cup \Gamma_\ell$ consisting of positive literals $E_i(n)$, is not an axiom in $\text{Diag}(\mathcal{P}, \emptyset) + (\text{prg})^\emptyset$.

Consider the case when the last inference is a $(\text{prg})^\emptyset$ for a $<_j$:

$$\frac{\{(\mathcal{E}^X, \mathcal{E}_\ell^X)_{j,m} \vdash^\beta \Gamma, \Gamma_\ell, E_j(\bar{m}) : m <_{j, \emptyset}^* n\}}{(\mathcal{E}^X, \mathcal{E}_\ell^X) \vdash^\alpha \Gamma, \Gamma_\ell} (\text{prg})^\emptyset$$

where $E_j(\bar{n})$ is in $\Gamma \cup \Gamma_\ell$, and $(\mathcal{E}^X, \mathcal{E}_\ell^X)_{j,m}$ denotes the sequence $(\mathcal{E}^X, \mathcal{E}_\ell^X)$ except \mathcal{E}_j^X is replaced by $\mathcal{E}_j^X \cup \{m\}$. We have (3), $n \in \mathcal{E}_j^X$.

First consider the case $j \neq \ell$. By the assumption we see that there exists an $m <_{j, \emptyset}^* n$ such that $(\mathcal{E}^X, \mathcal{E}_\ell^X)_{j,m} \vdash^\beta \Gamma, E_j(\bar{m}), E_j(\bar{n})$ does not hold. IH yields $(\mathcal{E}^X, \mathcal{E}_\ell^X) \vdash^\beta \Gamma_\ell$.

Second consider the case $j = \ell$. We see from the assumption that for each $m <_{\ell, \emptyset}^* n$, $(\mathcal{E}^X, \mathcal{E}_\ell^X)_{\ell,m} \vdash^\beta \Gamma_\ell, E_\ell(\bar{m})$ holds. Then an inference $(\text{prg})^\emptyset$ yields $(\mathcal{E}^X, \mathcal{E}_\ell^X) \vdash^\alpha \Gamma_\ell$. \square

(**Proof** of Theorem 3.5). Let us prove Theorem 3.5 by induction on β . Suppose that in the calculus $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP)$, $\vdash_\alpha^\beta \Phi, \Gamma$ holds in a derivation π for a sequent $\sigma : \Phi, \Gamma$ such that Φ is a Σ_1^1 -formulas with $\text{Diag}(\mathcal{P}, \emptyset) \not\models \bigvee \Phi$ and $\Gamma \subset \Delta(\mathcal{E}^X(\sigma); \Phi, \Gamma)$. Moreover the condition (3) is assumed to be enjoyed for each $(\text{prg})^\infty$ occurring in π . Let $\mathcal{E}^X = \mathcal{E}^X(\sigma)$.

If Φ, Γ is an axiom in $\text{Diag}(\mathcal{P}) + (\text{prg})^\infty + (WP)$, then $\Delta(\mathcal{E}^X; \Gamma)$ is an axiom in $\text{Diag}(\mathcal{P}, \emptyset) + (\text{prg})^\mathcal{E}$ since $\text{Diag}(\mathcal{P}, \emptyset) \not\models \bigvee \Phi$. For example, consider the case

when $\{\bar{E}_i(m), E_i(m)\} \subset \Phi \cup \Gamma$. Since Γ contains positive literals only, we may assume $\bar{E}_i(m) \in \Phi$ and $E_i(m) \in \Gamma$. From $\text{Diag}(\mathcal{P}, \emptyset) \models \bar{E}_i(m)$ we see that this is not the case.

Consider the last inference in the derivation showing $\vdash_\alpha^\beta \Phi, \Gamma$.

Case 1. The last inference is a $(prg)^\infty$. For $\gamma < \beta$ and $i = \lceil A \rceil$, we have for an $E_i(m) \in \Phi \cup \Gamma$

$$\frac{\{((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X \cup \{n\}) \vdash_\alpha^\gamma \Phi, n \not\prec_A^* m, \Gamma, E_i(n) : n \in \mathbb{N}\}}{((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X) \vdash_\alpha^\beta \Phi, \Gamma} (prg)^\infty$$

IH yields $((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X \cup \{n\}) \vdash^{F(\gamma, \alpha) + \gamma} \Delta_n$ for each $n <_A^{*\emptyset} m$, where $E_i(n) \in \Delta_n = \Delta((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X \cup \{n\}; \Phi, n \not\prec_A^* m, \Gamma, E_i(m), E_i(n))$, and $\Delta((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X \cup \{n\}; \Phi, \Gamma, E_i(m)) \cup \{E_i(n)\} = \Delta_n$. By the assumption we obtain (3), $m \in \mathcal{E}_i^X$. An inference $(prg)^\emptyset$ with $F(\gamma, \alpha) + \gamma < F(\beta, \alpha) + \beta$ yields $((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X) \vdash^{F(\beta, \alpha) + \beta} \Delta((\mathcal{E}_j^X)_{j \neq i}, \mathcal{E}_i^X; \Phi, \Gamma, E_i(m))$.

Case 2. The last inference is a (WP) . For $\gamma < \beta$, $\alpha_0 < \alpha$, $\text{Var}(A) \subset \text{Var}(\Gamma) \cup \{P\}$, $i = \lceil A \rceil$, and the variable E_i not occurring in Γ , we have

$$\frac{\{((\mathcal{E}_j^X)_{j \neq i}, \{n\}) \vdash_{\alpha_0}^\gamma \Phi, \Gamma, E_i(n) : n \in \mathbb{N}\} \quad \mathcal{E}^X \vdash_{\alpha_0}^\gamma \neg \text{TI}[\prec_{\mathbf{g}_A}], \Phi, \Gamma}{(\mathcal{E}_j^X)_{j \neq i} \vdash_\alpha^\beta \Phi, \Gamma} (WP)$$

For the left upper sequent we have for each $n \in \mathbb{N}$

$$((\mathcal{E}_j^X)_{j \neq i}, \{n\}) \vdash_{\alpha_0}^\gamma \Phi, \Gamma, E_i(n)$$

By IH we obtain $((\mathcal{E}_j^X)_{j \neq i}, \{n\}) \vdash^{F(\gamma, \alpha_0) + \gamma} \Delta((\mathcal{E}_j^X)_{j \neq i}; \Phi, \Gamma, E_i(n))$. Here note that $\Delta((\mathcal{E}_j^X)_{j \neq i}; \Phi, \Gamma) \cup \{E_i(n)\} = \Delta((\mathcal{E}_j^X)_{j \neq i}, \{n\}; \Phi, \Gamma, E_i(n))$. From Lemma 4.1 we see that either $(\mathcal{E}_j^X)_{j \neq i} \vdash^{F(\gamma, \alpha_0) + \gamma} \Delta(\mathcal{E}^X; \Phi, \Gamma)$ or $((\mathcal{E}_j^X)_{j \neq i}, \{n\}) \vdash^{F(\gamma, \alpha_0) + \gamma} E_i(n)$ holds. If $(\mathcal{E}_j^X)_{j \neq i} \vdash^{F(\gamma, \alpha_0) + \gamma} \Delta(\mathcal{E}^X; \Phi, \Gamma)$, then we are done.

Assume $((\mathcal{E}_j^X)_{j \neq i}, \{n\}) \vdash^{F(\gamma, \alpha_0) + \gamma} E_i(n)$, i.e., $\{n\} \vdash^{F(\gamma, \alpha_0) + \gamma} E_i(n)$ for every n . By Theorem 3.6 we obtain an embedding f from $<_A^\emptyset$ to $\omega^{F(\gamma, \alpha_0) + \gamma + 1}$, which yields an embedding from $<_{\mathbf{g}(A)}^\emptyset$ to $\delta := \mathbf{g}(\omega^{F(\gamma, \alpha_0) + \gamma + 1})$ by Proposition 2.2. Hence we see from $\text{WO}(\delta)$ that $\text{Diag}(\mathcal{P}, \emptyset) \models \text{TI}[\prec_{\mathbf{g}_A}, C]$ for any first-order formulas C . Therefore $\text{Diag}(\mathcal{P}, \emptyset) \models \neg \text{TI}[\prec_{\mathbf{g}_A}]$.

Second consider the right upper sequent. IH yields

$$(\mathcal{E}_j^X)_{j \neq i} \vdash^{F(\gamma, \alpha_0) + \gamma} \Delta(\mathcal{E}^X; \Phi, \Gamma)$$

On the other side Proposition 3.4.1 with (4) yields $F(\gamma, \alpha_0) + \gamma \leq F(\gamma, \alpha) + \gamma \leq F(\beta, \alpha) + \beta$. Hence the assertion $\mathcal{E}^X \vdash^{F(\beta, \alpha) + \beta} \Delta(\mathcal{E}^X; \Phi, \Gamma)$ follows.

Case 3. The last inference is other than $(prg)^\infty$ and (WP) .

Consider first the case when the last inference is a rule for existential second-order quantifier.

$$\frac{(\mathcal{E}_j^X)_{j \neq i} \vdash_\alpha^\gamma \neg \text{TI}[\prec_{\mathbf{g}_A}], \neg \text{TI}[\prec_{\mathbf{g}_A}, C], \Phi, \Gamma}{(\mathcal{E}_j^X)_{j \neq i} \vdash_\alpha^\beta \neg \text{TI}[\prec_{\mathbf{g}_A}], \Phi, \Gamma} (\exists_{1st}^2)$$

where $\gamma < \beta$ and C is a first-order formula such that $Var(C) \subset Var(A, \Gamma)$. IH yields $(\mathcal{E}_j^X)_{j \neq i} \vdash^{F(\gamma, \alpha) + \gamma} \Delta(\mathcal{E}^X; \Phi, \Gamma)$ for $\Delta(\mathcal{E}^X; \Phi, \Gamma) = \Delta(\mathcal{E}^X; \neg \text{TI}[\langle \mathbf{g}_A, C \rangle], \Phi, \Gamma)$ by $Var(A, C) \subset Var(\Phi, \Gamma)$.

Second consider the case when the last inference is a rule (\bigwedge) For $\gamma < \beta$ we have

$$\frac{\{\mathcal{E}^X \vdash_\alpha^\gamma A_n, \Phi, \Gamma : n \in \mathbb{N}\}}{\mathcal{E}^X \vdash_\alpha^\beta \Phi, \Gamma} (\bigwedge)$$

where $\bigwedge_n A_n$ is in the set Φ . Let n be the least number such that $\text{Diag}(\mathcal{P}, \emptyset) \not\models A_n$. IH yields $\mathcal{E}^X \vdash^{F(\gamma, \alpha) + \gamma} \Delta(\mathcal{E}^X; \Phi, \Gamma)$.

Other cases are similarly seen. \square

References

- [1] Afshari, B., Rathjen, M.: Reverse Mathematics and Well-ordering Principles: A pilot study. *Ann. Pure Appl. Logic* 160, 231-237(2009)
- [2] Arai, T.: Derivatives of normal functions and ω -models, *Arch. Math. Logic* 57, 649-664(2017)
- [3] Girard, J.-Y.: Proof theory and logical complexity, vol. 1. Bibliopolis, Napoli(1987)
- [4] Marcone, A., Montalbán, A.: The Veblen functions for computability theorists. *Jour. Symb. Logic* 76, 575-602(2011)
- [5] Rathjen, M.: ω -models and Well-ordering Principles. In: Tennant, N. (ed.) *In Foundational Adventures: Essays in Honor of Harvey M. Friedman*, pp. 179-212. College Publications, London(2014)
- [6] Rathjen, M., Weiermann, A.: Reverse Mathematics and Well-ordering Principles. In: Cooper, S., Sorbi, A. (eds.) *Computability in Context: Computation and Logic in the Real World*, pp. 351-370. Imperial College Press, London (2011)
- [7] Schütte, K.: Proof Theory. Springer, Berlin(1977)
- [8] Simpson, S. G. : Subsystems of second order arithmetic, 2nd edition, *Perspectives in Logic*, Cambridge UP, Cambridge(2009)
- [9] Takeuti, G.: A remark on Gentzen's paper "Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie". *Proc. Japan Acad.* 39, 263-269(1963)
- [10] Takeuti, G.: Proof Theory, second edition. North-Holland, Amsterdam (1987)