

Iterated self-applicable truth

GEL & MR

14th March 2024

Contents

1	Overview	1
2	A theory	1
3	From the thesis	2
3.1	A weak extension of the theory I	2
3.2	An impredicative theory extending I	5
3.3	Extending the well-ordering proofs	9
4	Theories with multiple self-applicable truth predicates	10
4.1	Consistency	16
4.2	An ordinal notation system for impredicative theories	20
4.3	Lower bounds on the proof-theoretic ordinal	26
4.4	Upper bounds on the proof-theoretic ordinal	33
5	Conclusion	40

1 Overview

Should *is true* be *is justifiable* or *verifiable*?

2 A theory

Language \mathcal{L}_Ω : extend language of PRA by binary predicate T . Assuming an ordinal notation system of a sufficiently large ordinal κ , and $\eta < \kappa$, $T_\eta s$ means $T(\ulcorner \eta \urcorner, s)$.

For $\beta < \kappa$, define a theory T_β extending HA by axioms, for each $\eta \leq \kappa$:

$$\forall \ulcorner A \urcorner \ulcorner B \urcorner (T_\eta \ulcorner A \urcorner \rightarrow B \urcorner \rightarrow (T_\eta \ulcorner A \urcorner \rightarrow T_\eta \ulcorner B \urcorner)) \quad (\text{I})$$

$$\forall \ulcorner A(x) \urcorner (\forall n T_\eta \ulcorner A(\dot{n}) \urcorner \rightarrow T_\eta \ulcorner \forall x A(x) \urcorner) \quad (\text{U})$$

$$\forall \xi \leq \beta (T_\eta \ulcorner T_\xi \dot{x} \urcorner \rightarrow T_\eta x) \quad (\text{D})$$

$$\forall \xi < \eta (T_\eta x \rightarrow T_\eta \ulcorner T_\xi \dot{x} \urcorner) \quad (\text{R})$$

and rules of inference

$$\begin{array}{rcl} A \vdash T_\beta \ulcorner A \urcorner & & (\text{Nec}) \\ T_\beta \ulcorner A \urcorner \vdash A & & (\text{Conec}) \end{array}$$

What gives the strength is that (D) ‘collapses’ all internal truth predicates to the current level (η). In contrast, (R) only permits ‘expanding’ the internal level by lower level predicates.

With $T_\beta^c = T_\beta + \text{LEM}$, the thesis argues that

Theorem 1. *For every $p < \omega$,*

$$\|T_p^c\| = \vartheta(\Omega^{p+1} \cdot \omega).$$

Theorem 2. *For $\omega \leq \beta < \vartheta\Omega^\Omega$,*

$$\vartheta(\Omega^\beta \cdot \omega) \leq \|T_\beta^c\| \leq \vartheta(\Omega^{\beta+1} \cdot \omega).$$

Theorem 3. *The limit of the autonomous progression of $\{T_\beta^c : \beta < \Omega\}$ is the large Veblen ordinal, $\vartheta\Omega^\Omega$.*

It seems likely that the same claims hold for (intuitionistic) T_β .

3 From the thesis

chap:ext

In this section we ask what effect adding other axioms of truth to a theory has on its proof-theoretic strength. Each of the nine theories is maximally consistent, so one cannot add further Optional Axioms. Candidates for inclusion will be sentences such as $T(\ulcorner T\text{-Imp} \urcorner)$ and $T(\ulcorner \forall\text{-Inf} \urcorner)$ which are derivable in E, F and G, but not I. Recall G and I have the same proof-theoretic strength, but I includes the axiom T-Del, while G favours T-Intro.

3.1 A weak extension of the theory I

Theorem 4. *The theory obtained by adding to I one of the two sentences $T(\ulcorner T\text{-Imp} \urcorner)$ and $T(\ulcorner \forall\text{-Inf} \urcorner)$ is proof-theoretically equivalent to I, whereas the theory obtained by adding both sentences forms a consistent theory which proves the same arithmetical statements as the theory $\text{PA} + \text{TI}(<\varphi_{30})$. Thus, the conjunction of the two sentences is independent over I.*

Theorem 4 will be dealt with in two parts. We begin with proposition 3 below, in which we prove the addition of only one of the axioms $T(\ulcorner T\text{-Imp} \urcorner)$ and $T(\ulcorner \forall\text{-Inf} \urcorner)$ has no effect on the truth-free consequences of I. Essentially, we repeat the analysis of I as in ??, embedding the extensions of I in suitably chosen extensions of \mathcal{J}_∞ , which we prove are closed under T-Elim. The key is showing that the same ordinal bounds used in the analysis of \mathcal{J}_∞ can be used to bound applications of T-Elim in these extensions. For this we utilise ???? and two subsystems of \mathcal{J}_∞ .

Suppose (\dagger) denotes one of the two sentences, $T(\ulcorner T\text{-Imp} \urcorner)$ and $T(\ulcorner \forall\text{-Inf} \urcorner)$, I^\dagger denotes $I + (\dagger)$ and $\mathcal{J}_\infty^\dagger$ is the infinitary system \mathcal{J}_∞ with (Ax.4) replaced by

$$(\text{Ax.4}^\dagger) \quad \mathcal{J}_\infty^\dagger \upharpoonright_k^\alpha \Gamma, T(s) \text{ whenever } s^\mathbb{N} = \ulcorner A \urcorner \text{ with } A \text{ an } \mathcal{L}_T\text{-sentence and } \mathcal{G}_\infty \upharpoonright_0^\alpha A,$$

where $\mathcal{G}_\infty^\dagger$ represents \mathcal{G}_∞ , if (\dagger) is $T(\ulcorner T\text{-Imp} \urcorner)$, and \mathcal{G}_∞^* , if (\dagger) is $T(\ulcorner \forall\text{-Inf} \urcorner)$. It should be clear that $\mathcal{J}_\infty^\dagger$ supports cut elimination. Moreover, ???? entail

$$\mathcal{G}_\infty^\dagger \upharpoonright_0^\alpha T(\ulcorner A \urcorner) \text{ implies } \mathcal{G}_\infty^\dagger \upharpoonright_0^\beta A \text{ for some } \beta < \alpha^{\varepsilon^+}. \quad (1)$$

{exteqn:G8-TEl1}

Lemma 1. For every sequent Γ and ordinal α , $\mathcal{G}_\infty^\dagger|_0^\alpha \Gamma$ implies $\mathcal{J}_\infty^\dagger|_0^\alpha \Gamma$.

Proof. Proceed by induction on α . The only non-trivial case is where an application of (T-Intro) was used within $\mathcal{G}_\infty^\dagger$. In that case the sentence $\top(\ulcorner A \urcorner)$ is in Γ for some \mathcal{L}_T -sentence A and $\mathcal{G}_\infty^\dagger|_0^\beta A$ holds for some $\beta < \alpha$, whence (Ax.4[†]) yields $\mathcal{J}_\infty^\dagger|_0^\alpha \Gamma$. \square

Define an \mathcal{L}_T -structure $\mathfrak{M}_\alpha^\dagger$ by

$$\mathfrak{M}_\alpha^\dagger = \langle \mathbb{N}, \{ \ulcorner B \urcorner : \mathcal{G}_\infty^\dagger|_0^{\varepsilon_\alpha} B \} \rangle.$$

Proposition 2. For every T -positive sequent Γ and for every ordinal α ,

$$\mathcal{J}_\infty^\dagger|_0^\alpha \Gamma \text{ implies } \mathfrak{M}_\alpha^\dagger \models \Gamma.$$

Proof. Suppose $\mathcal{J}_\infty^\dagger|_0^\alpha \Gamma$. That $\mathfrak{M}_\alpha^\dagger \models \bigvee \Gamma$ holds if Γ is an instance of (Ax.4[†]) is immediate given the choice of $\mathfrak{M}_\alpha^\dagger$. Moreover, by eq. (1) and the choice of ordinal bound in $\mathfrak{M}_\alpha^\dagger$, $\mathfrak{M}_\alpha^\dagger \models \text{T-Del} \wedge \text{T-Imp} \wedge \text{T-Rep}$, for every α . So, the case where the last rule of inference is one of (T-Del), (T-Rep), (T-Imp) may be deduced from the induction hypothesis. This leaves only (T- ω) to consider, which is easily dealt with since the function $\alpha \mapsto \varepsilon_\alpha$ is strictly increasing and continuous. \square

Proposition 3. The theory obtained by adding as an axiom to \mathbf{l} only one of the two sentences $\top(\ulcorner \text{T-Imp} \urcorner)$ and $\top(\ulcorner \forall\text{-Inf} \urcorner)$ proves no more arithmetical theorems than \mathbf{l} .

Proof. Combining lemma 1 and proposition 2, yields

$$\mathcal{J}_\infty^\dagger|_0^\alpha \top(\ulcorner A \urcorner) \text{ implies } \mathcal{J}_\infty^\dagger|_0^{\varepsilon_\alpha} A.,$$

and hence for every sentence A ,

$$\mathbf{l}^\dagger \vdash A \text{ implies } \mathcal{J}_\infty^\dagger|_0^\alpha A \text{ for some } \alpha < \varphi_{20},$$

whence we learn that \mathbf{l}^\dagger proves the same arithmetical statements as \mathbf{l} . \square

We now move to the second part of theorem 4; namely establishing the consistency and proof-theoretic ordinal of the theory \mathbf{l}^\ddagger , the finitary theory \mathbf{l} with as an additional axiom,

$$\top(\ulcorner \forall\text{-Inf} \urcorner) \wedge \top(\ulcorner \text{T-Imp} \urcorner),$$

hereby denoted as (\ddagger) . The consistency of \mathbf{l}^\ddagger can be derived from Friedman and Sheard's model construction for \mathbf{l} [?, §3]. However, it is not clear from the construction what the proof-theoretic strength of \mathbf{l}^\ddagger is. We will embed \mathbf{l}^\ddagger in a variant of \mathcal{J}_∞ , allowing a bound on the proof-theoretic ordinal to be easily obtained. Observe that (\ddagger) is derivable in the infinitary system $\mathcal{J}_\infty^\ddagger$ defined as \mathcal{J}_∞ (??) with (Ax.4) replaced by

$$(\text{Ax.4}^\ddagger) \quad \mathcal{J}_\infty^\ddagger|_k^\alpha \Gamma, \top(s) \text{ whenever } s^\mathbb{N} = \ulcorner A \urcorner \text{ with } A \text{ an } \mathcal{L}_T\text{-sentence, and } \mathcal{G}_\infty|_0^\alpha A.$$

$\mathcal{J}_\infty^\ddagger$, like its predecessors, supports cut-elimination as we have not introduced any further means to derive non T-positive sequents. Closure under T-Elim can be shown for this system in the same manner as before using the models

$$\mathfrak{M}_\alpha^\ddagger := \langle \mathbb{N}, \{ \ulcorner B \urcorner : \mathcal{G}_\infty \upharpoonright_0^{\varphi 2\alpha} B \} \rangle.$$

First, however, note

$$\begin{aligned} \mathcal{G}_\infty \upharpoonright_0^\alpha \Gamma &\text{ implies } \mathcal{J}_\infty \upharpoonright_0^\alpha \Gamma, \\ \mathcal{G}_\infty \upharpoonright_0^\alpha \text{T}(\ulcorner A \urcorner) &\text{ implies } \mathcal{G}_\infty \upharpoonright_0^{\varepsilon_\alpha} A. \end{aligned}$$

Lemma 4. *For every T-positive sequent Γ and for every ordinal α ,*

$$\mathcal{J}_\infty \upharpoonright_0^\alpha \Gamma \text{ implies } \mathfrak{M}_\alpha^\ddagger \models \bigvee \Gamma.$$

Proof. Suppose $\mathcal{J}_\infty \upharpoonright_0^\alpha \Gamma$. That $\mathfrak{M}_\alpha^\ddagger \models \bigvee \Gamma$ holds if Γ is an instance of (Ax.4 ‡) is immediate given the choice of $\mathfrak{M}_\alpha^\ddagger$. Moreover, by the T-Elimination theorem for \mathcal{G}_∞ (??) and the choice of ordinal bound in $\mathfrak{M}_\alpha^\ddagger$, $\mathfrak{M}_\alpha^\ddagger \models \text{T-Del} \wedge \text{T-Imp} \wedge \text{T-Rep}$, for every α , and thus the case where the last rule of inference is one of (T-Del), (T-Rep), (T-Imp) may be deduced from the induction hypothesis. This leaves only (T- ω) to consider, which is easily dealt with since the function $\alpha \mapsto \varphi 2\alpha$ is strictly increasing and continuous. \square

As a corollary one obtains that finite derivations within \mathbb{I}^\ddagger can be transformed into derivations in $\mathcal{J}_\infty^\ddagger$ bounded by the ordinal $\varphi 30$, and hence a bound on the proof-theoretic strength of \mathbb{I}^\ddagger .

Corollary 5. *\mathbb{I}^\ddagger proves no more arithmetical statements than $\text{PA} + \text{TI}(<\varphi 30)$ and hence is consistent.*

Proof. We embed \mathbb{I}^\ddagger into $\mathcal{J}_\infty^\ddagger$. If $\mathcal{J}_\infty \upharpoonright_0^\alpha \text{T}(\ulcorner A \urcorner)$, the preceding two lemmata imply $\mathcal{J}_\infty \upharpoonright_0^{\varphi 2\alpha} A$. Thus every arithmetical theorem of \mathbb{I}^\ddagger is cut-free derivable in $\mathcal{J}_\infty^\ddagger$ with some ordinal bound smaller than $\varphi 30$. \square

A lower bound on the strength of \mathbb{I}^\ddagger is easier to establish since axiom (\ddagger) allows one to deduce that \mathbb{I}^\ddagger proves a reflection principle over S_1 .

Proposition 6. $\mathbb{I}^\ddagger \vdash \forall x (\text{Bew}_{S_1}(x) \wedge \text{Sent}_{\mathcal{L}_1}(x) \implies \text{T}(x)).$

Proof. We argue informally within \mathbb{I}^\ddagger . Suppose A is an axiom of S_1 . If A is either T-Imp or \forall -Inf, $\text{T}(\ulcorner A \urcorner)$ is an axiom of \mathbb{I}^\ddagger and we are done; if A is arithmetical, $\mathbb{I}^\ddagger \vdash \text{T}(\ulcorner A \urcorner)$ by ??.¹ This leaves only the case A is the axiom $\forall x (\text{val}(x) \implies \text{T}(\ulcorner \text{val}(x) \urcorner))$ or $\forall x (\text{Ax}_{\text{PRA}}(x) \implies \text{T}(x))$ of Base_T . For the latter, ?? and T-Rep imply

$$\mathbb{I}^\ddagger \vdash \forall x (\text{T}(\ulcorner \neg \text{Ax}_{\text{PRA}}(x) \urcorner) \vee \text{T}(\ulcorner \text{T}(x) \urcorner)),$$

and hence

$$\mathbb{I}^\ddagger \vdash \text{T}(\ulcorner \forall x (\text{Ax}_{\text{PRA}}(x) \implies \text{T}(x)) \urcorner).$$

The former axiom is similar. The induction step is straightforward to establish as the presence of T-Del in \mathbb{I}^\ddagger formalises applications of T-Elim in S_1 . \square

¹In ?? it was shown $\text{Base}_T + \forall\text{-Inf} \vdash A \implies \text{T}(\ulcorner A \urcorner)$ for every arithmetical sentence A .

Proposition 6 provides the opportunity to reflect upon derivations in S_1 . In ?? we established $S_1 \vdash \text{TI}(\alpha, A) \implies \text{wo}(\varepsilon_\alpha)$, where $A(x)$ is the formula $\text{wo}(\varepsilon_x)$ (cf. the proof of ??). Thus, using T-Del we obtain

$$I^\ddagger \vdash \forall x(\text{wo}(x) \implies \text{wo}(\varepsilon_x)), \quad (2)$$

and hence

$$I^\ddagger \vdash \forall x(\text{wo}(\varphi\bar{2}x) \implies \text{wo}(\varphi\bar{2}(x+1))) \quad (3)$$

by induction. Equation (3) yields $I^\ddagger \vdash \text{Prog}A$, where $A(x)$ is the formula $\text{wo}(\varphi\bar{2}x)$ (the limit case being provided by ??). Thus, if $I^\ddagger \vdash \text{TI}(<\alpha)$ it follows that $I^\ddagger \vdash \text{wo}(\varphi\bar{2}\bar{\alpha})$ and so $I^\ddagger \vdash \text{TI}(<\varphi\bar{2}\bar{\alpha})$; hence I^\ddagger proves the schema of transfinite induction for all ordinals smaller than $\varphi 30$.

As the addition of (\ddagger) to I yields a stronger (but consistent) theory, (\ddagger) must be independent from I .

We are now in a position to complete the proof of theorem 4. Proposition 3 shows I formulated with only one of $\text{T}(\ulcorner \text{T-Imp} \urcorner)$ and $\text{T}(\ulcorner \forall\text{-Inf} \urcorner)$ has the same proof-theoretic strength as I , whereas the preceding remark and corollary 5 demonstrate that I extended by the conjunction of the two sentences has proof-theoretic ordinal $\varphi 30$.

3.2 An impredicative theory extending I

In the previous section we observed how adding the principles $\text{T}(\ulcorner \text{T-Imp} \urcorner)$ and $\text{T}(\ulcorner \forall\text{-Inf} \urcorner)$ to the theory I resulted in a proof-theoretically stronger theory. These two principles are also statements one might reasonably expect a self-applicable truth predicate to satisfy. We do not wish to embark on a discussion on the relative acceptability of these principles, but rather ask whether a similar scenario may occur with the other theories of truth considered here. The theories D and G both contain T-Intro, T-Imp and $\forall\text{-Inf}$ and hence within them one may also derive $\text{T}(\ulcorner \text{T-Imp} \urcorner)$ and $\text{T}(\ulcorner \forall\text{-Inf} \urcorner)$. Both theories, however, have proof-theoretic ordinal strictly smaller than that of I^\ddagger . The reason appears to be the lack of a principle similar to T-Del which plays an essential role in the proof-theoretic analysis of F , E and I^\ddagger (cf. ?? and proposition 6). On the other hand, comparing I to F also highlights the importance of T-Intro. The next lemma demonstrates that I^\ddagger is closed under a restricted form of T-Intro and suggests it is indeed the melding of T-Del and T-Intro that increases the proof-theoretic strength of E and F beyond that of I or G .

Lemma 7. $S_1 + \text{T-Intro}$ is a sub-theory of I^\ddagger but not of I .

Proof. Observe that, due to T-Rep in I^\ddagger proposition 6 may be generalised to obtain

$$I^\ddagger \vdash \forall x(\text{Bew}_{S_1+\text{T-Intro}}(x) \wedge \text{Sent}_{\mathcal{L}_1}(x) \implies \text{T}(x)). \quad (4)$$

Since I^\ddagger contains T-Elim, eq. (4) completes the first part of the lemma. The second part is straightforward: S_1 contains the axioms $\forall\text{-Inf}$ and T-Imp, hence $S_1 + \text{T-Intro} \vdash \text{T}(\ulcorner \forall\text{-Inf} \urcorner) \wedge \text{T}(\ulcorner \text{T-Imp} \urcorner)$, and so $S_1 + \text{T-Intro}$ cannot be a sub-theory of I due to theorem 4. \square

Lemma 7 shows that in forming deductions within I^\ddagger one may apply T-Intro provided one has not already used the axioms T-Rep and T-Del. The argument that it is only the combined effect of T-Intro and T-Del that increases the proof-theoretic strength of these theories is supported by the fact I^\ddagger , with its restricted T-Intro, is only slightly stronger than I and still significantly weaker than F . One

might therefore reasonably ask what effect adding the principle $T(\ulcorner T\text{-Del} \urcorner)$ to I^\ddagger has. Denote this new theory, $I + T(\ulcorner T\text{-Imp} \urcorner) + T(\ulcorner \forall\text{-Inf} \urcorner) + T(\ulcorner T\text{-Del} \urcorner)$, by S_3 . A glance at lemma 7 reveals

$$S_3 \vdash \forall x (\text{Bew}_{S_2}(x) \wedge \text{Sent}_{\mathcal{L}_T}(x) \implies T(x)), \quad (5)$$

so S_2 , and hence F , are sub-theories of S_3 . This marks a significant jump in proof-theoretic strength, provided of course S_3 is consistent. Consistency, though, is not obvious; after all eq. (5) shows S_3 contains $T\text{-Del}$, $T\text{-Rep}$, $T\text{-Elim}$ and closure under $T\text{-Intro}$ for proofs not involving $T\text{-Rep}$, which is extremely close to the inconsistent quadruple of Optional Axioms $T\text{-Del}$, $T\text{-Rep}$, $T\text{-Intro}$ and $T\text{-Elim}$. We will show S_3 is indeed consistent and the ability to apply $T\text{-Intro}$ to the axiom $T\text{-Rep}$ is essential to derive the inconsistency described. Consistency of S_3 will ultimately result from expanding the model construction for F outlined by Friedman and Sheard. Define a hierarchy of theories by recursion.

$$\begin{aligned} \text{Th}_0 &= \text{Base}_T, \\ \text{Th}_{n+1} &= \text{Base}_T + T\text{-Del} + \forall\text{-Inf} + \{T(\ulcorner A \urcorner) : \text{Th}_n \vdash_\omega A\}. \end{aligned}$$

In [?, §3] the authors show that the \mathcal{L}_T -structure $\langle \mathbb{N}, \{\ulcorner A \urcorner : \text{Th}_n \vdash_\omega A\} \rangle$ forms a model of Th_{n+1} for each $n \in \mathbb{N}$, whereby

$$\text{Th}_\infty := \{A : \text{Th}_n \vdash_\omega A \text{ for some } n\}$$

is a consistent theory closed under $T\text{-Intro}$ and $T\text{-Elim}$, and hence contains F .

We will extend this construction into the transfinite and utilise the resulting theory as an interpretation of the truth predicate of S_3 . To help visualise the reasoning behind extending the previous construction consider the following example. Suppose A is a sentence with $F \vdash A$; $T\text{-Intro}$ yields

$$\begin{aligned} F &\vdash T(\ulcorner A \urcorner), \\ F &\vdash T(\ulcorner T(\ulcorner A \urcorner) \urcorner), \end{aligned}$$

and in general

$$F \vdash T(f(\bar{k}, \ulcorner A \urcorner)) \quad (6)$$

where $f(0, n) = \ulcorner \bar{n} \urcorner$ and $f(m+1, n) = \ulcorner T(f(m, n)) \urcorner$, but

$$F \not\vdash \forall x T(f(x, \ulcorner A \urcorner)). \quad (7)$$

For the same reason we see that the theory Th_∞ defined above is a finite theory and not closed under the ω -rule. Let $T^\omega(y)$ denote $\forall x T(f(x, y))$. As a result of eq. (5) and $T\text{-Rep}$ one has $S_3 \vdash T^\omega(\ulcorner A \urcorner)$, whence $T\text{-Rep}$ and $\forall\text{-Inf}$ yields

$$S_3 \vdash T(\ulcorner T^\omega(\ulcorner A \urcorner) \urcorner), \quad (8)$$

and so

$$S_3 \vdash \forall x T(f(x, \ulcorner T^\omega(\ulcorner A \urcorner) \urcorner)), \quad (9)$$

that is $S_3 \vdash T^\omega(\ulcorner T^\omega(\ulcorner A \urcorner) \urcorner)$. In order to capture eq. (8) one would need to extend the model construction for F to $\text{Th}_\omega = \{A : \text{Th}_\infty \vdash_\omega A\}$. To manage eq. (9) one must construct theories

$$\begin{aligned} \text{Th}_{\omega+1} &= \text{Base}_T + T\text{-Del} + \forall\text{-Inf} + \{\ulcorner A \urcorner : \text{Th}_\omega \vdash A\}, \\ \text{Th}_{\omega+m+1} &= \text{Base}_T + T\text{-Del} + \forall\text{-Inf} + \{\ulcorner A \urcorner : \text{Th}_{\omega+m} \vdash A\}, \end{aligned}$$

and finally $\text{Th}_{\omega \cdot 2} = \{A : \exists n \text{Th}_{\omega+n} \vdash_{\omega} A\}$. Of course we are far from exhausting the possible iterations of the truth predicate in S_3 and as a result this construction must continue. It is problematic, however, to establish the closure of the theories Th_{α} under T-Elim when α is a limit ordinal. For example, to construct an \mathcal{L}_T -model $\langle \mathbb{N}, \mathfrak{X} \rangle$ of Th_{ω} one requires \mathfrak{X} to be closed under ω -logic (as $\text{Th}_{\omega} \vdash \forall\text{-Inf}$) and contain $\bigcup_n \text{Th}_n$, whence \mathfrak{X} extends Th_{ω} . Moreover, $\text{Th}_{\omega} \vdash \text{T-Del}$, so \mathfrak{X} must itself be closed under T-Elim, a criterion apparently at least as complex as what we are seeking to establish.

The problems described above can be avoided by borrowing the techniques used in the analysis of the infinitary system \mathcal{F}_{∞} (??). We will construct a transfinite hierarchy of theories of truth formulated in a Tait-style sequent calculus. We then construct models for only the T-positive part of these theories, allowing us to establish closure under T-Elim without the interpretation necessarily being closed under ω -logic.

Define a theory \mathcal{F}_{α} for $\alpha \in \mathbb{O}$ according to the following rules, in which Γ, Δ denote arbitrary sequents (finite sets of sentences). There are three axioms:

- Ax.1. $\mathcal{F}_{\alpha} \vdash \Gamma, A$ whenever A is a true atomic literal,
- Ax.2. $\mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(s), \neg\text{T}(s)$ for every closed term s ,
- Ax.3. $\mathcal{F}_{\alpha} \vdash \Gamma, \neg\text{T}(s)$ if $s^{\mathbb{N}}$ is not the code of an \mathcal{L}_T sentence.

The rules of this system are the usual arithmetical rules (\wedge), (\vee_0), (\vee_1), (\exists) together with the following six rules.

- $$\frac{\mathcal{F}_{\beta} \vdash \Gamma \quad \beta \leq \alpha}{\mathcal{F}_{\alpha} \vdash \Gamma, \Delta} \text{W}$$
- $$\frac{\mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(\ulcorner A \urcorner) \quad \mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(\ulcorner A \implies B \urcorner)}{\mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(\ulcorner B \urcorner)} \text{Imp}$$
- $$\frac{\mathcal{F}_{\beta} \vdash A \quad \beta < \alpha}{\mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(\ulcorner A \urcorner)} \text{Intro}$$
- $$\frac{\mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(\ulcorner \text{T}(s) \urcorner)}{\mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(s)} \text{Del}$$
- $$\frac{\mathcal{F}_{\alpha} \vdash \Gamma, \forall x \text{T}(\ulcorner A(\dot{x}) \urcorner)}{\mathcal{F}_{\alpha} \vdash \Gamma, \text{T}(\ulcorner \forall x A(x) \urcorner)} \text{U}$$
- $$\frac{\mathcal{F}_{\alpha} \vdash \Gamma, A(\bar{n}) \quad \text{for every } n \in \mathbb{N}}{\mathcal{F}_{\alpha} \vdash \Gamma, \forall x A(x)} \omega$$

By writing $\mathcal{F}_{\alpha} \vdash \Gamma$ we mean, of course, that one may derive $\mathcal{F}_{\alpha} \vdash \Gamma$ by means of the rules above. It is not hard to see that \mathcal{F}_{α} extends the system \mathcal{F}_{∞} introduced in ?? but we do not impose a bound on the height of a derivation. Moreover, the same argument establishes \mathcal{F}_{α} admits the cut rule.

Lemma 8 (Admissibility of cut). *$\mathcal{F}_{\alpha} \vdash \Gamma, A$ and $\mathcal{F}_{\alpha} \vdash \Delta, \neg A$ implies $\mathcal{F}_{\alpha} \vdash \Gamma, \Delta$ for every sentence A and ordinal α .*

We omit the technical results required to establish lemma 8 and related results and instead refer the reader to the relevant parts of ??.

The theory \mathcal{F}_n is simply a reformulation of Th_n , so $F \vdash A$ implies $\mathcal{F}_\omega \vdash A$ whenever A is a sentence. Define

$$\mathcal{F}_{<\alpha} = \{\ulcorner A \urcorner : \mathcal{F}_\beta \vdash A \text{ for some } \beta < \alpha.\}.$$

In fact, $F \vdash A$ implies $\ulcorner A \urcorner \in \mathcal{F}_{<\omega}$ if A an \mathcal{L}_T -sentence. Presenting the theories in this manner allows us to expand upon the proof of T-Elimination for \mathcal{F}_∞ (??) and consider derivations with transfinite T-rank.

Lemma 9.

1. $\ulcorner A \urcorner \in \mathcal{F}_{<\alpha}$ implies $\mathcal{F}_\alpha \vdash A$,
2. $\mathcal{F}_\alpha \vdash \Gamma$ and Γ a T-positive sequent implies $\langle \mathbb{N}, \mathcal{F}_{<\alpha} \rangle \models \bigvee \Gamma$,
3. $\mathcal{F}_\alpha \vdash T(\ulcorner A \urcorner)$ implies $\mathcal{F}_\alpha \vdash A$.

Proof. We argue by transfinite induction on α .

- (i). $\ulcorner A \urcorner \in \mathcal{F}_{<\alpha}$ implies $\mathcal{F}_\beta \vdash A$ for some $\beta < \alpha$, by definition, whence weakening yields $\mathcal{F}_\alpha \vdash A$.
- (ii). Here we have a further *subsidiary* induction on the height of the derivation $\mathcal{F}_\alpha \vdash \Gamma$. In the base case, only in (Ax.1) is Γ a T-positive sequent, and of course $\langle \mathbb{N}, \mathcal{F}_{<\alpha} \rangle \models \bigvee \Gamma \vee A$ for any true literal A . This leaves the induction step, so suppose $\mathcal{F}_\alpha \vdash \Gamma$ was derived via one of the rules of \mathcal{F}_α . If the last applied rule was T-Intro, $\langle \mathbb{N}, \mathcal{F}_{<\alpha} \rangle \models \bigvee \Gamma$ holds as, by definition, $\mathcal{F}_\beta \vdash A$ implies $\ulcorner A \urcorner \in \mathcal{F}_{<\alpha}$ if $\beta < \alpha$. In all other cases the premises of the last applied rule are also T-positive and so the subsidiary induction hypothesis can be applied. Weakening and the ω -rule are both immediate consequences of the induction hypothesis, leaving only the cases of T-Imp, T-Del and \forall -Inf to deal with. The main induction hypothesis for (ii) shows $\mathcal{F}_{<\alpha}$ is closed under T-Elim. By lemma 8 it is also closed under the cut rule, so $\langle \mathbb{N}, \mathcal{F}_{<\alpha} \rangle \models \text{T-Del} \wedge \text{T-Imp}$ and hence $\langle \mathbb{N}, \mathcal{F}_{<\alpha} \rangle \models \bigvee \Gamma$ follows from the subsidiary induction hypothesis if the last rule applied was T-Del or T-Imp. Suppose, therefore, that the last rule applied was \forall -Inf. Then α is not a limit ordinal and $\mathcal{F}_{<\alpha}$ is closed under the ω -rule,² so we deduce $\langle \mathbb{N}, \mathcal{F}_{<\alpha} \rangle \models \bigvee \Gamma$ from the subsidiary induction hypothesis, completing (ii).
- (iii). By (ii), $\mathcal{F}_\alpha \vdash T(\ulcorner A \urcorner)$ implies $\ulcorner A \urcorner \in \mathcal{F}_{<\alpha}$, whence (i) yields $\mathcal{F}_\alpha \vdash A$. □

Theorem 5. S_3 is a consistent theory.

Proof. Let $\mathcal{F}_{<\Omega} = \{\ulcorner A \urcorner : \mathcal{F}_\beta \vdash A \text{ for some } \beta < \Omega\}$, where Ω denotes the first uncountable ordinal. $\mathcal{F}_{<\Omega}$ provides a suitable interpretation of the truth predicate in S_3 . Define $\mathfrak{M} = \langle \mathbb{N}, \mathcal{F}_{<\Omega} \rangle$ and

$$\text{Th}_\Omega = \text{Base}_T + \text{T-Del} + \forall\text{-Inf} + \text{T-Rep} + \{T(\bar{x}) : x \in \mathcal{F}_{<\Omega}\}.$$

$\mathcal{F}_{<\Omega}$ contains PA_T and is closed under the cut rule, so $\mathfrak{M} \models \text{Base}_T$. Moreover, the previous lemma shows $\mathcal{F}_{<\Omega}$ is closed under T-Elim and T-Intro, whence $\mathfrak{M} \models \text{T-Del} \wedge \text{T-Rep}$. To deduce $\mathfrak{M} \models \text{Th}_\Omega$, all that remains is to show $\mathfrak{M} \models \forall\text{-Inf}$, that is, prove $\mathcal{F}_{<\Omega}$ is closed under ω -logic. Suppose, therefore, there are ordinals $\alpha_n < \Omega$ such that $\mathcal{F}_{\alpha_n} \vdash A(\bar{n})$ for each n . The set $\{\alpha_n : n < \omega\}$ is a countable set of countable ordinals and hence is bounded in Ω . In other words, there is some $\beta < \Omega$ such that $\alpha_n < \beta$ for every $n < \omega$. By weakening $\mathcal{F}_\beta \vdash A(\bar{n})$ for every n and thus $\ulcorner \forall x A(x) \urcorner \in \mathcal{F}_{<\Omega}$, as desired.

To conclude the proof, observe that $\ulcorner A \urcorner \in \mathcal{F}_{<\Omega}$ implies $\text{Th}_\Omega \vdash_\omega A$, whence Th_Ω is closed under T-Elim and thus S_3 is a sub-theory of Th_Ω . □

²Note \mathcal{F}_0 is vacuously closed under (ω) and if $\alpha = \beta + 1$, $\mathcal{F}_{<\alpha} = \{\ulcorner A \urcorner : \mathcal{F}_\beta \vdash A\}$ which is also closed under (ω) .

3.3 Extending the well-ordering proofs

Since S_3 is consistent we can concern ourselves with determining its proof-theoretic ordinal. Recall from ?? the formula $\text{wo}(y)$ denoting

$$\forall^\Gamma B(x)^\neg \text{T}(\neg \text{TI}(\dot{y}, B)^\neg),$$

and let $A(y)$ abbreviate the formula $\forall \alpha (\text{wo}(\alpha) \implies \text{wo}(\varphi y \alpha))$. In our analysis of S_2 (??) we presented a means to pass from $S_2 \vdash A(\bar{n})$ to $S_2 \vdash A(\overline{n+1})$. This process involved two crucial steps: showing for every n there is a formula $B_n(x)$ such that

$$S_2 \vdash A(\bar{n}) \implies \forall \alpha (\text{TI}(\alpha, B_n) \implies \text{wo}(\varphi(\overline{n+1})\alpha)); \quad (10)$$

and proving from eq. (10) and $S_2 \vdash A(\bar{n})$ we can deduce $S_2 \vdash A(\overline{n+1})$ using T-Intro, \forall -Inf and T-Del. Even within S_2 eq. (10) can be generalised to arbitrary $\beta < \Gamma_0$ in so far as we can pick a formula $B_\beta(x)$ for each $\beta < \Gamma_0$ such that

$$S_2 \vdash \forall \beta [A(\beta) \implies \forall \alpha (\text{TI}(\alpha, B_\beta) \implies \text{wo}(\varphi(\beta+1)\alpha))].$$

The theory S_3 also provides the ability to reflect upon derivations in S_2 , so one may expect

$$S_3 \vdash \forall \beta \exists^\Gamma B(x)^\neg [\text{T}(\neg A(\dot{\beta}) \implies \forall \alpha (\text{TI}(\alpha, B) \implies \text{wo}(\varphi(\dot{\beta}+1)\alpha))^\neg)]. \quad (11)$$

Since, $S_3 \vdash \text{T}(x \rightarrow y) \implies \text{T}(\neg \text{T}(\dot{x}) \implies \text{T}(\dot{y})^\neg)$ owing to T-Rep and $\text{T}(\neg \text{T-Imp}^\neg)$, eq. (11) entails

$$S_3 \vdash \forall \beta [\text{T}(\neg A(\dot{\beta})^\neg) \implies \text{T}(\neg \forall \alpha [\text{wo}(\alpha) \implies \text{wo}(\varphi(\dot{\beta}+1)\alpha)]^\neg)],$$

that is $S_3 \vdash \forall \beta [\text{T}(\neg A(\dot{\beta})^\neg) \implies \text{T}(\neg A(\dot{\beta}+1)^\neg)]$. This enables us to easily deduce $\text{wo}(\varphi \omega 0)$, but without a further application of T-Intro it is not clear that one could necessarily progress beyond this. As the next proposition demonstrates, we can proceed well beyond the strength of S_2 .

Proposition 10. $S_3 \vdash \text{Prog}C$, where $C(x)$ abbreviates the formula $\text{T}(\neg \forall \alpha (\text{wo}(\alpha) \implies \text{wo}(\varphi \dot{x} \alpha))^\neg)$.

Proof. This involves proving three cases:

1. $S_3 \vdash C(\bar{0})$, i.e. $\text{T}(\neg \forall \alpha (\text{wo}(\alpha) \implies \text{wo}(\varphi \dot{0} \alpha))^\neg)$;
2. $S_3 \vdash C(\beta) \implies C(\beta+1)$; and
3. $S_3 \vdash (\forall \alpha < \lambda C(\alpha)) \implies C(\lambda)$ for limit ordinals λ .

a) holds since S_2 is a sub-theory of S_3 . To show b) we begin by following the proof of ?? and argue within S_2 . Assume

$$\forall \alpha (\text{wo}(\alpha) \implies \text{wo}(\varphi \beta \alpha)).$$

Let $\varphi^0 \beta \alpha = \alpha$ and $\varphi^{k+1} \beta \alpha = \varphi \beta (\varphi^k \beta \alpha)$. By induction we may deduce $\forall k \text{wo}(\varphi^k \beta 0)$ and

$$\text{wo}(\varphi \beta' \alpha) \implies \forall k \text{wo}(\varphi^k \beta (\varphi \beta' \alpha + 1)), \quad (12)$$

where β' denotes $\beta + 1$. Since

$$\begin{aligned} \sup_{k \in \mathbb{N}} \varphi^k \beta 0 &= \varphi \beta' 0, \\ \sup_{k \in \mathbb{N}} \varphi^k \beta (\varphi \beta' \alpha + 1) &= \varphi \beta' (\alpha + 1), \end{aligned}$$

and $\text{wo}(\alpha)$ is progressive in α , we may deduce

$$\begin{aligned} & \text{wo}(\varphi\beta'0), \\ & (\forall\gamma < \alpha \text{ wo}(\varphi\beta'\gamma)) \implies \text{wo}(\varphi\beta'\alpha); \end{aligned}$$

i.e. $\text{Progwo}_{\beta'}$, where $\text{wo}_y(x)$ is the formula $\text{wo}(\varphi yx)$. Thus,

$$S_2 \vdash \forall\alpha(\text{wo}(\alpha) \implies \text{wo}(\varphi\beta\alpha)) \implies \text{Progwo}_{\beta'}. \quad (13) \quad \{\text{exteqn:S3wo6}\}$$

Stepping out of S_2 into the theory S_3 and using eq. (13) yields

$$S_3 \vdash \forall\beta[C(\beta) \implies T(\ulcorner \text{Progwo}_{\beta'} \urcorner)]. \quad (14) \quad \{\text{exteqn:S3wo7}\}$$

But $S_2 \vdash \forall\beta[T(\ulcorner \text{Progwo}_{\beta'} \urcorner) \implies \forall\alpha(\text{wo}(\alpha) \implies \text{wo}(\varphi\beta\alpha))]$, so eq. (14) (with an application of T-Rep) implies

$$S_3 \vdash \forall\beta[C(\beta) \implies T(\ulcorner \forall\alpha(\text{wo}(\alpha) \implies \text{wo}(\varphi\beta'\alpha)) \urcorner)],$$

that is, $S_3 \vdash \forall\beta(C(\beta) \implies C(\beta + 1))$.

Finally we show c). Again we start by arguing within the theory S_2 . Assume

$$\forall\beta < \lambda \forall\alpha(\text{wo}(\alpha) \implies \text{wo}(\varphi\beta\alpha)). \quad (15) \quad \{\text{exteqn:S3wo8}\}$$

By induction this implies $\text{wo}(\varphi\lambda\alpha) \implies \forall\beta < \lambda \forall k \text{ wo}(\varphi^k\beta(\varphi\lambda\alpha + 1))$, and so we deduce

$$\begin{aligned} & \text{wo}(\varphi\lambda 0), \\ & (\forall\gamma < \alpha \text{ wo}(\varphi\lambda\gamma)) \implies \text{wo}(\varphi\lambda\alpha); \end{aligned}$$

in other words, $S_2 \vdash [\forall\beta < \lambda \forall\alpha(\text{wo}(\alpha) \implies \text{wo}(\varphi\beta\alpha))] \implies \text{Progwo}_{\lambda}$. An application of T-Intro and T-Rep implies $S_3 \vdash (\forall\beta < \lambda C(\beta)) \implies C(\lambda)$, completing the proof. \square

Theorem 6. $\|S_3\| \geq \Gamma_0$.

Proof. Since S_3 contains S_0 we have $\text{wo}(\varphi\bar{0}\bar{0})$ and so also $\text{TI}(<1)$. Furthermore if $S_3 \vdash \text{TI}(<\alpha)$, the previous proposition implies $S_3 \vdash C(\bar{\gamma})$ for every $\gamma < \alpha$, and hence $S_3 \vdash \text{wo}(\varphi\gamma 0)$ for every $\gamma < \alpha$; thus also $S_3 \vdash \text{TI}(<\varphi\alpha 0)$. As Γ_0 denotes the least ordinal closed under the function $\alpha \mapsto \varphi\alpha 0$ we obtain the result. \square

It is possible to develop an infinitary calculus with the aim of deriving an upper bound on the strength of S_3 . This will be achieved in the succeeding sections but at the same time we will also consider abstracting the step that lead us from S_2 to S_3 and attempt to iterate it. The proof-theoretic strength of S_3 is determined in theorem 18.

4 Theories with multiple self-applicable truth predicates

Truth is often used as a means of reflection; a tool by which one may obtain principles, schemata etc. that were implicit, but not necessary explicit, in the acceptance of some axiomatic system. Feferman, for example, views the theory of truth Ref (see ??) as an operation which, when applied to a theory S , answers the question “which statements in the base language ... ought to be accepted if

one has accepted the basic axioms and rules of [S]?” [?, p. 2]. The Friedman-Sheard theories A to I can also be viewed as operations which have been applied to PRA: one adds to PRA a (new) predicate T, formalising the acceptance of PRA; on top of this one adds some subset of the Optional Axioms, for example adding \forall -Inf formalises the acceptance of ω -logic, while the axiom T-Del formalises closure under the rule T-Elim.

Viewing theories of truth as operations provides a natural way to describe the general processes behind their construction and allows one to possibly iterate the operation. In this section we will look at this specifically from the perspective of the Friedman-Sheard theory F.

One way of arguing for the naturalness of S_3 is to view it as formalising the acceptance of S_2 . Within S_3 one has T-Rep and T-Del, formalising the rules of inference T-Intro and T-Elim of S_2 , and thus

$$S_3 \vdash \forall x (\text{Bew}_{S_2}(x) \wedge \text{Sent}_{\mathcal{L}_1}(x) \implies T(x)).$$

However, S_2 contains the rule T-Intro, so it seems reasonable that the theory attempting to formalise its acceptance should also be closed under T-Intro. But the presence of T-Rep, T-Del and T-Elim in S_3 means this is not possible, so perhaps S_3 is not such a natural theory after all.

Since closure of F under \neg T-Elim is vacuous, F and S_2 are identical as theories. Therefore, S_3 can be seen as formalising the acceptance of F, although one might expect in this case, to also add

$$T(\ulcorner \neg T(\dot{x}) \urcorner) \implies T(\neg x)$$

as an axiom.³ Still, the resulting theory cannot be closed under T-Intro, as one would like.

If one were to stratify the language, in much the same way as one would to form a Tarskian hierarchy of truth predicates, the problem can be circumvented. Recall $F = \text{Base}_T + \text{T-Intro} + \text{T-Elim} + \neg\text{T-Elim} + \text{T-Del} + \forall\text{-Inf}$. Let F_0 denote F formulated with the predicate T_0 in place of T, and suppose T_1 is a (new) unary predicate symbol. The theory formalising acceptance of F_0 , which we shall denote by F_1 , would then contain the following axioms

$$\text{val}(x) \wedge \text{Sent}_{\mathcal{L}_{T_0}}(x) \implies T_1(x), \quad (16) \quad \{\text{exteqn:FFax0}\}$$

$$\text{Ax}_{F_0}(x) \wedge \text{Sent}_{\mathcal{L}_{T_0}}(x) \implies T_1(x), \quad (17) \quad \{\text{exteqn:FFax1}\}$$

$$T_1(x) \wedge T_1(x \rightarrow y) \implies T_1(y), \quad (18) \quad \{\text{exteqn:FFax2}\}$$

$$(\forall x T_1(\ulcorner A(\dot{x}) \urcorner)) \implies T_1(\ulcorner \forall x A(x) \urcorner), \quad (19) \quad \{\text{exteqn:FFaxw}\}$$

$$T_1(x) \implies T_1(\ulcorner T_0(\dot{x}) \urcorner), \quad (20) \quad \{\text{exteqn:FFax3}\}$$

$$T_1(\ulcorner T_0(\dot{x}) \urcorner) \implies T_1(x), \quad (21) \quad \{\text{exteqn:FFax4}\}$$

$$T_1(\ulcorner \neg T_0(\dot{x}) \urcorner) \implies T_1(\neg x). \quad (22) \quad \{\text{exteqn:FFax5}\}$$

Equations (16) and (17) state the acceptance of all axioms of F_0 (logical and non-logical), whereas eq. (18) formalises *modus ponens* in F_0 . These three also combine to imply the axioms of Base_{T_1} . Equations (20) to (22) express the acceptance of the rules T_0 -Intro, T_0 -Elim and $\neg T_0$ -Elim, respectively, in F_0 , while eq. (19) closes the predicate under ω -logic.

The predicate T_1 is viewed as an extension of the predicate T_0 and as such we would expect it to satisfy the relevant axioms of F, that is, we also have

$$T_1(\ulcorner T_1(\dot{x}) \urcorner) \implies T_1(x),$$

³As the \mathcal{L}_1 -structure \mathfrak{M} used in theorem 5 also satisfies the axiom $T(\ulcorner \neg T(\dot{x}) \urcorner) \implies T(\neg x)$, the extension of S_3 obtained by adding this axiom is also consistent.

and closure under T_1 -Elim, T_1 -Intro and $\neg T_1$ -Elim. Combining also the axioms of F_0 it is then easy to deduce

$$\forall x(\text{Bew}_{F_0}(x) \wedge \text{Sent}_{\mathcal{L}_{T_0}}(x) \implies T_1(x)). \quad (23)$$

We are happy with the thought of T_0 being a self-applicable truth predicate, and so far there is nothing to stop T_1 also being self-applicable. Moreover, T_0 may meaningfully occur in the scope of the predicate T_1 . Thus we have described the first step in a hierarchy of self-applicable truth predicates. But, should the predicate T_1 be allowed to occur in the scope of T_0 ? After all, the motivation behind working with theories that contain their own truth predicate is in their ability to reason about themselves. Since T_1 -Intro may apply to arbitrary sentences in \mathcal{L}_{T_0, T_1} , the question of whether or not T_0 can meaningfully apply to sentences containing T_1 is essentially decided by how we restrict the quantifiers in eqs. (18) to (22) (in particular eq. (20) pertaining to T_0 -Intro) for inclusion in F_1 : if we restrict them to range over only codes of \mathcal{L}_{T_0} -sentences we will have no non-trivial occurrences of this inter-applicability.

As is consistent with our earlier chapters, we view our theories as being presented in a Hilbert style deduction system, with certain axioms and rules of inference which are treated in their broadest sense. Namely, we consider a rule of inference of a theory S to be applicable to any extension of the language, logic or axioms of S . Thus, if one imagines the theory F_0 being first formulated in the language $\mathcal{L} \cup \{T_0, T_1\}$, and only then completing the reflection step to F_1 by adding the axioms and rules pertaining to T_1 , it seems natural to suppose the predicates T_0 and T_1 are inter-applicable. Since T_0 -Intro was applicable in F_0 to formulae containing the predicate T_1 , so should T_0 in F_1 . Thus, we expect F_1 to have the axiom T_0 -Imp,

$$\forall x \forall y [T_0(x) \wedge T_0(x \rightarrow y) \implies T_0(y)],$$

as opposed to its relativised form

$$\forall x \forall y [\text{Sent}_{\mathcal{L}_{T_0}}(x) \wedge \text{Sent}_{\mathcal{L}_{T_0}}(y) \implies (T_0(x) \wedge T_0(x \rightarrow y) \implies T_0(y))]; \quad (24)$$

and, more importantly, that F_1 contains the unrelativised axioms

$$\begin{aligned} \forall x [T_1(x) &\implies T_1(\ulcorner T_0(\dot{x}) \urcorner)], \\ \forall x [T_1(\ulcorner T_0(\dot{x}) \urcorner) &\implies T_1(x)], \\ \forall x [T_1(\ulcorner \neg T_0(\dot{x}) \urcorner) &\implies T_1(\neg x)]. \end{aligned} \quad (25)$$

This provides, for example,

$$\forall x [T_1(\ulcorner T_1(\dot{x}) \urcorner) \leftrightarrow T_1(\ulcorner T_0(\ulcorner T_1(\dot{x}) \urcorner) \urcorner)],$$

which, by T_1 -Intro and eq. (23), yields $T_0(\ulcorner T_1(\ulcorner A \urcorner) \urcorner)$ for every theorem A of F_0 . Thus our informal interpretation leads us to the curious situation where we have two self-applicable predicates T_0, T_1 which may also be applied to each other.

If, on the other hand, we had considered relativising the quantifiers as in eq. (24), T_1 may apply to the language \mathcal{L}_{T_0, T_1} whereas T_0 may only meaningfully apply to \mathcal{L}_{T_0} and one would obtain the base of a strict hierarchy of self-applicable truth predicates. In this case we can no longer argue that T_0 represents a truth predicate for the whole language but only of the sub-language \mathcal{L}_{T_0} . The reason for choosing a self-applicable notion of truth in the first place was that it may be treated as a truth predicate for the entire language, including any subsequent extension. Thus, T_0 *should* be applicable

to sentences involving the predicate T_1 and we find ourselves returning to the world of two inter-applicable truth predicates.

So far we have argued that in F_1 the interpretation of T_0 should be closed under T_0 -Elim and $\neg T_0$ -Elim while the interpretation of T_1 should be closed under T_0 -Intro, T_0 -Elim, $\neg T_0$ -Elim, T_1 -Elim and $\neg T_1$ -Elim. However, we desire T_1 to be viewed as an extension of T_0 so as to allow for closure of F_1 under a form of *truth introduction* and this fact should be recognised by the theory. That is, from the point of view of the theory F_1 , there should be no essential difference between the predicates T_0 and T_1 . Thus we propose to also include the principle

$$T_0(\ulcorner T_1(x) \urcorner) \implies T_0(x) \tag{26} \quad \{\text{exteqn:T1Del}\}$$

as an axiom of F_1 (in fact as an axiom of F_0). Equation (26) expresses that the interpretation of T_0 is closed under the rule T_1 -Elim. This is vacuously valid in the theory F_0 . It also serves to confirm the inter-applicability of the two predicates by allowing meaningful inferences regarding T_1 under a T_0 predicate. If we accept eq. (26) we should also accept $T_0(\ulcorner \neg T_1(x) \urcorner) \implies T_0(\neg x)$, but as we shall see, this axiom will be trivially satisfied by our model.

We have only described the step F_0 to F_1 , but one can imagine repeating this, first adding an additional truth predicate T_2 to F_1 and then reflecting upon it to form the theory F_2 .⁴ This process may be continued into the transfinite to form a hierarchy of theories, supporting a hierarchy of inter-applicable truth predicates.⁵

The description of F_1 and F_2 presented above is purely motivational and requires making explicit, which the next definition achieves. As we pass beyond finite iterations and consider the construction of theories F_ω , F_{ε_0} , *etc.*, one requires the ability to internally quantify over the indices of truth predicates in the language. We therefore need to fix some computable ordinal κ from the outset and only consider iterating the construction up to ordinals $\alpha < \kappa$. In fact to maintain as much similarity with our previous work as possible (for example to ensure a primitive recursive Gödel numbering) we will insist κ is primitive recursively definable. One could consider taking $\kappa = \Gamma_0$ and utilising the encoding chosen in ??, but as we shall see $\|F_1\| \geq \Gamma_0$ (see theorem 9 below) and so we will require the construction of a larger class of ordinals to perform a sufficient proof-theoretic analysis. Suitable choices for κ will ultimately depend on our analysis and have no substantial role in defining the theories. Since the precise definition of κ is not essential for the definition, we shall assume for the time being that κ represents some fixed primitive recursive ordinal.

Definition 11. For $\beta < \kappa$, let \mathcal{L}_β denote the language of PRA augmented by predicates T_ξ for each $\xi \leq \beta$. Let F_β be the theory formulated in the language \mathcal{L}_β , extending PA with the schema of induction

⁴A more precise way to describe the construction of F_2 may be given as follows: first imagine formulating F_0 with three predicates, T_0 , T_1 , and T_2 . One then formulates F_1 in this language by reflecting upon F_0 , and only then is F_2 formulated by reflecting upon F_1 . In this sense we view F_0 as not being formulated in a fixed language, but rather in a language that may be expanded as and when we see fit.

⁵The inter-applicability of the predicates calls into question whether what we obtain is truly a “hierarchy of truth predicates” as we describe. Our model construction, in the next section, will show that one can view the truth predicates as based on a hierarchy, although not, perhaps, in a manner one might first imagine.

for \mathcal{L}_β , and for each $\eta \leq \beta$

$$\begin{aligned}
& \forall x (T_\eta(x_1) \wedge T_\eta(x_1 \rightarrow x_2) \implies T_\eta(x_2)), & (T_\eta\text{-Imp}) & \{\text{exteqn:T-Imp}\} \\
& \forall^\top A(x)^\top [(\forall n T_\eta(\ulcorner A(\dot{n}) \urcorner)) \implies T_\eta(\ulcorner \forall x A(x) \urcorner)], & (T_\eta\text{-}\forall\text{-Inf}) & \{\text{exteqn:T-UInf}\} \\
& \forall \xi \leq \bar{\beta} \forall x (T_\eta(\ulcorner T_\xi(\dot{x}) \urcorner) \implies T_\eta(x)), & (T_\eta\text{-Del}) & \{\text{exteqn:T-dag1}\} \\
& \forall \xi < \bar{\eta} \forall x (T_\eta(x) \implies T_\eta(\ulcorner T_\xi(\dot{x}) \urcorner)), & (T_\eta^-\text{-Rep}) & \{\text{exteqn:T-dag2}\} \\
& \forall \xi \leq \bar{\beta} \forall x (T_\eta(\ulcorner \neg T_\xi(\dot{x}) \urcorner) \implies T_\eta(\ulcorner \neg x \urcorner)), & (\neg T_\eta\text{-Del}) & \{\text{exteqn:T-dag3}\}
\end{aligned}$$

as well as the rules of inference

$$\begin{aligned}
& \text{from } A \text{ infer } T_\beta(\ulcorner A \urcorner), & (T_\beta\text{-Intro}) & \{\text{exteqn:T-Intro}\} \\
& \text{from } \neg T_\beta(\ulcorner A \urcorner) \text{ infer } \neg A, & (\neg T_\beta\text{-Elim}) & \{\text{exteqn:notT-Elim}\} \\
& \text{from } T_\beta(\ulcorner A \urcorner) \text{ infer } A. & (T_\beta\text{-Elim}) & \{\text{exteqn:T-Elim}\}
\end{aligned}$$

Define $F_{<\beta} = \bigcup_{\eta < \beta} F_\beta$. Let us denote by F_β^n the collection of theorems of F_β provable with at most n (serial) applications of T_β -Intro and no restriction on the number of applications of T_β -Elim. Thus F_β^0 denotes the theory F_β without T_β -Intro, and F_β^n is a sub-theory of F_β^{n+1} for every n .

The theory F_0 is identical to F and the definition of F_β fits the informal description we gave of ‘ F viewed as an operation applied to $F_{<\beta}$ ’. Also, F_β^0 is a conservative extension of PA as, with no applications of T_β -Intro, all predicates T_η in F_β^0 may be interpreted trivially to show neither $F_\beta^0 \vdash T_\beta(s)$ nor $F_\beta^0 \vdash \neg T_\beta(s)$ may hold.

Although F_1 may be borne from a notion of truth and acceptance, it is by no means necessarily consistent. Indeed, with its multiple truth predicates and inter-applicability, the reader would be duly justified to view this construction with some scepticism. The process of reflection that led us to F_1 , however, is almost identical to that which describes the theory S_3 . As a result it may not be surprising to know that F_1 is consistent and in fact F_1^1 proves the same arithmetical statements as the theory S_3 . Unfortunately we do not at this time have the suitable machinery to prove their equivalence (this will have to wait until section 5), but the following remark should motivate the connection.

extrem:S3

Remark 1. It is natural to first consider embedding F_1^1 into S_3 . As the predicate T_1 can be interpreted vacuously in F_1^0 (as can T_0), one may expect the interpretation given by

$$T_1(s)^* = T_0(s)^* = T(g(s))$$

to suffice, where g is a primitive recursive function chosen, by the primitive recursion theorem, such that

$$\begin{aligned}
& g(x) = x, \text{ if } x \text{ is the code of an arithmetical literal,} \\
& g(\ulcorner T_1(s) \urcorner) = \ulcorner s \neq s \urcorner, \\
& g(\ulcorner T_0(s) \urcorner) = \ulcorner T(g(s)) \urcorner, \\
& g(\ulcorner A \circ B \urcorner) = g(\ulcorner A \urcorner) \circ g(\ulcorner B \urcorner), \text{ for } \circ \in \{\wedge, \vee\} \\
& g(\ulcorner Qx A(x) \urcorner) = \forall x g(\ulcorner A(x) \urcorner).
\end{aligned}$$

This interpretation will translate the axioms T_1 -Del, T_1 - \forall -Inf and T_1^- -Rep of F_1 to the axioms T -Del, \forall -Inf and T -Rep respectively of S_3 . Since S_3 does not have the axiom $\neg T$ -Del (nor can we expect to

derive it in S_3) one cannot derive the translation of $\neg T_1\text{-Del}$ from F_1^1 . Fortunately, the model construction of section 3.2 can easily incorporate the additional axiom $\neg T\text{-Del}$ and rule $\neg T\text{-Elim}$ with minimal adjustments: one may consistently add the rule pertaining to $\neg T\text{-Del}$ to each of the theories \mathcal{F}_α and expand lemma 9 to also prove $\mathcal{F}_\alpha \vdash \neg T(\ulcorner A \urcorner)$ implies $\mathcal{F}_\alpha \vdash \neg A$, thus showing $S_3 + \neg T\text{-Del} + \neg T\text{-Elim}$ is a consistent theory.

The problem with this interpretation manifests when dealing with applications of $T_1\text{-Elim}$ in F_1^1 . Suppose $F_1^1 \vdash T_1(\ulcorner A \urcorner)$. If A does not contain the predicate T_1 , $\ulcorner A^* \urcorner = g(\ulcorner A \urcorner)$ and an application of $T\text{-Elim}$ in S_3 suffices. If, however, A contains T_1 , the interpretation of $T_1(\ulcorner A \urcorner)$ and A under $*$ are very different; indeed, there will be sentences B for which $S_3 \vdash T(g(\ulcorner B \urcorner))$ but $S_3 \not\vdash B^*$ (for example, take B to be $\neg T_1(\ulcorner C \urcorner)$ where C is any statement provable in S_3 . Then $S_3 \vdash B^*$ implies $S_3 + \neg T\text{-Elim} \vdash \neg C$, which contradicts the consistency of $S_3 + \neg T\text{-Elim}$. On the other hand, $g(\ulcorner B \urcorner) = \ulcorner \neg(s \neq s) \urcorner$ for some s , so $S_3 \vdash T(g(\ulcorner B \urcorner))$ holds). If A were an axiom of F_1^0 though, $S_3 \vdash A^*$; moreover, if one views $T_1\text{-Imp}$, $T_1\text{-Del}$ and $T_1\text{-Rep}$ as a rule of inference, as in \mathcal{F} , one could deduce closure under $T_1\text{-Elim}$ by induction on the height of a derivation. Viewing the axioms of F_1^1 in this manner is reminiscent of the use of the infinitary system \mathcal{F}_∞ in the analysis of F ; thus it seems natural to delay a further investigation of this connection until we have first analysed F_1 in detail.

Although the addition of full $T_1\text{-Intro}$ to F_1^1 (forming F_1) creates a theory markedly stronger than S_3 , it is not straightforward to embed S_3 into F_1 . Such an embedding would require stratifying \mathcal{L}_T to involve the two predicates T_0 and T_1 of \mathcal{L}_1 . The most obvious way to proceed would appear to involve replacing the outermost predicate by T_1 and all others by T_0 , thus mapping the axioms $T(\ulcorner T\text{-Imp} \urcorner)$, $T(\ulcorner \forall\text{-Inf} \urcorner)$ and $T(\ulcorner T\text{-Del} \urcorner)$ of S_3 to theorems of F_1^1 . This could be achieved by picking a primitive recursive function f which recursively substitutes $T_0(f(s))$ for $T(s)$, and considering the interpretation $*$ of \mathcal{L}_T into \mathcal{L}_1 given by

$$(T(s))^* = T_1(f(s)).$$

This interpretation maps the axioms $\forall\text{-Inf}$, $T\text{-Imp}$, $T\text{-Del}$ and $T\text{-Rep}$ by instances of axioms $T_1\text{-}\forall\text{-Inf}$, $T_1\text{-Imp}$, $T_1\text{-Del}$ and $T_1\text{-Rep}$ respectively, but since $f(\ulcorner A \urcorner)$ need not equal $\ulcorner A^* \urcorner$, applications of $T\text{-Elim}$ in S_3 do not translate into inferences we can readily recognise as holding in F_1 .

We now move to the task of proving the consistency of F_β for arbitrary β . This will be obtained by generalising the consistency argument for S_3 to theories with multiple predicates and is presented in section 4.1 below. Following this we will perform a detailed ordinal analysis of these theories, first determining lower bounds on their proof-theoretic strength, and second upper bounds. Lower bounds on the strength of the theories F_β will be obtained in section 4.3 by extending the well-ordering results previously established for S_2 and S_3 . Upper bounds are determined in section 4.4 where we attempt to directly formalise the model constructions of section 4.1 in a manner similar to our analysis of F . However, much care will be required when dealing with a hierarchy of truth predicates and the transfinite iterations of $T\text{-Intro}$ caused by the interaction between axioms $T_\eta\text{-}\forall\text{-Inf}$ and $T_\eta\text{-Rep}$.

The axiom $T_\eta\text{-Del}$ implies the interpretation of the predicate T_η is closed under $T_\xi\text{-Elim}$ for every $\xi \leq \eta$. Likewise $T_\eta\text{-Rep}$ implies T_η is closed under $T_\xi\text{-Intro}$ for each $\xi < \eta$. From this observation we may deduce the following proposition, in which Bew_η abbreviates the formalised provability predicate for F_η .

Proposition 12. *For every $\eta < \beta$, F_β is closed under $T_\eta\text{-Intro}$ and $T_\eta\text{-Elim}$. Moreover,*

$$F_\beta^1 \vdash \forall x (\text{Bew}_\eta(x) \wedge \text{Sent}_{\mathcal{L}_\eta}(x) \implies T_\beta(x)).$$

Proof. All axioms of F_η are axioms of F_β^0 , and by T_β -Intro we have

$$F_\beta^1 \vdash \forall x (Ax_\eta(x) \wedge \text{Sent}_{\mathcal{L}_\eta}(x) \implies T_\beta(x)),$$

where $Ax_\eta(x)$ expresses that x is a sentential axiom of F_η . To complete the proof, we observe that F_β^1 has the axioms T_β -Imp, T_β -Del, T_β^- -Rep and $\neg T_\beta$ -Del formalising applications of *modus ponens*, T_η -Elim, T_η -Intro and $\neg T_\eta$ -Elim respectively in F_η . \square

4.1 Consistency

At first glance the theory F_β could easily look suspect, after all it contains T_η -Del, T_η -Elim, T_η -Intro and an axiom that appears extremely close to T_η -Rep, for each $\eta \leq \beta$. As the motivation behind the theories F_β comes from abstracting the transition from PA to F one might expect that if F_β is consistent, models of F_β may be constructed by extending models of F. This is indeed the case; moreover, the extension we defined for establishing the consistency of S_3 makes a suitable base from which to start the construction. We will only sketch the consistency argument as it will be subsumed by our work in section 4.4 where we determine an upper bound on the proof-theoretic strength of F_β .

Let Ω denote the first uncountable ordinal. For the remainder of this section Greek letters, ρ, σ, τ will be used to range over elements of \mathbb{O} , the class of all ordinals; letters α, β , etc., will range over countable ordinals; we reserve the letters η, ξ for indices of the truth predicates and so represent ordinals below κ . Suppose

$$\rho = \Omega^{\alpha_0} \cdot \beta_0 + \dots + \Omega^{\alpha_n} \cdot \beta_n,$$

with $\alpha_0 > \dots > \alpha_n$ and $\beta_i < \Omega$ for each $i \leq n$. We denote by $\rho|_\gamma$ the ordinal $\Omega^{\alpha_0} \cdot \beta_0 + \dots + \Omega^{\alpha_k} \cdot \beta_k$ where $k < n$ is the least such that $\alpha_k > \gamma \geq \alpha_{k+1}$, or $k = n$ if $\alpha_n > \gamma$. An ordinal ρ is called an Ω -limit if $\rho = \rho_0 + \Omega^\eta \cdot \alpha$ and either α or η is a limit ordinal.

Define a system \mathcal{F}_ρ^* for $\rho \in \mathbb{O}$, formulated in a Tait-style sequent calculus in the language \mathcal{L}_κ . The system has three axioms.

- Ax.1. $\mathcal{F}_\rho^* \vdash \Gamma, A$ whenever A is a true atomic literal,
- Ax.2 $_\eta$. $\mathcal{F}_\rho^* \vdash \Gamma, T_\eta(s), \neg T_\eta(s)$ for every closed term s and $\eta < \kappa$,
- Ax.3 $_\eta$. $\mathcal{F}_\rho^* \vdash \Gamma, \neg T_\eta(s)$ if $s^\mathbb{N}$ is not the code of an \mathcal{L}_κ -sentence.

The rules of this system are the usual arithmetical rules (\wedge), (\vee_0), (\vee_1) and (\exists), plus the following six rules for every $\eta < \kappa$.

- $$\frac{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta(s_0) \quad \mathcal{F}_\rho^* \vdash \Gamma, T_\eta(s_0 \multimap s_1)}{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta(s_1)} \text{Imp}_\eta$$
- $$\frac{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta(\ulcorner T_\xi s \urcorner) \quad \xi < \beta}{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta s} \text{Del}_\eta$$
- $$\frac{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta s \quad \xi < \eta}{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta(\ulcorner T_\xi s \urcorner)} \text{Rep}_\eta$$
- $$\frac{\mathcal{F}_\sigma^* \vdash A \quad \sigma < \rho|_\eta}{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta \ulcorner A \urcorner} \text{Intro}_\eta$$

$$\begin{aligned}
& \bullet \frac{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta \ulcorner A(\underline{n}) \urcorner \text{ for all } n \quad \rho|_\eta \text{ not an } \Omega\text{-limit}}{\mathcal{F}_\rho^* \vdash \Gamma, T_\eta \ulcorner \forall x A(x) \urcorner} U_\eta \\
& \bullet \frac{\mathcal{F}_\rho^* \vdash \Gamma, A(\underline{n}) \quad \text{for every } n \in \mathbb{N}}{\mathcal{F}_\rho^* \vdash \Gamma, \forall x A(x)} \omega
\end{aligned}$$

Moreover, for each ρ define the set of (codes of) \mathcal{L}_κ -sentences

$$\mathcal{F}_{<\rho}^* = \{ \ulcorner A \urcorner : \mathcal{F}_\sigma^* \vdash A \text{ for some } \sigma < \rho \}.$$

Before we proceed with analysing the role of the truth predicates in \mathcal{F}_ρ^* , it is important to note that a rule of *modus ponens*, or cut is lacking from our definition. However, it is not hard to show the cut rule is, in fact, admissible in \mathcal{F}_ρ^* : a derivation of the form $\Gamma, \neg T_\eta(s)$ with $\neg T_\eta(s)$ active *must* be an instance of an axiom, hence if one has derivations of $\Gamma, \neg T_\eta(s)$ and $\Gamma, T_\eta(s)$ one can easily obtain instead a derivation of Γ . This argument is essentially identical to that employed in the Cut-Elimination theorem for \mathcal{F}_∞ , ??.

Lemma 13 (Admissibility of cut). *Suppose $\eta < \kappa$, $\mathcal{F}_\rho^* \vdash \Gamma, T_\eta(s)$, and $\mathcal{F}_\rho^* \vdash \Delta, \neg T_\eta(s)$ for some ordinal ρ . Then $\mathcal{F}_\rho^* \vdash \Gamma, \Delta$.*

If $\alpha < \Omega$, $\alpha|_0 = 0$ and so no applications the rule T_η -Intro is not available in \mathcal{F}_α^* for any η . Thus, $\mathcal{F}_{<\Omega}^*$ is trivially closed under T_η -Elim for every η . Moreover, for every $\alpha < \Omega$, only applications of T_0 -Intro have been permitted in $\mathcal{F}_{\Omega \cdot \alpha}^*$. Thus, we can establish by induction on $n < \omega$ that the predicate T_0 in $\mathcal{F}_{\Omega \cdot n}^*$ may be interpreted as the theory $\mathcal{F}_{<\Omega \cdot n}^*$ and that $\mathcal{F}_{\Omega \cdot n}^*$ is closed under T_0 -Elim (this is identical to the argument showing $\mathcal{F}_\infty|_{n+1}^\alpha \vdash \ulcorner \Gamma A \urcorner$ for some α implies $\mathcal{F}_\infty|_n^\beta \vdash A$ for some β). Since all other truth predicates in $\mathcal{F}_{\Omega \cdot n}^*$ may be interpreted vacuously, we conclude $\mathcal{F}_{\Omega \cdot n}^*$ is closed under T_η -Elim for every $\eta < \kappa$. One may also interpret every predicate simply by \mathbb{N} , so, in fact, $\mathcal{F}_{\Omega \cdot n}^*$ is also closed under $\neg T_\eta$ -Elim for every η . Hence F_0^n may be interpreted in $\mathcal{F}_{\Omega \cdot n}^*$.

This suggests that for $n < \omega$, $\mathcal{F}_{\Omega \cdot n}^*$, like \mathcal{F}_n , reconstructs the theories Th_n used by Friedman and Sheard to prove the consistency of F. At the first limit ordinal, we obtain $\mathcal{F}_{<\Omega \cdot \omega}^*$, a set of \mathcal{L}_κ -sentences closed under T_0 -Intro, T_0 -Elim and $\neg T_0$ -Elim, and containing T_0 -Del, T_0^- -Rep (which holds vacuously), $\neg T_0$ -Del and all other axioms of F_0 .

To proceed with the analysis of F_1 , we first consider F_1^0 which, without the rule T_1 -Intro, is vacuously closed under T_1 -Elim and $\neg T_1$ -Elim. In F_1^1 , the situation differs from previous case; we need to interpret the predicate T_1 as a theory closed under ω -logic (due to T_1 - \forall -Inf), T_0 -Intro (due to T_1^- -Rep), T_0 -Elim and T_1 -Elim (due to T_1 -Del), as well as $\neg T_0$ -Elim and $\neg T_1$ -Elim (due to $\neg T_1$ -Del). Moreover, we need to find an interpretation of T_0 closed under ω -logic, T_0 -Elim and $\neg T_0$ -Elim, and now also T_1 -Elim and $\neg T_1$ -Elim (as implied by the axioms T_0 -Del and $\neg T_0$ -Del). The properties we established for the set $\mathcal{F}_{<\Omega}$ in section 3.2 motivate us to consider $\mathcal{F}_{<\Omega^2}^*$, a set closed under T_0 -Intro, T_0 -Elim, $\neg T_0$ -Elim and, by a similar argument as before, ω -logic (cf. the proof of theorem 5). For every $\alpha < \Omega$, the predicate T_1 may be interpreted vacuously in $\mathcal{F}_{<\Omega \cdot \alpha}^*$, so $\mathcal{F}_{<\Omega^2}^*$ is also closed under T_1 -Elim. Thus $\mathcal{F}_{<\Omega^2}^*$ provides a consistent interpretation of both predicates T_0 and T_1 in F_1^1 .

The next step is to consider T_1 in F_1^2 . Two applications of T_1 -Intro are permitted and one can derive sentences of the form $T_1(\ulcorner T_1(\ulcorner A \urcorner) \urcorner)$ whenever $F_1^0 \vdash A$, suggesting a shift to $\mathcal{F}_{\Omega^2}^*$, where one can derive $T_1(\ulcorner A \urcorner)$ whenever $A \in \mathcal{F}_{<\Omega^2}^*$, might yield a suitable interpretation for T_1 . However, $\mathcal{F}_{\Omega^2}^*$ is not closed under T_0 -Intro (only the systems $\mathcal{F}_{<\rho+\Omega \cdot \sigma}^*$ for limit ordinals σ are), leading us instead to consider $\mathcal{F}_{<\Omega^2+\Omega \cdot \omega}^*$ which is closed under T_0 -Intro, but not ω -logic; there will be sentences $A \in \mathcal{F}_{<\Omega^2}^*$ for which $\mathcal{F}_{\Omega^2+\Omega \cdot n}^* \vdash T_0(f(n, \ulcorner A \urcorner))$ for each $n < \omega$, where $f(0, n) = \ulcorner \bar{n} \urcorner$ and $f(m+1, n) = \ulcorner T_0(f(m, n)) \urcorner$, but the

sentence $\forall x T_0(f(x, \ulcorner A \urcorner))$ is not contained in $\mathcal{F}_{<\Omega^2+\Omega\cdot\omega}^*$. Indeed to obtain both closure under ω -logic and T_0 -Intro we must move to the theory $\mathcal{F}_{<\Omega^2\cdot 2}^*$. We also require the interpretation to be closed under T_0 -Elim. To manage this we repeat the same argument as before, but starting from $\mathcal{F}_{\Omega^2}^*$ in place of \mathcal{F}_{Ω}^* . We know $\mathcal{F}_{\Omega^2}^*$ is closed under T_0 -Elim since the predicate T_0 can be consistently interpreted as the set $\mathcal{F}_{<\Omega^2}^*$. This leads us to successively deduce the theories $\mathcal{F}_{\Omega^2+\Omega\cdot n}^*$ are closed under T_0 -Elim for each $n < \omega$. Note, we can still interpret T_1 in $\mathcal{F}_{\Omega^2+\Omega\cdot n}^*$ by the set $\mathcal{F}_{<\Omega^2}^*$ as there has been no further applications of T_1 -Intro. In $\mathcal{F}_{\Omega^2+\Omega\cdot\omega}^*$ we aim to interpret T_0 by $\mathcal{F}_{<\Omega^2+\Omega\cdot\omega}^*$, which unlike $\mathcal{F}_{\Omega^2+\Omega\cdot n}^*$ is not closed under ω -logic; however, $\mathcal{F}_{\Omega^2+\Omega\cdot\omega}^*$ is not closed under $(T_1\text{-}\forall\text{-Inf})$ so this does not pose a problem. Thus we may continue through the construction of $\mathcal{F}_{<\Omega^2\cdot 2}^*$ determining each theory $\mathcal{F}_{\Omega^2+\Omega\cdot\alpha}^*$ for $\alpha < \Omega$ is closed under T_0 -Elim.

The argument above highlights that the predicates T_1 and T_0 in F_1^n may be interpreted as the set $\mathcal{F}_{<\Omega^2\cdot n}^*$, and hence F_1 naturally embeds into $\mathcal{F}_{<\Omega^2\cdot\omega}^*$. If we wanted to proceed beyond this and construct models for F_2 , we could imagine constructing a sequence of systems

$$\mathcal{F}_{\Omega^2\cdot\omega}^*, \mathcal{F}_{\Omega^2\cdot\omega+\Omega}^*, \dots, \mathcal{F}_{\Omega^2\cdot\omega+\Omega\cdot\alpha}^*, \dots, \mathcal{F}_{\Omega^2\cdot(\omega+1)}^*, \dots, \mathcal{F}_{\Omega^2\cdot\alpha}^*, \dots$$

to obtain $\mathcal{F}_{<\Omega^3}^*$, an interpretation of the predicate T_2 in F_2^1 . The ability to recognise each theory $\mathcal{F}_{\Omega^2\cdot\alpha+\Omega\cdot\gamma}^*$ as closed under T_1 -Elim and T_0 -Elim, however, is essential for the interpretation of T_2 -Del in F_2^1 . As already argued, the set $\mathcal{F}_{<\Omega^2\cdot\alpha+\Omega\cdot\gamma}^*$ provides an interpretation of T_0 in $\mathcal{F}_{\Omega^2\cdot\alpha+\Omega\cdot\gamma}^*$; but unless γ is a limit ordinal, this need not be closed under T_0 -Intro, so cannot interpret the predicate T_1 . The answer is to interpret T_1 in $\mathcal{F}_{\Omega^2\cdot\alpha+\Omega\cdot\gamma}^*$ as the set $\mathcal{F}_{<\Omega^2\cdot\alpha}^*$ for every $\gamma < \Omega$. Only when we pass to $\mathcal{F}_{\Omega^2\cdot(\alpha+1)}^*$ do we alter the interpretation of T_1 (in this case it is changed to the set $\mathcal{F}_{<\Omega^2\cdot(\alpha+1)}^*$). It is for exactly this reason that the rule $T_\eta\text{-}\forall\text{-Inf}$ was restricted so as to apply to \mathcal{F}_ρ^* only if $\rho|_\eta$ is not an Ω -limit; the set $\mathcal{F}_{<\Omega^2\cdot\alpha}^*$ will not be closed under ω -logic if α is a limit ordinal.

Once one has constructed $\mathcal{F}_{<\Omega^3}^*$ and verified that it is closed under T_n -Elim for $n = 0, 1, 2$, one would then embark on the construction of a further sequence of systems

$$\mathcal{F}_{\Omega^3}^*, \dots, \mathcal{F}_{\Omega^3+\Omega\cdot\alpha}^*, \dots, \mathcal{F}_{\Omega^3+\Omega^2}^*, \dots, \mathcal{F}_{\Omega^3+\Omega^2\cdot 2}^*, \dots, \mathcal{F}_{\Omega^3+\Omega^2\cdot\alpha}^*, \dots, \mathcal{F}_{\Omega^3\cdot 2}^*, \dots$$

and subsequently $\mathcal{F}_{<\Omega^3\cdot\omega}^*$, a theory into which F_2 embeds. In general, we expect F_β to embed into $\mathcal{F}_{\Omega^{\beta+1}\cdot\omega}^*$ for each β .

The next lemma deals with the task of determining the theory \mathcal{F}_ρ^* is closed under T_η -Elim for every $\eta < \kappa$. Before that, however, we require the following result regarding the behaviour of Ω -limits.

Proposition 14. *If ρ is not an Ω -limit and $\sigma_n < \rho$ for every $n < \omega$,*

$$\sup_{n < \omega} \sigma_n < \rho.$$

Proof. Suppose ρ is not an Ω -limit and $\sigma_n < \rho$ for every $n < \omega$. Then $\rho > 0$ and there are ordinals ρ_0, α_0 such that $\rho = \rho_0 + \Omega^{\alpha_0} \cdot \Omega$. This means we can associate an ordinal $\delta_n < \Omega$ to each $n < \omega$ so that $\sigma_n < \rho_0 + \Omega^{\alpha_0} \cdot \delta_n$. The set $\{\delta_n : n < \omega\}$ is a countable set of countable ordinals, and hence is bounded in Ω , whence

$$\begin{aligned} \sup_n \sigma_n &\leq \sup_n \{\rho_0 + \Omega^{\alpha_0} \cdot \delta_n\} \\ &\leq \rho_0 + \Omega^{\alpha_0} \cdot (\sup_n \delta_n) \\ &< \rho_0 + \Omega^{\alpha_0} \cdot \Omega \\ &= \rho. \end{aligned}$$

□

A sequent Γ is called *T-positive* if all occurrences of a predicate T_η in Γ for any $\eta < \kappa$ are positive. Define, for each ordinal ρ , an \mathcal{L}_κ -structure \mathfrak{M}_ρ according to the following criterion.

$$\mathfrak{M}_\rho \models T_\eta(s) \text{ iff } s^\mathbb{N} \in \mathcal{F}_{<\rho|\eta}^*.$$

Theorem 7 (T-Elimination theorem). *Suppose $\rho \in \mathbb{O}$.*

1. *For every T-positive sequent Γ , $\mathcal{F}_\rho^* \vdash \Gamma$ implies $\mathfrak{M}_\rho \models \bigvee \Gamma$;*
2. *For any $\eta < \kappa$, $\mathcal{F}_\rho^* \vdash T_\eta(s)$ implies there is a sentence A with $s^\mathbb{N} = \ulcorner A \urcorner$ and $\mathcal{F}_\rho^* \vdash A$;*
3. *For any $\eta < \kappa$, $\mathcal{F}_\rho^* \vdash \neg T_\eta(s)$ implies there is a sentence A with $s^\mathbb{N} = \ulcorner A \urcorner$ and $\mathcal{F}_\rho^* \vdash \neg A$.*

Proof. We proceed by transfinite induction on ρ . For (i), one has a *subsidiary* induction on the height of the derivation. The base case is easy to deal with. For the induction step we argue according to the last rule applied in the derivation $\mathcal{F}_\rho^* \vdash \Gamma$. Whichever rule was applied, the sequent(s) in the premise must also be T-positive and we may apply the subsidiary induction hypothesis to them.

If the last rule was one of the arithmetical rules, that is, (\vee_i) , (\wedge) , (ω) or (\exists) , $\mathfrak{M}_\rho \models \bigvee \Gamma$ is an immediate consequence of the subsidiary induction hypothesis, and in the case of the weakening rule, $\mathfrak{M}_\rho \models \bigvee \Gamma$ follows from the fact that Γ is T-positive. If the last applied rule was T_η -Intro, $T_\eta(\ulcorner A \urcorner)$ is contained in Γ and $\mathcal{F}_\sigma^* \vdash A$ for some $\sigma < \rho|\eta$, so $\mathfrak{M}_\rho \models T_\eta(\ulcorner A \urcorner)$. For the remaining rules, the subsidiary induction hypothesis implies $\mathfrak{M}_\rho \models \bigvee \Gamma \vee (A_0 \wedge A_1)$ for some suitable choice of A_0, A_1 . Of course, if $\mathfrak{M}_\rho \models \bigvee \Gamma$ we are done, so we may assume $\mathfrak{M}_\rho \models A_0 \wedge A_1$.

T_η -Imp. If the last rule applied was T_η -Imp, we may assume A_0 is $T_\eta(s_0)$ and A_1 is $T_\eta(s_0 \rightarrow s_1)$, while Γ contains $T_\eta(s_1)$. By the above, we may assume $\mathfrak{M}_\rho \models T_\eta(s_0) \wedge T_\eta(s_0 \rightarrow s_1)$. Thus, $s_0^\mathbb{N}$ and $s_1^\mathbb{N}$ are Gödel numbers of \mathcal{L}_κ -sentences, say B_0 and B_1 respectively, and there is some $\sigma < \rho|\eta$ so that $\mathcal{F}_\sigma^* \vdash B_0$ and $\mathcal{F}_\sigma^* \vdash \neg B_0, B_1$. Admissibility of the cut rule (lemma 13) yields $\mathcal{F}_\sigma^* \vdash B_1$, and hence $\mathfrak{M}_\rho \models T_\eta(s_1)$.

T_η -Del. In the case the last applied rule is T_η -Del, we may identify A_0 as $T_\eta(\ulcorner T_\xi(s) \urcorner)$ for some $\xi < \kappa$ and term s ; moreover, $T_\eta(s)$ is contained in Γ . $\mathfrak{M}_\rho \models T_\eta(\ulcorner T_\xi(s) \urcorner)$ implies $\mathcal{F}_\sigma^* \vdash T_\xi(s)$ for some $\sigma < \rho|\eta$. Since $\sigma < \rho$, the *main* induction hypothesis may be applied, whence $s^\mathbb{N} = \ulcorner A \urcorner$ for some A and $\mathcal{F}_\sigma^* \vdash A$. Thus $\mathfrak{M}_\rho \models T_\eta(s)$ and $\mathfrak{M}_\rho \models \bigvee \Gamma$.

T_η^- -Rep. Here we have $\mathfrak{M}_\rho \models T_\eta(s)$ and $T_\eta(\ulcorner T_\xi(s) \urcorner)$ is in Γ for some $\xi < \eta$. By definition this implies $s^\mathbb{N} = \ulcorner A \urcorner$ for some sentence A and $\mathcal{F}_\sigma^* \vdash A$ for some $\sigma < \rho|\eta$, whence $\mathcal{F}_{\sigma+\Omega^{\xi+1}}^* \vdash T_\xi(s)$ is derivable. But since $\xi < \eta$ and $\sigma < \rho|\eta$, we have $\sigma + \Omega^{\xi+1} < \rho|\eta$, and so $\mathfrak{M}_\rho \models \bigvee \Gamma$.

$\neg T_\eta$ -Del. This case is similar to T_η -Del above; by the main induction hypothesis we know, for every $\sigma < \rho$ and $\xi < \kappa$, that \mathcal{F}_σ^* is closed under $\neg T_\xi$ -Elim, thus $\mathfrak{M}_\rho \models T_\eta(\ulcorner \neg T_\xi(s) \urcorner)$ implies $\mathfrak{M}_\rho \models T_\eta(\ulcorner \neg s \urcorner)$ as desired.

T_η - \forall -Inf. The assumption is that $\mathfrak{M}_\rho \models \forall x T_\eta(\ulcorner A(\dot{x}) \urcorner)$. This entails the existence of, for every $n < \omega$, an ordinal $\sigma_n < \rho|\eta$ such that $\mathcal{F}_{\sigma_n}^* \vdash A(\bar{n})$. Weakening and the ω -rule yields $\mathcal{F}_\sigma^* \vdash \forall x A(x)$, where $\sigma = \sup_n \sigma_n$, but one need not in general have $\sigma < \rho|\eta$.⁶ Due to the restriction on applications of T_η - \forall -Inf, however, $\rho|\eta$ is not an Ω -limit, thus by proposition 14, $\sigma < \rho|\eta$ and so $\mathfrak{M}_\rho \models T_\eta(\ulcorner \forall x A(x) \urcorner)$, whence $\mathfrak{M}_\rho \models \bigvee \Gamma$.

⁶For example, suppose $\rho|\eta = \rho_0 + \Omega^\xi$ and ξ is a limit ordinal. If $\sigma_n = \rho_0 + \Omega^{\xi_n}$, where $\xi = \sup_n \xi_n$ and $\xi_n < \xi$ for every $n < \omega$, one has $\sigma_n < \rho|\eta$, but $\sup_n \sigma_n = \rho|\eta$.

This completes the proof of (i).

(ii) is now a consequence of (i). If $\mathcal{F}_\rho^* \vdash T_\eta(s)$, (i) implies $\mathfrak{M}_\rho \models T_\eta(s)$, whence $s^\mathbb{N} = \ulcorner A \urcorner$ for some \mathcal{L}_κ -sentence A and $\mathcal{F}_\sigma^* \vdash A$ for some $\sigma < \rho|_\eta$. By weakening, $\mathcal{F}_\rho^* \vdash A$, as desired.

Observe that in the case of every rule of inference in the system \mathcal{F}_ρ^* , T-positive premises yield T-positive consequents. Therefore $\mathcal{F}_\eta^* \vdash \Gamma$ implies $\bigvee \Gamma$ is satisfied in the *everything is true* \mathcal{L}_κ structure, so $\mathcal{F}_\rho^* \vdash \neg T_\eta(s)$ is impossible and (iii) holds vacuously. \square

Proposition 15. *Let A be any axiom of F_β . Then $\mathcal{F}_{\Omega^{\beta+1}}^* \vdash A$.*

Proof. One can derive each of the axioms via the corresponding rule and Ax.2 $_\eta$, as in ???. In the case of $T_\eta\text{-}\forall\text{-Inf}$ note $\Omega^{\beta+1}|_\eta$ is not an Ω -limit for any $\eta \leq \beta$. \square

Theorem 8. *The theory F_β is consistent for every $\beta < \kappa$.*

Proof. Lemma 13, theorem 7 and the previous proposition provide the means to deduce, by induction on n , that F_β^n embeds into $\mathcal{F}_{\Omega^{\beta+1}.n}^*$. Thus every sentential theorem of F_β is contained in $\mathcal{F}_{<\Omega^\beta.\omega}^*$. However, clearly the empty sequent is not derivable in \mathcal{F}_ρ^* for any ρ , so F_β must be consistent. \square

4.2 An ordinal notation system for impredicative theories

To carry out an ordinal analysis of F_β we require the current set of ordinal terms, OT, to be extended to cover a larger segment of the ordinals. We will make use of an ordinal notation system for the Bachmann-Howard ordinal introduced by Rathjen and Weiermann [?]. This ordinal has proved significant in the analysis of certain impredicative systems such as the theory of inductive definitions, ID₁ [?]. It will turn out that the theories F_β are substantially weaker than ID₁, but this notation system is still a natural one to consider. The key to generating notations for characteristic ordinals beyond Γ_0 is the use of constructions referencing certain ‘external points’. In our case the ‘external point’ will be Ω , the first uncountable ordinal.

In order to generate unique representations for ordinals we will introduce a normal form for non- ε -ordinals, based on the Cantor normal form. We write $\alpha =_{\text{NF}} \omega^\gamma + \delta$ if $\alpha = \omega^\gamma + \delta$ and either $\delta = 0$ and $\gamma < \alpha$, or $\delta = \omega^{\delta_1} + \dots + \omega^{\delta_k}$, $\gamma \geq \delta_1 \geq \dots \geq \delta_k$ and $k \geq 1$. Let $\varepsilon_{\Omega+1}$ be the first ε -ordinal larger than Ω . For each $\alpha < \varepsilon_{\Omega+1}$ we denote by α^* the largest ε -ordinal below Ω used in the normal form presentation for α ; that is,

1. $0^* = \Omega^* = 0$,
2. $\alpha^* = \alpha$, if $\alpha < \Omega$ is an ε -ordinal,
3. $\alpha^* = \max\{\gamma^*, \delta^*\}$, if $\alpha =_{\text{NF}} \omega^\gamma + \delta$.

Define sets of ordinals $C_k(\alpha, \beta)$, and a function $\vartheta: \mathbb{O} \rightarrow \Omega$ by transfinite recursion on $\alpha \in \mathbb{O}$ as follows.

- (C1) $\{0, \Omega\} \cup \beta \subseteq C_k(\alpha, \beta)$,
- (C2) $\gamma, \delta \in C_k(\alpha, \beta)$ and $\xi =_{\text{NF}} \omega^\gamma + \delta$ implies $\xi \in C_{k+1}(\alpha, \beta)$,
- (C3) $\xi \in C_k(\alpha, \beta)$ and $\xi < \alpha$ implies $\vartheta\xi \in C_{k+1}(\alpha, \beta)$,
- (C4) $C(\alpha, \beta) = \bigcup_{k < \omega} C_k(\alpha, \beta)$,

$$(\vartheta_1) \quad \vartheta\alpha = \min\{\xi < \Omega : C(\alpha, \xi) \cap \Omega \subseteq \xi \wedge \alpha \in C(\alpha, \xi)\}.$$

The next two propositions shed some light on the role the function ϑ plays in generating initial segments of \mathbb{O} .

Proposition 16. $\vartheta\alpha$ is defined for every $\alpha < \varepsilon_{\Omega+1}$.

Proof. Let $\gamma_0 = \alpha^* + 1$. By rules (C1) and (C2) we may deduce $\alpha \in C(\alpha, \gamma_0)$. Suppose $\gamma_k < \Omega$ has been defined. As $C(\alpha, \gamma_k)$ has a countable definition, it contains only countably many elements; thus $C(\alpha, \gamma_k) \cap \Omega$ is bounded in Ω . Let $\gamma_{k+1} < \Omega$ be such that $C(\alpha, \gamma_0) \cap \Omega \subseteq \gamma_{k+1}$ and define $\gamma = \sup_{k < \omega} \gamma_k$. Since $\{\gamma_k : k < \omega\}$ is a countable set of countable ordinals, it too must be bounded in Ω , so $\gamma < \Omega$. Since $\alpha < \gamma_0 \leq \gamma$, also $\alpha^* \in C(\alpha, \gamma)$. Finally,

$$C(\alpha, \gamma) \subseteq \bigcup_{n < \omega} C(\alpha, \gamma_n),$$

so $C(\alpha, \gamma) \cap \Omega \subseteq \gamma$ and $\vartheta\alpha \leq \gamma$ by (ϑ_1) . \square

The argument in the proof above provides a means to approximate the ordinal $\vartheta\alpha$ from below. Define $\gamma_0 = \alpha^* + 1$ and $\gamma_{m+1} = \min\{\xi < \Omega : C(\alpha, \gamma_m) \cap \Omega \subseteq \xi\}$; then $\vartheta\alpha \leq \sup_m \gamma_m$. By (C2) it is clear that each γ_m is a limit ordinal, whence we may deduce

$$[(\forall \delta \in C(\alpha, \gamma_m) \cap \Omega) \forall \xi < \delta F(\xi)] \implies \forall \delta < \gamma_{m+1} F(\delta)$$

for every formula F .

The function ϑ works by “collapsing” ordinals below $\varepsilon_{\Omega+1}$ into countable ordinals, thus allowing one to represent ordinals beyond Γ_0 . Moreover, the condition “ $\alpha \in C(\alpha, \xi)$ ” in (ϑ_1) ensures the function $\alpha \mapsto \vartheta\alpha$ is strictly increasing on Ω . Thus, unlike the Veblen functions φ_α , one never has $\vartheta\beta = \vartheta\vartheta\beta$. In the following proposition we show ϑ is in fact injective on $\varepsilon_{\Omega+1}$ and show that relation ‘ $\alpha < \beta$ ’ may be decided purely on the normal form presentation for α and β .

Proposition 17. For all ordinals α, β and γ the following holds.

1. $\alpha \in C(\alpha, \vartheta\alpha)$,
2. $\vartheta\alpha = C(\alpha, \vartheta\alpha) \cap \Omega$ and $\vartheta\alpha \notin C(\alpha, \vartheta\alpha)$,
3. $\vartheta\alpha$ is an ε -ordinal,
4. $\gamma \in C(\alpha, \beta)$ if and only if $\gamma^* \in C(\alpha, \beta)$,
5. $\beta < \vartheta\alpha$ if and only if $\beta < \Omega$ and $\beta^* < \vartheta\alpha$,
6. $\alpha^* < \vartheta\alpha$,
7. $\vartheta\alpha = \vartheta\beta$ if and only if $\alpha = \beta$,
8. $\vartheta\alpha < \vartheta\beta$ if and only if $(\alpha < \beta \wedge \alpha^* < \vartheta\beta) \vee (\beta < \alpha \wedge \vartheta\alpha \leq \beta^*)$.

Proof. Proposition 16 ensures $\vartheta\alpha$ is defined for every $\alpha < \varepsilon_{\Omega+1}$, so (i) is a direct consequence of $(\vartheta 1)$. By (C1), $\vartheta\alpha \subseteq C(\alpha, \vartheta\alpha)$ whence (ii) also follows from $(\vartheta 1)$. (ii) then implies (iii) since, as a result of (C2), $\gamma < \vartheta\alpha$ only if $\omega^\gamma < \vartheta\alpha$.

(iv). Suppose $\gamma \in C_k(\alpha, \beta)$. If $k = 0$, $\gamma^* \in C(\alpha, \beta)$ is immediate by (C1), so suppose $k > 0$. We show $\gamma^* \in C(\alpha, \beta)$ by examining the normal form of γ . If γ is either 0 or Ω , $\gamma^* = 0 \in C(\alpha, \beta)$ by (C1), and if γ is an ε -ordinal, $\gamma^* = \gamma$ and we are done. Otherwise $\gamma =_{\text{NF}} \omega^\xi + \delta$ and $\delta, \xi \in C_{k-1}(\alpha, \beta)$. The induction hypothesis implies $\delta^*, \xi^* \in C(\alpha, \beta)$, whence $\gamma^* \in C(\alpha, \beta)$. The converse direction holds by a similar argument.

(v) is an immediate consequence of (iv) and (ii); (vi) holds on account of (i), (ii) and (iv). To show (vii) suppose $\vartheta\alpha = \vartheta\beta$ but $\alpha < \beta$. Then $C(\alpha, \vartheta\alpha) \subseteq C(\beta, \vartheta\beta)$, so $\alpha \in C(\beta, \vartheta\beta)$ by (i), whence $\vartheta\alpha \in C(\beta, \vartheta\beta) \cap \Omega$ by (C3). Thus $\vartheta\beta \in C(\beta, \vartheta\beta) \cap \Omega$ contradicting (ii).

(viii). Assume $\alpha < \beta$. By (vi), $\vartheta\alpha < \vartheta\beta$ implies $\alpha^* < \vartheta\beta$. Also $\alpha^* < \vartheta\beta$ implies $\alpha^* \in C(\beta, \vartheta\beta)$, whence $\vartheta\alpha \in C(\beta, \vartheta\beta) \cap \Omega$ and so $\vartheta\alpha < \vartheta\beta$ by (ii). Thus

$$\alpha < \beta \implies (\vartheta\alpha < \vartheta\beta \leftrightarrow \alpha^* < \vartheta\beta). \quad (27) \quad \text{\texttt{\texttt{exteqn:theta1.}}}$$

Now suppose $\beta < \alpha$. By the same argument we obtain

$$\beta < \alpha \implies (\vartheta\beta < \vartheta\alpha \leftrightarrow \beta^* < \vartheta\alpha),$$

and so, by (vii),

$$\beta < \alpha \implies (\vartheta\alpha < \vartheta\beta \leftrightarrow \vartheta\alpha \leq \beta^*). \quad (28) \quad \text{\texttt{\texttt{exteqn:theta1.}}}$$

Combining eqs. (27) and (28) gives (viii). \square

We can now proceed with defining a primitive recursive set of ordinal terms for use in the later analysis of F_β .

Definition 18. Define a subset $\text{OT}_\Omega \subseteq \mathbb{N}$, an encoding τ of ordinals into OT_Ω and a rank function $|\cdot|$ on ordinals by recursion according to the following rules.

1. $\tau(0) = 0 \in \text{OT}_\Omega$, $\tau(\Omega) = \langle 0, 1 \rangle \in \text{OT}_\Omega$, and $|0| = |\Omega| = 0$,
2. If $\alpha = \vartheta\alpha_0$ and $\tau(\alpha_0) \in \text{OT}_\Omega$, $\tau(\alpha) = \langle 1, \tau(\alpha_0) \rangle \in \text{OT}_\Omega$ and $|\alpha| = |\alpha_0| + 1$,
3. If $\alpha =_{\text{NF}} \omega^\gamma + \delta$ and $\tau(\gamma), \tau(\delta) \in \text{OT}_\Omega$, $\tau(\alpha) = \langle 2, \tau(\gamma), \tau(\delta) \rangle \in \text{OT}_\Omega$ and $|\alpha| = \max\{|\gamma|, |\delta|\} + 1$.

It should be noted that the definition of $x \in \text{OT}_\Omega$ and $|\alpha|$ are primitive recursive.

We now want to define an ordering $<_\vartheta$ on OT_Ω such that $\tau(\alpha) <_\vartheta \tau(\beta)$ if and only if $\alpha < \beta$. Conditions (iii) and (vii) of proposition 17 ensures every ordinal built up from the constants 0, Ω and functions $\alpha, \beta \mapsto \omega^\alpha + \beta$ and $\alpha \mapsto \vartheta\alpha$ has a unique representation. We may therefore dispense with the function τ and identify members of OT_Ω with the ordinals they represent, as was the case with OT .

Define the relation $\alpha <_\vartheta \beta$ on OT_Ω by recursion on the value of $|\alpha| + |\beta|$. The conditions involved in comparing two ordinals $\vartheta\xi_0$ and $\vartheta\xi_1$ will be taken from (viii) of proposition 17. Let $\alpha <_\vartheta \beta$ if and only if one of the following conditions hold.

1. $\alpha = 0$ and $\beta \neq 0$;
2. $\alpha =_{\text{NF}} \omega^\gamma + \delta$ and either:

- a) $\beta = \Omega$ and $\gamma <_{\mathcal{O}} \beta$,
 - b) $\beta =_{\text{NF}} \omega^{\gamma_0} + \delta_0$ and $\gamma <_{\mathcal{O}} \gamma_0$, or $\gamma = \gamma_0 \wedge \delta <_{\mathcal{O}} \delta_0$, or
 - c) $\beta = \vartheta \xi$ and $\gamma <_{\mathcal{O}} \beta$;
3. $\alpha = \vartheta \xi$ and either:
- a) $\beta = \Omega$,
 - b) $\beta =_{\text{NF}} \omega^{\gamma} + \delta$ and $\alpha \leq_{\mathcal{O}} \gamma$,⁷ or
 - c) $\beta = \vartheta \eta$ and either, $\xi <_{\mathcal{O}} \eta \wedge \xi^* <_{\mathcal{O}} \beta$, or $\eta <_{\mathcal{O}} \xi \wedge \alpha \leq_{\mathcal{O}} \eta^*$.
- ($\gamma \leq_{\mathcal{O}} \delta$ abbreviates $\gamma <_{\mathcal{O}} \delta$ or $\gamma = \delta$.)

Since the function $\alpha \mapsto \alpha^*$ is primitive recursive, the relation $<_{\mathcal{O}}$ is also primitive recursive.

Before we proceed with the analysis of F_{β} , we will show how the ordinals $\alpha \geq \Omega$ in OT_{Ω} enable the generation of the φ_{α} functions for $\alpha < \Gamma_0$ and that OT_{Ω} properly extends OT.

Let $\Omega \cdot 0 = 0$ and if $\alpha = \omega^{\alpha_0} + \dots + \omega^{\alpha_n}$ and $\alpha_0 \geq \dots \geq \alpha_n$, define

$$\begin{aligned}\Omega \cdot \alpha &= \omega^{\Omega+\alpha_0} + \dots + \omega^{\Omega+\alpha_n}, \\ \Omega^{\beta} \cdot \alpha &= \omega^{\Omega \cdot \beta + \alpha_0} + \dots + \omega^{\Omega \cdot \beta + \alpha_n}.\end{aligned}$$

Proposition 19. $\text{OT}_{\Omega} \cap \Omega$ forms an initial segment of the ordinals.

Proof. Suppose $\delta \in \text{OT}_{\Omega} \cap \Omega$. We prove $\alpha \in \text{OT}_{\Omega}$ for every $\alpha < \delta$ by transfinite induction on α . Suppose

$$\forall \xi < \alpha (\xi \in \text{OT}_{\Omega}). \quad (29)$$

If $\alpha = 0$ or $\alpha =_{\text{NF}} \omega^{\gamma} + \delta$, eq. (29), (C1) and (C2) imply $\alpha \in \text{OT}_{\Omega}$, so assume α is an ε -ordinal. In search of a contradiction, assume further that $\alpha \notin \text{OT}_{\Omega}$. We prove $\alpha = \vartheta \xi$ for some $\xi \in \text{OT}_{\Omega}$. Pick β to be the least ordinal in $\{\gamma \leq \delta : \gamma \in \text{OT}_{\Omega} \wedge \alpha < \gamma\}$. Then β must be an ε -ordinal as otherwise $\beta = \omega^{\gamma_0} + \gamma_1$ for some $\gamma_0, \gamma_1 < \beta$, whence $\alpha \leq \max\{\gamma, \delta\} \in \text{OT}_{\Omega} \cap \Omega$. Since $\beta \in \text{OT}_{\Omega}$, $\beta = \vartheta \xi$ for some $\xi \in \text{OT}_{\Omega}$. Now, as $\xi^* < \vartheta \xi$ we have $\xi^* < \alpha$. Thus $\xi \in C(\xi, \alpha)$.

Moreover, we claim $C(\xi, \alpha) \cap \Omega \subseteq \alpha$. The argument proceeds by induction on the definition of $\eta \in C(\xi, \alpha) \cap \Omega$. In case $\eta = 0$ or $\eta =_{\text{NF}} \omega^{\gamma} + \delta$, it is immediate that $\eta < \alpha$, so suppose $\eta = \vartheta \zeta$. Then $\zeta < \xi$ and, since

$$\eta \in C(\xi, \alpha) \cap \Omega \subseteq C(\xi, \vartheta \xi) \cap \Omega = \vartheta \xi,$$

$\eta < \beta$. By eq. (29) and the fact $\xi \in \text{OT}_{\Omega}$, we obtain $C(\xi, \alpha) \subseteq \text{OT}_{\Omega}$, so $\eta = \vartheta \zeta \in \text{OT}_{\Omega}$, whence $\eta < \alpha$ by the choice of β .

Thus we have shown $\xi \in C(\xi, \alpha)$ and $C(\xi, \alpha) \cap \Omega \subseteq \alpha$, whence $\vartheta \xi \leq \alpha$ by ($\vartheta 1$), and $\alpha = \beta \in \text{OT}_{\Omega}$. \square

Proposition 20.

1. $\vartheta \Omega$ is the least ordinal closed under the function $\gamma \mapsto \vartheta \gamma$.

⁷Recall that if γ is an ε -ordinal, $\beta =_{\text{NF}} \omega^{\gamma} + \delta$ only if $\delta > 0$, whence $\vartheta \xi <_{\mathcal{O}} \beta$ if $\vartheta \xi = \gamma$.

2. $\vartheta\Omega^2$ is the least ordinal closed under the function $\gamma \mapsto \vartheta(\Omega \cdot \gamma)$.

3. $\vartheta\Omega^3$ is the least ordinal closed under the function $\gamma \mapsto \vartheta(\Omega^2 \cdot \gamma)$.

Proof. (i). Let $\gamma_0 = 0$ and $\gamma_{m+1} = \vartheta\gamma_m$; we claim

$$\sup_{n < \omega} \gamma_n = \vartheta\Omega,$$

whence it is clear (i) holds. $\gamma_0 < \vartheta\Omega$ holds trivially and, if $\gamma_m < \vartheta\Omega$, we have $(\gamma_m)^* < \vartheta\Omega$ and so $\gamma_{m+1} = \vartheta\gamma_m < \vartheta\Omega$ by (viii) of proposition 17; thus $\sup_n \gamma_n \leq \vartheta\Omega$.

To show $\vartheta\Omega \leq \sup_n \gamma_n$ we prove $\beta < \vartheta\Omega$ implies $\beta < \gamma_m$ for some m by induction on the rank of β . Suppose $\beta < \vartheta\Omega$. Both $\vartheta\Omega$ and $\sup_n \gamma_n$ are ε -ordinals, so the case $\beta =_{\text{NF}} \omega^{\beta_0} + \beta_1$, holds by the induction hypothesis. If $\beta = \vartheta\xi$ we deduce $\xi < \Omega$ and $\xi^* < \vartheta\Omega$, since $\beta < \vartheta\Omega$ and $\Omega^* = 0$. ξ^* has rank strictly less than β , so the induction hypothesis yields an $m < \omega$ such that $\xi^* < \gamma_m$. Proposition 17 (v) then entails $\xi < \gamma_m$. Moreover, since $\gamma_m = (\gamma_m)^* < \vartheta\gamma_m$, by proposition 17 (vi), $\gamma_m < \gamma_{m+1}$, so $\xi^* < \vartheta\gamma_m$. Therefore $\beta = \vartheta\xi < \vartheta\gamma_m = \gamma_{m+1}$.

(ii). Let $\gamma_0 = 0$ and $\gamma_{m+1} = \vartheta(\Omega \cdot \gamma_m)$; we claim

$$\sup_{n < \omega} \gamma_n = \vartheta\Omega^2,$$

whence (ii) holds. Let $\alpha = \sup_n \gamma_n$. Naturally, $\gamma_0 < \vartheta\Omega^2$, and if $\gamma_m < \vartheta\Omega^2$,

$$(\Omega \cdot \gamma_m)^* = \gamma_m^* < \vartheta\Omega^2,$$

so $\gamma_{m+1} < \vartheta\Omega^2$ by proposition 17 (viii). Thus $\alpha \leq \vartheta\Omega^2$.

To show the converse, we prove $\beta < \vartheta\Omega^2$ implies $\beta < \alpha$ by induction on the rank of β . Suppose $\beta < \vartheta\Omega^2$ and $\beta = \vartheta\xi$ for some ξ . As $(\Omega^2)^* = 0$ and $\beta < \vartheta\Omega^2$, proposition 17 (viii) implies $\xi < \Omega^2$ and $\xi^* < \vartheta\Omega^2$, whence the induction hypothesis implies $\xi^* < \gamma_m$ for some m . Since $\xi < \Omega^2$, there are $\delta_0, \delta_1 < \Omega$ such that $\xi = \Omega \cdot \delta_0 + \delta_1$, whence $\delta_0^* \leq \xi^* < \gamma_m$. So $\xi < \Omega \cdot \gamma_m$ and $\beta < \gamma_{m+1}$ by proposition 17 (viii).

(iii) involves a near identical argument as (ii). Pick $\gamma_0 = 0$ and $\gamma_{m+1} = \vartheta(\Omega^2 \cdot \gamma_m)$. That $\sup_n \gamma_n \leq \vartheta\Omega^3$ is easily established using proposition 17. For the converse direction, $\vartheta\Omega^3 \leq \sup_n \gamma_n$, we suppose $\beta < \vartheta\Omega^3$ and seek to determine $\beta < \gamma_m$ for some m . If $\beta = \vartheta\xi < \vartheta\Omega^3$, we may assume $\xi^* < \gamma_m$ for some m , whence $\xi < \Omega^2 \cdot \gamma_m$, and so $\beta < \gamma_{m+1}$. \square

Proposition 20 allows us to identify some characteristic ordinals in terms of both the Veblen and ϑ functions.

Corollary 21.

1. $\alpha < \varphi_{20}$ implies $\varepsilon_\alpha = \vartheta\alpha$.
2. $\vartheta\Omega = \varphi_{20}$.
3. $\vartheta\Omega^2 = \Gamma_0$.
4. $\vartheta(\Omega^2 + \Omega)$ is the least fixed point of the function $\xi \mapsto \Gamma_\xi$, which enumerates the class $\{\xi : \xi = \varphi\xi 0\}$.

Proof. (i) is argued by transfinite induction on $\alpha < \varphi 20$. Suppose $\vartheta\beta = \varepsilon_\beta$ for every $\beta < \alpha$. We will begin by showing a) $C(\alpha, \varepsilon_\alpha) \cap \Omega \subseteq \varepsilon_\alpha$, and b) $\alpha \in C(\alpha, \varepsilon_\alpha)$, allowing us to deduce $\vartheta\alpha \leq \varepsilon_\alpha$ by ($\vartheta 1$).

a) is shown by a further induction on the construction of $C(\alpha, \varepsilon_\alpha)$. It is trivial that $\beta \in C_0(\alpha, \varepsilon_\alpha) \cap \Omega$ entails $\beta < \varepsilon_\alpha$, and that ε_α is closed under applications of rule (C2). Moreover, the induction hypothesis implies that for $\beta < \alpha$, $\vartheta\beta < \varepsilon_\alpha$, thus (C3) is also dealt with, and $C(\alpha, \varepsilon_\alpha) \cap \Omega \subseteq \varepsilon_\alpha$.

b) uses the fact $\alpha < \varphi 20$, whence $\alpha < \varepsilon_\alpha$ and $\alpha \in C(\alpha, \varepsilon_\alpha)$ by (C1).

To see $\varepsilon_\alpha \leq \vartheta\alpha$, assume otherwise. Then $\vartheta\alpha = \varepsilon_\beta$ for some $\beta < \alpha$ by proposition 17 (iii). The induction hypothesis yields $\vartheta\alpha = \vartheta\beta$, contradicting $\beta < \alpha$.

(ii) is an immediate consequence of (i) and proposition 20.

(iii). The proof for (i) above can be extended to $\alpha > \varphi 20$, but then one can at best show $\varepsilon_\alpha \leq \vartheta\alpha \leq \varepsilon_{\alpha+1}$ for $\alpha < \Omega$.⁸ One can then prove

$$\varphi 2\alpha \leq \vartheta(\Omega + \alpha) \leq \varphi 2(\alpha + 1)$$

for $\alpha < \Omega$ by transfinite induction on α , using the definition of ϑ . This can easily be extended to deduce, in general,

$$\varphi\alpha\beta \leq \vartheta(\Omega \cdot \alpha + \beta) \leq \varphi(\alpha + 1)(\beta + 1)$$

for $\alpha, \beta < \Gamma_0$, from which proposition 20 (ii) implies $\vartheta\Omega^2 = \Gamma_0$.

(iv). Let Δ_0 denote the least fixed point of the function $\xi \mapsto \Gamma_\xi$. Following from (iii) above, $\vartheta(\Omega^2 + \alpha) = \Gamma_\alpha$ for $\alpha < \Delta_0$. Since $\vartheta(\Omega^2 + \Omega)$ is the least ordinal closed under the function $\alpha \mapsto \vartheta(\Omega^2 + \alpha)$, we deduce $\vartheta(\Omega^2 + \Omega) = \Delta_0$. \square

In this notation system, $\vartheta\Omega^3$ represents the *Ackermann ordinal*, $\vartheta\Omega^\Omega$ denotes the *Veblen ordinal* and $\vartheta\varepsilon_{\Omega+1}$ is the *Bachmann-Howard ordinal* where

$$\vartheta\varepsilon_{\Omega+1} = \sup\{\vartheta\Omega, \vartheta\Omega^\Omega, \vartheta\Omega^{\Omega^\Omega}, \dots\} = \sup_{\alpha \in \text{OT}_\Omega} \vartheta\alpha.$$

Having established an ordinal notation system suitable for the analysis of the theories F_β , we may now fix the language of F_β . Since the proof-theoretic strength of each theory F_β with $\beta < \vartheta\varepsilon_{\Omega+1}$ will not exceed $\vartheta\varepsilon_{\Omega+1}$, we may pick $\kappa = \vartheta\varepsilon_{\Omega+1}$ and suppose the theories F_β are formulated in the language \mathcal{L}_κ .

We require a few further technical results about ordinals before we can proceed with the analysis. Suppose $\beta = \Omega^{\alpha_0} \cdot \beta_0 + \dots + \Omega^{\alpha_n} \cdot \beta_n$ such that $\alpha_0 > \dots > \alpha_n$ and $\beta_i < \Omega$ for each $i \leq n$. Recall from the previous section that $\beta|_\gamma$ denotes the ordinal $\Omega^{\alpha_0} \cdot \beta_0 + \dots + \Omega^{\alpha_k} \cdot \beta_k$ where $k < n$ is the least such that $\alpha_k > \gamma \geq \alpha_{k+1}$, or $k = n$ if $\alpha_n > \gamma$.

The following observations are immediate consequences of the definition.

Proposition 22. For all ordinals $\alpha, \beta < \varepsilon_{\Omega+1}$ and $\gamma, \delta < \Omega$,

1. $\gamma < \delta$ implies $\alpha|_\gamma \leq \alpha|_\delta$,
2. $\alpha < \beta$ implies $\alpha|_\gamma \leq \beta|_\gamma$.
3. $\delta \leq \gamma$ implies $(\alpha|_\gamma)|_\delta = \alpha|_\gamma$,
4. $\beta < \alpha|_\gamma$ if and only if $\beta + \Omega^{\gamma+1} \leq \alpha$,
5. $\beta < \alpha|_\gamma$ and $\delta \leq \gamma$ implies $\beta + \Omega^\delta < \alpha|_\gamma$.

⁸ $\vartheta\alpha = \varepsilon_{\alpha+1}$ for $\alpha = \varphi 20$ for example.

4.3 Lower bounds on the proof-theoretic ordinal

We will now establish lower bounds for the theories F_β . This will be achieved by extending the well-ordering proofs used in our analysis of F (??) and S_3 (section 3.3). Recall $\kappa = \vartheta_{\varepsilon_{\Omega+1}}$.

F_0 is identical to F , for which an optimal lower bound was established in ??. However, because of the change in ordinal notation system and the reflective nature of the theories F_β it will be useful to provide a new proof of the result. For each $\xi < \kappa$ let $\text{wo}_\xi(x)$ denote the formula

$$\forall^\top A(x)^\top \forall y < x \text{ T}_\xi(\top \text{TI}(\dot{y}, A)^\top).$$

Let $F_0(\rho)$ denote the formula $\text{wo}_0(\rho^*) \wedge \forall \sigma < \rho [\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta \sigma)]$ and $F_0^\rho(\alpha)$ denote $\alpha < \Omega \implies F_0(\rho + \alpha)$. We begin with a technical lemma.

Lemma 23. $F_0^1 \vdash \forall \rho [\text{T}_0(\top F_0(\dot{\rho})^\top) \implies \text{wo}_0(\vartheta \rho)]$.

Proof. Argue within F_0^1 , and assume

$$\text{T}_0(\top F_0(\dot{\rho})^\top). \tag{30}$$

Let $\gamma_0 = \rho^* + 1$ and $\gamma_{m+1} = C(\rho, \gamma_m) \cap \Omega$. Moreover, let $\text{wo}_0^1(x)$ denote $\text{T}_0(\top \text{wo}_0(\dot{x})^\top)$. Equation (30) implies $\text{wo}_0^1(\gamma_0)$ and $\forall \sigma < \rho [\text{wo}_0^1(\sigma^*) \implies \text{wo}_0^1(\vartheta \sigma)]$, whence it is easy to deduce $\forall m [\text{wo}_0^1(\gamma_m) \implies \text{wo}_0^1(\gamma_{m+1})]$ and thus $\text{wo}_0^1(\vartheta \rho)$. By T_0 -Del, $\text{wo}_0(\vartheta \rho)$ holds. \square

Lemma 24. For every $m < \omega$, $F_0^m \vdash F_0(\Omega \cdot \bar{m})$.

Proof. $F_0^0 \vdash F_0(\bar{0})$ holds vacuously, so suppose $m = n + 1 > 0$ and

$$F_0^n \vdash F_0(\Omega \cdot \bar{n}). \tag{31}$$

The first step is to establish $F_0^n \vdash \text{Prog}_{F_0^{\Omega \cdot \bar{n}}}$. Argue informally within F_0^n , assuming $\forall v < \mu F_0^{\Omega \cdot \bar{n}}(v)$ for some μ , that is,

$$\forall v < \mu [\text{wo}_0(v^*) \wedge \forall \sigma < \Omega \cdot \bar{n} + v [\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta \sigma)]]. \tag{32}$$

We want to show $\text{wo}_0(\mu^*)$ and $\forall \sigma < \Omega \cdot \bar{n} + \mu [\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta \sigma)]$. The former is obvious since the fact $\text{wo}_0(\varepsilon_\alpha)$ is progressive in α is provable in F_0^1 (cf. ??). To prove the latter, assume $\text{wo}_0(\sigma^*)$ for some $\sigma < \Omega \cdot \bar{n} + \mu$. If $\mu = 0$ or is a limit ordinal, $\text{wo}_0(\vartheta \sigma)$ is immediate given eq. (32). Otherwise $\mu = v + 1$ for some v , whence we may assume $\sigma = \Omega \cdot \bar{n} + v$. Let $\gamma_0 = \sigma^* + 1$ and $\gamma_{m+1} = C(\sigma, \gamma_m) \cap \Omega$. Then $\gamma_m < \Omega$ for each m and

$$\vartheta \sigma \leq \sup_m \gamma_m$$

by proposition 16. $\text{wo}_0(\gamma_0)$ is a consequence of $\text{wo}_0(\sigma^*)$, so suppose

$$\text{wo}_0(\gamma_m) \tag{33}$$

with the aim of showing $\text{wo}_0(\gamma_{m+1})$ by induction on the recursive definition of $C(\sigma, \gamma_m)$. Assume $\text{wo}_0(\alpha)$ holds for every $\alpha \in C_k(\sigma, \gamma_m) \cap \Omega$ and suppose $\beta \in C_{k+1}(\sigma, \gamma_m) \cap \Omega$.

(C1). $\beta \leq \gamma_m$, so $\text{wo}_0(\beta)$ is a result of eq. (33).

(C2). $\beta =_{\text{NF}} \omega^\delta + \eta$, and $\delta, \eta \in C_k(\sigma, \gamma_m)$. Since also $\delta, \eta < \Omega$ the induction hypothesis yields $\text{wo}_0(\delta) \wedge \text{wo}_0(\eta)$ and so $\text{wo}_0(\beta)$.

(C3). $\beta =_{\text{NF}} \vartheta \xi$ and $\xi \in C_k(\sigma, \gamma_m) \cap \sigma$. Thus, $\xi^* \in C_k(\sigma, \gamma_m) \cap \Omega$ and therefore $\text{wo}_0(\xi^*)$ by the induction hypothesis. If $\xi < \Omega \cdot n$, $\text{wo}_0(\vartheta \xi)$ is a consequence of eq. (31), otherwise $\Omega \cdot n \leq \xi < \sigma$ and $\text{wo}_0(\vartheta \xi)$ is implied by eq. (32).

Thus we may deduce $\forall \alpha < \gamma_{m+1} \text{wo}_0(\alpha)$, hence $\text{wo}_0(\gamma_{m+1})$, and so $\text{wo}_0(\vartheta \sigma)$, concluding the proof of

$$F_0^n \vdash \text{Prog} F_0^{\Omega \cdot \bar{n}}. \quad (34) \quad \{\text{exteqn:4.2}\}$$

An application of T_0 -Intro entails $F_0^m \vdash T_0(\ulcorner \text{Prog} F_0^{\Omega \cdot \bar{n}} \urcorner)$, so

$$F_0^m \vdash \forall \alpha [\text{wo}_0(\alpha) \implies T_0(\ulcorner F_0(\Omega \cdot \bar{n} + \dot{\alpha}) \urcorner)],$$

and hence, by lemma 23,

$$F_0^m \vdash \forall \alpha (\text{wo}_0(\alpha) \implies \text{wo}_0 \vartheta (\Omega \cdot \bar{n} + \alpha)). \quad (35) \quad \{\text{exteqn:4.6}\}$$

To obtain $F_0^m \vdash F_0(\Omega \cdot \bar{m})$ and complete the proof we argue within F_0^m . Firstly, $\text{wo}_0((\Omega \cdot \bar{m})^*)$ holds trivially as $(\Omega \cdot \bar{m})^* = 0$. Secondly, if $\sigma < \Omega \cdot \bar{m}$, we have either $\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta \sigma)$ by eq. (31), or $\sigma = \Omega \cdot \bar{n} + \zeta$ for some $\zeta < \Omega$, whence $\text{wo}_0(\sigma^*) \implies \text{wo}_0(\zeta)$ and $\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta \sigma)$ results from eq. (35). \square

Corollary 25. $\|F_0^m\| \geq \vartheta(\Omega \cdot m)$ and $\|F_0\| \geq \vartheta(\Omega \cdot \omega)$.

Proof. Let $\gamma_0 = 1$ and $\gamma_{k+1} = C(\Omega \cdot m, \gamma_k) \cap \Omega$. Then $\vartheta(\Omega \cdot m) \leq \sup_{k < \omega} \gamma_k$ and $F_0^m \vdash \text{wo}_0(\gamma_0)$ holds. Moreover, if $F_0^m \vdash \text{wo}_0(\bar{\alpha})$ for every $\alpha < \gamma_k$ we may deduce $F_0^m \vdash \text{wo}_0(\bar{\alpha})$ for every $\alpha < \gamma_{k+1}$ by induction on the definition of γ_{k+1} thus: suppose $\beta \in C_{k+1}(\Omega \cdot m, \gamma_k) \cap \Omega$ and $F_0^m \vdash \text{wo}_0(\bar{\alpha})$ for every $\alpha \in C_k(\Omega \cdot m, \gamma_k) \cap \Omega$. If β was enumerated into $C_{k+1}(\Omega \cdot m, \gamma_k)$ by either (C1) or (C2), $F_0^m \vdash \text{wo}_0(\bar{\beta})$ is easily obtained from the induction hypothesis. If, however, $\beta = \vartheta \xi$ for some $\xi \in C_k(\Omega \cdot m, \gamma_k) \cap \Omega \cdot m$, $\xi = \Omega \cdot n + \alpha$ for some $n < m$, $\alpha \in C_k(\Omega \cdot m, \gamma_k) \cap \Omega$ and $F_0^m \vdash \text{wo}_0(\bar{\xi}^*)$ by the induction hypothesis, whence lemma 24 implies $F_0^m \vdash \text{wo}_0(\bar{\beta})$.

Since $\vartheta(\Omega \cdot m) \leq \sup_{k < \omega} \gamma_k$, we obtain $F_0^m \vdash \text{wo}_0(\bar{\alpha})$ for every $\alpha < \vartheta(\Omega \cdot m)$ and so $F_0^m \vdash T_1(< \vartheta(\Omega \cdot m))$ by an application of T_0 -Elim. \square

We will now extend the well-ordering proof above to determine lower bounds on the strength of each theory F_β . This will be done in stages, first for $\beta = 1$, then for arbitrary $\beta < \omega$ and finally for transfinite levels of the hierarchy. In doing so we will find ourselves migrating from the function $\alpha \mapsto \vartheta \alpha$ to the function $\alpha \mapsto \vartheta(\Omega \cdot \alpha)$, and eventually to functions $\alpha \mapsto \vartheta(\Omega^\beta \cdot \alpha)$.

Before proceeding directly with F_1 we require a slightly more general form of lemma 24. As its proof makes no explicit use of the fact m is finite, nor any application of T_0 -Intro in showing $\text{Prog} F_0^{\Omega \cdot \bar{m}}$ given $F_0(\Omega \cdot \bar{m})$, we may readily deduce the following generalisation.

Proposition 26. $F_0^1 \vdash \forall \rho [F_0(\rho) \implies \text{Prog} F_0^\rho]$.

Proof. Argue inside F_0^1 and assume $F_0(\rho)$ and $\forall v < \mu F_0^\rho(v)$, that is,

$$\text{wo}_0(\rho^*), \quad (36) \quad \{\text{exteqn:F03.1}\}$$

$$\forall \sigma < \rho [\text{wo}_0(\sigma^*) \implies \text{wo}_0 \vartheta \sigma], \quad (37) \quad \{\text{exteqn:F03.2}\}$$

$$\forall v < \mu (\text{wo}_0(v^*) \wedge \forall \sigma < \rho + v [\text{wo}_0(\sigma^*) \implies \text{wo}_0 \vartheta \sigma]), \quad (38) \quad \{\text{exteqn:F03.3}\}$$

for some $\mu < \Omega$; we want to prove $\text{wo}_0(\mu^*)$ and $\forall \tau < \rho + \mu[\text{wo}_0(\tau^*) \implies \text{wo}_0(\vartheta\tau)]$. The former holds immediately given eq. (38) so assume

$$\text{wo}_0(\tau^*), \quad (39)$$

for some $\tau < \rho + \mu$. We will prove $\text{wo}_0(\vartheta\tau)$.

Let $\gamma_0 = \tau^* + 1$ and $\gamma_{m+1} = C(\tau, \gamma_m) \cap \Omega$. That $\text{wo}_0(\gamma_m)$ holds for each m will be established by induction on m . From eq. (39) one has $\text{wo}_0(\gamma_0)$. Assume $\text{wo}_0(\gamma_m)$. In order to show $\text{wo}_0(\gamma_{m+1})$, assume $\text{wo}_0(\alpha)$ for every $\alpha \in C_k(\tau, \gamma_m) \cap \Omega$. Pick an arbitrary $\alpha_0 \in C_{k+1}(\tau, \gamma_m) \cap \Omega$. If α_0 was enumerated into the set by either (C1) or (C2), $\text{wo}_0(\alpha_0)$ is immediate. Otherwise $\alpha_0 = \vartheta\xi$ for some $\xi \in C_k(\tau, \gamma_m) \cap \tau$ and, since $\xi^* \in C_k(\tau, \gamma_m) \cap \Omega$, we have $\text{wo}_0(\xi^*)$. If $\xi < \rho$, eq. (37) provides $\text{wo}_0(\vartheta\xi)$. Otherwise $\tau = \rho + \nu$ for some $\nu < \mu$ and $\text{wo}_0(\vartheta\xi)$ holds due to eq. (38). Either way $\text{wo}_0(\alpha_0)$, and so $\text{wo}_0(\alpha)$ for every $\alpha < \gamma_{m+1}$, whence $\text{wo}_0(\gamma_{m+1})$.

Since $\vartheta\tau \geq \sup_m \gamma_m$ we obtain $\text{wo}_0(\vartheta\tau)$. \square

Proposition 26 plays a key role in the analysis of F_0^m and also F_1^1 . Lemma 23 entails

$$\begin{aligned} F_0^1 \vdash T_0(\ulcorner \text{Prog} F_0^{\dot{\rho}} \urcorner) &\implies \forall \alpha[\text{wo}_0(\alpha) \implies T_0(\ulcorner F_0^{\dot{\rho}}(\dot{\alpha}) \urcorner)] \\ &\implies \forall \alpha[\text{wo}_0(\alpha) \implies \text{wo}_0(\vartheta(\rho + \alpha))], \end{aligned}$$

so $F_1^1 \vdash \forall \rho[T_1(\ulcorner F_0(\dot{\rho}) \urcorner) \implies T_1(\ulcorner F_0(\dot{\rho} + \Omega) \urcorner)]$. This amounts to proving

$$F_1^1 \vdash \forall \rho[F_1(\rho) \implies \text{Prog} F_1^\rho] \quad (40)$$

where $F_1(\rho)$ is the formula $T_1(\ulcorner F_0(\rho) \urcorner)$ and $F_1^\rho(\alpha)$ denotes $\alpha < \Omega \implies F_1(\rho + \Omega \cdot \alpha)$. Equation (40) is sufficient to deduce a lower bound on the strength of the theory F_1^1 .

Corollary 27. $\|F_1^1\| \geq \vartheta\Omega^2$.

Proof. Since $F_0 \vdash F_0(\bar{0})$, eq. (40) implies

$$F_1^1 \vdash \text{Prog} F_1^{\bar{0}}. \quad (41)$$

Let $\sigma_0 = 1$ and $\sigma_{m+1} = \vartheta(\Omega \cdot \sigma_m)$. By proposition 20 (ii), $\vartheta\Omega^2 = \sup_m \sigma_m$, so it suffices to show $F_1^1 \vdash \text{wo}_1(\bar{\sigma}_m)$ for each m . This is trivial for $m = 0$; for $m = n + 1$ argue within F_1^1 assuming $\text{wo}_1(\bar{\sigma}_n)$. Then $\text{wo}_1(\bar{\sigma}_n + 1)$ and so $T_1(\ulcorner F_0(\Omega \cdot \bar{\sigma}_n) \urcorner)$ by an application of T_1 -Elim and eq. (41). Lemma 23 yields $T_1(\ulcorner \text{wo}_0(\vartheta(\Omega \cdot \bar{\sigma}_n)) \urcorner)$ and so $\text{wo}_1(\bar{\sigma}_m)$ holds. \square

Within F_1^2 , the above proof may be replicated under a T_1 predicate, allowing one to reach ordinals beyond $\vartheta\Omega^2$, as the next proposition demonstrates.

Lemma 28. For each m , $F_1^{m+1} \vdash F_1(\Omega^2 \cdot \bar{m})$ and $F_1^{m+1} \vdash \text{Prog} F_1^{\Omega^2 \cdot \bar{m}}$.

Proof. $F_1^1 \vdash F_1(\bar{0})$ holds trivially, so suppose $m = n + 1$ and $F_1^m \vdash F_1(\Omega^2 \cdot \bar{n})$. Equation (40) yields $F_1^m \vdash \text{Prog} F_1^{\Omega^2 \cdot \bar{n}}$, whence an application of T_1 -Intro and T_1^- -Rep implies

$$F_1^{m+1} \vdash T_1(\ulcorner \forall \alpha[\text{wo}_0(\alpha) \implies T_0(\ulcorner F_1(\Omega^2 \cdot \bar{n} + \Omega \cdot \dot{\alpha}) \urcorner)] \urcorner). \quad (42)$$

However, arguing within F_0 , from $F_0(\rho)$ one obtains $\text{wo}_0(\vartheta\rho)$, so eq. (42) entails

$$F_1^{m+1} \vdash T_1(\ulcorner \forall \alpha[\text{wo}_0(\alpha) \implies \text{wo}_0(\vartheta(\Omega^2 \cdot \bar{n} + \Omega \cdot \dot{\alpha})) \urcorner],$$

and thus $F_1^{m+1} \vdash F_1(\Omega^2 \cdot \bar{m})$ as required. \square

Theorem 9. Suppose $m < \omega$. Then every theorem of $\text{PA} + \text{TI}(<\vartheta(\Omega^2 \cdot m))$ is derivable in F_1^m . Moreover, every theorem of $\text{PA} + \text{TI}(<\vartheta(\Omega^2 \cdot \omega))$ is derivable in F_1 .

Proof. Since $\vartheta 0 = \varepsilon_0$ and F_1^0 extends PA, the case $m = 0$ holds, so suppose $m = n+1 > 0$. If $F_1^m \vdash \text{wo}_0(\bar{\alpha})$, lemma 28 implies

$$F_1^m \vdash \text{wo}_0 \vartheta(\Omega^2 \cdot \bar{n} + \Omega \cdot \bar{\alpha}),$$

whereby if $\sigma_0 = 1$ and $\sigma_{k+1} = \vartheta(\Omega^2 \cdot n + \Omega \cdot \sigma_k)$, $F_1^m \vdash \text{wo}_0(\bar{\sigma}_k)$ for every k .

Thus we require to show $\vartheta(\Omega^2 \cdot m) \leq \sup_k \sigma_k$. This is proved by induction on the rank of $\alpha < \vartheta(\Omega^2 \cdot m)$. If $\alpha =_{\text{NF}} \omega^\gamma + \delta < \vartheta(\Omega^2 \cdot m)$, the induction hypothesis immediately implies $\alpha < \sigma_k$ for some k . Otherwise

$$\alpha = \vartheta \xi < \vartheta(\Omega^2 \cdot m)$$

and there are two cases to consider:

1. $\xi < \Omega^2 \cdot m$ and $\xi^* < \vartheta(\Omega^2 \cdot m)$; or
2. $\xi > \Omega^2 \cdot m$ but $\vartheta \xi \leq (\Omega^2 \cdot m)^*$.

b) cannot hold since $(\Omega^2 \cdot m)^* = 0$, so $\xi^* < \vartheta(\Omega^2 \cdot m)$. As ξ^* has rank strictly smaller than α the induction hypothesis implies $\xi^* < \sigma_k$ for some k . But then $\xi < \Omega^2 \cdot n + \Omega \cdot \sigma_k$ and $\alpha < \sigma_{k+1}$.

The second part of the theorem is easily established using the fact $\vartheta(\Omega^2 \cdot \omega) = \sup_k \vartheta(\Omega^2 \cdot k)$. \square

We can now turn our attention to the theories F_p for $p < \omega$. Lemma 28 essentially shows $F_1 \vdash F_1(\Omega^2 \cdot \bar{\alpha})$ implies $F_1 \vdash F_1(\Omega^2 \cdot (\bar{\alpha} + 1))$. This can be extended to show $F_1 \vdash \forall v < \bar{\mu} F_1(\Omega^2 \cdot v)$ implies $F_1 \vdash F_1(\Omega^2 \cdot \bar{\mu})$, whence

$$F_2^1 \vdash \text{Prog} F_2^0 \tag{43}$$

where $F_2^0(\alpha)$ is the formula $\alpha < \Omega \wedge T_2(\ulcorner F_0(\rho + \Omega^2 \cdot \alpha) \urcorner)$.

Equation (43) suffices to deduce a lower bound for F_2^1 and acts as the base step in the analysis of F_2 and ultimately F_p , which follows a generalised form of the procedure used in lemma 28.

Let $F_p(\rho)$, for $0 < p < \omega$, be the formula $T_p(\ulcorner F_0(\rho) \urcorner)$, that is

$$T_p(\ulcorner \text{wo}_0(\rho^*) \wedge \forall \sigma < \dot{\rho}[\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta \sigma)] \urcorner),$$

and denote by $F_p^\rho(\alpha)$ the formula $\alpha < \Omega \wedge F_p(\rho + \Omega^{\bar{p}} \cdot \alpha)$.

Lemma 29. For each $p < \omega$, $F_p^1 \vdash \forall \rho[F_p(\rho) \implies \text{Prog} F_p^\rho]$ and, for $m < \omega$, $F_p^{m+1} \vdash F_p(\Omega^{\bar{p}+1} \cdot \bar{m})$.

Proof. The proof proceeds by induction on $p < \omega$. The case of $p = 0$ has already been shown in proposition 26 so suppose $p = q + 1 > 0$. Argue informally within F_p^1 . Assume

$$F_p(\rho), \tag{44}$$

$$\forall v < \mu F_p^\rho(v) \tag{45}$$

{exteqn:Fpwell11

{exteqn:Fpwell12

for some $\mu < \Omega$. If $\mu = 0$, of course $F_p^\rho(\mu)$ holds by eq. (44), and if μ is a limit ordinal, eq. (45) implies $T_p(\ulcorner \forall v < \dot{\mu} F_0(\rho + \Omega^{\bar{p}} \cdot \mu) \urcorner)$, whence $F_p^\rho(\mu)$ is immediate. This leaves only the case in which μ is a successor ordinal. But for every ordinal τ ,

$$\begin{aligned}
F_p(\tau) &\implies T_p(\ulcorner F_q(\dot{\tau}) \urcorner), \\
&\implies T_p(\ulcorner \text{Prog} F_q^{\dot{\tau}} \urcorner), \\
&\implies T_p(\ulcorner T_0(\ulcorner \text{Prog} F_q^{\dot{\tau}} \urcorner) \urcorner), \\
&\implies T_p(\ulcorner \forall \alpha [\text{wo}_0(\alpha) \implies T_0(\ulcorner F_q^{\dot{\tau}}(\dot{\alpha}) \urcorner)] \urcorner), \\
&\implies T_p(\ulcorner \forall \alpha [\text{wo}_0(\alpha) \implies T_0(\ulcorner F_0(\dot{\tau} + \Omega^{\bar{q}} \cdot \dot{\alpha}) \urcorner)] \urcorner), \\
&\implies T_p(\ulcorner \forall \alpha [\text{wo}_0(\alpha) \implies \text{wo}_0(\vartheta(\dot{\tau} + \Omega^{\bar{q}} \cdot \dot{\alpha})) \urcorner] \urcorner).
\end{aligned} \tag{46}$$

The second implication holds on account of the induction hypothesis, while the final holds due to T_0 -Del and lemma 23. Given that if $\tau \leq \sigma < \tau + \Omega^p$ there exists some $\zeta < \Omega$ such that $\sigma < \tau + \Omega^q \cdot \zeta$ and $\text{wo}_0(\sigma^*) \implies \text{wo}_0(\zeta)$ (pick $\zeta = (\sigma^*)^{\varepsilon^+}$), eq. (46) entails $F_p(\tau) \implies T_p(\ulcorner F_0(\dot{\tau} + \Omega^{\bar{p}}) \urcorner)$, that is $\forall \tau [F_p^\tau(\alpha) \implies F_p^\tau(\alpha + 1)]$.

For the second part, the case $m = 0$ is immediate, so suppose $m = n + 1$ and $F_p^m \vdash F_p(\Omega^{\bar{p}+1} \cdot \bar{n})$. Then $F_p^{m+1} \vdash T_p(\ulcorner \text{Prog} F_p^{\Omega^{\bar{p}+1} \cdot \bar{n}} \urcorner)$, from which we deduce

$$F_p^{m+1} \vdash T_p(\ulcorner \forall \alpha [\text{wo}_0(\alpha) \implies T_0(\ulcorner F_p(\Omega^{\bar{p}+1} \cdot \bar{n} + \Omega^{\bar{p}} \cdot \dot{\alpha}) \urcorner)] \urcorner)$$

and hence obtain $F_p^{m+1} \vdash F_p(\Omega^{\bar{p}+1} \cdot \bar{m})$. \square

Theorem 10. For every $m < \omega$ and $p < \omega$, $F_p^m \vdash \text{TI}(<\vartheta(\Omega^{p+1} \cdot m))$ and $F_p \vdash \text{TI}(<\vartheta(\Omega^{p+1} \cdot \omega))$.

Proof. For every p , the base case, $m = 0$, is immediate since F_p^0 extends PA formulated in the language \mathcal{L}_p . Otherwise $m = n + 1 > 0$ and the previous lemma shows $F_p^m \vdash \text{Prog} F_p^{\Omega^{\bar{p}+1} \cdot \bar{n}}$. Given $F_p^m \vdash \text{wo}_p(\bar{\alpha})$, one obtains $F_p^m \vdash F_p(\Omega^{\bar{p}+1} \cdot \bar{n} + \Omega^{\bar{p}} \cdot \bar{\alpha})$, and so

$$F_p^m \vdash \text{wo}_p(\vartheta(\Omega^{\bar{p}+1} \cdot \bar{n} + \Omega^{\bar{p}} \cdot \bar{\alpha})),$$

by unravelling the definition of F_p and lemma 23.

Let $\sigma_0 = 1$ and $\sigma_{k+1} = \vartheta(\Omega^{p+1} \cdot n + \Omega^p \cdot \sigma_k)$. The previous paragraph establishes $F_p^m \vdash \text{TI}(<\sigma_k)$ for every k , so all that remains is to show $\vartheta(\Omega^{p+1} \cdot m) \leq \sup_k \sigma_k$, which proceeds by induction on the rank of $\alpha < \vartheta(\Omega^{p+1} \cdot m)$. If $\alpha = 0$ we are done, and if $\alpha =_{\text{NF}} \omega^\gamma + \delta$, the induction hypothesis implies $\alpha < \sigma_k$ for some k . Thus, suppose

$$\alpha = \vartheta \xi < \vartheta(\Omega^{p+1} \cdot m)$$

for which there are two cases to consider.

1. $\xi < \Omega^{p+1} \cdot m$ and $\xi^* < \vartheta(\Omega^{p+1} \cdot m)$; or
2. $\xi > \Omega^{p+1} \cdot m$ but $\vartheta \xi \leq (\Omega^{p+1} \cdot m)^*$.

Since $(\Omega^{p+1} \cdot m)^* = 0$, b) is impossible, and so $\xi^* < \sigma_k$ for some k by the induction hypothesis. Then $\xi < \Omega^{p+1} \cdot n + \Omega^p \cdot \sigma_k$, whence $\alpha < \vartheta(\Omega^{p+1} \cdot n + \Omega^p \cdot \sigma_k) = \sigma_{k+1}$.

The second part of the theorem is an immediate consequence of the fact

$$\vartheta(\Omega^{p+1} \cdot \omega) = \sup_{k < \omega} \vartheta(\Omega^{p+1} \cdot k).$$

\square

Finally, we extend the well-ordering proofs to theories F_β for $\beta \geq \omega$. For $\beta = \omega$ this involves generalising the above proof so that one may derive

$$F_\omega \vdash \forall p < \omega \, T_\omega(\ulcorner \text{Prog} F_p^{\bar{0}} \urcorner), \quad (47)$$

whence $F_\omega \vdash \forall p < \omega \, T_\omega(\ulcorner \forall \alpha [\text{wo}_0(\alpha) \implies \text{wo}_0(\vartheta(\Omega^p \cdot \alpha))] \urcorner)$ and also

$$F_\omega \vdash \forall \sigma < \Omega^\omega [\text{wo}_\omega(\sigma^*) \implies \text{wo}_\omega(\vartheta\sigma)].$$

Equation (47) is not difficult to establish as the proof of lemma 29 is clearly uniform in $p < \omega$ and thus can be repeated under a T_ω predicate in F_ω . But in order to lift the construction to arbitrary $\beta \in \text{OT}_\Omega \cap \Omega$ we require a more general argument.

For each $\beta \in \text{OT}_\Omega$ let $G_\beta(\rho)$ denote the formula

$$T_\beta(\ulcorner F_0(\dot{\rho}) \urcorner),$$

and let $G_\beta^\rho(\alpha)$ abbreviate $\alpha < \Omega \wedge \forall \eta < \bar{\beta} \, G_\beta(\rho + \Omega^\eta \cdot \alpha)$.

Proposition 30. *For each $\beta \in \text{OT}_\Omega \cap \Omega$, $F_\beta^1 \vdash \forall \rho [G_\beta(\rho) \implies \text{Prog} G_\beta^\rho]$ and for every $m < \omega$, $F_\beta^{m+1} \vdash G_\beta(\Omega^{\bar{\beta}} \cdot \bar{m})$.*

The proof of proposition 30 is by transfinite induction on β and requires, for a given β , the following technical lemmata.

Lemma 31. $F_\beta^1 \vdash \forall \rho [G_\beta(\rho) \implies \forall \eta < \bar{\beta} \, T_\beta(\ulcorner \text{Prog} G_\eta^\rho \urcorner)]$.

Lemma 32. $F_\beta^1 \vdash T_\beta(\ulcorner \forall \tau \forall \eta < \bar{\beta} [F_0(\tau) \wedge T_0(\ulcorner \text{Prog} G_\eta^\tau \urcorner) \implies F_0(\tau + \Omega^\eta)] \urcorner)$.

Proof. The two lemmata hold vacuously when $\beta = 0$. For $\beta > 0$

$$F_\beta^1 \vdash G_\beta(\rho) \leftrightarrow \forall \eta < \bar{\beta} \, T_\beta(\ulcorner G_\eta(\dot{\rho}) \urcorner),$$

so the first lemma would result from replicating the proof of (the transfinite induction hypothesis of) proposition 30 under a T_β predicate. This is possible as the proof of the proposition, which is presented below, is uniform in $\eta < \bar{\beta}$.

In order to establish lemma 32, argue within F_β^1 under the scope of a T_β predicate. Fix $\eta < \bar{\beta}$, some arbitrary τ and assume

$$T_0(\ulcorner \text{Prog} G_\eta^\tau \urcorner), \quad (48)$$

$$F_0(\tau). \quad (49)$$

Equation (48) entails

$$\forall \alpha [\text{wo}_0(\alpha) \implies \forall \xi < \eta \, T_0(\ulcorner G_\eta(\dot{\tau} + \Omega^\xi \cdot \dot{\alpha}) \urcorner)],$$

so $\forall \alpha [\text{wo}_0(\alpha) \implies \forall \xi < \eta \, T_0(\ulcorner F_0(\dot{\tau} + \Omega^\xi \cdot \dot{\alpha}) \urcorner)]$ by T_0 -Del and

$$\forall \xi < \eta \, \forall \alpha [\text{wo}_0(\alpha) \implies \text{wo}_0(\vartheta(\tau + \Omega^\xi \cdot \alpha))] \quad (50)$$

by lemma 23. If $\tau \leq \sigma < \tau + \Omega^\eta$, there exists some $\zeta < \Omega$ and $\xi < \eta$ such that $\sigma < \tau + \Omega^\xi \cdot \zeta$ and $\text{wo}_0(\sigma^*) \implies \text{wo}_0(\zeta)$ (pick $\zeta = (\sigma^*)^{\varepsilon+}$), whence $\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta\sigma)$ results from eq. (50) as $\vartheta\sigma < \vartheta(\tau + \Omega^\xi \cdot \zeta)$. If, however, $\sigma < \tau$, eq. (49) implies $\text{wo}_0(\sigma^*) \implies \text{wo}_0(\vartheta\sigma)$. Thus, $F_0(\tau + \Omega^\eta)$ holds. \square

of proposition 30. Argue informally within F_β^1 . Fix some arbitrary ρ and assume $G_\beta(\rho)$. Lemma 31 yields $\forall \eta < \bar{\beta} \top_\beta(\ulcorner \text{Prog} G_\eta^\rho \urcorner)$ from which, using \top_β^- -Rep and lemma 32, one obtains $\forall \eta < \bar{\beta} \top_\beta(\ulcorner F_0(\rho + \Omega^\eta) \urcorner)$, that is, $\forall \eta < \bar{\beta} G_\beta(\rho + \Omega^\eta)$. Therefore

$$F_\beta^1 \vdash \forall \rho [G_\beta^\rho(\alpha) \implies G_\beta^\rho(\alpha + 1)]. \quad (51) \quad \text{\texttt{\{exteqn:Fbetaw\}}}$$

If μ is a limit ordinal, $\forall v < \mu G_\beta^\rho(v)$ and \top_β - \forall -Inf implies

$$\forall \eta < \bar{\beta} \top_\beta(\ulcorner \forall v < \dot{\mu} F_0(\dot{\rho} + \Omega^\eta \cdot v) \urcorner),$$

from which $\top_\beta(\ulcorner F_0(\dot{\rho} + \Omega^\eta \cdot \dot{\mu}) \urcorner)$ is easily deduced. Thus eq. (51) entails

$$F_\beta^1 \vdash \forall \rho [G_\beta(\rho) \implies \text{Prog} G_\beta^\rho].$$

The second half of proposition 30 is proved by induction on $m < \omega$. The case $m = 0$ holds since $F_0(\bar{0})$ is vacuously true. If $m = n + 1$ and $F_\beta^m \vdash G_\beta(\Omega^{\bar{\beta}} \cdot \bar{n})$,

$$F_\beta^{m+1} \vdash \top_\beta(\ulcorner \text{Prog} G_\beta^{\Omega^{\bar{\beta}} \cdot \bar{n}} \urcorner),$$

so $F_\beta^{m+1} \vdash G_\beta(\Omega^{\bar{\beta}} \cdot \bar{m})$ by an argument similar to lemma 32. □

Much as in the finite case, proposition 30 suffices to obtain lower bounds on the theories F_β^m . In theorem 11 below we prove $\|F_\beta\| \geq \vartheta(\Omega^\beta \cdot \omega)$. This is achieved by showing the proof-theoretic ordinal of F_β^m is closed under the operations $\alpha \mapsto \vartheta(\Omega^\eta \cdot \alpha)$ for every $\eta < \beta$ and extends the strength of F_β^n for $n < m$. The method will only work if β is not “too large”, that is $\vartheta \Omega^\beta$ is indeed the *least* ordinal closed under the above operations. By taking β to be no larger than the Veblen ordinal $\vartheta \Omega^\Omega$, we can ensure that this is the case.

Theorem 11. *For every $\beta < \vartheta \Omega^\Omega$ and every $m < \omega$, $F_\beta^m \vdash \text{TI}(< \vartheta(\Omega^\beta \cdot m))$ and $F_\beta \vdash \text{TI}(< \vartheta(\Omega^\beta \cdot \omega))$.*

Proof. For every $\beta < \vartheta \Omega^\Omega$, the base case, $m = 0$, is immediate since F_β^0 extends PA formulated in the language \mathcal{L}_κ . Otherwise $m = n + 1 > 0$ and proposition 30 implies $F_\beta^m \vdash \text{Prog} G_\beta^{\Omega^{\bar{\beta}} \cdot \bar{n}}$. Then, given $F_\beta^m \vdash \text{wo}_\beta(\bar{\alpha})$, we obtain $F_\beta^m \vdash \forall \eta < \bar{\beta} G_\beta(\Omega^{\bar{\beta}} \cdot \bar{n} + \Omega^\eta \cdot \bar{\alpha})$, and so

$$F_\beta^m \vdash \forall \eta < \bar{\beta} \text{wo}_\beta \vartheta(\Omega^{\bar{\beta}} \cdot \bar{n} + \Omega^\eta \cdot \bar{\alpha}),$$

by unravelling the definition of G_β .

Let $\sigma_0 = 1$ and $\sigma_{k+1} = \sup_{\eta < \beta} \vartheta(\Omega^\beta \cdot n + \Omega^\eta \cdot \sigma_k)$. The previous paragraph establishes $F_\beta^m \vdash \text{TI}(< \sigma_k)$ for every k . Therefore, all that remains is to show $\vartheta(\Omega^\beta \cdot m) \leq \sup_k \sigma_k$, which proceeds by induction on the rank of $\alpha < \vartheta(\Omega^\beta \cdot m)$. If $\alpha = 0$ we are done, and if $\alpha =_{\text{NF}} \omega^\gamma + \delta$, the induction hypothesis implies $\alpha < \sigma_k$ for some k . Thus, suppose

$$\alpha = \vartheta \xi < \vartheta(\Omega^\beta \cdot m).$$

There are two cases to consider.

1. $\xi < \Omega^\beta \cdot m$ and $\xi^* < \vartheta(\Omega^\beta \cdot m)$; or

2. $\xi > \Omega^\beta \cdot m$ but $\vartheta\xi \leq (\Omega^\beta \cdot m)^*$.

a) entails $\xi^* < \sigma_k$ for some k by the induction hypothesis (since ξ^* has rank strictly smaller than α). Then $\xi < \Omega^\beta \cdot m + \Omega^\eta \cdot \sigma_k$ for some $\eta < \beta$, whence $\alpha < \vartheta(\Omega^\beta \cdot n + \Omega^\eta \cdot \sigma_k) \leq \sigma_{k+1}$.

To manage b) one utilises $\beta < \vartheta(\Omega^\Omega)$. Since $\vartheta\xi \leq \beta^*$ and $(\Omega^\Omega)^* = 0$, we have

$$\vartheta\xi < \vartheta(\Omega^\Omega)$$

and $\xi < \Omega^\Omega$. But then $\beta \leq \xi^* < \Omega$ contradicting $\vartheta\xi \leq \beta^*$.

The lower bound on F_β is an immediate consequence of the fact

$$\vartheta(\Omega^\beta \cdot \omega) = \sup_{k < \omega} \vartheta(\Omega^\beta \cdot k).$$

□

4.4 Upper bounds on the proof-theoretic ordinal

We now move to the task of determining an upper bound on the strength of the theory F_β . The ultimate aim is to demonstrate that any arithmetical theorem of F_β is derivable in the theory $PA + TI(<\delta)$ for a suitable ordinal δ . To achieve this it is natural to attempt to formalise the consistency proof provided in section 4.1. For it to be successful, however, we must show that the arguments involved can be carried out within the confines of $PA + TI(<\delta)$. The main difficulty is with formalising the T-Elimination theorem, theorem 7, since:

- The proof makes free use of transfinite induction beyond the first uncountable ordinal.
- Moreover, each subsidiary induction appears to also require transfinite induction beyond δ (albeit this time bounded in Ω) in order to support later steps in the main induction.
- In particular, the case in which the last rule employed was T_η - \forall -Inf makes use of non-finitistic properties of Ω ; namely proposition 14.

By introducing bounds on the height of the derivations in \mathcal{F}_ρ^* (for a fixed ordinal $\rho \in OT_\Omega$), we observe that δ can be chosen so as to bound all uses of the subsidiary transfinite induction hypothesis described in the second point. Moreover, δ can be chosen in such a way that all references to the main induction hypothesis in theorem 7 are for ordinals of the form $\Omega^{\alpha_0} \cdot \beta_0 + \Omega^{\alpha_1} \cdot \beta_1 + \dots + \Omega^{\alpha_n} \cdot \beta_n$ where $\alpha_i, \beta_i < \delta$ for every $i \leq n$; the upshot being that a formalised version of theorem 7 needs only countable induction on some suitable (lexicographical-style) ordering.

The third point, however, has no immediate alternative; we need the regularity of Ω to ensure the set $\mathcal{F}_{<\rho}^*$ is closed under the ω -rule if ρ is not an Ω -limit, a property required to establish that \mathcal{F}_ρ^* is closed under T_η -Elim. Suppose $\rho = \rho' + \Omega^{\xi+1}$ is not an Ω -limit. If $A(x)$ is a formula of \mathcal{L}_κ and $\mathcal{F}_{\rho_n}^* \vdash A(\bar{n})$ with $\rho_n < \rho$ for every n , we want to obtain an ordinal $\sigma < \rho$ such that $\mathcal{F}_\sigma^* \vdash \forall x A(x)$. As each application of T_η -Intro must increase the height of a derivation, introducing bounds on the height of a derivation indirectly yields bounds on its T-rank. That is, if $\mathcal{F}_\tau^* \vdash B$ with height α , one can prove the existence of an ordinal $\tau_0 \leq \tau$ such that $\mathcal{F}_{\tau_0}^* \vdash B$ and τ_0 is no larger than the maximum of α and all η such that T_η -Intro was utilised in the derivation $\mathcal{F}_\tau^* \vdash B$. Thus, if α is an ordinal bounding the height of the derivation $\mathcal{F}_{\rho_n}^* \vdash A(\bar{n})$ for every n , we may observe the ordinal ρ_n is “bounded” by α

for every n , and so effectively find an ordinal $\gamma < \Omega$ depending on α and ρ , so that if $\sigma = \rho' + \Omega^\xi \cdot \gamma$, $\rho_n < \sigma$ for every n and, hence, $\mathcal{F}_\sigma^* \vdash \forall x A(x)$.

Although at present these arguments are somewhat vague, they should explain the motivation behind a number of the technical results to follow. In particular, the remarks we made in relation to the third point are made explicit in the Condensation lemma (lemma 35).

It will be important to compare ordinals greater than Ω by their construction as well as their size. Therefore, we introduce a new ordering $<^*$, defined on OT_Ω by

$$\rho <^* \sigma \leftrightarrow (\rho < \sigma \wedge \rho^* \leq \sigma^*).$$

$\rho \leq^* \sigma$ abbreviates $\rho <^* \sigma \vee \rho = \sigma$.

Definition 33. Fix $\kappa = \vartheta_{\varepsilon_{\Omega+1}}$. Define $\mathcal{T}_\infty \frac{\alpha}{\rho} \Gamma$ for $\alpha, \rho \in \text{OT}_\Omega$ with $\alpha < \Omega$ and Γ an \mathcal{L}_κ -sequent, according to the following rules: (Ax.1), (\wedge) , (\vee) , (ω) and (\exists) as in ??, along with the following for each $\eta < \kappa$ and each $\beta, \gamma < \alpha \in \text{OT}_\Omega \cap \Omega$.

$$(\text{Ax.2}_\eta) \frac{\alpha}{\rho} \Gamma, \neg T_\eta(s_0), T_\eta(s_1) \text{ if } s_0^{\mathbb{N}} = s_1^{\mathbb{N}}.$$

$$(\text{Ax.3}_\eta) \frac{\alpha}{\rho} \Gamma, \neg T_\eta(s) \text{ if } s^{\mathbb{N}} \text{ is not the Gödel number of an } \mathcal{L}_\kappa\text{-sentence.}$$

$$(\text{T}_\eta\text{-Imp}) \frac{\beta}{\rho} \Gamma, T_\eta(s_0), \frac{\gamma}{\rho} \Gamma, T_\eta(s_1) \text{ and } s_0^{\mathbb{N}} = (s_1 \rightarrow s_2)^{\mathbb{N}} \text{ implies } \frac{\alpha}{\rho} \Gamma, T_\eta(s_2).$$

$$(\text{T}_\eta\text{-Intro}) \frac{\alpha}{\rho} \Gamma, T_\eta(s) \text{ whenever } \frac{\beta}{\sigma} A, \text{ with } \sigma \# \Omega^{\eta+1} \leq^* \rho \text{ and } s^{\mathbb{N}} = \ulcorner A \urcorner.$$

$$\text{extinf:dag1} \quad (\text{T}_\eta\text{-Del}) \frac{\beta}{\rho} \Gamma, T_\eta(s), s^{\mathbb{N}} = \ulcorner T_\xi(s_0) \urcorner \text{ and } \xi < \kappa \text{ implies } \frac{\alpha}{\rho} \Gamma, T_\eta(s_0).$$

$$(\neg T_\eta\text{-Del}) \frac{\beta}{\rho} \Gamma, T_\eta(s), s^{\mathbb{N}} = \ulcorner \neg T_\xi(s_0) \urcorner \text{ and } \xi < \kappa \text{ implies } \frac{\alpha}{\rho} \Gamma, T_\eta(s_1), \text{ if } s_1^{\mathbb{N}} = (\neg s_0)^{\mathbb{N}}.$$

$$\text{extinf:dag2} \quad (\text{T}_\eta^-\text{-Rep}) \frac{\beta}{\rho} \Gamma, T_\eta(s), \xi < \eta \text{ implies } \frac{\alpha}{\rho} \Gamma, T_\eta(s_0) \text{ if } s^{\mathbb{N}} = \ulcorner T_\xi(s_0) \urcorner \text{ and } s_0^{\mathbb{N}} \text{ is the Gödel number of an } \mathcal{L}_\kappa\text{-sentence.}$$

$$(\text{T}_\eta\text{-}\omega) \frac{\beta}{\rho} \Gamma, T_\eta(\ulcorner A(\bar{n}) \urcorner) \text{ for every } n \in \mathbb{N}, \text{ where } A(x) \text{ is a formula of } \mathcal{L}_\kappa \text{ with at most } x \text{ free, implies } \frac{\alpha}{\rho} \Gamma, T_\eta(s) \text{ whenever } s^{\mathbb{N}} = \ulcorner \forall x A(x) \urcorner \text{ and } \rho|_\eta \text{ is not an } \Omega\text{-limit.}$$

$$(\text{Weakening}) \frac{\beta}{\sigma} \Gamma \text{ implies } \frac{\alpha}{\rho} \Gamma, \Delta \text{ whenever } \sigma \leq^* \rho \text{ and } \Delta \text{ is an } \mathcal{L}_\Gamma\text{-sequent.}$$

We refer to ρ as the *T-rank* and α as the *height* of the derivation $\mathcal{T}_\infty \frac{\alpha}{\rho} \Gamma$. For a fixed η , the rules $(\text{T}_\eta\text{-Imp})$ to $(\text{T}_\eta\text{-}\omega)$, are collectively referred to as *T_η-rules*. The collection as a whole (for any η) is referred to as the *T-rules*. In each application of a rule or axiom above we specify the *active*, *minor* and *side* formulae as usual.

Often when working with hierarchical systems one needs an extended notion of the rank of a formula. In this case, however, the basic definition of rank suffices and as such our work in \mathcal{T}_∞ is greatly simplified. Define the *rank* of an \mathcal{L}_κ -sentence A , denoted $|A|$, as in ?? with the addition that $|T_\eta(s)| = |\neg T_\eta(s)| = 0$ for every $\eta < \kappa$.

We observe the following straightforward results for \mathcal{T}_∞ and, henceforth make free use of them without mention.

Proposition 34.

1. $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, A$ implies $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma$ if A is a false arithmetical literal.
2. $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, \mathsf{T}_\eta(s)$ implies $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma$ if $s^\mathbb{N}$ is not the code of an \mathcal{L}_κ -sentence.
3. $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, \mathsf{T}_\eta(s), \mathsf{T}_\eta(t)$ and $s^\mathbb{N} = t^\mathbb{N}$ implies $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, \mathsf{T}_\eta(s)$.
4. $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, \neg \mathsf{T}_\eta(s), \neg \mathsf{T}_\eta(t)$ and $s^\mathbb{N} = t^\mathbb{N}$ implies $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, \neg \mathsf{T}_\eta(s)$.
5. $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, A(t)$ and $s^\mathbb{N} = t^\mathbb{N}$ implies $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, A(s)$.

Proof. The proof is essentially identical to the proof of ??, and hence omitted. \square

Although we did not formulate \mathcal{T}_∞ with a cut rule, it is easy to see that such a rule is admissible. The proof proceeds in the same manner as the cut elimination theorems of ?? (cf. ??????).

Theorem 12 (Admissibility of cut). *For every $\alpha, \beta < \Omega$ and every $\sigma \in \text{OT}_\Omega$, $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma, A$ and $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Delta, \neg A$ imply $\mathcal{T}_\infty \frac{\beta}{\sigma} \Gamma, \Delta$ for some $\beta < \alpha^{\varepsilon^+}$.*⁹

Proof. By the definition of \mathcal{T}_∞ , $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Delta, \neg \mathsf{T}_\eta(s)$ holds with $\neg \mathsf{T}_\eta(s)$ active only if $\Delta, \neg \mathsf{T}_\eta(s)$ is an instance of either (Ax. $_{2\eta}$) or (Ax. $_{3\eta}$). If further, $\mathcal{T}_\infty \frac{\beta}{\sigma} \Gamma, \mathsf{T}_\eta(s)$ holds, it is easy to deduce in fact $\mathcal{T}_\infty \frac{\alpha \# \beta}{\sigma} \Gamma, \Delta$. Thus the claim holds when A is either $\mathsf{T}_\eta(s)$ or $\neg \mathsf{T}_\eta(s)$ for some $\eta < \kappa$ and term s .

The other cases are proved by induction $|A|$ and their arguments are identical to those used in the Reduction Lemma for \mathcal{T}_∞ , ??. \square

The remainder of this section is concerned with proving the equivalent form of the T-Elimination theorem for \mathcal{T}_ρ^* , theorem 7, and showing the argument can be formalised within the theory $\text{PA} + \text{TI}(<\delta)$ for some suitable δ . We begin with the Condensation lemma, replacing the use of proposition 14.

Lemma 35 (Condensation lemma). *Let $\Theta^*(\sigma)$ denote the least ordinal η^* such that $\Omega^{\eta^+1} > \sigma$. If $\mathcal{T}_\infty \frac{\alpha}{\sigma} \Gamma$ holds for some α, σ and Γ , there is a $\rho \leq \sigma$ with $\rho^* \leq \max\{\alpha^*, \Theta^*(\sigma)\}$ such that $\mathcal{T}_\infty \frac{\alpha}{\rho} \Gamma$.*

Proof. We proceed by transfinite induction on the height of the derivation $\frac{\alpha}{\sigma} \Gamma$. In the case of axioms one may always take $\rho = 0$ and for any rule other than $(\mathsf{T}_\eta\text{-}\omega)$ or $(\mathsf{T}_\eta\text{-Intro})$ one may apply the induction hypothesis and leave ρ unchanged. In the case $(\mathsf{T}_\eta\text{-Intro})$ was the last rule applied one obtains a formula $\mathsf{T}_\eta(\ulcorner A \urcorner)$ in Γ and a derivation $\frac{\alpha_0}{\tau} A$ for some $\alpha_0 < \alpha$ with $\tau \# \Omega^{\eta^+1} \leq^* \sigma$. The induction hypothesis yields an ordinal $\tau_0 \leq \tau$ fulfilling $\tau_0^* \leq \max\{\alpha_0^*, \Theta^*(\tau)\}$ such that $\frac{\alpha_0}{\tau_0} A$, whence we observe taking $\rho = \tau_0 \# \Omega^{\eta^+1}$ suffices, since

$$\begin{aligned} \rho^* &\leq \max\{\tau_0^*, \eta^*\} \\ &\leq \max\{\alpha_0^*, \Theta^*(\tau), \eta^*\} \\ &\leq \max\{\alpha^*, \Theta^*(\sigma)\}. \end{aligned}$$

In the case of $(\mathsf{T}_\eta\text{-}\omega)$ suppose $\sigma|_\eta$ is not an Ω -limit, Γ contains the sentence $\mathsf{T}_\eta(\ulcorner \forall x A(x) \urcorner)$ and there exists $\alpha_0 < \alpha$ such that $\frac{\alpha_0}{\sigma} \Gamma, \mathsf{T}_\eta(\ulcorner A(\bar{n}) \urcorner)$ for every n . The induction hypothesis yields a $\tau_0 \leq \sigma$ such that $\frac{\alpha_0}{\tau_0} \Gamma, \mathsf{T}_\eta(\ulcorner A(\bar{n}) \urcorner)$ with $\tau_0^* \leq \max\{\alpha_0^*, \Theta^*(\sigma)\}$. If τ_0 is not an Ω -limit take $\rho = \tau_0$, otherwise pick $\rho = \tau_0 + \Omega^\xi$, where ξ is such that $\sigma|_\eta = \sigma_0 + \Omega^\xi$ for some $\sigma_0 < \sigma$. In either case $\rho \leq \sigma$ and

$$\rho^* = \max\{\tau_0^*, \xi^*\} \leq \max\{\alpha^*, \Theta^*(\sigma)\}.$$

\square

⁹Recall α^{ε^+} is the least ε -ordinal greater than α .

We will now define a collection of standard \mathcal{L}_κ -structures \mathfrak{M}_ρ indexed by ordinals $\alpha < \Omega$ and $\rho < \Omega^\kappa$. These play the role of the structures \mathfrak{M}_ρ used in theorem 7. Set

$$\mathfrak{M}_\rho \models T_\eta(s) \text{ iff there is an } \mathcal{L}_\kappa\text{-sentence } A \text{ with } s^\mathbb{N} = \ulcorner A \urcorner \text{ and ordinals } \beta, \sigma \text{ such that } \sigma^*, \beta < \vartheta\rho, \sigma < \rho|_\eta \text{ and } \mathcal{T}_\infty \upharpoonright_\sigma^\beta A.$$

Proposition 36. *If Γ is a T-positive sequent and $\rho \leq^* \sigma$, $\mathfrak{M}_\rho \models \bigvee \Gamma$ implies $\mathfrak{M}_\sigma \models \bigvee \Gamma$.*

Proof. Suppose $\rho \leq^* \sigma$. Since $\rho^* < \vartheta\sigma$, we have $\vartheta\rho \leq \vartheta\sigma$ by proposition 17 (viii), which coupled with the fact $\rho|_\eta \leq \sigma|_\eta$ entails $\mathfrak{M}_\rho \models T_\eta(s)$ implies $\mathfrak{M}_\sigma \models T_\eta(s)$ for every $\eta < \kappa$. The remaining cases are now easily verified. \square

The next proposition will enable us to pass from derivations in \mathcal{T}_∞ to satisfaction in the structures \mathfrak{M}_ρ for T-positive sequents.

Lemma 37. *The following hold for every $\alpha < \Omega$, $\eta < \kappa$ and $\rho < \Omega^\kappa$.*

1. $\mathfrak{M}_\rho \models \forall x \forall y (T_\eta(x) \wedge T_\eta(x \rightarrow y) \implies T_\eta(y))$.
2. $\mathfrak{M}_\rho \models \forall \xi < \bar{\eta} \forall x (T_\eta(x) \implies T_\eta(\ulcorner T_\xi(\dot{x}) \urcorner))$.
3. $\mathfrak{M}_\rho \models \forall \xi < \bar{\kappa} \forall x (T_\eta(\ulcorner T_\xi(\dot{x}) \urcorner) \implies T_\eta(x))$ if, for every $\sigma < \rho$ and β such that $\sigma^*, \beta < \vartheta\rho$, $\mathcal{T}_\infty \upharpoonright_\sigma^\beta \Gamma$ implies $\mathfrak{M}_{\sigma+\beta} \models \bigvee \Gamma$.
4. $\mathfrak{M}_\rho \models \forall \xi < \bar{\kappa} \forall \ulcorner A \urcorner (T_\eta(\ulcorner \neg T_\xi(\ulcorner A \urcorner) \urcorner) \implies T_\eta(\ulcorner \neg A \urcorner))$.
5. If $\rho|_\eta$ is not an Ω -limit, $\mathfrak{M}_\rho \models \forall x T_\eta(\ulcorner A(\dot{x}) \urcorner)$ and $\rho <^* \rho'$, $\mathfrak{M}_{\rho'} \models T_\eta(\ulcorner \forall x A(x) \urcorner)$.

Proof. 1. We observe that if $\mathfrak{M}_\rho \models T_\eta(\ulcorner A \urcorner) \wedge T_\eta(\ulcorner A \implies B \urcorner)$, there is an $\alpha < \vartheta\rho$ and $\sigma < \rho|_\eta$ such that

$$\mathcal{T}_\infty \upharpoonright_\sigma^\alpha A \text{ and } \mathcal{T}_\infty \upharpoonright_\sigma^\alpha \neg A, B$$

holds and $\sigma^* < \vartheta\rho$. Admissibility of the cut rule, theorem 12, yields $\mathcal{T}_\infty \upharpoonright_\sigma^\gamma B$ for some $\gamma < \alpha^{\varepsilon^+}$. But $\rho > \Omega$, so $\gamma < \vartheta\alpha < \vartheta\rho$ and hence $\mathfrak{M}_\rho \models T_\eta(\ulcorner B \urcorner)$ by definition.

2. Fix some $\eta < \kappa$. Note $(\rho|_\eta)^* < \vartheta\rho$ for every η , since $\rho^* < \vartheta\rho$. Also $\rho < \Omega^\kappa$, so $\eta^* \leq (\rho|_\eta)^*$. If $\rho|_\eta = 0$, 2 holds vacuously. Otherwise, suppose $\mathcal{T}_\infty \upharpoonright_\sigma^\beta A$ for some \mathcal{L}_κ -sentence A , some β and $\sigma < \rho|_\eta$ such that $\beta, \sigma^* < \vartheta\rho$. Then $\sigma \# \Omega^{\xi+1} < \rho|_\eta$ for every $\xi < \eta$, and $(\sigma \# \Omega^{\xi+1})^* \leq \max\{\sigma^*, \eta^*\} < \vartheta\rho$. So $\mathfrak{M}_\rho \models \forall \xi < \bar{\eta} T_\eta(\ulcorner T_\xi(\ulcorner A \urcorner) \urcorner)$ as required.

3. Suppose $\mathfrak{M}_\rho \models T_\eta(\ulcorner T_\xi(s) \urcorner)$ and $\xi < \kappa$; we want to show $\mathfrak{M}_\rho \models T_\eta(s)$. By definition there are ordinals α and $\sigma < \rho|_\eta$ such that $\sigma^*, \alpha < \vartheta\rho$ and

$$\mathcal{T}_\infty \upharpoonright_\sigma^\alpha T_\xi(s).$$

The additional assumption yields $\mathfrak{M}_{\sigma+\alpha} \models T_\xi(s)$, whence $s^\mathbb{N} = \ulcorner A \urcorner$ for some A and there is a β and $\tau < \sigma|_\xi$ such that $\tau^*, \beta < \vartheta(\sigma + \alpha)$ and

$$\mathcal{T}_\infty \upharpoonright_\tau^\beta A.$$

Since $\sigma^*, \alpha < \vartheta\rho$ and $\sigma + \alpha < \rho$,

$$\vartheta(\sigma + \alpha) < \vartheta\rho,$$

and thus $\tau^*, \beta < \vartheta\rho$. Moreover, $\tau < \rho|_\eta$ so $\mathfrak{M}_\rho \models \mathsf{T}_\eta(s)$ follows by definition.

4 holds vacuously as $\mathcal{T}_\infty|_\sigma^\beta \neg \mathsf{T}_\xi(s)$ cannot hold for any β, σ and term s .

5. Suppose $\mathfrak{M}_\rho \models \forall x \mathsf{T}_\eta(\ulcorner A(x) \urcorner)$, $\rho < \rho'$ and $\rho|_\eta$ is not an Ω -limit. By definition, $\rho|_\eta > 0$ and one obtains ordinals α_n and τ_n for each $n \in \mathbb{N}$ such that

$$\begin{aligned} \tau_n &< \rho|_\eta, \\ \alpha_n, \tau_n^* &< \vartheta\rho, \\ \mathcal{T}_\infty|_{\tau_n}^{\alpha_n} A(\bar{n}). \end{aligned}$$

Our aim is to find ordinals σ, γ such that

$$\sigma < \rho'|_\eta, \tag{52}$$

$$\gamma, \sigma^* < \vartheta\rho', \tag{53}$$

$$\mathcal{T}_\infty|_\sigma^{\gamma+1} \forall x A(x). \tag{54}$$

As $\rho|_\eta$ is not an Ω -limit there exists $\xi, \rho_0 < \rho$ such that $\rho|_\eta = \rho_0 + \Omega^{\xi+1}$. Moreover, the Condensation lemma 35 provides ordinals $\sigma_n \leq \tau_n$ such that

$$\sigma_n^* \leq \max\{\alpha_n^*, \Theta^*(\tau_n)\}, \tag{55}$$

$$\mathcal{T}_\infty|_{\sigma_n}^{\alpha_n} A(\bar{n}), \tag{56}$$

where $\Theta^*(\tau_n)$ is defined as in the statement of lemma 35. Let $\gamma = \vartheta\rho$ and $\sigma = \rho_0 + \Omega^\xi \cdot \vartheta\rho$. Then eq. (52) clearly holds; moreover,

$$\begin{aligned} \sigma^* &\leq \max\{\rho_0^*, \xi^*, \vartheta\rho\}, \\ &= \vartheta\rho, \end{aligned} \tag{57}$$

so $\sigma^* = \vartheta\rho < \vartheta\rho'$. Therefore eq. (53) also holds.

To establish eq. (54) and complete the proof, we will show $\sigma_n \leq^* \sigma$ and apply (Weakening) to eq. (56). Fix an arbitrary $n \in \mathbb{N}$. By eq. (55),

$$\begin{aligned} \sigma_n^* &\leq \max\{\alpha_n^*, \Theta^*(\tau_n)\}, \\ &\leq \max\{\alpha_n^*, \tau_n^*\}, \\ &< \vartheta\rho. \end{aligned}$$

Moreover, $\sigma_n < \sigma$ because either $\sigma_n < \rho_0 < \sigma$ or there is some $\gamma_0 < \Omega$ such that

$$\rho_0 + \Omega^\xi \cdot \gamma_0 \leq \sigma_n < \rho_0 + \Omega^\xi \cdot (\gamma_0 + 1),$$

whence $\gamma_0^* \leq \sigma_n^* < \vartheta\rho$, so $\gamma_0 < \vartheta\rho$ and $\sigma_n < \sigma$. Hence $\mathfrak{M}_{\rho'} \models \mathsf{T}_\eta(\ulcorner \forall x A(x) \urcorner)$ as desired. \square

Theorem 13 (T-Elimination theorem). *For every $\rho \in \mathsf{OT}_\Omega$ and $\alpha \in \mathsf{OT}_\Omega \cap \Omega$,*

1. $\mathcal{T}_\infty|_\rho^\alpha \Gamma$ implies $\mathfrak{M}_{\rho+\alpha} \models \bigvee \Gamma$ if Γ is a T -positive sequent;

2. $\mathcal{T}_\infty \frac{\alpha}{\rho} \top_\eta(s)$ implies there is an \mathcal{L}_\top -sentence A with $s^\mathbb{N} = \ulcorner A \urcorner$ such that $\mathcal{T}_\infty \frac{\beta}{\sigma} A$ for some $\beta, \sigma < \rho|_\eta$ with $\sigma^*, \beta < \vartheta(\rho + \alpha)$;
3. $\mathcal{T}_\infty \frac{\alpha}{\rho} \neg \top_\eta(s)$ implies there is an \mathcal{L}_\top -sentence A with $s^\mathbb{N} = \ulcorner A \urcorner$ such that $\mathcal{T}_\infty \frac{\alpha}{\rho} \neg A$.

Proof. By main transfinite induction on ρ with a subsidiary transfinite induction on α . We begin with (i).

First observe that since \mathfrak{M}_ρ is standard all arithmetical rules and axioms are dealt with easily. Proposition 36 deals with (Weakening). Moreover, no instances of (Ax.2 $_\eta$) and (Ax.3 $_\eta$) are T-positive, so Γ is not an instance of these axioms. Thus we need only consider the T-rules. Suppose $\mathcal{T}_\infty \frac{\alpha}{\rho} \Gamma$ is derived via a \top_η -rule. If the last rule applied was (\top_η -Intro), there is an \mathcal{L}_κ -sentence A and term s with $s^\mathbb{N} = \ulcorner A \urcorner$ so that $\top_\eta(s)$ is in Γ , and $\mathcal{T}_\infty \frac{\beta}{\sigma} A$ holds for some $\beta < \alpha$ and σ such that $\sigma \# \Omega^{\eta+1} \leq^* \rho$. But then $\sigma < \rho|_\eta$ and $\beta, \sigma^* < \vartheta\rho$, so $\mathfrak{M}_{\rho+\alpha} \models \bigvee \Gamma$.

For the remaining rules, lemma 37 entails $\mathfrak{M}_{\rho+\alpha} \models \bigvee \Gamma$ is a consequence of the subsidiary induction hypothesis and, in the case of (\top_η -Del), the main induction hypothesis.

(ii) follows directly as a result of (i) and the definition of $\mathfrak{M}_{\rho+\alpha}$.

It is easy to check that $\mathcal{T}_\infty \frac{\alpha}{\rho} \neg \top_\eta(s)$ can only hold if $s^\mathbb{N}$ is not the code of some \mathcal{L}_κ -sentence, whence (iii) holds vacuously. \square

Having established that \mathcal{T}_∞ is also closed under \top_η -Elim for every $\eta < \kappa$ we can proceed with embedding each F_η into \mathcal{T}_∞ . This is the focus of the following proposition and theorem.

Proposition 38. For each axiom A of F_β , $\mathcal{T}_\infty \frac{\alpha}{\Omega^{\beta+1} \cdot n} A$, for some $\alpha < \varepsilon_0$ and every $n < \omega$.

Lemma 39 (Embedding lemma for \mathcal{T}_∞). Fix some $\beta < \vartheta\varepsilon_{\Omega+1}$ and suppose $F_\beta \vdash A$ with A a sentence of \mathcal{L}_β . Then there are ordinals α and ρ such that $\mathcal{T}_\infty \frac{\alpha}{\rho} A$, $\rho < \Omega^{\beta+1} \cdot \omega$, $\alpha, \rho^* < \vartheta(\Omega^{\beta+1} \cdot \omega)$ and $\rho = \rho|_0$.

Proof. We proceed by induction on the length of the derivation $F_\beta \vdash A$.

Case I. $\beta = 0$. This essentially follows from the embedding of F in \mathcal{T}_∞ , ??, since $\vartheta(\Omega \cdot \omega) = \varphi\omega 0$. Proposition 38 deals with the axioms of F_0 . According to the induction hypothesis, an application of *modus ponens* in F_0 means one has

$$\mathcal{T}_\infty \frac{\alpha_0}{\Omega \cdot m_0} B, \quad \mathcal{T}_\infty \frac{\alpha_1}{\Omega \cdot m_1} \neg B, A$$

for some sentence B , ordinals $\alpha_0, \alpha_1 < \vartheta(\Omega \cdot (m+1))$, and $m_0, m_1 < \omega$. Weakening implies

$$\mathcal{T}_\infty \frac{\gamma}{\Omega \cdot m} B, \quad \mathcal{T}_\infty \frac{\gamma}{\Omega \cdot m} \neg B, A$$

where $\gamma = \max\{\alpha_0, \alpha_1\}$ and $m = \max\{m_0, m_1\}$, whence $\mathcal{T}_\infty \frac{\alpha}{\Omega \cdot m} A$ holds for some $\alpha < \gamma^{\varepsilon+}$. Since $\gamma < \vartheta(\Omega \cdot \omega) < \Omega$, $\vartheta\gamma < \vartheta(\Omega \cdot \omega)$ and so $\alpha < \vartheta(\Omega \cdot \omega)$ whence we are done.

An application of \top_0 -Intro in F_0 corresponds to an application of the rule (\top_0 -Intro) in \mathcal{T}_∞ . It is an immediate consequence of the induction hypothesis that in this case the desired bounds are maintained.

This leaves only applications of \top_0 -Elim and $\neg\top_0$ -Elim in F_0 to consider, for which we employ theorem 13. $\neg\top_0$ -Elim is immediate, so suppose $F_0 \vdash \top_0(\ulcorner A \urcorner)$ for some \mathcal{L}_κ -sentence A . The induction hypothesis implies

$$\mathcal{T}_\infty \frac{\alpha}{\Omega \cdot m} \top_0(\ulcorner A \urcorner) \tag{58}$$

for some $\alpha < \vartheta(\Omega \cdot \omega)$ and $m < \omega$. By theorem 13 one obtains some $\gamma < \vartheta(\Omega \cdot m + \alpha)$ and $n < m$ such that

$$\mathcal{T}_\infty \upharpoonright_{\Omega \cdot n}^\gamma A$$

holds. But $\Omega \cdot m + \alpha < \Omega \cdot \omega$ and $\alpha < \vartheta(\Omega \cdot \omega)$, so $\gamma < \vartheta(\Omega \cdot \omega)$.

Case II. $\beta > 0$. Proposition 38 deals with the axioms of F_β . According to the induction hypothesis, an application of *modus ponens* in F_β means one has

$$\mathcal{T}_\infty \upharpoonright_\sigma^{\alpha_0} B, \quad \mathcal{T}_\infty \upharpoonright_\tau^{\alpha_1} \neg B, A$$

for some sentence B , ordinals $\alpha_0, \alpha_1 < \vartheta(\Omega^{\beta+1} \cdot \omega)$, and $\sigma, \tau < \Omega^{\beta+1} \cdot \omega$ such that $\sigma^*, \tau^* < \vartheta(\Omega^{\beta+1} \cdot \omega)$, $\sigma|_0 = \sigma$ and $\tau|_0 = \tau$. Applying (Weakening) yields

$$\mathcal{T}_\infty \upharpoonright_\rho^\alpha B, \quad \mathcal{T}_\infty \upharpoonright_\rho^\alpha \neg B, A$$

where $\rho = \sigma \# \tau$ and $\alpha = \max\{\alpha_0, \alpha_1\}$, whence one obtains $\mathcal{T}_\infty \upharpoonright_\rho^\gamma A$ for some $\gamma < \alpha^{\varepsilon+}$. We take $\rho = \sigma \# \tau$ in place of $\rho = \max\{\sigma, \tau\}$ to ensure $\sigma, \tau \leq^* \rho$ as is required for (Weakening). Since $\rho < \Omega^{\beta+1} \cdot \omega$ and $\rho^* = \max\{\sigma^*, \tau^*\}$, we have $\rho^*, \gamma < \vartheta(\Omega^{\beta+1} \cdot \omega)$.

If an application of T_β -Intro was utilised in F_β , A is $\mathsf{T}_\beta(\ulcorner B \urcorner)$ for some sentence B and the induction hypothesis yields

$$\mathcal{T}_\infty \upharpoonright_\rho^\alpha B$$

for some $\alpha < \vartheta(\Omega^{\beta+1} \cdot \omega)$ and $\rho < \Omega^{\beta+1} \cdot \omega$. An application of the rule (T_β -Intro) in \mathcal{T}_∞ yields

$$\mathcal{T}_\infty \upharpoonright_{\rho \# \Omega^{\beta+1}}^{\alpha+1} \mathsf{T}_\beta(\ulcorner B \urcorner).$$

$\rho \# \Omega^{\beta+1} < \Omega^{\beta+1} \cdot \omega$ if and only if $\rho < \Omega^{\beta+1} \cdot \omega$, and

$$\begin{aligned} (\rho \# \Omega^{\beta+1})^* &= \max\{\rho^*, \beta^*\} \\ &\leq \vartheta(\Omega^{\beta+1} \cdot \omega), \end{aligned}$$

so the desired bounds are maintained.

This leaves only applications of T_β -Elim and $\neg\mathsf{T}_\beta$ -Elim in F_β to consider, for which we employ theorem 13. The case of $\neg\mathsf{T}_\beta$ -Elim is immediate given the induction hypothesis, so suppose $F_\beta \vdash \mathsf{T}_\beta(\ulcorner A \urcorner)$ for some \mathcal{L}_κ -sentence A . The induction hypothesis entails

$$\mathcal{T}_\infty \upharpoonright_\rho^\alpha \mathsf{T}_\beta(\ulcorner A \urcorner) \tag{59}$$

for some $\rho < \Omega^{\beta+1} \cdot \omega$ and $\alpha, \rho^* < \vartheta(\Omega^{\beta+1} \cdot \omega)$. By theorem 13 (ii) one obtains some γ and σ such that $\sigma^*, \gamma < \vartheta(\rho + \alpha)$, $\sigma < \rho|_\beta$ and

$$\mathcal{T}_\infty \upharpoonright_\sigma^\gamma A$$

holds. $\sigma < \Omega^{\beta+1} \cdot \omega$ and, since $\rho^*, \alpha < \vartheta(\Omega^{\beta+1} \cdot \omega)$, one has $\vartheta(\rho + \alpha) < \vartheta(\Omega^{\beta+1} \cdot \omega)$, whence $\gamma, \sigma^* < \vartheta(\Omega^{\beta+1} \cdot \omega)$. \square

Lemma 39 allows us to easily determine an upper bound on the strength of the theories F_β . By making use of the remarks at the beginning of this section, we see that the next theorem is provable with an almost identical argument to ??.

Theorem 14. *Every arithmetical theorem of F_β , for $\beta < \vartheta_{\Omega+1}$, is derivable in $\text{PA} + \text{TI}(<\vartheta(\Omega^{\beta+1} \cdot \omega))$.*

5 Conclusion

By combining the work of sections 4.3 and 4.4 we may determine the strength of the theory F_β for every $\beta < \vartheta\Omega^\Omega$.

Theorem 15. *For every $p < \omega$*

$$\|F_p\| = \vartheta(\Omega^{p+1} \cdot \omega).$$

Theorem 16. *For every $\beta \geq \omega$ with $\beta < \vartheta\Omega^\Omega$,*

$$\vartheta(\Omega^\beta \cdot \omega) \leq \|F_\beta\| \leq \vartheta(\Omega^{\beta+1} \cdot \omega).$$

Proof. Theorem 14 provides the upper bounds for the two theorems. The lower bound for theorem 15 results from theorem 10, whereas the lower bound of theorem 16 is a consequence of theorem 11 \square

When presented with some hierarchy of theories $\{T_\alpha : \alpha \in \mathbb{O}\}$ it is natural to ask the limit of the corresponding autonomous progression, that is, the least ordinal not in the set X_T generated by the operation $\|T_0\| \subseteq X_T$ and $\alpha \in X_T$ implies $\|T_\alpha\| \subseteq X_T$. Autonomous progressions of ramified theories such as RA_α are well-studied, as are those obtained by iterating reflection principles. In the case $T_\alpha = F_\alpha$ this is not difficult to determine given theorem 16.

Theorem 17. *The limit of the autonomous progression defined from $\{F_\beta : \beta < \Omega\}$ is the large Veblen ordinal, $\vartheta\Omega^\Omega$.*

Proof. Let $\sigma_0 = 0$ and $\sigma_{m+1} = \vartheta(\Omega^{\sigma_m})$. Theorem 11 implies $\|F_{\sigma_m}^1\| \geq \sigma_{m+1}$, while theorem 16 entails $\|F_{\sigma_m}\| < \sigma_{m+2}$, so $X_F = \sup_m \sigma_m$. It remains to show

$$\vartheta(\Omega^\Omega) = \sup_{m < \omega} \sigma_m.$$

Since $\sigma_m < \Omega$ for every m , we have $\Omega^{\sigma_m} < \Omega^\Omega$. So $\sigma_m < \vartheta(\Omega^\Omega)$ implies $\sigma_{m+1} < \vartheta(\Omega^\Omega)$. Thus $\vartheta(\Omega^\Omega) \geq \sup_m \sigma_m$ is established by induction on m .

For the converse direction we prove $\alpha < \vartheta(\Omega^\Omega)$ implies $\alpha < \sigma_m$ for some m by induction on the rank of α . If $\alpha =_{NF} \omega^\gamma + \delta$ for some γ, δ , one easily obtains $\alpha < \sigma_m$ by the induction hypothesis, so suppose

$$\alpha = \vartheta\xi < \vartheta(\Omega^\Omega),$$

for which there are two cases to consider:

1. $\xi < \Omega^\Omega$ and $\xi^* < \vartheta(\Omega^\Omega)$; or
2. $\xi > \Omega^\Omega$ but $\vartheta\xi \leq (\Omega^\Omega)^*$.

Since $(\Omega^\Omega)^* = 0$, the latter is impossible. From the former, however, one obtains $\alpha < \vartheta(\Omega^\Omega)$ via the induction hypothesis. \square

Our motivation for defining the theory F_β as we did stemmed from the idea of formalising the acceptance of F . The theory S_3 with just one truth predicate appears to almost achieve this, but the general inability to close S_3 under the rule T-Intro means the truth predicate no longer satisfies the same principles as it did in F . This lead us to consider stratifying the language, viewing the original predicate of F , now T_0 , as the base level and gradually extending the language by including predicates T_1, T_2 , etc. in such a way that each predicate T_η in the language locally satisfies the same axioms and rules as T_0 .

The analysis of the theories F_β reveals that stratification of the language did not lead us as far from the world of a single self-applicable truth predicate as might have first appeared. Indeed, theorems 7 and 8 show the truth predicates of F_β^1 may be treated as identical; they can all be interpreted as the set $\mathcal{F}_{<\Omega^{\beta+1}}^*$. Within F_β^2 they may all be interpreted as the set $\mathcal{F}_{<\Omega^{\beta+1.2}}^*$ and, in general, all truth predicates in F_β can be interpreted as the set $\mathcal{F}_{<\Omega^{\beta+2}}^*$ (one cannot simply use $\mathcal{F}_{<\Omega^{\beta+1},\omega}^*$ for the interpretation of T_β in F_β as the set is not closed under the ω -rule, whereas $\mathcal{F}_{<\Omega^{\beta+2}}^*$ is, as well being closed under T_η -Elim, T_η -Intro for every $\eta \leq \beta$.) The upshot is that we may view each predicate T_η as “extending” the base predicate T_0 as well as T_ξ for $\xi < \eta$. It would be interesting to determine whether the theory F_β can be rewritten in some natural type-free form.

The model constructions employed in the previous section for the analysis of F_1 allow us to obtain an upper bound for the theory S_3 introduced in section 3.2. Essentially, we stratify the language \mathcal{L}_T as described in remark 1, interpreting the outermost truth predicate by T_1 and all others by T_0 , but by first embedding S_3 in an infinitary theory formulated without T-Elim, we avoid the problems relating to T_0 -Elim and T_1 -Elim.

Theorem 18. S_3 proves the same arithmetical statements as F_1^1 and hence has proof-theoretic ordinal Γ_0 .

Proof. We define an infinitary proof system \mathcal{S}_∞ based on \mathcal{T}_∞ into which we may embed S_3 . Let $*$ be the interpretation of \mathcal{L}_0 into \mathcal{L}_T that recursively interprets the predicate T_0 as T and otherwise commutes with all connectives and quantifiers. Define $\mathcal{S}_\infty \models^\alpha \Gamma$ according to the rules (Ax.1), (Ax.2), (Ax.3), (\wedge) , (\vee_i) , (ω) , (\exists) , (T-Imp), (T-Del), (T-Rep), (T- ω), and the following additional rule

$$(Ax.4)^{\mathcal{F}} \quad \mathcal{T}_\infty \models_{\Omega, \gamma}^\beta A \quad \text{for some } \gamma, \beta < \Omega \quad \text{implies} \quad \mathcal{S}_\infty \models^\alpha \Gamma, T(\ulcorner A^* \urcorner) \quad \text{for any } \alpha > \max\{\beta, \gamma\}.$$

The Condensation lemma entails, for $\gamma < \Omega$, that

$$\mathcal{T}_\infty \models_{\Omega, \gamma}^\alpha \Gamma \quad \text{implies} \quad \mathcal{S}_\infty \models^\alpha \Gamma^*. \quad (60) \quad \{\text{exteqn:S81}\}$$

Define a sequence of \mathcal{L}_T -structures

$$\mathfrak{M}_\alpha = \langle \mathbb{N}, \{\ulcorner A^* \urcorner : \mathcal{T}_\infty \models_{\Omega, \gamma}^\beta A \text{ for some } \gamma < \alpha \text{ and } \beta < \vartheta(\Omega \cdot \alpha)\} \rangle.$$

We claim

$$\mathcal{S}_\infty \models^\alpha \Gamma \quad \text{implies} \quad \mathfrak{M}_\alpha \models \bigvee \Gamma \quad (61) \quad \{\text{exteqn:S82}\}$$

whenever Γ is T-positive. The proof proceeds by transfinite induction on α . If Γ is an instance of $(Ax.4)^{\mathcal{F}}$, $\mathfrak{M}_\alpha \models \bigvee \Gamma$ holds by definition, while if Γ is derived through an application of (T-Rep), it follows from the induction hypothesis and closure of \mathcal{T}_∞ under $(T_0$ -Intro). If the last applied rule is (T- ω), $\mathfrak{M}_\alpha \models \bigvee \Gamma$ holds by an application of (ω) in \mathcal{T}_∞ and the fact $\vartheta(\Omega \cdot \alpha)$ is increasing in α .

Furthermore, by the definition of the function ϑ , we have $\vartheta(\Omega \cdot \delta + \beta) < \vartheta(\Omega \cdot \alpha)$ whenever $\delta < \alpha$ and $\beta < \vartheta(\Omega \cdot \alpha)$; thus theorem 13 implies

$$\mathfrak{M}_\alpha \models \forall^\top A^\top (T(\top T(\top A^\top)^\top) \implies T(\top A^\top)),$$

and we may deduce $\mathfrak{M}_\alpha \models \bigvee \Gamma$ from the induction hypothesis if the last rule applied was (T-Del).

Equation (61) can now be utilised along with eq. (60) to conclude

$$\mathcal{S}_\infty \models^\alpha T(\top A^\top) \text{ implies } \mathcal{S}_\infty \models^{\vartheta(\Omega \cdot \alpha)} A.$$

Since $\vartheta\Omega^2$ is the least ordinal closed under the function $\alpha \mapsto \vartheta(\Omega \cdot \alpha)$ (see proposition 20) we may deduce $S_3 \vdash A$ implies $\mathcal{S}_\infty \models^\alpha A$ for some $\alpha < \vartheta\Omega^2$ for any sentence A , whence $\|S_3\| \leq \vartheta\Omega^2$. Finally, note $\vartheta\Omega^2 = \Gamma_0$ by corollary 21. \square

On a final note, one may reasonably consider applying this idea to the other Friedman-Sheard theories (or indeed to other theories of self-referential truth). The theory E is ω -inconsistent (*cf.* ??), so a definition of E_1 cannot consistently contain the axiom $T_1\text{-}\forall\text{-Inf}$. Thus, since the predicate T_1 could not adhere to the rules governing T_0 , this process could not be applied to E . Iterating the operation pertaining to I , however, does lead to a hierarchy of consistent theories. I does not contain the rule $T\text{-Intro}$, so one would not naturally include the principle

$$T_1(x) \implies T_1(\top T_0(\dot{x})^\top)$$

as an axiom of I_β . Without such an axiom, though, there would be no dependency between the two predicates: any occurrence of T_0 under T_1 may be treated vacuous, as can any occurrence of T_1 under T_0 . In an attempt to resolve the matter one could add to I_1 the axiom $T_1(\top A^\top)$ whenever A is an axiom of I_0 (I formulated with the predicate T_0). This would appear a natural choice as it agrees with the construction of F_1 . It would be interesting to ascertain the strength of the resulting iteration and compare it to our results regarding F_β .