# Iterated self-applicable truth

## GEL & MR

### 23rd April 2024

## Contents

## 1 Overview

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

## 2 A theory

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

For $\beta < \kappa$, define a theory $\mathsf{T}_\beta$ extending HA by axioms, for each $\eta \le \kappa$:

$$\forall \ulcorner A \urcorner \ulcorner B \urcorner (T_\eta \ulcorner A \to B \urcorner \to (T_\eta \ulcorner A \urcorner \to T_\eta \ulcorner B \urcorner)) \tag{I}$$

$$\forall \ulcorner A(x) \urcorner (\forall n T_\eta \ulcorner A(\dot{n}) \urcorner \to T_\eta \ulcorner \forall x A(x) \urcorner \tag{U}$$

$$\forall \xi \le \beta (T_\eta \ulcorner T_\xi \dot{x} \urcorner \to T_\eta x) \tag{D}$$

$$\forall \xi < \eta (T_\eta x \to T_\eta \ulcorner T_\xi \dot{x} \urcorner) \tag{R}$$

and rules of inference

$$A \vdash T_\beta \ulcorner A \urcorner \qquad \text{(Nec)}$$
$$T_\beta \ulcorner A \urcorner \vdash A \qquad \text{(Conec)}$$

What gives the strength is that (D) 'collapses' all internal truth predicates to the current level ($\eta$). In contrast, (R) only permits 'expanding' the internal level by lower level predicates.

We argue that

**Theorem 1.** *For every $p < \omega$,*
$$\|T_p\| = \vartheta(\Omega^{p+1} \cdot \omega).$$

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

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

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

It seems likely that the same claims hold for (intuitionistic) $T_\beta$.

The ordinal analysis of $T_\beta$ theories is derived from the analysis of a classical extension of the theories in the first author's PhD thesis [**?**]. Some arguments require changing to accommodate the intuitionistic base (T-elimination), and some have been streamlined (?).

## 3  Theories with multiple self-applicable truth predicates — material from the thesis

chap:ext

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 $\omega$-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

$$S_3 \vdash \forall x (\mathsf{Bew}_{S_2}(x) \wedge \mathsf{Sent}_{\mathscr{L}_T}(x) \to T(x)).$$

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 ¬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

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

as an axiom.[1] 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 $F = \mathsf{Base}_T + \text{T-Intro} + \text{T-Elim} + \neg\text{T-Elim} + \text{T-Del} + \forall\text{-Inf}$. Let $\mathsf{HST}_0$ 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 $\mathsf{HST}_0$, which we shall denote by $\mathsf{HST}_1$, would then contain the following axioms

$$\mathsf{val}(x) \wedge \mathsf{Sent}_{\mathscr{L}_{T_0}}(x) \rightarrow T_1(x), \tag{1} \quad \texttt{\{eqn:FFax0\}}$$

$$\mathsf{Ax}_{\mathsf{HST}_0}(x) \wedge \mathsf{Sent}_{\mathscr{L}_{T_0}}(x) \rightarrow T_1(x), \tag{2} \quad \texttt{\{eqn:FFax1\}}$$

$$T_1(x) \wedge T_1(x \dot{\rightarrow} y) \rightarrow T_1(y), \tag{3} \quad \texttt{\{eqn:FFax2\}}$$

$$(\forall x\, T_1(\ulcorner A(\dot{x}) \urcorner)) \rightarrow T_1(\ulcorner \forall x A(x) \urcorner), \tag{4} \quad \texttt{\{eqn:FFaxw\}}$$

$$T_1(x) \rightarrow T_1(\ulcorner T_0(\dot{x}) \urcorner), \tag{5} \quad \texttt{\{eqn:FFax3\}}$$

$$T_1(\ulcorner T_0(\dot{x}) \urcorner) \rightarrow T_1(x), \tag{6} \quad \texttt{\{eqn:FFax4\}}$$

$$T_1(\ulcorner \neg T_0(\dot{x}) \urcorner) \rightarrow T_1(\dot{\neg} x). \tag{7} \quad \texttt{\{eqn:FFax5\}}$$

(**??**) state the acceptance of all axioms of $\mathsf{HST}_0$ (logical and non-logical), whereas (3) formalises *modus ponens* in $\mathsf{HST}_0$. These three also combine to imply the axioms of $\mathsf{Base}_{T_1}$. (**??**) express the acceptance of the rules $\mathsf{nec}_0$, $\mathsf{conec}_0$ and $\neg\mathsf{conec}_0$, respectively, in $\mathsf{HST}_0$, while (4) closes the predicate under $\omega$-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

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

and closure under $\mathsf{conec}_1$, $\mathsf{nec}_1$ and $\neg\mathsf{conec}_1$. Combining also the axioms of $\mathsf{HST}_0$ it is then easy to deduce

$$\forall x (\mathsf{Bew}_{\mathsf{HST}_0}(x) \wedge \mathsf{Sent}_{\mathscr{L}_{T_0}}(x) \rightarrow T_1(x)). \tag{8} \quad \texttt{\{eqn:F1reflecti}}$$

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 $\mathsf{nec}_1$ may apply to arbitrary sentences in $\mathscr{L}_{T_0,T_1}$, 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 (**??**) (in particular (5) pertaining to $\mathsf{nec}_0$) for inclusion in $\mathsf{HST}_1$: if we restrict them to range over only codes of $\mathscr{L}_{T_0}$-sentences we will have no non-trivial occurrences of this inter-applicability.

---

[1] As the $\mathscr{L}_T$-structure $\mathfrak{M}$ used in **??** also satisfies the axiom $T(\ulcorner \neg T(\dot{x}) \urcorner) \rightarrow T(\dot{\neg} x)$, the extension of $S_3$ obtained by adding this axiom is also consistent.

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 $\mathsf{S}$ to be applicable to any extension of the language, logic or axioms of $\mathsf{S}$. Thus, if one imagines the theory $\mathsf{HST}_0$ being first formulated in the language $\mathscr{L} \cup \{\mathsf{T}_0, \mathsf{T}_1\}$, and only then completing the reflection step to $\mathsf{HST}_1$ by adding the axioms and rules pertaining to $\mathsf{T}_1$, it seems natural to suppose the predicates $\mathsf{T}_0$ and $\mathsf{T}_1$ are inter-applicable. Since $\mathrm{nec}_0$ was applicable in $\mathsf{HST}_0$ to formulae containing the predicate $\mathsf{T}_1$, so should $\mathsf{T}_0$ in $\mathsf{HST}_1$. Thus, we expect $\mathsf{HST}_1$ to have the axiom $\mathsf{T}_0$-Imp,

$$\forall x \forall y [\mathsf{T}_0(x) \wedge \mathsf{T}_0(x \,\dot{\to}\, y) \to \mathsf{T}_0(y)],$$

as opposed to its relativised form

$$\forall x \forall y [\mathsf{Sent}_{\mathscr{L}_{\mathsf{T}_0}}(x) \wedge \mathsf{Sent}_{\mathscr{L}_{\mathsf{T}_0}}(y) \to (\mathsf{T}_0(x) \wedge \mathsf{T}_0(x \,\dot{\to}\, y) \to \mathsf{T}_0(y))]; \tag{9}$$

and, more importantly, that $\mathsf{HST}_1$ contains the unrelativised axioms

$$\forall x [\mathsf{T}_1(x) \to \mathsf{T}_1(\ulcorner \mathsf{T}_0(\dot{x}) \urcorner)],$$
$$\forall x [\mathsf{T}_1(\ulcorner \mathsf{T}_0(\dot{x}) \urcorner) \to \mathsf{T}_1(x)], \tag{10}$$
$$\forall x [\mathsf{T}_1(\ulcorner \neg \mathsf{T}_0(\dot{x}) \urcorner) \to \mathsf{T}_1(\dot{\neg} x)].$$

This provides, for example,

$$\forall x [\mathsf{T}_1(\ulcorner \mathsf{T}_1(\dot{x}) \urcorner) \leftrightarrow \mathsf{T}_1(\ulcorner \mathsf{T}_0(\ulcorner \mathsf{T}_1(\dot{x}) \urcorner) \urcorner)],$$

which, by $\mathrm{nec}_1$ and (8), yields $\mathsf{T}_0(\ulcorner \mathsf{T}_1(\ulcorner A \urcorner) \urcorner)$ for every theorem $A$ of $\mathsf{HST}_0$. Thus our informal interpretation leads us to the curious situation where we have two self-applicable predicates $\mathsf{T}_0, \mathsf{T}_1$ which may also be applied to each other.

If, on the other hand, we had considered relativising the quantifiers as in (9), $\mathsf{T}_1$ may apply to the language $\mathscr{L}_{\mathsf{T}_0, \mathsf{T}_1}$ whereas $\mathsf{T}_0$ may only meaningfully apply to $\mathscr{L}_{\mathsf{T}_0}$ and one would obtain the base of a strict hierarchy of self-applicable truth predicates. In this case we can no longer argue that $\mathsf{T}_0$ represents a truth predicate for the whole language but only of the sub-language $\mathscr{L}_{\mathsf{T}_0}$. 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, $\mathsf{T}_0$ *should* be applicable to sentences involving the predicate $\mathsf{T}_1$ and we find ourselves returning to the world of two inter-applicable truth predicates.

So far we have argued that in $\mathsf{HST}_1$ the interpretation of $\mathsf{T}_0$ should be closed under $\mathrm{conec}_0$ and $\neg\mathrm{conec}_0$ while the interpretation of $\mathsf{T}_1$ should be closed under $\mathrm{nec}_0, \mathrm{conec}_0, \neg\mathrm{conec}_0$, $\mathrm{conec}_1$ and $\neg\mathrm{conec}_1$. However, we desire $\mathsf{T}_1$ to be viewed as an extension of $\mathsf{T}_0$ so as to allow for closure of $\mathsf{HST}_1$ 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 $\mathsf{HST}_1$, there should be no essential difference between the predicates $\mathsf{T}_0$ and $\mathsf{T}_1$. Thus we propose to also include the principle

$$\mathsf{T}_0(\ulcorner \mathsf{T}_1(\dot{x}) \urcorner) \to \mathsf{T}_0(x) \tag{11}$$

as an axiom of $\mathsf{HST}_1$ (in fact as an axiom of $\mathsf{HST}_0$). (11) expresses that the interpretation of $\mathsf{T}_0$ is closed under the rule $\mathrm{conec}_1$. This is vacuously valid in the theory $\mathsf{HST}_0$. 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 (11) we should also accept $T_0(\ulcorner \neg T_1(\dot{x}) \urcorner) \to T_0(\dot{\neg} x)$, but as we shall see, this axiom will be trivially satisfied by our model.

We have only described the step $HST_0$ to $HST_1$, but one can imagine repeating this, first adding an additional truth predicate $T_2$ to $HST_1$ and then reflecting upon it to form the theory $HST_2$.[2] This process may be continued into the transfinite to form a hierarchy of theories, supporting a hierarchy of inter-applicable truth predicates.[3]

The description of $HST_1$ and $HST_2$ 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 $HST_\omega$, $HST_{\varepsilon_0}$, *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 $\kappa$ 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 $\kappa$ 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 6 below) and so we will require the construction of a larger class of ordinals to perform a sufficient proof-theoretic analysis. Suitable choices for $\kappa$ will ultimately depend on our analysis and have no substantial role in defining the theories. Since the precise definition of $\kappa$ is not essential for the definition, we shall assume for the time being that $\kappa$ represents some fixed primitive recursive ordinal.

**Definition 3.1.** For $\beta < \kappa$, let $\mathscr{L}_\beta$ denote the language of PRA augmented by predicates $T_\xi$ for each $\xi \leq \beta$. Let $HST_\beta$ be the theory formulated in the language $\mathscr{L}_\beta$, extending PA with the schema of induction for $\mathscr{L}_\beta$, and for each $\eta \leq \beta$

$$\forall x(T_\eta(x_1) \wedge T_\eta(x_1 \dot{\to} x_2) \to T_\eta(x_2)), \tag{$\mathrm{imp}_\eta$} \quad \text{\{ax:imp\}}$$

$$\forall \ulcorner A(x) \urcorner [(\forall n T_\eta(\ulcorner A(\dot{n}) \urcorner)) \to T_\eta(\ulcorner \forall x\, A(x) \urcorner)], \tag{$\mathrm{uni}_\eta$} \quad \text{\{ax:uni\}}$$

$$\forall \xi \leq \bar{\beta}\, \forall x(T_\eta(\ulcorner T_\xi(\dot{x}) \urcorner) \to T_\eta(x)), \tag{$\mathrm{del}_\eta$} \quad \text{\{ax:del\}}$$

$$\forall \xi < \bar{\eta}\, \forall x(T_\eta(x) \to T_\eta(\ulcorner T_\xi(\dot{x}) \urcorner)), \tag{$\mathrm{rep}_\eta$} \quad \text{\{ax:rep\}}$$

as well as the rules of inference

$$\text{from } A \text{ infer } T_\beta(\ulcorner A \urcorner), \tag{$\mathrm{nec}_\beta$} \quad \text{\{ax:nec\}}$$

$$\text{from } T_\beta(\ulcorner A \urcorner) \text{ infer } A. \tag{$\mathrm{conec}_\beta$} \quad \text{\{ax:conec\}}$$

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

The theory $HST_0$ is identical to $F$ and the definition of $HST_\beta$ fits the informal description we gave of 'F viewed as an operation applied to $HST_{<\beta}$'. Also, $HST_\beta^0$ is a conservative extension

---

[2] A more precise way to describe the construction of $HST_2$ may be given as follows: first imagine formulating $HST_0$ with three predicates, $T_0$, $T_1$, and $T_2$. One then formulates $HST_1$ in this language by reflecting upon $HST_0$, and only then is $HST_2$ formulated by reflecting upon $HST_1$. In this sense we view $HST_0$ as not being formulated in a fixed language, but rather in a language that may be expanded as and when we see fit.

[3] The 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.

of HA as, with no applications of ($\text{nec}_\beta$), all predicates $T_\eta$ in $\text{HST}^0_\beta$ may be interpreted trivially to show neither $\text{HST}^0_\beta \vdash T_\beta s$ nor $\text{HST}^0_\beta \vdash \neg T_\beta s$ may hold.

Although $\text{HST}_1$ 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 $\text{HST}_1$, however, is almost identical to that which describes the theory $S_3$. As a result it may not be surprising to know that $\text{HST}_1$ is consistent and in fact $\text{HST}^1_1$ 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 8), but the following remark should motivate the connection.

**Remark 1.** It is natural to first consider embedding $\text{HST}^1_1$ into $S_3$. As the predicate $T_1$ can be interpreted vacuously in $\text{HST}^0_1$ (as can $T_0$), one may expect the interpretation given by

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

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

$$g(x) = x, \text{ if } x \text{ is the code of an arithmetical literal,}$$
$$g(\ulcorner T_1 s \urcorner) = \ulcorner s \neq s \urcorner,$$
$$g(\ulcorner T_0 s \urcorner) = \ulcorner T(gs) \urcorner,$$
$$g(\ulcorner A \circ B \urcorner) = g(\ulcorner A \urcorner) \mathbin{\dot{\circ}} g(\ulcorner B \urcorner), \text{ for } \circ \in \{\wedge, \vee\}$$
$$g(\ulcorner QxA(x) \urcorner) = \dot{\forall} x\, g(\ulcorner A(x) \urcorner).$$

This interpretation will translate the axioms $\text{del}_1$, $\text{uni}_1$ and $\text{rep}_1$ of $\text{HST}_1$ 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 $\neg\text{del}_1$ from $\text{HST}^1_1$. Fortunately, the model construction of **??** can easily incorporate the additional axiom $\neg$T-Del and rule $\neg$T-Elim with minimal adjustments: one may consistently add the rule pertaining to $\neg$T-Del to each of the theories $\mathscr{T}_\alpha$ and expand **??** to also prove $\mathscr{T}_\alpha \vdash \neg T(\ulcorner A \urcorner)$ implies $\mathscr{T}_\alpha \vdash \neg A$, thus showing $S_3 + \neg$T-Del $+ \neg$T-Elim is a consistent theory.

The problem with this interpretation manifests when dealing with applications of $\text{conec}_1$ in $\text{HST}^1_1$. Suppose $\text{HST}^1_1 \vdash T_1(\ulcorner A \urcorner)$. If $A$ does not contain the predicate $T_1$, $\ulcorner A^* \urcorner = g(\ulcorner A \urcorner)$ and an application of T-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 $S_3 \nvdash B^*$ (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-Elim $\vdash \neg C$, which contradicts the consistency of $S_3 + \neg$T-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 $\text{HST}^0_1$ though, $S_3 \vdash A^*$; moreover, if one views $T_1$-Imp, $\text{del}_1$ and $T_1$-Rep as a rule of inference, as in $\mathscr{T}$, one could deduce closure under $\text{conec}_1$ by induction on the height of a derivation. Viewing the axioms of $\text{HST}^1_1$ in this manner is reminiscent of the use of the infinitary system $\mathscr{F}_\infty$ in the analysis of $F$; thus it seems natural to delay a further investigation of this connection until we have first analysed $\text{HST}_1$ in detail.

Although the addition of full $\text{nec}_1$ to $\text{HST}^1_1$ (forming $\text{HST}_1$) creates a theory markedly stronger than $S_3$, it is not straightforward to embed $S_3$ into $\text{HST}_1$. Such an embedding would require stratifying $\mathscr{L}_T$ to involve the two predicates $T_0$ and $T_1$ of $\mathscr{L}_1$. 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 $T(\ulcorner\text{T-Imp}\urcorner)$, $T(\ulcorner\forall\text{-Inf}\urcorner)$ and $T(\ulcorner\text{T-Del}\urcorner)$ of $S_3$ to theorems of $\text{HST}_1^1$. This could be achieved by picking a primitive recursive function $f$ which recursively substitutes $T_0(fs)$ for $Ts$, and considering the interpretation $^*$ of $\mathscr{L}_T$ into $\mathscr{L}_1$ given by

$$(Ts)^* = T_1(fs).$$

This interpretation maps the axioms $\forall$-Inf, T-Imp, T-Del and T-Rep by instances of axioms $\text{uni}_1$, $T_1$-Imp, $\text{del}_1$ and $\text{rep}_1$ respectively, but since $f(\ulcorner A\urcorner)$ need not equal $\ulcorner A^*\urcorner$, applications of T-Elim in $S_3$ do not translate into inferences we can readily recognise as holding in $\text{HST}_1$.

We now move to the task of proving the consistency of $\text{HST}_\beta$ for arbitrary $\beta$. This will be obtained by generalising the consistency argument for $S_3$ to theories with multiple predicates and is presented in section 4 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 $\text{HST}_\beta$ will be obtained in section 6 by extending the well-ordering results previously established for $S_2$ and $S_3$. Upper bounds are determined in section 7 where we attempt to directly formalise the model constructions of section 4 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-Intro caused by the interaction between axioms $\text{uni}_\eta$ and $\text{rep}_\eta$.

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

<span style="float:left">prop:FFreflec</span>**Proposition 3.2.** *For every $\eta < \beta$, $\text{HST}_\beta$ is closed under $\text{nec}_\eta$ and $\text{conec}_\eta$. Moreover,*

$$\text{HST}_\beta^1 \vdash \forall x(\text{Bew}_\eta(x) \wedge \text{Sent}_{\mathscr{L}_\eta}(x) \rightarrow T_\beta(x)).$$

*Proof.* All axioms of $\text{HST}_\eta$ are axioms of $\text{HST}_\beta^0$, and by $\text{nec}_\beta$ we have

$$\text{HST}_\beta^1 \vdash \forall x(\text{Ax}_\eta(x) \wedge \text{Sent}_{\mathscr{L}_\eta}(x) \rightarrow T_\beta(x)),$$

where $\text{Ax}_\eta(x)$ expresses that $x$ is a sentential axiom of $\text{HST}_\eta$. To complete the proof, we observe that $\text{HST}_\beta^1$ has the axioms $\text{imp}_\beta$, $\text{del}_\beta$, $T_\beta^-$-Rep and $\neg\text{del}_\beta$ formalising applications of *modus ponens*, $\text{conec}_\eta$, $\text{nec}_\eta$ and $\neg\text{conec}_\eta$ respectively in $\text{HST}_\eta$. $\qquad\square$

## 4 Consistency

<span style="float:left">sec:Fbconsis</span>At first glance the theory $\text{HST}_\beta$ could easily look suspect, after all it contains $\text{del}_\eta$, $\text{conec}_\eta$, $\text{nec}_\eta$ and an axiom that appears extremely close to $T_\eta$-Rep, for each $\eta \leq \beta$. As the motivation behind the theories $\text{HST}_\beta$ comes from abstracting the transition from PA to F one might expect that if $\text{HST}_\beta$ is consistent, models of $\text{HST}_\beta$ 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 7 where we determine an upper bound on the proof-theoretic strength of $\text{HST}_\beta$.

Let $\Omega$ denote the first uncountable ordinal. For the remainder of this section Greek letters, $\rho, \sigma, \tau$ will be used to range over elements of $\mathbb{O}$, the class of all ordinals; letters $\alpha, \beta$, etc., will range over countable ordinals; we reserve the letters $\eta, \xi$ for indices of the truth predicates and so represent ordinals below $\kappa$. Suppose

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

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

**Definition 4.1** (Sequent). Let $\kappa \in \mathbb{O}$ be fixed. A ($\kappa$-)sequent is an expression $\Gamma \Rightarrow A$ where $\Gamma \cup \{A\}$ is a finite of sentences of $\mathscr{L}_\kappa$.

**Definition 4.2** (Semiformal calculus). For each $\rho \in \mathbb{O}$, a Tait-style sequent calculus $\vdash_\rho$ is introduced for $\kappa$-sequents, defined by the following rules.<span style="color:magenta">One- or two-sided?</span>

- Ax.1. $\vdash_\rho \Gamma \Rightarrow A$ whenever $A$ is a true atomic literal,

- Ax.2$_\eta$. $\vdash_\rho \Gamma, \mathsf{T}_\eta s \Rightarrow \mathsf{T}_\eta s$ for every closed term $s$ and $\eta < \kappa$,

- Ax.3$_\eta$. $\vdash_\rho \Gamma, \mathsf{T}_\eta s \Rightarrow A$ if $s^\mathbb{N}$ is not the code of an $\mathscr{L}_\kappa$-sentence.

- The usual arithmetical rules for $\wedge$, $\vee$ and $\exists$,

- The $\omega$ rule: $\dfrac{\vdash_\rho \Gamma \Rightarrow A(\underline{n}) \quad \text{for every } n < \omega}{\vdash_\rho \Gamma \Rightarrow \forall x A(x)} \omega$

- The following six rules for every $\eta < \kappa$.

$\text{imp}_\eta \; \dfrac{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta s \quad \vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta(s \dot{\rightarrow} t)}{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta t} \text{imp}_\eta$

$\text{del}_\eta \; \dfrac{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta(\ulcorner \mathsf{T}_\xi s \urcorner) \quad \xi < \beta}{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta s} \text{del}_\eta$

$\text{rep}_\eta \; \dfrac{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta s \quad \xi < \eta}{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta(\ulcorner \mathsf{T}_\xi s \urcorner)} \text{rep}_\eta$

$\text{nec}_\eta \; \dfrac{\vdash_\sigma \Gamma \Rightarrow A \quad \sigma < \rho|_\eta}{\vdash_\rho \Delta, \mathsf{T}_\eta \ulcorner \Gamma \urcorner \Rightarrow \mathsf{T}_\eta \ulcorner A \urcorner} \text{nec}_\eta$

$\text{uni}_\eta \; \dfrac{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta \ulcorner A(\underline{n}) \urcorner \text{ for all } n \quad \rho|_\eta \text{ not an } \Omega\text{-limit}}{\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta \ulcorner \forall x A(x) \urcorner} \text{uni}_\eta$

Moreover, for each $\rho$ define the set of $\mathscr{L}_\kappa$-sentences

$$\mathscr{T}_{<\rho} = \{A : \vdash_\sigma \Rightarrow A \text{ for some } \sigma < \rho\}.$$

Before we proceed with analysing the role of the truth predicates in $\mathscr{T}_\rho$, 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 admissible: a derivation of the form $\mathsf{T}_\eta s, \Gamma \Rightarrow A$ with $\mathsf{T}_\eta s$ principal *must* be an instance of an axiom, hence given derivations of $\mathsf{T}_\eta s, \Gamma \Rightarrow A$ and $\Delta \Rightarrow \mathsf{T}_\eta s$ one can easily obtain instead a derivation of $\Gamma$. This argument is essentially identical to that employed in the Cut Elimination theorem in [1] for the analysis of the theory $\mathsf{F}$ (see [1, Thm. 3.7]).

**Lemma 4.3** (Admissibility of cut). *Suppose* $\vdash_\rho \Gamma \Rightarrow \mathsf{T}_\eta s$, *and* $\vdash_\rho \Delta, \mathsf{T}_\eta s \Rightarrow A$. *Then* $\vdash_\rho \Gamma, \Delta \Rightarrow A$.

If $\rho < \Omega$, then $\rho|_0 = 0$ and so no applications the rule $\mathrm{nec}_\eta$ is available in $\vdash_\rho$ for any $\eta$. Thus, $\mathscr{T}_{<\Omega}$ is trivially closed under $\mathrm{conec}_\eta$ for every $\eta$. <span style="color:purple">Simple but not 'trivial'</span> Moreover, for every $\rho < \Omega$, only applications of $\mathrm{nec}_0$ have been permitted in $\vdash_{\Omega\cdot\rho}$. Thus, we can establish by induction on $n < \omega$ that in derivations in $\vdash_{\Omega\cdot n}$ the predicate $\mathsf{T}_0$ may be interpreted by the $\mathscr{T}_{<\Omega\cdot n}$ which happens to be closed under $\mathrm{conec}_0$. Since all other truth predicates that can occur in $\vdash_{\Omega\cdot n}$-derivable sequents may be interpreted vacuously, we conclude $\vdash_{\Omega\cdot n}$ is closed under $\mathrm{conec}_\eta$ for every $\eta < \kappa$. Hence $\mathsf{HST}_0^n$ may be interpreted in $\mathscr{T}_{\Omega\cdot n}$.

This suggests that for $n < \omega$, $\mathscr{T}_{\Omega\cdot n}$, like $\mathscr{T}_n$, reconstructs the theories $\mathrm{Th}_n$ used by Friedman and Sheard to prove the consistency of $\mathsf{F}$. At the first limit ordinal, we obtain $\mathscr{T}_{<\Omega\cdot\omega}$, a set of $\mathscr{L}_\kappa$-sentences closed under $\mathrm{nec}_0$ and $\mathrm{conec}_0$, and containing $\mathrm{del}_0$, $\mathrm{rep}_0$ (which holds vacuously) and all other axioms of $\mathsf{HST}_0$.

To proceed with the analysis of $\mathsf{HST}_1$, we first consider $\mathsf{HST}_1^0$ which, without the rule $\mathrm{nec}_1$, is vacuously closed under $\mathrm{conec}_1$ and $\neg\mathrm{conec}_1$. In $\mathsf{HST}_1^1$, the situation differs from previous case; we need to interpret the predicate $\mathsf{T}_1$ as a theory closed under $\omega$-logic (due to $\mathrm{uni}_1$), $\mathrm{nec}_0$ (due to $\mathrm{rep}_1$), $\mathrm{conec}_0$ and $\mathrm{conec}_1$ (due to $\mathrm{del}_1$). Moreover, we need to find an interpretation of $\mathsf{T}_0$ closed under $\omega$-logic, $\mathrm{conec}_0$, and now also $\mathrm{conec}_1$ (as implied by the axiom $\mathrm{del}_0$). The properties we established for the set $\mathscr{T}_{<\Omega}$ in **??** motivate us to consider $\mathscr{T}_{<\Omega^2}$, a set closed under $\mathrm{nec}_0$, $\mathrm{conec}_0$ and, by a similar argument as before, $\omega$-logic (*cf.* the proof of **??**). For every $\alpha < \Omega$, the predicate $\mathsf{T}_1$ may be interpreted vacuously in $\mathscr{T}_{<\Omega\cdot\alpha}$, so $\mathscr{T}_{<\Omega^2}$ is also closed under $\mathrm{conec}_1$. Thus $\mathscr{T}_{<\Omega^2}$ provides a consistent interpretation of both predicates $\mathsf{T}_0$ and $\mathsf{T}_1$ in $\mathsf{HST}_1^1$.

The next step is to consider $\mathsf{T}_1$ in $\mathsf{HST}_1^2$. Two applications of $\mathrm{nec}_1$ are permitted and one can derive sentences of the form $\mathsf{T}_1(\ulcorner\mathsf{T}_1(\ulcorner A\urcorner)\urcorner)$ whenever $\mathsf{HST}_1^0 \vdash A$, suggesting a shift to $\mathscr{T}_{\Omega^2}$, where one can derive $\mathsf{T}_1(\ulcorner A\urcorner)$ whenever $A \in \mathscr{T}_{<\Omega^2}$, might yield a suitable interpretation for $\mathsf{T}_1$. However, $\mathscr{T}_{\Omega^2}$ is not closed under $\mathrm{nec}_0$ (only the systems $\mathscr{T}_{<\rho+\Omega\cdot\sigma}$ for limit ordinals $\sigma$ are), leading us instead to consider $\mathscr{T}_{<\Omega^2+\Omega\cdot\omega}$ which is closed under $\mathrm{nec}_0$, but not $\omega$-logic; there will be sentences $A \in \mathscr{T}_{<\Omega^2}$ for which $\mathscr{T}_{\Omega^2+\Omega\cdot n} \vdash \mathsf{T}_0(f(n, \ulcorner A\urcorner))$ for each $n < \omega$, where $f(0, n) = \ulcorner\bar{n}\urcorner$ and $f(m+1, n) = \ulcorner\mathsf{T}_0(f(m, n))\urcorner$, but the sentence $\forall x\, \mathsf{T}_0(f(x, \ulcorner A\urcorner)$ is not contained in $\mathscr{T}_{<\Omega^2+\Omega\cdot\omega}$. Indeed to obtain both closure under $\omega$-logic and $\mathrm{nec}_0$ we must move to the theory $\mathscr{T}_{<\Omega^2\cdot 2}$. We also require the interpretation to be closed under $\mathrm{conec}_0$. To manage this we repeat the same argument as before, but starting from $\mathscr{T}_{\Omega^2}$ in place of $\mathscr{T}_\Omega$. We know $\mathscr{T}_{\Omega^2}$ is closed under $\mathrm{conec}_0$ since the predicate $\mathsf{T}_0$ can be consistently interpreted as the set $\mathscr{T}_{<\Omega^2}$. This leads us to successively deduce the theories $\mathscr{T}_{\Omega^2+\Omega\cdot n}$ are closed under $\mathrm{conec}_0$ for each $n < \omega$. Note, we can still interpret $\mathsf{T}_1$ in $\mathscr{T}_{\Omega^2+\Omega\cdot n}$ by the set $\mathscr{T}_{<\Omega^2}$ as there has been no further applications of $\mathrm{nec}_1$. In $\mathscr{T}_{\Omega^2+\Omega\cdot\omega}$ we aim to interpret $\mathsf{T}_0$ by $\mathscr{T}_{<\Omega^2+\Omega\cdot\omega}$, which unlike $\mathscr{T}_{\Omega^2+\Omega\cdot n}$ is not closed under $\omega$-logic; however, $\mathscr{T}_{\Omega^2+\Omega\cdot\omega}$ is not closed under ($\mathrm{uni}_1$) so this does not pose a problem. Thus we may continue through the construction of $\mathscr{T}_{<\Omega^2\cdot 2}$ determining each theory $\mathscr{T}_{<\Omega^2+\Omega\cdot\alpha}$ for $\alpha < \Omega$ is closed under $\mathrm{conec}_0$.

The argument above highlights that the predicates $\mathsf{T}_1$ and $\mathsf{T}_0$ in $\mathsf{HST}_1^n$ may be interpreted as the set $\mathscr{T}_{<\Omega^2\cdot n}$, and hence $\mathsf{HST}_1$ naturally embeds into $\mathscr{T}_{<\Omega^2\cdot\omega}$. If we wanted to proceed beyond this and construct models for $\mathsf{HST}_2$, we could imagine constructing a sequence of systems

$$\mathscr{T}_{\Omega^2\cdot\omega}, \mathscr{T}_{\Omega^2\cdot\omega+\Omega}, \ldots, \mathscr{T}_{\Omega^2\cdot\omega+\Omega\cdot\alpha}, \ldots, \mathscr{T}_{\Omega^2\cdot(\omega+1)}, \ldots, \mathscr{T}_{\Omega^2\cdot\alpha}, \ldots$$

to obtain $\mathscr{T}_{<\Omega^3}$, an interpretation of the predicate $\mathsf{T}_2$ in $\mathsf{HST}_2^1$. The ability to recognise each theory $\mathscr{T}_{\Omega^2\cdot\alpha+\Omega\cdot\gamma}$ as closed under $\mathrm{conec}_1$ and $\mathrm{conec}_0$, however, is essential for the interpretation of $\mathsf{T}_2$-Del in $\mathsf{HST}_2^1$. As already argued, the set $\mathscr{T}_{<\Omega^2\cdot\alpha+\Omega\cdot\gamma}$ provides an interpretation of $\mathsf{T}_0$ in

$\mathcal{T}_{\Omega^2 \cdot \alpha + \Omega \cdot \gamma}$; but unless $\gamma$ is a limit ordinal, this need not be closed under $\mathrm{nec}_0$, so cannot interpret the predicate $\mathsf{T}_1$. The answer is to interpret $\mathsf{T}_1$ in $\mathcal{T}_{\Omega^2 \cdot \alpha + \Omega \cdot \gamma}$ as the set $\mathcal{T}_{<\Omega^2 \cdot \alpha}$ for *every* $\gamma < \Omega$. Only when we pass to $\mathcal{T}_{\Omega^2 \cdot (\alpha+1)}$ do we alter the interpretation of $\mathsf{T}_1$ (in this case it is changed to the set $\mathcal{T}_{<\Omega^2 \cdot (\alpha+1)}$). It is for exactly this reason that the rule $\mathrm{uni}_\eta$ was restricted so as to apply to $\mathcal{T}_\rho$ only if $\rho|_\eta$ is not an $\Omega$-limit; the set $\mathcal{T}_{<\Omega^2 \cdot \alpha}$ will not be closed under $\omega$-logic if $\alpha$ is a limit ordinal.

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

$$\mathcal{T}_{\Omega^3}, \ldots, \mathcal{T}_{\Omega^3 + \Omega \cdot \alpha}, \ldots, \mathcal{T}_{\Omega^3 + \Omega^2}, \ldots, \mathcal{T}_{\Omega^3 + \Omega^2 \cdot 2}, \ldots, \mathcal{T}_{\Omega^3 + \Omega^2 \cdot \alpha}, \ldots, \mathcal{T}_{\Omega^3 \cdot 2}, \ldots$$

and subsequently $\mathcal{T}_{<\Omega^3 \cdot \omega}$, a theory into which $\mathsf{HST}_2$ embeds. In general, we expect $\mathsf{HST}_\beta$ to embed into $\mathcal{T}_{\Omega^{\beta+1} \cdot \omega}$ for each $\beta$.

The next lemma deals with the task of determining the theory $\mathcal{T}_\rho$ is closed under $\mathrm{conec}_\eta$ for every $\eta < \kappa$. Before that, however, we require a result on the behaviour of $\Omega$-limits.

<span style="font-variant: small-caps">lem:ordOlimit</span> **Proposition 4.4.** *If $\rho$ is not an $\Omega$-limit and $\sigma_n < \rho$ for every $n < \omega$,*

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

*Proof.* Suppose $\rho$ is not an $\Omega$-limit and $\sigma_n < \rho$ for every $n < \omega$. Then $\rho > 0$ and there are ordinals $\rho_0, \alpha_0$ 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 $\Omega$, whence

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

$\square$

A sequent $\Gamma$ is called *T-positive* if all occurrences of a predicate $\mathsf{T}_\eta$ in $\Gamma$ for any $\eta < \kappa$ are positive. Define, for each ordinal $\rho$, an $\mathscr{L}_\kappa$-structure $\mathfrak{M}_\rho$ according to the following criterion.

$$\mathfrak{M}_\rho \models \mathsf{T}_\eta s \text{ iff } s^{\mathbb{N}} \in \mathcal{T}_{<\rho|_\eta}.$$

<span style="font-variant: small-caps">thm:FFTElim</span> **Theorem 4** (T-Elimination theorem). *Suppose $\rho \in \mathbb{O}$.*

1. *For every T-positive sequent $\Gamma$, $\vdash_\rho \Gamma$ implies $\mathfrak{M}_\rho \models \bigvee \Gamma$;*

2. *For any $\eta < \kappa$, $\vdash_\rho \mathsf{T}_\eta s$ implies there is a sentence $A$ with $s^{\mathbb{N}} = \ulcorner A \urcorner$ and $\vdash_\rho A$;*

*Proof.* We proceed by transfinite induction on $\rho$. 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 $\vdash_\rho \Gamma$. 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, $(\vee_i)$, $(\wedge)$, $(\omega)$ 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 $\Gamma$ is T-positive. If the last applied rule was $\mathrm{nec}_\eta$, $\mathsf{T}_\eta(\ulcorner A \urcorner)$ is contained in $\Gamma$ and $\vdash_\sigma A$ for some $\sigma < \rho|_\eta$, so $\mathfrak{M}_\rho \models \mathsf{T}_\eta(\ulcorner A \urcorner)$. 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 $\mathfrak{M}_\rho \models A_0 \wedge A_1$.

$\mathrm{imp}_\eta$. If the last rule applied was $\mathrm{imp}_\eta$, we may assume $A_0$ is $\mathsf{T}_\eta(s_0)$ and $A_1$ is $\mathsf{T}_\eta(s_0 \dot{\to} s_1)$, while $\Gamma$ contains $\mathsf{T}_\eta(s_1)$. By the above, we may assume $\mathfrak{M}_\rho \models \mathsf{T}_\eta(s_0) \wedge \mathsf{T}_\eta(s_0 \dot{\to} s_1)$. Thus, $s_0^{\mathbb{N}}$ and $s_1^{\mathbb{N}}$ are Gödel numbers of $\mathscr{L}_\kappa$-sentences, say $B_0$ and $B_1$ respectively, and there is some $\sigma < \rho|_\eta$ so that $\vdash_\sigma B_0$ and $\vdash_\sigma \neg B_0, B_1$. Admissibility of the cut rule (lemma 4.3) yields $\vdash_\sigma B_1$, and hence $\mathfrak{M}_\rho \models \mathsf{T}_\eta(s_1)$.

$\mathrm{del}_\eta$. In the case the last applied rule is $\mathrm{del}_\eta$, we may identify $A_0$ as $\mathsf{T}_\eta(\ulcorner \mathsf{T}_\xi s \urcorner)$ for some $\xi < \kappa$ and term $s$; moreover, $\mathsf{T}_\eta s$ is contained in $\Gamma$. $\mathfrak{M}_\rho \models \mathsf{T}_\eta(\ulcorner \mathsf{T}_\xi s \urcorner)$ implies $\vdash_\sigma \mathsf{T}_\xi s$ 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 $\vdash_\sigma A$. Thus $\mathfrak{M}_\rho \models \mathsf{T}_\eta s$ and $\mathfrak{M}_\rho \models \bigvee \Gamma$.

$\mathrm{rep}_\eta$. Here we have $\mathfrak{M}_\rho \models \mathsf{T}_\eta s$ and $\mathsf{T}_\eta(\ulcorner \mathsf{T}_\xi s \urcorner)$ is in $\Gamma$ for some $\xi < \eta$. By definition this implies $s^{\mathbb{N}} = \ulcorner A \urcorner$ for some sentence $A$ and $\vdash_\sigma A$ for some $\sigma < \rho|_\eta$, whence $\mathscr{T}_{\sigma + \Omega^{\xi+1}} \vdash \mathsf{T}_\xi s$ is derivable. But since $\xi < \eta$ and $\sigma < \rho|_\eta$, we have $\sigma + \Omega^{\xi+1} < \rho|_\eta$, and so $\mathfrak{M}_\rho \models \bigvee \Gamma$.

$\mathrm{uni}_\eta$. The assumption is that $\mathfrak{M}_\rho \models \forall x\, \mathsf{T}_\eta(\ulcorner A(\dot{x}) \urcorner)$. This entails the existence of, for every $n < \omega$, an ordinal $\sigma_n < \rho|_\eta$ such that $\mathscr{T}_{\sigma_n} \vdash A(\bar{n})$. Weakening and the $\omega$-rule yields $\vdash_\sigma \forall x\, A(x)$, where $\sigma = \sup_n \sigma_n$, but one need not in general have $\sigma < \rho|_\eta$.[4] Due to the restriction on applications of $\mathrm{uni}_\eta$, however, $\rho|_\eta$ is not an $\Omega$-limit, thus by proposition 4.4, $\sigma < \rho|_\eta$ and so $\mathfrak{M}_\rho \models \mathsf{T}_\eta(\ulcorner \forall x A(x) \urcorner)$, whence $\mathfrak{M}_\rho \models \bigvee \Gamma$.

This completes the proof of (i).

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

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

**Proposition 4.5.** *Let $A$ be any axiom of $\mathsf{HST}_\beta$. Then $\mathscr{T}_{\Omega^{\beta+1}} \vdash A$.*

*Proof.* One can derive each of the axioms via the corresponding rule and Ax.2$_\eta$, as in **??**. In the case of $\mathrm{uni}_\eta$ note $\Omega^{\beta+1}|_\eta$ is not an $\Omega$-limit for any $\eta \le \beta$. $\qquad \square$

**Theorem 5.** *The theory $\mathsf{HST}_\beta$ is consistent for every $\beta < \kappa$.*

*Proof.* Lemma 4.3, theorem 4 and the previous proposition provide the means to deduce, by induction on $n$, that $\mathsf{HST}_\beta^n$ embeds into $\mathscr{T}_{\Omega^{\beta+1} \cdot n}$. Thus every sentential theorem of $\mathsf{HST}_\beta$ is contained in $\mathscr{T}_{<\Omega^\beta \cdot \omega}$. However, clearly the empty sequent is not derivable in $\mathscr{T}_\rho$ for any $\rho$, so $\mathsf{HST}_\beta$ must be consistent. $\qquad \square$

---

[4] For example, suppose $\rho|_\eta = \rho_0 + \Omega^\xi$ and $\xi$ 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 $\sup_n \sigma_n = \rho|_\eta$.

# 5 An ordinal notation system for impredicative theories

To carry out an ordinal analysis of $\mathsf{HST}_\beta$ 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, $\mathsf{ID}_1$ [?]. It will turn out that the theories $\mathsf{HST}_\beta$ are substantially weaker than $\mathsf{ID}_1$, but this notation system is still a natural one to consider. The key to generating notations for characteristic ordinals beyond $\Gamma_0$ is the use of constructions referencing certain 'external points'. In our case the 'external point' will be $\Omega$, the first uncountable ordinal.

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

1. $0^* = \Omega^* = 0$,

2. $\alpha^* = \alpha$, if $\alpha < \Omega$ is an $\varepsilon$-ordinal,

3. $\alpha^* = \max\{\gamma^*, \delta^*\}$, if $\alpha =_{\mathrm{NF}} \omega^\gamma + \delta$.

[Elaborate that the following is constructive.]

Define sets of ordinals $C_k(\alpha, \beta)$, and a function $\vartheta \colon \mathbb{O} \to \Omega$ 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 $\xi =_{\mathrm{NF}} \omega^\gamma + \delta$ 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) $\vartheta\alpha = \min\{\xi < \Omega : C(\alpha, \xi) \cap \Omega \subseteq \xi \ \wedge \ \alpha \in C(\alpha, \xi)\}$.

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

**Proposition 5.1.** *$\vartheta\alpha$ 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 $\Omega$. 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 $\Omega$, 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 ($\vartheta$1). $\qquad\square$

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 $\gamma_m$ is a limit ordinal, whence we may deduce

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

for every formula $F$.

The function $\vartheta$ works by "collapsing" ordinals below $\varepsilon_{\Omega+1}$ into countable ordinals, thus allowing one to represent ordinals beyond $\Gamma_0$. Moreover, the condition "$\alpha \in C(\alpha, \xi)$" in ($\vartheta$1) ensures the function $\alpha \mapsto \vartheta\alpha$ is strictly increasing on $\Omega$. Thus, unlike the Veblen functions $\varphi_\alpha$, one never has $\vartheta\beta = \vartheta\vartheta\beta$. In the following proposition we show $\vartheta$ is in fact injective on $\varepsilon_{\Omega+1}$ and show that relation '$\alpha < \beta$' may be decided purely on the normal form presentation for $\alpha$ and $\beta$.

**Proposition 5.2.** *For all ordinals $\alpha$, $\beta$ and $\gamma$ 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 $\varepsilon$-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 5.1 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 $\gamma$. If $\gamma$ is either 0 or $\Omega$, $\gamma^* = 0 \in C(\alpha, \beta)$ by (C1), and if $\gamma$ is an $\varepsilon$-ordinal, $\gamma^* = \gamma$ and we are done. Otherwise $\gamma =_{\mathrm{NF}} \omega^\xi + \delta$ 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

$$\alpha < \beta \rightarrow (\vartheta\alpha < \vartheta\beta \leftrightarrow \alpha^* < \vartheta\beta). \tag{12}$$

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

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

and so, by (vii),

$$\beta < \alpha \rightarrow (\vartheta\alpha < \vartheta\beta \leftrightarrow \vartheta\alpha \leq \beta^*). \tag{13}$$

Combining (**??**) gives (viii). □

We can now proceed with defining a primitive recursive set of ordinal terms for use in the later analysis of $\mathsf{HST}_\beta$.

**Definition 5.3.** Define a subset $\mathrm{OT}_\Omega \subseteq \mathbb{N}$, an encoding $\tau$ of ordinals into $\mathrm{OT}_\Omega$ and a rank function $|.|$ on ordinals by recursion according to the following rules.

1. $\tau(0) = 0 \in \mathrm{OT}_\Omega$, $\tau(\Omega) = \langle 0, 1 \rangle \in \mathrm{OT}_\Omega$, and $|0| = |\Omega| = 0$,

2. If $\alpha = \vartheta\alpha_0$ and $\tau(\alpha_0) \in \mathrm{OT}_\Omega$, $\tau(\alpha) = \langle 1, \tau(\alpha_0) \rangle \in \mathrm{OT}_\Omega$ and $|\alpha| = |\alpha_0| + 1$,

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

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

We now want to define an ordering $<_\vartheta$ on $\mathrm{OT}_\Omega$ such that $\tau(\alpha) <_\vartheta \tau(\beta)$ if and only if $\alpha < \beta$. Conditions (iii) and (vii) of proposition 5.2 ensures every ordinal built up from the constants $0, \Omega$ and functions $\alpha, \beta \mapsto \omega^\alpha + \beta$ and $\alpha \mapsto \vartheta\alpha$ has a unique representation. We may therefore dispense with the function $\tau$ and identify members of $\mathrm{OT}_\Omega$ with the ordinals they represent, as was the case with OT.

Define the relation $\alpha <_\vartheta \beta$ on $\mathrm{OT}_\Omega$ 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 5.2. Let $\alpha <_\vartheta \beta$ if and only if one of the following conditions hold.

1. $\alpha = 0$ and $\beta \neq 0$;

2. $\alpha =_{\mathrm{NF}} \omega^\gamma + \delta$ and either:

    a) $\beta = \Omega$ and $\gamma <_\vartheta \beta$,

    b) $\beta =_{\mathrm{NF}} \omega^{\gamma_0} + \delta_0$ and $\gamma <_\vartheta \gamma_0$, or $\gamma = \gamma_0 \wedge \delta <_\vartheta \delta_0$, or

    c) $\beta = \vartheta\xi$ and $\gamma <_\vartheta \beta$;

3. $\alpha = \vartheta\xi$ and either:

    a) $\beta = \Omega$,

    b) $\beta =_{\mathrm{NF}} \omega^\gamma + \delta$ and $\alpha \leq_\vartheta \gamma$,[5] or

    c) $\beta = \vartheta\eta$ and either, $\xi <_\vartheta \eta \wedge \xi^* <_\vartheta \beta$, or $\eta <_\vartheta \xi \wedge \alpha \leq_\vartheta \eta^*$.

    ($\gamma \leq_\vartheta \delta$ abbreviates $\gamma <_\vartheta \delta$ or $\gamma = \delta$.)

Since the function $\alpha \mapsto \alpha^*$ is primitive recursive, the relation $<_\vartheta$ is also primitive recursive.

Before we proceed with the analysis of $\mathsf{HST}_\beta$, we will show how the ordinals $\alpha \geq \Omega$ in $\mathrm{OT}_\Omega$ enable the generation of the $\varphi_\alpha$ functions for $\alpha < \Gamma_0$ and that $\mathrm{OT}_\Omega$ properly extends OT.

---

[5]Recall that if $\gamma$ is an $\varepsilon$-ordinal, $\beta =_{\mathrm{NF}} \omega^\gamma + \delta$ only if $\delta > 0$, whence $\vartheta\xi <_\vartheta \beta$ if $\vartheta\xi = \gamma$.

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

$$\Omega \cdot \alpha = \omega^{\Omega + \alpha_0} + \cdots + \omega^{\Omega + \alpha_n},$$
$$\Omega^\beta \cdot \alpha = \omega^{\Omega \cdot \beta + \alpha_0} + \cdots + \omega^{\Omega \cdot \beta + \alpha_n}.$$

**Proposition 5.4.** $\mathrm{OT}_\Omega \cap \Omega$ *forms an initial segment of the ordinals.*

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

$$\forall \xi < \alpha (\xi \in \mathrm{OT}_\Omega). \tag{14}$$

If $\alpha = 0$ or $\alpha =_{\mathrm{NF}} \omega^\gamma + \delta$, (14), (C1) and (C2) imply $\alpha \in \mathrm{OT}_\Omega$, so assume $\alpha$ is an $\varepsilon$-ordinal. In search of a contradiction, assume further that $\alpha \notin \mathrm{OT}_\Omega$. We prove $\alpha = \vartheta\xi$ for some $\xi \in \mathrm{OT}_\Omega$. Pick $\beta$ to be the least ordinal in $\{\gamma \leq \delta : \gamma \in \mathrm{OT}_\Omega \wedge \alpha < \gamma\}$. Then $\beta$ must be an $\varepsilon$-ordinal as otherwise $\beta = \omega^{\gamma_0} + \gamma_1$ for some $\gamma_0, \gamma_1 < \beta$, whence $\alpha \leq \max\{\gamma, \delta\} \in \mathrm{OT}_\Omega \cap \Omega$. Since $\beta \in \mathrm{OT}_\Omega$, $\beta = \vartheta\xi$ for some $\xi \in \mathrm{OT}_\Omega$. 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 $\eta =_{\mathrm{NF}} \omega^\gamma + \delta$, 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 (14) and the fact $\xi \in \mathrm{OT}_\Omega$, we obtain $C(\xi, \alpha) \subseteq \mathrm{OT}_\Omega$, so $\eta = \vartheta\zeta \in \mathrm{OT}_\Omega$, whence $\eta < \alpha$ by the choice of $\beta$.

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 $\alpha = \beta \in \mathrm{OT}_\Omega$. $\square$

**Proposition 5.5.**

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

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 5.2; 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 $\beta$. Suppose $\beta < \vartheta\Omega$. Both $\vartheta\Omega$ and $\sup_n \gamma_n$ are $\varepsilon$-ordinals, so the case $\beta =_{\mathrm{NF}} \omega^{\beta_0} + \beta_1$, holds by the induction hypothesis. If $\beta = \vartheta\xi$ we deduce $\xi < \Omega$ and $\xi^* < \vartheta\Omega$, since $\beta < \vartheta\Omega$ and $\Omega^* = 0$. $\xi^*$ has rank strictly less than $\beta$, so the induction hypothesis yields an $m < \omega$ such that $\xi^* < \gamma_m$. Proposition 5.2 (v) then entails $\xi < \gamma_m$. Moreover, since $\gamma_m = (\gamma_m)^* < \vartheta\gamma_m$, by proposition 5.2 (vi), $\gamma_m < \gamma_{m+1}$, so $\xi^* < \vartheta\gamma_m$. Therefore $\beta = \vartheta\xi < \vartheta\gamma_m = \gamma_{m+1}$.

15

(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 5.2 (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 $\beta$. Suppose $\beta < \vartheta\Omega^2$ and $\beta = \vartheta\xi$ for some $\xi$. As $(\Omega^2)^* = 0$ and $\beta < \vartheta\Omega^2$, proposition 5.2 (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 5.2 (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 5.2. 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 $\beta < \gamma_{m+1}$. $\qquad\square$

Proposition 5.5 allows us to identify some characteristic ordinals in terms of both the Veblen and $\vartheta$ functions.

**Corollary 5.6.**

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 $\varepsilon_\alpha$ 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 5.2 (iii). The induction hypothesis yields $\vartheta\alpha = \vartheta\beta$, contradicting $\beta < \alpha$.

(ii) is an immediate consequence of (i) and proposition 5.5.

(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 $\alpha < \Omega$.[6] One can then prove

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

---

[6] $\vartheta\alpha = \varepsilon_{\alpha+1}$ for $\alpha = \varphi 20$ for example.

for $\alpha < \Omega$ by transfinite induction on $\alpha$, using the definition of $\vartheta$. 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 5.5 (ii) implies $\vartheta\Omega^2 = \Gamma_0$.

(iv). Let $\Delta_0$ 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 $\vartheta(\Omega^2 + \Omega) = \Delta_0$. □

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

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

Having established an ordinal notation system suitable for the analysis of the theories $\mathsf{HST}_\beta$, we may now fix the language of $\mathsf{HST}_\beta$. Since the proof-theoretic strength of each theory $\mathsf{HST}_\beta$ 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 $\mathsf{HST}_\beta$ are formulated in the language $\mathscr{L}_\kappa$.

We require a few further technical results about ordinals before we can proceed with the analysis. Suppose $\beta = \Omega^{\alpha_0} \cdot \beta_0 + \cdots + \Omega^{\alpha_n} \cdot \beta_n$ such that $\alpha_0 > \cdots > \alpha_n$ and $\beta_i < \Omega$ for each $i \leq n$. Recall from the previous section that $\beta|_\gamma$ denotes the ordinal $\Omega^{\alpha_0} \cdot \beta_0 + \cdots + \Omega^{\alpha_k} \cdot \beta_k$ 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 5.7.** *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 $\beta + \Omega^\delta < \alpha|_\gamma$.*

# 6 Lower bounds on the proof-theoretic ordinal

We will now establish lower bounds for the theories $\mathsf{HST}_\beta$. This will be achieved by extending the well-ordering proofs used in our analysis of $\mathsf{F}$ (**??**) and $\mathsf{S}_3$ (**??**). Recall $\kappa = \vartheta\varepsilon_{\Omega+1}$.

[Remark that case distinctions below are all decidable]

$\mathsf{HST}_0$ is identical to $\mathsf{F}$, 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 $\mathsf{HST}_\beta$ it will be useful to provide a new proof of the result. For each $\xi < \kappa$ let $\mathsf{wo}_\xi(x)$ denote the formula

$$\forall \ulcorner A(x) \urcorner \forall y \prec x \, \mathsf{T}_\xi(\ulcorner \mathsf{TI}(\dot{y}, A) \urcorner).$$

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

**Lemma 6.1.** $\mathsf{HST}_0^1 \vdash \forall \rho [\mathsf{T}_0(\ulcorner F_0(\dot\rho)\urcorner) \to \mathsf{wo}_0(\vartheta\rho)]$.

*Proof.* Argue within $\mathsf{HST}_0^1$, and assume

$$\mathsf{T}_0(\ulcorner F_0(\dot\rho)\urcorner). \tag{15}$$

Let $\gamma_0 = \rho^* + 1$ and $\gamma_{m+1} = C(\rho, \gamma_m) \cap \Omega$. Moreover, let $\mathsf{wo}_0^1(x)$ denote $\mathsf{T}_0(\ulcorner \mathsf{wo}_0(\dot x)\urcorner)$. (15) implies $\mathsf{wo}_0^1(\gamma_0)$ and $\forall \sigma < \rho[\mathsf{wo}_0^1(\sigma^*) \to \mathsf{wo}_0^1(\vartheta\sigma))]$, whence it is easy to deduce $\forall m[\mathsf{wo}_0^1(\gamma_m) \to \mathsf{wo}_0^1(\gamma_{m+1})]$ and thus $\mathsf{wo}_0^1(\vartheta\rho)$. By $\mathsf{del}_0$, $\mathsf{wo}_0(\vartheta\rho)$ holds. $\square$

**Lemma 6.2.** *For every $m < \omega$, $\mathsf{HST}_0^m \vdash F_0(\Omega \cdot \bar m)$.*

*Proof.* $\mathsf{HST}_0^0 \vdash F_0(\bar 0)$ holds vacuously, so suppose $m = n + 1 > 0$ and

$$\mathsf{HST}_0^n \vdash F_0(\Omega \cdot \bar n). \tag{16}$$

The first step is to establish $\mathsf{HST}_0^n \vdash \mathsf{Prog}F_0^{\Omega\cdot\bar n}$. Argue informally within $\mathsf{HST}_0^n$, assuming $\forall \nu < \mu\, F_0^{\Omega\cdot\bar n}(\nu)$ for some $\mu$, that is,

$$\forall \nu < \mu(\mathsf{wo}_0(\nu^*) \wedge \forall \sigma < \Omega \cdot \bar n + \nu[\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\vartheta\sigma)]). \tag{17}$$

We want to show $\mathsf{wo}_0(\mu^*)$ and $\forall \sigma < \Omega \cdot \bar n + \mu[\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\vartheta\sigma)]$. The former is obvious since the fact $\mathsf{wo}_0(\varepsilon_\alpha)$ is progressive in $\alpha$ is provable in $\mathsf{HST}_0^1$ (*cf.* **??**). To prove the latter, assume $\mathsf{wo}_0(\sigma^*)$ for some $\sigma < \Omega \cdot \bar n + \mu$. If $\mu = 0$ or is a limit ordinal, $\mathsf{wo}_0(\vartheta\sigma)$ is immediate given (17). Otherwise $\mu = \nu + 1$ for some $\nu$, whence we may assume $\sigma = \Omega \cdot \bar n + \nu$. Let $\gamma_0 = \sigma^* + 1$ and $\gamma_{m+1} = C(\sigma, \gamma_m) \cap \Omega$. Then $\gamma_m < \Omega$ for each $m$ and

$$\vartheta\sigma \le \sup_m \gamma_m$$

by proposition 5.1. $\mathsf{wo}_0(\gamma_0)$ is a consequence of $\mathsf{wo}_0(\sigma^*)$, so suppose

$$\mathsf{wo}_0(\gamma_m) \tag{18}$$

with the aim of showing $\mathsf{wo}_0(\gamma_{m+1})$ by induction on the recursive definition of $C(\sigma, \gamma_m)$. Assume $\mathsf{wo}_0(\alpha)$ 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). $\beta \le \gamma_m$, so $\mathsf{wo}_0(\beta)$ is a result of (18).

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

(C3). $\beta =_{\mathrm{NF}} \vartheta\xi$ and $\xi \in C_k(\sigma, \gamma_m) \cap \sigma$. Thus, $\xi^* \in C_k(\sigma, \gamma_m) \cap \Omega$ and therefore $\mathsf{wo}_0(\xi^*)$ by the induction hypothesis. If $\xi < \Omega \cdot n$, $\mathsf{wo}_0(\vartheta\xi)$ is a consequence of (16), otherwise $\Omega \cdot n \le \xi < \sigma$ and $\mathsf{wo}_0(\vartheta\xi)$ is implied by (17).

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

$$\mathsf{HST}_0^n \vdash \mathsf{Prog}F_0^{\Omega\cdot\bar n}. \tag{19}$$

An application of $\mathsf{nec}_0$ entails $\mathsf{HST}_0^m \vdash \mathsf{T}_0(\ulcorner \mathsf{Prog}F_0^{\Omega\cdot\bar n}\urcorner)$, so

$$\mathsf{HST}_0^m \vdash \forall \alpha[\mathsf{wo}_0(\alpha) \to \mathsf{T}_0(\ulcorner F_0(\Omega \cdot \bar n + \dot\alpha)\urcorner)],$$

18

and hence, by lemma 6.1,

$$\mathsf{HST}_0^m \vdash \forall \alpha(\mathsf{wo}_0(\alpha) \to \mathsf{wo}_0 \vartheta(\Omega \cdot \bar{n} + \alpha)). \tag{20}$$

To obtain $\mathsf{HST}_0^m \vdash F_0(\Omega \cdot \bar{m})$ and complete the proof we argue within $\mathsf{HST}_0^m$. Firstly, $\mathsf{wo}_0((\Omega \cdot \bar{m})^*)$ holds trivially as $(\Omega \cdot \bar{m})^* = 0$. Secondly, if $\sigma < \Omega \cdot \bar{m}$, we have either $\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\vartheta \sigma)$ by (16), or $\sigma = \Omega \cdot \bar{n} + \zeta$ for some $\zeta < \Omega$, whence $\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\zeta)$ and $\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\vartheta \sigma)$ results from (20). □

**Corollary 6.3.** $\|\mathsf{HST}_0^m\| \geq \vartheta(\Omega \cdot m)$ *and* $\|\mathsf{HST}_0\| \geq \vartheta(\Omega \cdot \omega)$.

*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 $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\gamma_0})$ holds. Moreover, if $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\alpha})$ for every $\alpha < \gamma_k$ we may deduce $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\alpha})$ for every $\alpha < \gamma_{k+1}$ by induction on the definition of $\gamma_{k+1}$ thus: suppose $\beta \in C_{k+1}(\Omega \cdot m, \gamma_k) \cap \Omega$ and $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\alpha})$ for every $\alpha \in C_k(\Omega \cdot m, \gamma_k) \cap \Omega$. If $\beta$ was enumerated into $C_{k+1}(\Omega \cdot m, \gamma_k)$ by either (C1) or (C2), $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\beta})$ 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 $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\xi}^*)$ by the induction hypothesis, whence lemma 6.2 implies $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\beta})$.

Since $\vartheta(\Omega \cdot m) \leq \sup_{k<\omega} \gamma_k$, we obtain $\mathsf{HST}_0^m \vdash \mathsf{wo}_0(\bar{\alpha})$ for every $\alpha < \vartheta(\Omega \cdot m)$ and so $\mathsf{HST}_0^m \vdash \mathsf{TI}(<\vartheta(\Omega \cdot m))$ by an application of $\mathsf{conec}_0$. □

We will now extend the well-ordering proof above to determine lower bounds on the strength of each theory $\mathsf{HST}_\beta$. 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 $\mathsf{HST}_1$ we require a slightly more general form of lemma 6.2. As its proof makes no explicit use of the fact $m$ is finite, nor any application of $\mathsf{nec}_0$ in showing $\mathsf{Prog} F_0^{\Omega \cdot \bar{m}}$ given $F_0(\Omega \cdot \bar{m})$, we may readily deduce the following generalisation.

**Proposition 6.4.** $\mathsf{HST}_0^1 \vdash \forall \rho[F_0(\rho) \to \mathsf{Prog} F_0^\rho]$.

*Proof.* Argue inside $\mathsf{HST}_0^1$ and assume $F_0(\rho)$ and $\forall \nu < \mu \, F_0^\rho(\nu)$, that is,

$$\mathsf{wo}_0(\rho^*), \tag{21}$$

$$\forall \sigma < \rho[\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0 \vartheta \sigma], \tag{22}$$

$$\forall \nu < \mu(\mathsf{wo}_0(\nu^*) \land \forall \sigma < \rho + \nu[\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0 \vartheta \sigma]), \tag{23}$$

for some $\mu < \Omega$; we want to prove $\mathsf{wo}_0(\mu^*)$ and $\forall \tau < \rho + \mu[\mathsf{wo}_0(\tau^*) \to \mathsf{wo}_0(\vartheta \tau)]$. The former holds immediately given (23) so assume

$$\mathsf{wo}_0(\tau^*), \tag{24}$$

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

Let $\gamma_0 = \tau^* + 1$ and $\gamma_{m+1} = C(\tau, \gamma_m) \cap \Omega$. That $\mathsf{wo}_0(\gamma_m)$ holds for each $m$ will be established by induction on $m$. From (24) one has $\mathsf{wo}_0(\gamma_0)$. Assume $\mathsf{wo}_0(\gamma_m)$. In order to show $\mathsf{wo}_0(\gamma_{m+1})$, assume $\mathsf{wo}_0(\alpha)$ 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 $\alpha_0$ was enumerated into the set by either (C1) or (C2), $\mathsf{wo}_0(\alpha_0)$ 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 $\mathsf{wo}_0(\xi^*)$. If $\xi < \rho$, (22) provides $\mathsf{wo}_0(\vartheta\xi)$. Otherwise $\tau = \rho + \nu$ for some $\nu < \mu$ and $\mathsf{wo}_0(\vartheta\xi)$ holds due to (23). Either way $\mathsf{wo}_0(\alpha_0)$, and so $\mathsf{wo}_0(\alpha)$ for every $\alpha < \gamma_{m+1}$, whence $\mathsf{wo}_0(\gamma_{m+1})$.

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

Proposition 6.4 is plays a key role in the analysis of $\mathsf{HST}_0^m$ and also $\mathsf{HST}_1^1$. Lemma 6.1 entails

$$\mathsf{HST}_0^1 \vdash \mathsf{T}_0(\ulcorner \mathsf{Prog}F_0^{\dot\rho} \urcorner) \to \forall\alpha[\mathsf{wo}_0(\alpha) \to \mathsf{T}_0(\ulcorner F_0^{\dot\rho}(\dot\alpha) \urcorner)]).$$
$$\to \forall\alpha[\mathsf{wo}_0(\alpha) \to \mathsf{wo}_0\vartheta(\rho + \alpha)]),$$

so $\mathsf{HST}_1^1 \vdash \forall\rho[\mathsf{T}_1(\ulcorner F_0(\dot\rho) \urcorner) \to \mathsf{T}_1(\ulcorner F_0(\dot\rho + \Omega) \urcorner)]$. This amounts to proving

$$\mathsf{HST}_1^1 \vdash \forall\rho[F_1(\rho) \to \mathsf{Prog}F_1^\rho] \tag{25} \quad \text{\{eqn:F11.1\}}$$

where $F_1(\rho)$ is the formula $\mathsf{T}_1(\ulcorner F_0(\rho) \urcorner)$ and $F_1^\rho(\alpha)$ denotes $\alpha < \Omega \to F_1(\rho + \Omega \cdot \alpha)$. (25) is sufficient to deduce a lower bound on the strength of the theory $\mathsf{HST}_1^1$.

**Corollary 6.5.** $\|\mathsf{HST}_1^1\| \geq \vartheta\Omega^2$.

*Proof.* Since $\mathsf{HST}_0 \vdash F_0(\bar 0)$, (25) implies

$$\mathsf{HST}_1^1 \vdash \mathsf{Prog}F_1^{\bar 0}. \tag{26} \quad \text{\{eqn:F11lower1\}}$$

Let $\sigma_0 = 1$ and $\sigma_{m+1} = \vartheta(\Omega \cdot \sigma_m)$. By proposition 5.5 (ii), $\vartheta\Omega^2 = \sup_m \sigma_m$, so it suffices to show $\mathsf{HST}_1^1 \vdash \mathsf{wo}_1(\bar\sigma_m)$ for each $m$. This is trivial for $m = 0$; for $m = n + 1$ argue within $\mathsf{HST}_1^1$ assuming $\mathsf{wo}_1(\bar\sigma_n)$. Then $\mathsf{wo}_1(\bar\sigma_n + 1)$ and so $\mathsf{T}_1(\ulcorner F_0(\Omega \cdot \bar\sigma_n) \urcorner)$ by an application of $\mathsf{conec}_1$ and (26). Lemma 6.1 yields $\mathsf{T}_1(\ulcorner \mathsf{wo}_0\vartheta(\Omega \cdot \bar\sigma_n) \urcorner)$ and so $\mathsf{wo}_1(\bar\sigma_m)$ holds. □

Within $\mathsf{HST}_1^2$, the above proof may be replicated under a $\mathsf{T}_1$ predicate, allowing one to reach ordinals beyond $\vartheta\Omega^2$, as the next proposition demonstrates.

**Lemma 6.6.** *For each $m$, $\mathsf{HST}_1^{m+1} \vdash F_1(\Omega^2 \cdot \bar m)$ and $\mathsf{HST}_1^{m+1} \vdash \mathsf{Prog}F_1^{\Omega^2 \cdot \bar m}$.*

*Proof.* $\mathsf{HST}_1^1 \vdash F_1(\bar 0)$ holds trivially, so suppose $m = n + 1$ and $\mathsf{HST}_1^m \vdash F_1(\Omega^2 \cdot \bar n)$. (25) yields $\mathsf{HST}_1^m \vdash \mathsf{Prog}F_1^{\Omega^2 \cdot \bar n}$, whence an application of $\mathsf{nec}_1$ and $\mathsf{T}_1^-$-Rep implies

$$\mathsf{HST}_1^{m+1} \vdash \mathsf{T}_1(\ulcorner \forall\alpha[\mathsf{wo}_0(\alpha) \to \mathsf{T}_0(\ulcorner F_1(\Omega^2 \cdot \bar n + \Omega \cdot \dot\alpha) \urcorner)] \urcorner). \tag{27} \quad \text{\{eqn:F1well.1\}}$$

However, arguing within $\mathsf{HST}_0$, from $F_0(\rho)$ one obtains $\mathsf{wo}_0(\vartheta\rho)$, so (27) entails

$$\mathsf{HST}_1^{m+1} \vdash \mathsf{T}_1(\ulcorner \forall\alpha[\mathsf{wo}_0(\alpha) \to \mathsf{wo}_0\vartheta(\Omega^2 \cdot \bar n + \Omega \cdot \dot\alpha)] \urcorner),$$

and thus $\mathsf{HST}_1^{m+1} \vdash F_1(\Omega^2 \cdot \bar m)$ as required. □

**Theorem 6.** *Suppose $m < \omega$. Then every theorem of $\mathsf{PA} + \mathsf{TI}(<\vartheta(\Omega^2 \cdot m))$ is derivable in $\mathsf{HST}_1^m$. Moreover, every theorem of $\mathsf{PA} + \mathsf{TI}(<\vartheta(\Omega^2 \cdot \omega))$ is derivable in $\mathsf{HST}_1$.*

*Proof.* Since $\vartheta 0 = \varepsilon_0$ and $\mathsf{HST}_1^0$ extends $\mathsf{PA}$, the case $m = 0$ holds, so suppose $m = n + 1 > 0$. If $\mathsf{HST}_1^m \vdash \mathsf{wo}_0(\bar{\alpha})$, lemma 6.6 implies

$$\mathsf{HST}_1^m \vdash \mathsf{wo}_0 \vartheta(\Omega^2 \cdot \bar{n} + \Omega \cdot \bar{\alpha}),$$

whereby if $\sigma_0 = 1$ and $\sigma_{k+1} = \vartheta(\Omega^2 \cdot n + \Omega \cdot \sigma_k)$, $\mathsf{HST}_1^m \vdash \mathsf{wo}_0(\bar{\sigma}_k)$ 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 $\alpha =_{\mathrm{NF}} \omega^\gamma + \delta < \vartheta(\Omega^2 \cdot m)$, 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 $\xi^*$ has rank strictly smaller than $\alpha$ 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 $\mathsf{HST}_p$ for $p < \omega$. Lemma 6.6 essentially shows $\mathsf{HST}_1 \vdash F_1(\Omega^2 \cdot \bar{\alpha})$ implies $\mathsf{HST}_1 \vdash F_1(\Omega^2 \cdot (\bar{\alpha} + 1))$. This can be extended to show $\mathsf{HST}_1 \vdash \forall \nu < \bar{\mu} \, F_1(\Omega^2 \cdot \nu)$ implies $\mathsf{HST}_1 \vdash F_1(\Omega^2 \cdot \bar{\mu})$, whence

$$\mathsf{HST}_2^1 \vdash \mathsf{Prog} F_2^{\bar{0}} \tag{28}$$
<span style="float:right">{eqn:Fbwell1}</span>

where $F_2^\rho(\alpha)$ is the formula $\alpha < \Omega \wedge \mathsf{T}_2(\ulcorner F_0(\rho + \Omega^2 \cdot \alpha)\urcorner)$.

(28) suffices to deduce a lower bound for $\mathsf{HST}_2^1$ and acts as the base step in the analysis of $\mathsf{HST}_2$ and ultimately $\mathsf{HST}_p$, which follows a generalised form of the procedure used in lemma 6.6.

Let $F_p(\rho)$, for $0 < p < \omega$, be the formula $\mathsf{T}_p(\ulcorner F_0(\dot{\rho})\urcorner)$, that is

$$\mathsf{T}_p(\ulcorner \mathsf{wo}_0(\rho^*) \wedge \forall \sigma < \dot{\rho}[\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\vartheta \sigma)]\urcorner),$$

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

<span style="float:left">Fpwellordering1</span> **Lemma 6.7.** *For each $p < \omega$, $\mathsf{HST}_p^1 \vdash \forall \rho[F_p(\rho) \to \mathsf{Prog} F_p^\rho]$ and, for $m < \omega$, $\mathsf{HST}_p^{m+1} \vdash F_p(\Omega^{\bar{p}+1} \cdot \bar{m})$.*

*Proof.* The proof proceeds by induction on $p < \omega$. The case of $p = 0$ has already been shown in proposition 6.4 so suppose $p = q + 1 > 0$. Argue informally within $\mathsf{HST}_p^1$. Assume

$$F_p(\rho), \tag{29}$$
<span style="float:right">{eqn:Fpwell1}</span>

$$\forall \nu < \mu \, F_p^\rho(\nu) \tag{30}$$
<span style="float:right">{eqn:Fpwell2}</span>

for some $\mu < \Omega$. If $\mu = 0$, of course $F_p^\rho(\mu)$ holds by (29), and if $\mu$ is a limit ordinal, (30) implies $\mathsf{T}_p(\ulcorner \forall \nu < \dot\mu \, F_0(\rho + \Omega^{\bar p} \cdot \mu) \urcorner)$, whence $F_p^\rho(\mu)$ is immediate. This leaves only the case in which $\mu$ is a successor ordinal. But for every ordinal $\tau$,

$$
\begin{aligned}
F_p(\tau) &\to \mathsf{T}_p(\ulcorner F_q(\dot\tau) \urcorner), \\
&\to \mathsf{T}_p(\ulcorner \mathsf{Prog} F_q^{\dot\tau} \urcorner), \\
&\to \mathsf{T}_p(\ulcorner \mathsf{T}_0(\ulcorner \mathsf{Prog} F_q^{\dot\tau} \urcorner) \urcorner), \\
&\to \mathsf{T}_p(\ulcorner \forall \alpha [\mathsf{wo}_0(\alpha) \to \mathsf{T}_0(\ulcorner F_q^{\dot\tau}(\dot\alpha) \urcorner)] \urcorner), \\
&\to \mathsf{T}_p(\ulcorner \forall \alpha [\mathsf{wo}_0(\alpha) \to \mathsf{T}_0(\ulcorner F_0(\dot\tau + \Omega^{\bar q} \cdot \dot\alpha) \urcorner)] \urcorner), \\
&\to \mathsf{T}_p(\ulcorner \forall \alpha [\mathsf{wo}_0(\alpha) \to \mathsf{wo}_0(\vartheta(\dot\tau + \Omega^{\bar q} \cdot \alpha))] \urcorner).
\end{aligned}
\tag{31}
$$

<span style="float:right;color:gray;">{eqn:Fpwell3}</span>

The second implication holds on account of the induction hypothesis, while the final holds due to $\mathsf{del}_0$ and lemma 6.1. Given that if $\tau \le \sigma < \tau + \Omega^p$ there exists some $\zeta < \Omega$ such that $\sigma < \tau + \Omega^q \cdot \zeta$ and $\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\zeta)$ (pick $\zeta = (\sigma^*)^{\varepsilon+}$), (31) entails $F_p(\tau) \to \mathsf{T}_p(\ulcorner F_0(\dot\tau + \Omega^{\bar p}) \urcorner)$, <span style="color:gray;">define $\alpha^{\varepsilon+}$</span> that is $\forall \tau [F_p^\tau(\alpha) \to F_p^\tau(\alpha + 1)]$.

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

$$
\mathsf{HST}_p^{m+1} \vdash \mathsf{T}_p(\ulcorner \forall \alpha [\mathsf{wo}_0(\alpha) \to \mathsf{T}_0(\ulcorner F_p(\Omega^{\bar p+1} \cdot \bar n + \Omega^{\bar p} \cdot \dot\alpha) \urcorner)] \urcorner)
$$

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

<span style="color:gray;">m:Fplowerbound</span> **Theorem 7.** *For every $m < \omega$ and $p < \omega$, $\mathsf{HST}_p^m \vdash \mathsf{TI}(< \vartheta(\Omega^{p+1} \cdot m))$ and $\mathsf{HST}_p \vdash \mathsf{TI}(< \vartheta(\Omega^{p+1} \cdot \omega))$.*

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

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

by unravelling the definition of $F_p$ and lemma 6.1.

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

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 $\mathsf{HST}_\beta$ for $\beta \geq \omega$. For $\beta = \omega$ this involves generalising the above proof so that one may derive

$$\mathsf{HST}_\omega \vdash \forall p < \omega\, \mathsf{T}_\omega(\ulcorner \mathsf{Prog} F_p^{\bar{0}} \urcorner), \tag{32}$$

whence $\mathsf{HST}_\omega \vdash \forall p < \omega\, \mathsf{T}_\omega(\ulcorner \forall \alpha[\mathsf{wo}_0(\alpha) \to \mathsf{wo}_0 \vartheta(\Omega^p \cdot \alpha)] \urcorner)$ and also

$$\mathsf{HST}_\omega \vdash \forall \sigma < \Omega^\omega[\mathsf{wo}_\omega(\sigma^*) \to \mathsf{wo}_\omega(\vartheta\sigma))].$$

(32) is not difficult to establish as the proof of lemma 6.7 is clearly uniform in $p < \omega$ and thus can be repeated under a $\mathsf{T}_\omega$ predicate in $\mathsf{HST}_\omega$. But in order to lift the construction to arbitrary $\beta \in \mathrm{OT}_\Omega \cap \Omega$ we require a more general argument.

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

$$\mathsf{T}_\beta(\ulcorner F_0(\dot\rho) \urcorner),$$

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

**Proposition 6.8.** *For each* $\beta \in \mathrm{OT}_\Omega \cap \Omega$, $\mathsf{HST}_\beta^1 \vdash \forall \rho[G_\beta(\rho) \to \mathsf{Prog} G_\beta^\rho]$ *and for every* $m < \omega$, $\mathsf{HST}_\beta^{m+1} \vdash G_\beta(\Omega^{\bar\beta} \cdot \bar m)$.

The proof of proposition 6.8 is by transfinite induction on $\beta$ and requires, for a given $\beta$, the following technical lemmata.

**Lemma 6.9.** $\mathsf{HST}_\beta^1 \vdash \forall \rho[G_\beta(\rho) \to \forall \eta < \bar\beta\, \mathsf{T}_\beta(\ulcorner \mathsf{Prog} G_{\dot\eta}^{\dot\rho} \urcorner)]$.

**Lemma 6.10.** $\mathsf{HST}_\beta^1 \vdash \mathsf{T}_\beta(\ulcorner \forall \tau \forall \eta < \bar\beta\, [F_0(\tau) \wedge \mathsf{T}_0(\ulcorner \mathsf{Prog} G_{\dot\eta}^{\dot\tau} \urcorner) \to F_0(\tau + \Omega^\eta)] \urcorner)$.

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

$$\mathsf{HST}_\beta^1 \vdash G_\beta(\rho) \leftrightarrow \forall \eta < \bar\beta\, \mathsf{T}_\beta(\ulcorner G_\eta(\dot\rho) \urcorner),$$

so the first lemma would result from replicating the proof of (the transfinite induction hypothesis of) proposition 6.8 under a $\mathsf{T}_\beta$ predicate. This is possible as the proof of the proposition, which is presented below, is uniform in $\eta < \beta$.

In order to establish lemma 6.10, argue within $\mathsf{HST}_\beta^1$ *under* the scope of a $\mathsf{T}_\beta$ predicate. Fix $\eta < \bar\beta$, some arbitrary $\tau$ and assume

$$\mathsf{T}_0(\ulcorner \mathsf{Prog} G_{\dot\eta}^{\dot\tau} \urcorner), \tag{33}$$

$$F_0(\tau). \tag{34}$$

(33) entails

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

23

so $\forall\alpha[\mathsf{wo}_0(\alpha) \to \forall\xi < \eta\, \mathsf{T}_0(\ulcorner F_0(\dot\tau + \Omega^{\dot\xi} \cdot \dot\alpha)\urcorner)]$ by $\mathsf{del}_0$ and

$$\forall\xi < \eta\forall\alpha[\mathsf{wo}_0(\alpha) \to \mathsf{wo}_0(\vartheta(\tau + \Omega^\xi \cdot \alpha))] \tag{35}$$

by lemma 6.1. If $\tau \le \sigma < \tau + \Omega^\eta$, there exists some $\zeta < \Omega$ and $\xi < \eta$ such that $\sigma < \tau + \Omega^\xi \cdot \zeta$ and $\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\zeta)$ (pick $\zeta = (\sigma^*)^{\varepsilon+}$), whence $\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\vartheta\sigma)$ results from (35) as $\vartheta\sigma < \vartheta(\tau + \Omega^\xi \cdot \zeta)$. If, however, $\sigma < \tau$, (34) implies $\mathsf{wo}_0(\sigma^*) \to \mathsf{wo}_0(\vartheta\sigma)$. Thus, $F_0(\tau + \Omega^\eta)$ holds. $\qquad\square$

*of proposition 6.8.* Argue informally within $\mathsf{HST}^1_\beta$. Fix some arbitrary $\rho$ and assume $G_\beta(\rho)$. Lemma 6.9 yields $\forall\eta < \bar\beta\, \mathsf{T}_\beta(\ulcorner \mathsf{Prog}G^\rho_\eta\urcorner)$ from which, using $\mathsf{T}^-_\beta$-Rep and lemma 6.10, one obtains $\forall\eta < \bar\beta\, \mathsf{T}_\beta(\ulcorner F_0(\rho + \Omega^\eta)\urcorner)$, that is, $\forall\eta < \bar\beta\, G_\beta(\rho + \Omega^\eta)$. Therefore

$$\mathsf{HST}^1_p \vdash \forall\rho[G^\rho_\beta(\alpha) \to G^\rho_\beta(\alpha + 1)]. \tag{36}$$

If $\mu$ is a limit ordinal, $\forall\nu < \mu\, G^\rho_\beta(\nu)$ and $\mathsf{uni}_\beta$ implies

$$\forall\eta < \bar\beta\, \mathsf{T}_\beta(\ulcorner \forall\nu < \dot\mu\, F_0(\dot\rho + \Omega^{\dot\eta} \cdot \nu)\urcorner),$$

from which $\mathsf{T}_\beta(\ulcorner F_0(\dot\rho + \Omega^{\dot\eta} \cdot \dot\mu)\urcorner)$ is easily deduced. Thus (36) entails

$$\mathsf{HST}^1_\beta \vdash \forall\rho[G_\beta(\rho) \to \mathsf{Prog}G^\rho_\beta].$$

The second half of proposition 6.8 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 $\mathsf{HST}^m_\beta \vdash G_\beta(\Omega^{\bar\beta} \cdot \bar n)$,

$$\mathsf{HST}^{m+1}_\beta \vdash \mathsf{T}_\beta(\ulcorner \mathsf{Prog}G^{\Omega^{\bar\beta} \cdot \bar n}_\beta \urcorner),$$

so $\mathsf{HST}^{m+1}_\beta \vdash G_\beta(\Omega^{\bar\beta} \cdot \bar m)$ by an argument similar to lemma 6.10. $\qquad\square$

Much as in the finite case, proposition 6.8 suffices to obtain lower bounds on the theories $\mathsf{HST}^m_\beta$. In theorem 8 below we prove $\|\mathsf{HST}_\beta\| \ge \vartheta(\Omega^\beta \cdot \omega)$. This is achieved by showing the proof-theoretic ordinal of $\mathsf{HST}^m_\beta$ is closed under the operations $\alpha \mapsto \vartheta(\Omega^\eta \cdot \alpha)$ for every $\eta < \beta$ and extends the strength of $\mathsf{HST}^n_\beta$ for $n < m$. The method will only work if $\beta$ is not "too large", that is $\vartheta\Omega^\beta$ is indeed the *least* ordinal closed under the above operations. By taking $\beta$ to be no larger than the Veblen ordinal $\vartheta\Omega^\Omega$, we can ensure that this is the case.

**Theorem 8.** *For every $\beta < \vartheta\Omega^\Omega$ and every $m < \omega$, $\mathsf{HST}^m_\beta \vdash \mathsf{TI}(<\vartheta(\Omega^\beta \cdot m))$ and $\mathsf{HST}_\beta \vdash \mathsf{TI}(<\vartheta(\Omega^\beta \cdot \omega))$.*

*Proof.* For every $\beta < \vartheta\Omega^\Omega$, the base case, $m = 0$, is immediate since $\mathsf{HST}^0_\beta$ extends PA formulated in the language $\mathscr{L}_\kappa$. Otherwise $m = n + 1 > 0$ and proposition 6.8 implies $\mathsf{HST}^m_\beta \vdash \mathsf{Prog}G^{\Omega^{\bar\beta} \cdot \bar n}_\beta$. Then, given $\mathsf{HST}^m_\beta \vdash \mathsf{wo}_\beta(\bar\alpha)$, we obtain $\mathsf{HST}^m_\beta \vdash \forall\eta < \bar\beta\, G_\beta(\Omega^{\bar\beta} \cdot \bar n + \Omega^{\bar\eta} \cdot \bar\alpha)$, and so

$$\mathsf{HST}^m_\beta \vdash \forall\eta < \bar\beta\, \mathsf{wo}_\beta\,\vartheta(\Omega^{\bar\beta} \cdot \bar n + \Omega^{\bar\eta} \cdot \bar\alpha))),$$

24

by unravelling the definition of $G_\beta$.

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 $\mathsf{HST}^m_\beta \vdash \mathsf{TI}(<\sigma_k)$ 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 $\alpha =_{\mathrm{NF}} \omega^\gamma + \delta$, 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 $\xi^*$ has rank strictly smaller than $\alpha$). 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 $\mathsf{HST}_\beta$ is an immediate consequence of the fact

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

$\square$

# 7 Upper bounds on the proof-theoretic ordinal

We present a ordinal function $f$ on OT such that for all $\rho \in \mathrm{OT}$, all arithmetical theorems of $\mathsf{HST}_\rho$ are derivable in the classical theory $\mathsf{PA}+\mathsf{TI}(<f(\rho))$. By a slightly more involved argument, following the method of [1], the role of Peano arithmetic can be replaced by Heyting arithmetic, HA.

The proof proceeds by formalising the consistency proof of section 4 within the confines of the theory $\mathsf{PA} + \mathsf{TI}(<f(\rho))$. The main technical difficulty is in formalising the co-Necessitation theorem, theorem 4, for two reasons: (1) the proof appeals to transfinite induction beyond the first uncountable ordinal and a subsidiary induction on all countable ordinals, (2) the case in which the last rule employed was $\mathrm{uni}_\eta$ employs non-finitistic properties of $\Omega$, specifically proposition 4.4.

The first concern is directly alleviated by parameterising the height of derivations in $\mathscr{T}_\rho$ from which a suitable choice of $f$ ensures that applications of the subsidiary induction hypothesis are bounded below $f(\rho)$ and all references to the main induction hypothesis in the co-necessitation theorem are for ordinals of the form $\Omega^{\alpha_0} \cdot \beta_0 + \cdots + \Omega^{\alpha_n} \cdot \beta_n$ where $\alpha_i, \beta_i < f(\delta)$ for all $i$.

The second point, however, has no immediate alternative; regularity of $\Omega$ is needed to ensure that the set $\mathscr{T}_{<\rho}$ is closed under the $\omega$-rule in the case that $\rho$ is not an $\Omega$-limit, as this property is required to establish that $\mathscr{T}_\rho$ is closed under $\mathrm{conec}_\eta$. Suppose, for instance, that

$$\frac{A \text{ true literal}}{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, A}\text{Ax1} \qquad \frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, A \quad \big|\!\!\frac{\alpha}{\rho}\, \Gamma, B}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, A \wedge B}\wedge \qquad \frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, A(\bar n) \quad \text{all } n < \omega}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, \forall x A(x)}\omega$$

$$\frac{s^{\mathbb{N}} = t^{\mathbb{N}} \in \text{Sent}}{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, \neg \mathsf{T}_\eta s, \mathsf{T}_\eta t}\text{Ax2} \qquad \frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, A, B}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, A \vee B}\vee \qquad \frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, A(s)}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, \exists x A(x)}\exists$$

Figure 1: Logical rules of the infinitary calculus; in each rule $\beta > \alpha$.

$$\frac{s^{\mathbb{N}} \notin \text{Sent}_\kappa}{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, \neg \mathsf{T}_\eta s}\text{Ax3} \qquad \frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, \mathsf{T}_\eta r \quad \big|\!\!\frac{\alpha}{\rho}\, \Gamma, \mathsf{T}_\eta(r \,\dot{\to}\, s)}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, \mathsf{T}_\eta s}\text{imp}_\eta \qquad \frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, \mathsf{T}_\eta \ulcorner \mathsf{T}_\xi s \urcorner \quad s^{\mathbb{N}} = t^{\mathbb{N}}}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, \mathsf{T}_\eta t}\text{del}_\eta$$

$$\frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, \mathsf{T}_\eta s \quad t^{\mathbb{N}} = \ulcorner \mathsf{T}_\xi s \urcorner^{\mathbb{N}} \in \text{Sent}_\eta}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, \mathsf{T}_\eta t}\text{rep}_\eta \qquad \frac{\big|\!\!\frac{\alpha}{\sigma}\, A \quad \sigma \,\#\, \Omega^{\eta+1} \leq^* \rho \quad t^{\mathbb{N}} = \ulcorner A \urcorner^{\mathbb{N}}}{\big|\!\!\frac{\alpha}{\rho}\, \mathsf{T}_\eta t}\text{nec}_\eta$$

$$\frac{\big|\!\!\frac{\alpha}{\rho}\, \Gamma, \mathsf{T}_\eta \ulcorner A(\bar n) \urcorner \quad \text{for all } n < \omega, \rho|_\eta \text{ not an } \Omega\text{-limit and } s^{\mathbb{N}} = \ulcorner \forall x A(x) \urcorner^{\mathbb{N}}}{\big|\!\!\frac{\beta}{\rho}\, \Gamma, \mathsf{T}_\eta t}\text{uni}_\eta$$

Figure 2: Truth rules of the infinitary calculus; in each case $\beta > \alpha$.

$\rho = \rho' + \Omega^{\xi+1}$, which is not an $\Omega$-limit. If $A(x)$ is a formula of $\mathscr{L}_\kappa$ and $\vdash_{\rho_n} A(\bar n)$ with $\rho_n < \rho$ for every $n$, we want to obtain an ordinal $\sigma < \rho$ such that $\vdash_\sigma \forall x A(x)$. As each application of $\text{nec}_\eta$ in a derivation contributes to the height of a derivation, a bound on the height of a derivation implicitly bounds its T-rank. That is, if $\vdash_\tau B$ with height bound $\alpha$, it will follow that $\vdash_\sigma B$ for some $\sigma \leq \tau$ such that $\sigma^* \leq \alpha$. In particular, if $\alpha$ bounds the height of the derivation $\vdash_{\rho_n} A(\bar n)$ for each $n$, there exists an ordinal $\beta < \Omega$ effectively computed from $\rho$ and $\alpha$ such that $\vdash_\sigma \forall x A(x)$ where $\sigma = \rho' + \Omega^\xi \cdot \beta$.

Although the above arguments are still vague, the reader may find them helpful motivation for a number of the technical results to follow, in particular in the Bounding Lemma (lemma 7.4).

In what follows it will be important to compare ordinals beyond $\Omega$ by their construction as well in addition to their magnitude, for which we introduce a new ordering $<^*$, defined on $\text{OT}_\Omega$ by

$$\rho <^* \sigma \text{ iff } \rho < \sigma \text{ and } \rho^* \leq \sigma^*, \qquad \rho \leq^* \sigma \text{ iff } \rho <^* \sigma \text{ or } \rho = \sigma.$$

**Definition 7.1.** Fix $\kappa = \vartheta \varepsilon_{\Omega+1}$. We define $\big|\!\!\frac{\alpha}{\rho}\, \Gamma$ for $\alpha, \rho \in \text{OT}_\Omega$ with $\alpha < \Omega$ and $\mathscr{L}_\kappa$-sequent $\Gamma$ according to the inference rules in figs. 1 and 2.

Given $\big|\!\!\frac{\alpha}{\rho}\, \Gamma$, we refer to $\Gamma$ as being *derivable* with *T-rank* $\rho$ and *height* $\alpha$. For a fixed $\eta$, the rules $(\text{imp}_\eta)$ to $(\text{uni}_\eta)$ are collectively referred to as $\mathsf{T}_\eta$-*rules*. The collection as a whole (for all $\eta$) is referred to as the *Truth rules*. In each application of a rule or axiom above we specify the *principal*, *minor* and *side* formulae as standard.

The following lemma is an immediate consequence of the definition.

**Lemma 7.2** (Weakening). *If* $\big|\!\!\frac{\alpha}{\rho}\, \Gamma$ *then* $\big|\!\!\frac{\beta}{\sigma}\, \Gamma, \Delta$ *for all* $\Delta$, $\beta \geq \alpha$ *and* $\sigma$ *satisfying* $\rho \leq^* \sigma$.

26

The *rank* of an $\mathscr{L}_\kappa$-sentence $A$, denoted $|A|$, is defined straightforwardly as $|A| = 0$ for $A$ an atomic or negated atomic formula (including $A \in \{\mathsf{T}_\eta s, \neg\mathsf{T}_\eta s \mid \eta < \kappa\}$), $|A * B| = \max\{|A|, |B|\} + 1$ for $* \in \{\wedge, \vee\}$ and $|QxA(x)| = |A(\bar{0})| + 1$ for $Q \in \{\forall, \exists\}$.

Verification of the next proposition is straightforward and, hence, ommitted.

**Proposition 7.3.**

1. *For every $\mathscr{L}_\kappa$-sentence $A$,* $\vdash^{\omega}_{0} \neg A, A$

2. *If $A$ is a false equation, then* $\vdash^{\alpha}_{\sigma} \Gamma, A$ *implies* $\vdash^{\alpha}_{\sigma} \Gamma$ *.*

3. *If $s^{\mathbb{N}} \notin \mathsf{Sent}_\kappa$, then* $\vdash^{\alpha}_{\sigma} \Gamma, \mathsf{T}_\eta s$ *implies* $\vdash^{\alpha}_{\sigma} \Gamma$ *.We use assumptions about coding.*

4. *If* $\vdash^{\alpha}_{\sigma} \Gamma, A(s), A(t)$ *and $s^{\mathbb{N}} = t^{\mathbb{N}}$ then* $\vdash^{\alpha}_{\sigma} \Gamma, A(s)$.

For $\alpha < \Omega$, $\alpha^{\varepsilon+}$ denotes the least $\varepsilon$-number greater than $\alpha$, i.e.,

$$\alpha^{\varepsilon+} = \sup\{\alpha + 1, \omega^{\alpha+1}, \omega^{\omega^{\alpha+1}}, \dots\}.$$

The design of the calculus means that the cut rule is admissible with the same bounds as the arithmetical fragment.

**Theorem 9** (Admissibility of cut). *For every $\alpha, \beta < \Omega$ and every $\sigma \in \mathrm{OT}_\Omega$, if $\vdash^{\alpha}_{\sigma} \Gamma, A$ and $\vdash^{\alpha}_{\sigma} \Delta, \neg A$ then $\vdash^{\beta}_{\sigma} \Gamma, \Delta$ for some $\beta < \alpha^{\varepsilon+}$.*

*Proof.* By the definition, the literals $\neg\mathsf{T}_\eta s$ are principal in only two rules, both axioms. Thus, given $\vdash^{\alpha}_{\sigma} \Gamma, \mathsf{T}_\eta s$ and $\vdash^{\alpha}_{\sigma} \Delta, \neg\mathsf{T}_\eta s$, if the latter holds by virtue of axiom (ax2) with $\neg\mathsf{T}_\eta s$ principal, then $s^{\mathbb{N}} \notin \mathsf{Sent}_\kappa$ and $\vdash^{\alpha}_{\sigma} \Gamma, \Delta$ holds by the previous proposition. If, on the other hand, $\vdash^{\alpha}_{\sigma} \Delta, \neg\mathsf{T}_\eta s$ is an instance of (ax2), then $\vdash^{\alpha}_{\sigma} \Gamma, \Delta$ is either itself an instance of (ax2), or can be obtained from $\vdash^{\alpha}_{\sigma} \Gamma, \mathsf{T}_\eta s$ substitution of equal terms (also proposition 7.3). The remaining cases follow by standard cut-admissibility argument for $\omega$-logic; see, for example, [1]. $\qquad\square$

What remains is to establish a counterpart of the co-necessitation theorem from theorem 4, and confirm that the argument relative to proofs of T-rank $\rho$ can be formalised within the theory $\mathsf{PA} + \mathsf{TI}(<\delta)$ for suitable $\delta$. The first step is bounding property on T-ranks referred to at the beginning of the section.

**Lemma 7.4** (Bounding lemma). *Let $\Theta\sigma$ denote the least $\eta$ such that $\Omega^{\eta+1} > \sigma$. If $\vdash^{\alpha}_{\sigma} \Gamma$ then there exists $\rho \leq \sigma$ such that $\rho^* \leq \max\{\alpha^*, (\Theta\sigma)^*\}$ and $\vdash^{\alpha}_{\rho} \Gamma$.*

*Proof.* By transfinite induction on $\alpha$ in $\vdash^{\alpha}_{\sigma} \Gamma$. In the case of axioms it is always possible to choose $\rho = 0$ and for any rule other than (uni$_\eta$) or (nec$_\eta$) the choice of $\rho$ given by the induction hypothesis suffices for the conclusion. Suppose, therefore, that $\vdash^{\alpha}_{\sigma} \Gamma$ is witnessed by an instance of (nec$_\eta$), meaning that a formula $\mathsf{T}_\eta \ulcorner A \urcorner$ occurs in $\Gamma$ and $\vdash^{\alpha_0}_{\tau} A$ for some $\alpha_0 < \alpha$ and $\tau$ satisfying $\tau \# \Omega^{\eta+1} \leq^* \sigma$. The induction hypothesis yields $\rho_0 \leq \tau$ fulfilling $\rho_0^* \leq \max\{\alpha_0^*, (\Theta\tau)^*\}$ such that $\vdash^{\alpha_0}_{\rho_0} A$, whence we observe setting $\rho = \rho_0 \# \Omega^{\eta+1}$ suffices, as

$$\begin{aligned}
\rho^* &\leq \max\{\rho_0^*, \eta^*\} \\
&\leq \max\{\alpha_0^*, (\Theta\tau)^*, \eta^*\} \\
&\leq \max\{\alpha^*, (\Theta\sigma)^*\}.
\end{aligned}$$

For the case of (uni$_\eta$) suppose that $\sigma|_\eta$ is not an $\Omega$-limit, $\Gamma$ contains the sentence $\mathsf{T}_\eta\ulcorner\forall x\, A(x)\urcorner$ and there exists $\alpha_0 < \alpha$ such that $\vdash^{\alpha_0}_{\sigma} \Gamma, \mathsf{T}_\eta\ulcorner A(\bar{n})\urcorner$ for every $n$. The induction hypothesis yields a $\rho_0 \leq \sigma$ such that $\vdash^{\alpha_0}_{\rho_0} \Gamma, \mathsf{T}_\eta\ulcorner A(\bar{n})\urcorner$ with $\rho_0^* \leq \max\{\alpha_0^*, (\Theta\sigma)^*\}$. If $\rho_0$ is not an $\Omega$-limit take $\rho = \rho_0$, otherwise define $\rho = \rho_0 + \Omega^\xi$, where $\xi$ is such that $\sigma|_\eta = \sigma_0 + \Omega^\xi$ for some $\sigma_0 < \sigma$. In either case $\rho \leq \sigma$ and $\rho^* = \max\{\rho_0^*, \xi^*\} \leq \max\{\alpha^*, (\Theta\sigma)^*\}$. □

The next step is to define a collection of standard $\mathscr{L}_\kappa$-structures $\mathfrak{N}_\rho = \langle\, \mathbb{N}, (\mathscr{T}_{\rho,\eta})_{\eta\in\mathrm{OT}} \,\rangle$ for $\rho < \Omega^\kappa$ as surrogates for the structures $\mathfrak{M}_\rho$ from theorem 4. The interpretation of the predicate $\mathsf{T}_\eta$ in $\mathfrak{N}_\rho$ is given by the set

$$\mathscr{T}_{\rho,\eta} = \big\{\, \ulcorner A\urcorner \mid \text{there exist } \alpha, \sigma \text{ such that } \max\{\sigma^*, \alpha\} < \vartheta\rho, \ \sigma < \rho|_\eta \text{ and } \vdash^\alpha_\sigma A \,\big\}.$$

<div style="color: gray">prop:T8mono</div>

**Proposition 7.5.** *If $A$ is T-positive, $\rho \leq^* \sigma$ and $\mathfrak{N}_\rho \models A$ then $\mathfrak{N}_\sigma \models A$.*

<div style="color: magenta; text-align: right">Defined T-positive?</div>

*Proof.* Suppose $\rho \leq^* \sigma$. As $\rho^* < \vartheta\sigma$, we have $\vartheta\rho \leq \vartheta\sigma$ by proposition 5.2 (viii), which coupled with the fact that $\rho|_\eta \leq \sigma|_\eta$ implies that $\mathscr{T}_{\rho,\eta} \subseteq \mathscr{T}_{\sigma,\eta}$ for every $\eta < \kappa$. That $\mathfrak{N}_\rho \models A$ then $\mathfrak{N}_\sigma \models A$ for every T-positive sentence $A$ follows by induction. □

The structure $\mathfrak{N}_\rho$ is so chosen to provide a model of axioms of the theory $\mathsf{HST}_\rho$ for certain $\rho$. This claim witnessed by the next result.

<div style="color: gray">lem:T8Mprop</div>

**Lemma 7.6.** *The following hold for every $\alpha < \Omega$, $\eta < \kappa$ and $\rho < \Omega^\kappa$.*

<div style="color: gray">eqn:T8propii</div>

    *1.* $\mathfrak{N}_\rho \models \forall x \forall y (\mathsf{T}_\eta x \wedge \mathsf{T}_\eta (x \dot{\to} y) \to \mathsf{T}_\eta y)$.

<div style="color: gray">eqn:T8propiii</div>

    *2.* $\mathfrak{N}_\rho \models \forall x (\mathsf{T}_\eta x \to \forall \xi < \bar{\eta}\, \mathsf{T}_\eta \ulcorner \mathsf{T}_\xi \dot{x}\urcorner)$.

<div style="color: magenta; text-align: right">More coding</div>

    *3.* $\mathfrak{N}_\rho \models \forall \xi < \bar{\kappa} \forall x (\mathsf{T}_\eta \ulcorner \mathsf{T}_\xi \dot{x}\urcorner \to \mathsf{T}_\eta x)$ *if, for every $\sigma < \rho$ and $\beta$ such that $\max\{\sigma^*, \beta\} < \vartheta\rho$,*

<div style="color: gray">eqn:T8propiv</div>

$$\vdash^\beta_\sigma \Gamma \ \text{implies} \ \mathfrak{N}_{\sigma+\beta} \models \bigvee \Gamma.$$

<div style="color: gray">eqn:T8propvi</div>

    *4. If $\rho|_\eta$ is not an $\Omega$-limit, $\mathfrak{N}_\rho \models \forall x\, \mathsf{T}_\eta \ulcorner A(\dot{x})\urcorner$ and $\rho <^* \sigma$, then $\mathfrak{N}_\sigma \models \mathsf{T}_\eta \ulcorner \forall x\, A(x)\urcorner$.*

*Proof.* 1. Observe that if $\mathfrak{N}_\rho \models \mathsf{T}_\eta \ulcorner A\urcorner \wedge \mathsf{T}_\eta \ulcorner A \to B\urcorner$, then there is an $\alpha < \vartheta\rho$ and $\sigma < \rho|_\eta$ such that $\sigma^* < \vartheta\rho$ and

$$\vdash^\alpha_\sigma A \ \text{and} \ \vdash^\alpha_\sigma \neg A, B.$$

Admissibility of cut, theorem 9, implies $\vdash^\gamma_\sigma B$ for some $\gamma < \alpha^{\varepsilon+}$. As $\rho > \Omega$, we have $\gamma < \vartheta\alpha < \vartheta\rho$ and hence $\mathfrak{N}_\rho \models \mathsf{T}_\eta \ulcorner B\urcorner$. As $\mathfrak{N}_\rho$ is standard, the quantified form holds also.

2. Fix some $\eta < \kappa$. Since $\rho^* < \vartheta\rho$ it follows that $(\rho|_\eta)^* < \vartheta\rho$. If $\rho|_\eta = 0$, then $\mathscr{T}_{\rho,\eta} = \emptyset$ and the claim holds vacuously. Otherwise, $\eta^* \leq (\rho|_\eta)^*$ and suppose

$$\vdash^\alpha_\sigma A \ \text{where} \ \sigma < \rho|_\eta \text{ and } \alpha, \sigma^* < \vartheta\rho.$$

Fix an arbitrary $\xi < \eta$. Clearly, $\sigma \,\#\, \Omega^{\xi+1} < \rho|_\eta$, whence

$$(\sigma \,\#\, \Omega^{\xi+1})^* \leq \max\{\sigma^*, \eta^*\} < \vartheta\rho,$$

so $\vdash^{\alpha+1}_{\sigma\#\Omega^{\xi+1}} \mathsf{T}_\xi \ulcorner A\urcorner$, and $\mathsf{T}_\xi \ulcorner A\urcorner \in \mathscr{T}_{\rho,\eta}$ as required.

28

3. Suppose $\mathfrak{N}_\rho \models \mathsf{T}_\eta \ulcorner \mathsf{T}_\xi s \urcorner$ and $\xi < \kappa$. We claim that $\mathfrak{N}_\rho \models \mathsf{T}_\eta s$. By definition there exist $\alpha$ and $\sigma < \rho|_\eta$ satisfying $\max\{\sigma^*, \alpha\} < \vartheta\rho$ and

$$\vdash^{\alpha}_{\sigma} \mathsf{T}_\xi s.$$

The additional assumption implies that $\mathfrak{N}_{\sigma+\alpha} \models \mathsf{T}_\xi s$, whence $s^{\mathbb{N}} = \ulcorner A \urcorner$ for some $A$ and there is a $\beta$ and $\tau < \sigma|_\xi$ satisfying

$$\max\{\tau^*, \beta\} < \vartheta(\sigma + \alpha) \text{ and } \vdash^{\beta}_{\tau} A.$$

From $\max\{\sigma^*, \alpha\} < \vartheta\rho$ and $\sigma + \alpha < \rho$, we deduce

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

and thus $\max\{\tau^*, \beta\} < \vartheta\rho$. As, moreover, $\tau < \rho|_\eta$ we conclude $\mathfrak{N}_\rho \models \mathsf{T}_\eta s$.

4. Suppose $\mathfrak{N}_\rho \models \forall x\, \mathsf{T}_\eta \ulcorner A(\dot{x}) \urcorner$, that $\rho|_\eta$ is not an $\Omega$-limit. Let $\rho'$ be such that $\rho <^* \rho'$. As $\mathscr{T}_{\rho,\eta}$ is non-empty, $\rho|_\eta > 0$ and there exist ordinals $\alpha_n$ and $\rho_n$ for each $n < \omega$ such that

$$\rho_n < \rho|_\eta,$$
$$\max\{\alpha_n, \rho_n^*\} < \vartheta\rho,$$
$$\vdash^{\alpha_n}_{\rho_n} A(\bar{n}).$$

As $\rho|_\eta$ is not an $\Omega$-limit there exists $\xi$ and $\tau$ such that $\rho|_\eta = \tau + \Omega^{\xi+1}$. Let $\sigma = \tau + \Omega^\xi \cdot \vartheta\rho$. As $\rho_n^* < \sigma^*$ Weakening implies that

$$\vdash^{\vartheta\rho}_{\sigma} A(\bar{n}) \text{ for every } n < \omega,$$

whereby $\vdash^{\vartheta\rho+1}_{\sigma} \forall x A(x)$ follows. Clearly, $\sigma < \rho|_\eta \leq \rho'|_\eta$ and $\sigma^* \leq \max\{\rho^*, \vartheta\rho\} < \vartheta\rho'$. Therefore, $\mathfrak{N}_{\rho'} \models \mathsf{T}_\eta \ulcorner \forall x\, A(x) \urcorner$ as desired. $\qquad\square$

lem:T8fi5dleh

**Theorem 10** (T-Elimination theorem). *For every $\rho \in \mathrm{OT}_\Omega$ and $\alpha \in \mathrm{OT}_\Omega \cap \Omega$,*

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

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

3. $\mathscr{T}_\infty \vdash^{\alpha}_{\rho} \neg \mathsf{T}_\eta s$ *implies there is an $\mathscr{L}_\mathsf{T}$-sentence $A$ with $s^{\mathbb{N}} = \ulcorner A \urcorner$ such that $\mathscr{T}_\infty \vdash^{\alpha}_{\rho} \neg A$.*

*Proof.* By *main* transfinite induction on $\rho$ with a *subsidiary* transfinite induction on $\alpha$. We begin with (i).

First observe that since $\mathfrak{N}_\rho$ is standard all arithmetical rules and axioms are dealt with easily. Proposition 7.5 deals with (Weakening). Moreover, no instances of (Ax.$2_\eta$) and (Ax.$3_\eta$) are T-positive, so $\Gamma$ is not an instance of these axioms. Thus we need only consider the T-rules. Suppose $\mathscr{T}_\infty \vdash^{\alpha}_{\rho} \Gamma$ is derived via a $\mathsf{T}_\eta$-rule. If the last rule applied was (nec$_\eta$), there is an $\mathscr{L}_\kappa$-sentence $A$ and term $s$ with $s^{\mathbb{N}} = \ulcorner A \urcorner$ so that $\mathsf{T}_\eta s$ is in $\Gamma$, and $\mathscr{T}_\infty \vdash^{\beta}_{\sigma} A$ holds for some $\beta < \alpha$ and $\sigma$ 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 7.6 entails $\mathfrak{N}_{\rho+\alpha} \models \bigvee \Gamma$ is a consequence of the subsidiary induction hypothesis and, in the case of $(\text{del}_\eta)$, the main induction hypothesis.

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

It is easy to check that $\mathscr{T}_\infty \vdash^\alpha_\rho \neg T_\eta s$ can only hold if $s^\mathbb{N}$ is not the code of some $\mathscr{L}_\kappa$-sentence, whence (iii) holds vacuously. $\qquad\square$

Having established that $\mathscr{T}_\infty$ is also closed under $\text{conec}_\eta$ for every $\eta < \kappa$ we can proceed with embedding each $\mathsf{HST}_\eta$ into $\mathscr{T}_\infty$. This is the focus of the following proposition and theorem.

<span style="font-size:small">prop:T8embed</span> **Proposition 7.7.** *For each axiom $A$ of $\mathsf{HST}_\beta$, $\mathscr{T}_\infty \vdash^{\alpha}_{\Omega^{\beta+1}\cdot n} A$, for some $\alpha < \varepsilon_0$ and every $n < \omega$.*

<span style="font-size:small">thm:T8Embed</span> **Lemma 7.8** (Embedding lemma for $\mathscr{T}_\infty$). *Fix some $\beta < \vartheta\varepsilon_{\Omega+1}$ and suppose $\mathsf{HST}_\beta \vdash A$ with $A$ a sentence of $\mathscr{L}_\beta$. Then there are ordinals $\alpha$ and $\rho$ such that $\mathscr{T}_\infty \vdash^\alpha_\rho A$, $\rho < \Omega^{\beta+1}\cdot\omega$, $\alpha, \rho^* < \vartheta(\Omega^{\beta+1}\cdot\omega)$ and $\rho = \rho|_0$.*

*Proof.* We proceed by induction on the length of the derivation $\mathsf{HST}_\beta \vdash A$.

**Case I.** $\beta = 0$. This essentially follows from the embedding of $\mathsf{F}$ in $\mathscr{F}_\infty$, **??**, since $\vartheta(\Omega\cdot\omega) = \varphi\omega 0$. Proposition 7.7 deals with the axioms of $\mathsf{HST}_0$. According to the induction hypothesis, an application of *modus ponens* in $\mathsf{HST}_0$ means one has

$$\mathscr{T}_\infty \vdash^{\alpha_0}_{\Omega\cdot m_0} B, \qquad\qquad \mathscr{T}_\infty \vdash^{\alpha_1}_{\Omega\cdot m_1} \neg B, A$$

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

$$\mathscr{T}_\infty \vdash^{\gamma}_{\Omega\cdot m} B, \qquad\qquad \mathscr{T}_\infty \vdash^{\gamma}_{\Omega\cdot m} \neg B, A$$

where $\gamma = \max\{\alpha_0, \alpha_1\}$ and $m = \max\{m_0, m_1\}$, whence $\mathscr{T}_\infty \vdash^{\alpha}_{\Omega\cdot m} A$ 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 $\text{nec}_0$ in $\mathsf{HST}_0$ corresponds to an application of the rule $(\text{nec}_0)$ in $\mathscr{T}_\infty$. It is an immediate consequence of the induction hypothesis that in this case the desired bounds are maintained.

This leaves only applications of $\text{conec}_0$ and $\neg\text{conec}_0$ in $\mathsf{HST}_0$ to consider, for which we employ theorem 10. $\neg\text{conec}_0$ is immediate, so suppose $\mathsf{HST}_0 \vdash T_0(\ulcorner A \urcorner)$ for some $\mathscr{L}_\kappa$-sentence $A$. The induction hypothesis implies

$$\mathscr{T}_\infty \vdash^{\alpha}_{\Omega\cdot m} T_0(\ulcorner A \urcorner) \tag{37}$$

<span style="font-size:small">{eqn:T8embed3}</span>

for some $\alpha < \vartheta(\Omega \cdot \omega)$ and $m < \omega$. By theorem 10 one obtains some $\gamma < \vartheta(\Omega \cdot m + \alpha)$ and $n < m$ such that

$$\mathscr{T}_\infty \vdash^{\gamma}_{\Omega\cdot n} A$$

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 7.7 deals with the axioms of $\mathsf{HST}_\beta$. According to the induction hypothesis, an application of *modus ponens* in $\mathsf{HST}_\beta$ means one has

$$\mathscr{T}_\infty \vdash^{\alpha_0}_{\sigma} B, \qquad\qquad \mathscr{T}_\infty \vdash^{\alpha_1}_{\tau} \neg B, A$$

30

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

$$\mathscr{T}_\infty \vdash^{\alpha}_{\rho} B, \qquad\qquad\qquad \mathscr{T}_\infty \vdash^{\alpha}_{\rho} \neg B, A$$

where $\rho = \sigma \# \tau$ and $\alpha = \max\{\alpha_0, \alpha_1\}$, whence one obtains $\mathscr{T}_\infty \vdash^{\gamma}_{\rho} A$ 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 $\mathrm{nec}_\beta$ was utilised in $\mathsf{HST}_\beta$, $A$ is $\mathsf{T}_\beta(\ulcorner B \urcorner)$ for some sentence $B$ and the induction hypothesis yields

$$\mathscr{T}_\infty \vdash^{\alpha}_{\rho} B$$

for some $\alpha < \vartheta(\Omega^{\beta+1} \cdot \omega)$ and $\rho < \Omega^{\beta+1} \cdot \omega$. An application of the rule $(\mathrm{nec}_\beta)$ in $\mathscr{T}_\infty$ yields

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

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

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

so the desired bounds are maintained.

This leaves only applications of $\mathrm{conec}_\beta$ and $\neg\mathrm{conec}_\beta$ in $\mathsf{HST}_\beta$ to consider, for which we employ theorem 10. The case of $\neg\mathrm{conec}_\beta$ is immediate given the induction hypothesis, so suppose $\mathsf{HST}_\beta \vdash \mathsf{T}_\beta(\ulcorner A \urcorner)$ for some $\mathscr{L}_\kappa$-sentence $A$. The induction hypothesis entails

$$\mathscr{T}_\infty \vdash^{\alpha}_{\rho} \mathsf{T}_\beta(\ulcorner A \urcorner) \tag{38}$$

for some $\rho < \Omega^{\beta+1} \cdot \omega$ and $\alpha, \rho^* < \vartheta(\Omega^{\beta+1} \cdot \omega)$. By theorem 10 (ii) one obtains some $\gamma$ and $\sigma$ such that $\sigma^*, \gamma < \vartheta(\rho + \alpha)$, $\sigma < \rho|_\beta$ and

$$\mathscr{T}_\infty \vdash^{\gamma}_{\sigma} A$$

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 $\gamma, \sigma^* < \vartheta(\Omega^{\beta+1} \cdot \omega)$. $\qquad\square$

Lemma 7.8 allows us to easily determine an upper bound on the strength of the theories $\mathsf{HST}_\beta$. 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 11.** *Every arithmetical theorem of* $\mathsf{HST}_\beta$, *for* $\beta < \vartheta\varepsilon_{\Omega+1}$, *is derivable in* $\mathsf{PA} + \mathsf{TI}(<\vartheta(\Omega^{\beta+1} \cdot \omega))$.

# 8 Conclusion

We have introduced a hierarchy $(\mathsf{HST}_\alpha)_{\alpha \in \mathbb{O}}$ of intuitionistic theories of self-applicable truth indexed by ordinals from a fixed elementary ordinal notation system $\mathbb{O}$ and argued that $\mathsf{HST}_\alpha$ formalises an intensional acceptance of the theories $\mathsf{HST}_\beta$ for $\beta < \alpha$.

Presented with such a hierarchy of theories $(\mathsf{T}_\alpha)_{\alpha \in \mathbb{O}}$ it is natural to ask the limit of the corresponding autonomous progression, that is, the least ordinal not in the set $X_\mathsf{T}$ generated by the operation $\|\mathsf{T}_0\| \subseteq X_\mathsf{T}$ and $\alpha \in X_\mathsf{T}$ implies $\|\mathsf{T}_\alpha\| \subseteq X_\mathsf{T}$. Autonomous progressions of ramified theories such as $\mathsf{RA}_\alpha$ are well-studied, as are those obtained by iterating reflection principles [?]. In the case $\mathsf{T}_\alpha = \mathsf{HST}_\alpha$ this is not difficult to determine given theorem 13.

Combining the results of the previous section we determine the strength of the theory $\mathsf{HST}_\beta$ for every $\beta < \vartheta\Omega^\Omega$.

**Theorem 12.** *For every $p < \omega$, $\|\mathsf{HST}_p\| = \vartheta(\Omega^{p+1} \cdot \omega)$.*

**Theorem 13.** *For every $\beta \geq \omega$ with $\beta < \vartheta\Omega^\Omega$,*

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

*Proof.* Theorem 11 provides both upper bounds for the two theorems. The lower bounds are a corollary of theorem 7 and theorem 13 respectively. $\square$

**Theorem 14.** *The limit of the autonomous progression defined from $\{\mathsf{HST}_\beta : \beta < \Omega\}$ is the large Veblen ordinal, $\vartheta\Omega^\Omega$.*

*Proof.* Let $\sigma_0 = 0$ and $\sigma_{m+1} = \vartheta(\Omega^{\sigma_m})$. Theorem 8 implies $\|\mathsf{HST}^1_{\sigma_m}\| \geq \sigma_{m+1}$, while theorem 13 entails $\|\mathsf{HST}_{\sigma_m}\| < \sigma_{m+2}$, so $X_\mathsf{F} = \sup_m \sigma_m$. 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 $\alpha$. If $\alpha =_\mathrm{NF} \omega^\gamma + \delta$ for some $\gamma, \delta$, 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. $\square$

32

Our motivation for defining the theory $\mathsf{HST}_\beta$ as we did stemmed from the idea of formalising the acceptance of $\mathsf{F}$. The theory $\mathsf{S}_3$ with just one truth predicate appears to almost achieve this, but the general inability to close $\mathsf{S}_3$ under the rule T-Intro means the truth predicate no longer satisfies the same principles as it did in $\mathsf{F}$. This lead us to consider stratifying the language, viewing the original predicate of $\mathsf{F}$, now $\mathsf{T}_0$, as the base level and gradually extending the language by including predicates $\mathsf{T}_1$, $\mathsf{T}_2$, etc. in such a way that each predicate $\mathsf{T}_\eta$ in the language locally satisfies the same axioms and rules as $\mathsf{T}_0$.

The analysis of the theories $\mathsf{HST}_\beta$ 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, theorem 4 and **??** show the truth predicates of $\mathsf{HST}_\beta^1$ may be treated as identical; they can all be interpreted as the set $\mathscr{T}_{<\Omega^{\beta+1}}$. Within $\mathsf{HST}_\beta^2$ they may all be interpreted as the set $\mathscr{T}_{<\Omega^{\beta+1}\cdot 2}$ and, in general, all truth predicates in $\mathsf{HST}_\beta$ can be interpreted as the set $\mathscr{T}_{<\Omega^{\beta+2}}$ (one cannot simply use $\mathscr{T}_{<\Omega^{\beta+1}\cdot\omega}$ for the interpretation of $\mathsf{T}_\beta$ in $\mathsf{HST}_\beta$ as the set is not closed under the $\omega$-rule, whereas $\mathscr{T}_{<\Omega^{\beta+2}}$ is, as well being closed under $\mathrm{conec}_\eta$, $\mathrm{nec}_\eta$ for every $\eta \leq \beta$.) The upshot is that we may view each predicate $\mathsf{T}_\eta$ as "extending" the base predicate $\mathsf{T}_0$ as well as $\mathsf{T}_\xi$ for $\xi < \eta$. It would be interesting to determine whether the theory $\mathsf{HST}_\beta$ can be rewritten in some natural type-free form.

The model constructions employed in the previous section for the analysis of $\mathsf{HST}_1$ allow us to obtain an upper bound for the theory $\mathsf{S}_3$ introduced in **??**. Essentially, we stratify the language $\mathscr{L}_\mathsf{T}$ as described in remark 1, interpreting the outermost truth predicate by $\mathsf{T}_1$ and all others by $\mathsf{T}_0$, but by first embedding $\mathsf{S}_3$ in an infinitary theory formulated without T-Elim, we avoid the problems relating to $\mathrm{conec}_0$ and $\mathrm{conec}_1$.

<span style="font-variant: small-caps; color: gray;">thm:S3strength</span> **Theorem 15.** $\mathsf{S}_3$ *proves the same arithmetical statements as* $\mathsf{HST}_1^1$ *and hence has proof-theoretic ordinal* $\Gamma_0$.

*Proof.* We define an infinitary proof system $\mathscr{S}_\infty$ based on $\mathscr{T}_\infty$ into which we may embed $\mathsf{S}_3$. Let $^*$ be the interpretation of $\mathscr{L}_0$ into $\mathscr{L}_\mathsf{T}$ that recursively interprets the predicate $\mathsf{T}_0$ as $\mathsf{T}$ and otherwise commutes with all connectives and quantifiers. Define $\mathscr{S}_\infty \vdash^\alpha \Gamma$ according to the rules (Ax.1), (Ax.2), (Ax.3), ($\wedge$), ($\vee_i$), ($\omega$), ($\exists$), (T-Imp), (T-Del), (T-Rep), (T-$\omega$), and the following additional rule

$(\mathrm{Ax.4})^\mathscr{T}$ $\quad \mathscr{T}_\infty \vdash^\beta_{\Omega\cdot\gamma} A$ for some $\gamma, \beta < \Omega$ implies $\mathscr{S}_\infty \vdash^\alpha \Gamma, \mathsf{T}(\ulcorner A^* \urcorner)$ for any $\alpha > \max\{\beta, \gamma\}$.

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

$$\mathscr{T}_\infty \vdash^\alpha_{\Omega\cdot\gamma} \Gamma \text{ implies } \mathscr{S}_\infty \vdash^\alpha \Gamma^*. \tag{39}$$

<span style="color: gray;">{eqn:S81}</span>

Define a sequence of $\mathscr{L}_\mathsf{T}$-structures

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

We claim

$$\mathscr{S}_\infty \vdash^\alpha \Gamma \text{ implies } \mathfrak{M}_\alpha \models \bigvee \Gamma \tag{40}$$

<span style="color: gray;">{eqn:S82}</span>

whenever $\Gamma$ is T-positive. The proof proceeds by transfinite induction on $\alpha$. If $\Gamma$ is an instance of $(\mathrm{Ax.4})^\mathscr{T}$, $\mathfrak{M}_\alpha \models \bigvee \Gamma$ holds by definition, while if $\Gamma$ is derived through an application of

(T-Rep), it follows from the induction hypothesis and closure of $\mathscr{T}_\infty$ under (nec$_0$). If the last applied rule is (T-$\omega$), $\mathfrak{M}_\alpha \models \bigvee \Gamma$ holds by an application of ($\omega$) in $\mathscr{T}_\infty$ and the fact $\vartheta(\Omega \cdot \alpha)$ is increasing in $\alpha$. Furthermore, by the definition of the function $\vartheta$, we have $\vartheta(\Omega \cdot \delta + \beta) < \vartheta(\Omega \cdot \alpha)$ whenever $\delta < \alpha$ and $\beta < \vartheta(\Omega \cdot \alpha)$; thus theorem 10 implies

$$\mathfrak{M}_\alpha \models \forall \ulcorner A \urcorner (\mathsf{T}(\ulcorner \mathsf{T}(\ulcorner A \urcorner) \urcorner) \to \mathsf{T}(\ulcorner A \urcorner)),$$

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

(40) can now be utilised along with (39) to conclude

$$\mathscr{S}_\infty \vdash^\alpha \mathsf{T}(\ulcorner A \urcorner) \text{ implies } \mathscr{S}_\infty \vdash^{\vartheta(\Omega \cdot \alpha)} A.$$

Since $\vartheta \Omega^2$ is the least ordinal closed under the function $\alpha \mapsto \vartheta(\Omega \cdot \alpha)$ (see proposition 5.5) we may deduce $\mathsf{S}_3 \vdash A$ implies $\mathscr{S}_\infty \vdash^\alpha A$ for some $\alpha < \vartheta \Omega^2$ for any sentence $A$, whence $\|\mathsf{S}_3\| \leq \vartheta \Omega^2$. Finally, note $\vartheta \Omega^2 = \Gamma_0$ by corollary 5.6. $\qquad\square$

## References

LeiRat10  [1] Leigh, G.E. and M. Rathjen, An ordinal analysis of theories of self-applicable truth. *Arch. Math. Logic* (2010) 49:213–247

Lei-thesis  [2] *Proof-theoretic investigations into the Friedman–Sheard theories and other theories of truth*, Ph.D. thesis, School of Mathematics, University of Leeds, 2010.