# Epsilon Substitution for $ID_1$

Henry Towsner

May 24, 2005

**Abstract**

Hilbert's epsilon substitution method provides a technique for showing that a theory is consistent by producing progressively more accurate computable approximations to the non-computable components of a proof. If it can be shown that this process eventually halts with a sufficiently good approximation, the theory is consistent.

Here we produce a new formulation of the method for the theory $ID_1$ of inductive definitions which simplifies the proof given in [Ara03], and prove termination using the cut-elimination method of [MTB96].

## 1 Introduction

Hilbert introduced the epsilon calculus, c.f. [HB70] to provide a method for proving the consistency of arithmetic and analysis. In place of the usual quantifiers, a symbol $\epsilon$ is added, allowing terms of the form $\epsilon x\phi[x]$, which are interpreted as "some $x$ such that $\phi[x]$ holds, if such a number exists." When there is no $x$ satisfying $\phi[x]$, we allow $\epsilon x\phi[x]$ to take an arbitrary value (usually 0). Then the existential quantifier can be defined by

$$\exists x\phi[x] \Leftrightarrow \phi[\epsilon x\phi[x]]$$

and the universal quantifier by

$$\forall x\phi[x] \Leftrightarrow \phi[\epsilon x\neg\phi[x]]$$

Hilbert proposed a method for transforming non-finitistic proofs in this epsilon calculus into finitistic proofs by assigning numerical values to all the epsilon terms, making the proof entirely combinatorial.

The difficulty centers on the critical formulas, axioms of the form

$$\phi[t] \rightarrow \phi[\epsilon x\phi[x]]$$

In Hilbert's method, the (finite) list of critical formulas which appear in a proof is considered. Then a finite series of functions—$\epsilon$-substitutions—is defined, each providing values for some of the of the epsilon terms appearing in the list. Each of these substitutions will satisfy some, but not necessarily all, of the critical formulas appearing in proof. At each step we take the simplest unsatisfied critical formula and update an appropriate epsilon term so that it becomes true. The resulting series of substitutions is called the $H$-process.

The $H$-process can only halt when the final substitution makes every critical formula, and therefore every formula of the proof, true. Then, when every epsilon term in a proof of some formula is replaced by its value under the substitution, we have a purely numerical proof of the new formula. Typically, we start with a proof of $\exists x\phi[x]$ where $\phi$ is quantifier free. At the conclusion of the $H$-process, we have $\phi[n]$ and a proof that $n$ does, in fact, satisfy $\phi$.

If it can be shown that this process terminates for every (finite) set of starting formulas, then we have also proven the consistency of our theory.

After several attempts, Ackermann [Ack40] eventually proved that the provess terminates for first order arithmetic, and therefore that a substitution of numerical values for all infinitary terms can be found using a finite process.

For several decades, little work was done applying this technique to more powerful theories, in preference to other techniques for proving the consistency of theories, chiefly cut-elimination ([Gen34],[Gen36]). These techniques were extended to the impredicative theory of $\Pi^1_1$-comprehension in [Tak67]. This work led to the development of impredicative theories like $ID_1$ and their detailed analysis in [Kre63], [Fef70], and [BFPS81].

Grigori Mints developed a different technique for proving that the $H$-process terminates, using a cut-elimination argument applied to an ad hoc system of deductions. As laid out in [Min94], this technique shows that a cut-free derivation of the empty sequent in this system encodes the $H$-process, and the well-foundedness of the derivation implies that the process must terminate.

This technique was applied to more powerful theories, including elementary analysis [MTB96], ramified analysis [MT99], and the hyperarithmetical hierarchy [Ara02b]. This last paper also gives a proof using Ackermann's technique.

The most recent work has focused on extending either method to impredicative systems, specifically to theories of inductive definitions like $ID_1$. Based on the extended definition of an $\epsilon$-substitution given in [Min03], a proof using Ackermann's technique was given in [Ara03], and a slightly simplified version was explained more clearly, but without full detail, in [Ara].

In this paper we prove that the $H$-process for $ID_1$ terminates using the cut-elimination method. The transition to $ID_1$ from a system like first order arithmetic raises several separate issues which must be addressed. The first is the presence of transfinite ordinals, which means that we cannot recursively verify that a given solution to $\phi[x]$ is minimal. We resolve this in the same way as [Ara03], in that when $\phi[t]$ is true, we simply take the value of $t$ under the current substitution to be the new value for $\epsilon\nu.\phi[\nu]$, rather than selecting the least number satisfying $t$.

An immediate consequence of this is that even when a non-default value for $\epsilon x.\phi[x]$ is correct, we may change it to some other (smaller) non-default value. This requires some changes to the cut-elimination proof; the necessary modifications were made in [Tow03], which applies Mints' technique to Peano Arithmetic with transfinite induction.

The presence of impredicativity requires some kind of collapsing argument during cut-elimination. Despite the differences in the systems, the collapsing argument here is very similar to the standard cut-elimination argument for $ID_1$ as given in [Poh89].

Unlike Ackermann's method, the proof by cut-elimination requires carefully distinguishing which expressions are assigned default values because we have not yet considered their value, and which have been actually decided to have the default value, at least temporarily. With simpler systems this doesn't matter, but in $ID_1$, the monotonicity of our inductive predicate means we require some additional work. In particular, $(n \in I^{<\alpha}, \bot) \in S$ implies that we can assume $n \in I^{<\beta}$ is also false for $\beta < \alpha$. Consequently, we need to modify the $H$-rule by redefining the truncation $\Theta_{\leq r}$ of $\Theta$ to rank $r$ so that when we remove formulas of rank greater than $r$, we retain the implications that formula had for formulas of rank $\leq r$.

The presence of formulas $n \in I^{\alpha}$ and $n \in I^{<\alpha}$ in our substitutions creates one last oddity which cut-elimination must deal with. When we decide how to evaluate $n \in I^{<\alpha}$, we could create a cut with two premises, one corresponding to the claim that $n \in I^{<\alpha}$ is true, and the other to the claim that $n \in I^{<\alpha}$ is false. But this proves problematic, since we would like to ensure that $n \in I^{<\alpha}$ is true only when $n \in I^{\beta}$ is true for some $\beta < \alpha$. We resolve this by introducing an $FCut$ inference, a modified cut-rule whose subderivations correspond to $(n \in I^{<\alpha}, ?)$ and $(n \in I^{\beta}, \top)$ for each $\beta < \alpha$. Moreover, if we already have some $(n \in I^{<\gamma}, ?)$ in our sequent with $\gamma < \alpha$, we only allow $\beta$ in the range $[\gamma, \alpha)$. This inference does not have a welldefined cut-rank, since different premises add pairs with different ranks, so we develop a method of eliminating a cut in pieces as the cut-rank of the derivation decreases.

# 2 Ordinals

We use the system of ordinals $\psi(\Omega^{\Gamma})$ developed in [Poh89]. The ordinals of this system are generated from $0$ and $\Omega$ using $+$, the Veblen function $\varphi$, and the function $\psi$. We are mostly interested in the collapsing function $D$ [Poh89][Section 24], whose primary properties are that $D\alpha < \Omega$ for all $\alpha$, $D\alpha = \alpha$ when $\alpha < \Omega$, and the property that $D\alpha < D\beta$ iff either:

1. $\beta < \alpha$ and there is some $\eta \in SC(\beta)$ such that $D\alpha \leq \eta$, or
2. $\alpha < \beta$ and for every $\eta \in SC(\alpha)$, $\eta < D\beta$

where $SC(\alpha)$ is defined by:

1. $SC(0) = SC(\Omega) = \emptyset$
2. $SC(\alpha + \beta) = SC(\varphi\alpha\beta) = SC(\alpha) \cup SC(\beta)$
3. $SC(D\alpha) = \{D\alpha\}$ if $\alpha \geq \Omega$ and $SC(D\alpha) = SC(\alpha)$ otherwise

$\Omega$ and ordinals of the form $D\alpha$ for $\alpha \geq \Omega$ are strong critical, meaning that if $\alpha$ is strongly critical and $\beta, \gamma < \alpha$ then $\beta + \gamma, \varphi\beta\gamma < \alpha$.

**Definition 2.1.** *We say $\alpha \ll \beta$, $\alpha$ is* essentially less than *$\beta$ if $\alpha < \beta$ and $D\alpha < D\beta$.*

*We say $\alpha \underline{\ll} \beta$ iff $\alpha \leq \beta$ and $D\alpha \leq D\beta$.*

The following properties are proved in [Poh89], or follow immediately from results there:

**Lemma 2.1.** *1. $\alpha \ll \beta$ implies $\Omega + \alpha \ll \Omega + \beta$*

*2. If $n, m < \omega$ and $\alpha < \beta$ then $\alpha + n \ll \beta + m$ iff $\alpha \ll \beta$*

*3. If $\Omega \leq \alpha$ then $SC(\alpha) < D\alpha$*

We will mostly be interested in ordinals of the form $D(\alpha + \gamma)$ where $\alpha \geq \Omega > \gamma$, and if $\gamma \neq 0$ then $\gamma = D(\alpha' + \gamma')$ where $\alpha \leq \alpha'$. When we write $\xi = D(\alpha + \gamma)$, we always assume that $\gamma < \Omega \leq \alpha$, even when not explicitly stated.

The following ad hoc function on ordinals will be useful:

**Definition 2.2.** *$\alpha_r$ for $r < \omega$ is given by induction on $r$:*

$$\alpha_r(\eta) = \begin{cases} \eta & \text{if } r = 0 \\ \alpha^{\omega \,\#\, \alpha_s(\eta) \,\#\, \alpha_s(\eta) + 2} & \text{if } r = s + 1 \end{cases}$$

# 3 $ID_1\epsilon$

Our system is similar to the one given in [Ara] and [Ara03].

There are two types in $ID_1\epsilon$, numbers and ordinals, denoted $N$ and $O$ respectively. We let $\iota \in \{N, O\}$.

The language of $ID_1\epsilon$ consists of:

1. N-variables $x, y, z, \ldots$
2. O-variables $\xi, \eta, \zeta$
3. The $0$-ary function constants $0^\iota$
4. A $0$-ary function constant $\xi$ for each ordinal in $Ord$
5. The unary function $S$
6. Predicate constants for every $n$-ary computable predicate on numbers, including $=^N$, $<^N$, $add$, and $prod$
7. Binary predicate constants $<^O$ and $=^O$
8. Binary predicate constants $I$ and $I^<$
9. Unary predicate constant $I^{<\Omega}$
10. Propositional connectives $\neg$, $\wedge$, and $\rightarrow$
11. Propositional constants $\perp$ and $\top$
12. The epsilon symbol $\epsilon$

We typically use $x, y, z$ for $N$-variables, $\xi, \eta, \zeta$ for $O$-variables, and $\nu$ for variables which may be either. We use $m, n$ for $N$-terms (and also meta-language numbers), $\alpha, \beta$ for $O$-terms, $s, t$ for terms which may be either.

The expressions of $ID_1\epsilon$ include $\iota$-terms for $\iota \in \{N, O\}$, and formulas, and are defined inductively by:

1. Each $\iota$-variable is an $\iota$-term
2. $0^\iota$ is an $\iota$ term
3. Each ordinal $\alpha$ is an $O$-term
4. $\perp$ and $\top$ are formulas
5. If $t$ is an $N$-term then $St$ is an $N$-term
6. If $t_1, \ldots, t_n$ are $N$-terms and $R$ is an $n$-ary predicate constant then $Rt_1 \cdots t_n$ is a formula
7. If $s$ and $t$ are $O$-terms then $s_1 =^O s_2$ and $s_1 <^O s_2$ are formulas
8. If $\phi$ and $\psi$ are formulas then $\neg\phi$, $\phi \wedge \psi$, and $\phi \to \psi$ are formulas
9. If $s$ is an $O$-term and $t$ is an $N$-term then $Ist$, $I^{<}st$, and $I^{<\Omega}t$ are formulas written $t \in I^s$, $t \in I^{<s}$, and $t \in I^{<\Omega}$ respectively
10. If $\phi$ is a formula and $x$ an $N$-variable occuring free in $\phi$ then $\epsilon x < \omega\phi[x]$ is an $N$-term
11. If $\phi$ is a formula and $\eta$ an $O$-variable occuring free in $\phi$ then $\epsilon\eta < \Omega\phi[\eta]$ is an $O$-term
12. If $t$ is an $N$-term, $s$ an $O$-term, and $\eta$ an $O$-variable not occuring free in $t$ or $s$ then $\epsilon\eta < s[t \in I^\eta]$ is an $O$-term abbreviated $s\{t\}$.

Note that the bounds in (10) and (11) are true of all values in the range; they are included only to make the notation uniform. The restriction on the first $j+k$ parameters to $\{\}$ predidicates is neccessary to make sure that our reduction relation is confluent. (It is probably possible to avoid this by putting tighter restrictions on our reduction relation, but since we have no need for a more general definition, this suffices.)

The axioms of $ID_1\epsilon$ consist of:

1. All propositional tautologies
2. Substitution instances of quantifier free defining axioms for $n$-ary predicates, including the axioms for the linear orders $<^N$ and $<^O$
3. Equality axioms $t = t$, $s = t \to \phi(s) \to \phi(t)$ where $s$ and $t$ are $N$- or $O$-terms of the same type
4. $\neg St = 0$ and $Ss = St \to s = t$

5. $\neg t \in I^{<0}$

6. Critical formulas:

   **Pred** $\neg s = 0 \to s = S\epsilon x < \omega(s = Sx)$

   **Epsilon Axiom** $\phi[t] \wedge t <^{\iota} s \to (\epsilon\nu < s\phi[s]) \leq t \wedge \phi[\epsilon\nu < s\phi[s]]$

   **Inductive Definition Axiom** $t \in I^s \leftrightarrow A[I^{<s}, t]$

   **Inductive Minimality Axiom** $s > 0 \to (t \in I^{<s} \leftrightarrow t \in I^{s\{t\}})$

   **Closure** $A[I^{<\Omega}, t] \to t \in I^{<\Omega}$

We consider $t \in I^{<\Omega} \leftrightarrow t \in I^{\Omega\{t\}}$ to be an instance of the Inductive Minimality Axiom (where $\Omega > 0$ has been taken to be trivially true, and omitted since it is not a formula of our language). We consider $\phi[t] \to \phi[\epsilon\xi < \Omega\phi[\xi]]$ and $\phi[t] \to \phi[\epsilon x < \omega\phi[x]]$ to be instances of the Epsilon Axiom where $\xi <^{O} \Omega$ or $x <^{N} \omega$ is taken to be trivially true. Note that the Inductive Definition Axiom cannot apply with $s = \Omega$, since it would not be a formula of our language.

The only rule is modus ponens: $\dfrac{\phi \qquad \phi \to \psi}{\psi}$

To save notation, we require that $A$ have the form $\neg B(\epsilon x < \omega B(x))$ and $B$ contains no $\epsilon$-subterms, nor any predicates $I$, $I^{<}$, or $I^{<\Omega}$. We interpret $I^{<\Omega}$ to be the least fixed point of $A$. It is known that restricting $A$ to $\Pi_1$ formulas does not weaken the system [Mos74]. We are further restricting our formula to a single initial quantifier to simplify our notation.

# 4 $\epsilon$-Substitutions

**Definition 4.1.** *We define the depth of an expression to be the number of closed non-computable expressions it contains:*

1. $d(\nu) = d(0^{\iota}) = d(\alpha) = d(\bot) = d(\top) = 0$

2. $d(St) = d(t)$

3. $d(pe_1 \cdots e_n) = \Sigma_{i=1}^{n} d(e_i)$

4. $d(s_1 =^{O} s_2) = d(s_1 <^{O} s_2) = d(s_1) + d(s_2)$

5. $d(\neg\phi) = d(\phi)$

6. $d(\phi \wedge \psi) = d(\phi \to \psi) = d(\phi) + d(\psi)$

7. $d(t \in I^s) = d(I^{<s}) = \begin{cases} 1 + d(t) + d(s) & \textit{if } s \textit{ and } t \textit{ are closed} \\ d(t) + d(s) & \textit{otherwise} \end{cases}$

8. $d(t \in I^{<\Omega}) = \begin{cases} 1 + d(t) & \textit{if } t \textit{ is closed} \\ d(t) & \textit{otherwise} \end{cases}$

9. $d(\epsilon\nu\phi[\nu]) = \begin{cases} 1 + d(\phi) & \textit{if } \phi \textit{ is closed} \\ d(\phi) & \textit{otherwise} \end{cases}$

**Definition 4.2.** *An expression $e$ is* canonical *if it is closed, $d(e) = 1$, and it has one of the forms $\epsilon\nu < t\phi$, $t \in I^s$, or $t \in I^{<s}$.*

*An expression $e$ is* simple *if it is closed and $d(e) = 0$.*

**Definition 4.3.**

$$\iota(e) = \begin{cases} B & \text{if } e \text{ is a formula} \\ N & \text{if } e \text{ is an } N\text{-term} \\ \alpha & \text{if } e \text{ is the } O\text{-term } \epsilon\xi < \alpha\phi[\xi] \\ \Omega & \text{if } e \text{ is an } O\text{-term } \epsilon\xi < t\phi[\xi] \text{ and there is no } \alpha < \Omega \text{ such that } t = \alpha \end{cases}$$

$$V^\iota = \begin{cases} \mathbb{N} \setminus \{0\} & \iota = N \\ \{\top\} & \iota = B \\ \{\beta \mid 0 < \beta < \alpha\} & \iota = \alpha \end{cases}$$

*Also $0^B = \bot$*

**Definition 4.4.** *An $\epsilon$-substitution is a function $S$ such that:*

**Domain** *The domain is a set of canonical expressions*

**Range** $S(e) \in V^{\iota(e)} \cup \{?\}$ *for* $e \in \mathrm{dom}(S)$

**Monotonicity 1** *If* $(n \in I^\alpha, \top) \in S$ *and* $\alpha <^O \beta \leq \Omega$ *then* $(n \in I^\beta, ?), (n \in I^{<\beta}, ?) \notin S$

**Monotonicity 2** *If* $(n \in I^{<\alpha}, \top) \in S$ *then* $(n \in I^\beta, \top) \in S$ *for some* $\beta <^O \alpha$

**Parsimony 1** *If* $(e, u) \in S$ *and* $u \neq ?$ *then* $u \neq 0^{\iota(e)}$

**Parsimony 2** *If* $(\epsilon\nu <^{\iota(\nu)} \alpha.\phi[\nu], u) \in S$ *and* $u \neq ?$ *then* $u <^{\iota(\nu)} \alpha$

**Definition 4.5.**

$$\begin{aligned} S^* = S \quad & \cup\{(n \in I^\beta, \top) \mid \exists\alpha <^O \beta[(n \in I^\alpha, \top) \in S]\} \\ & \cup\{(n \in I^\beta, ?) \mid \exists\alpha >^O \beta[n \in I^\alpha \in \mathrm{dom}(S)]\} \\ & \cup\{(n \in I^\beta, ?) \mid \exists\alpha >^O \beta[(n \in I^{<\alpha}, ?) \in \mathrm{dom}(S)]\} \\ & \cup\{(n \in I^{<\beta}, \top) \mid \exists\alpha <^O \beta[(n \in I^\alpha, \top) \in S]\} \\ & \cup\{(n \in I^{<\beta}, ?) \mid \exists\alpha \geq^O \beta[n \in I^\alpha \in \mathrm{dom}(S)]\} \\ & \cup\{(n \in I^{<\beta}, ?) \mid \exists\alpha >^O \beta[(n \in I^{<\alpha}, ?) \in \mathrm{dom}(S)]\} \end{aligned}$$

*is called the* completion *of $S$.*

$$\begin{aligned} \overline{S} = S^* \quad & \cup\{(e, ?) \mid e \text{ is a canonical } \epsilon \text{ term not in } \mathrm{dom}(S)\} \\ & \cup\{(n \in I^\beta, ?) \mid \neg\exists\alpha[(n \in I^\alpha, \top) \in S]\} \\ & \cup\{(n \in I^{<\beta}, ?) \mid \neg\exists\alpha[(n \in I^\alpha, \top) \in S]\} \end{aligned}$$

*is called the* standard extension *of $S$.*

The completion internalizes the monotonicity of $I^\beta$ by assigning values to all $n \in I^\beta$ and $n \in I^{<\beta}$ whenever $n \in I^\alpha$ is in $S$. The standard extension assigns default values to all canonical expressions left undecided by $S^*$.

**Definition 4.6.**

$$\mathrm{rng}(S) = \{S(e) \mid e \in \mathrm{dom}(S)\}$$

**Definition 4.7.** *We say an $\epsilon$-substitution $S$ is* finitary *if the following conditions are satisfied:*

**Finite** *$S$ is finite*

$I?$**-free** *If $S(e) = ?$ then $e$ is not of the form $n \in I^\beta$*

**Parsimony 3** *If $(n \in I^\alpha, \top), (n \in I^\beta, \top) \in S$ then $\alpha = \beta$*

**Parsimony 4** *$(n \in I^{<\alpha}, \top) \notin S$ for any $n$ or $\alpha$*

Generally we will be interested in working with the completions of finitary substitutions. The completion expands the information in a finitary substitution to include all values for $n \in I^\beta$ and $n \in I^{<\beta}$ which we can directly infer from values already present. Unlike in [Ara03], we need to distinguish the completion from the standard extension because cut-elimination requires that we keep track of which expressions have been decided, even when they retain a default value. Similarly, our definition of finitary is slightly different, since we allow ? in the range of $S$ except for formulas $n \in I^\beta$. This is because when we wish to indicate $\neg n \in I^\beta$, we include $(n \in I^{<\beta+1}, ?)$ rather than $(n \in I^\beta, ?)$.

When we have $(n \in I^\alpha, \top) \in S$, we generally interpret this as $n \in I^\alpha \setminus I^{<\alpha}$.

# 5 Computations with $\epsilon$-Substitutions

**Definition 5.1.**
1. *If $(e, u) \in S^*$ then $e \hookrightarrow^1_S u$*
2. *If $(e, ?) \in S^*$ then $e \hookrightarrow^1_S 0^{\iota(e)}$*
3. *If $t \hookrightarrow^1_S t'$ then $St \hookrightarrow^1_S t'$*
4. *If $1 \leq i \leq n$ and $e_k \hookrightarrow^1_S e'_i$ then $Re_1 \cdots e_i \cdots e_n \hookrightarrow^1_S Re_1 \cdots e'_i \cdots e_n$*
5. *If $t \hookrightarrow^1_S t'$ then $t =^O s \hookrightarrow^1_S t' =^O s$, $t <^O s \hookrightarrow^1_S t' <^O s$, $s =^O t \hookrightarrow^1_S s =^0 t'$, $s <^O t \hookrightarrow^1_S s <^O t'$*
6. *If $\phi \hookrightarrow^1_S \phi'$ then $\neg\phi \hookrightarrow^1_S \neg\phi'$, $\phi \wedge \psi \hookrightarrow^1_S \phi' \wedge \psi$, $\psi \wedge \phi \hookrightarrow^1_S \psi \wedge \phi'$, $\phi \rightarrow \psi \hookrightarrow^1_S \phi' \rightarrow \psi$, $\psi \rightarrow \phi \hookrightarrow^1_S \psi \rightarrow \phi'$*
7. *If $\phi \hookrightarrow^1_S \phi'$ then $\epsilon\nu < s\phi \hookrightarrow^1_S \epsilon\nu < s\phi'$*
8. *If $t \hookrightarrow^1_S t'$ then $\epsilon\nu < t\phi \hookrightarrow^1_s \epsilon\nu < t'\phi$, $t \in I^s \hookrightarrow^1_S t' \in I^s$, $t \in I^{<s} \hookrightarrow^1_S t' \in I^{<s}$, $t \in I^{<\Omega} \hookrightarrow^1_S t' \in I^{<\Omega}$, and $t \in \{t_1, \ldots, t_j \mid s_1, \ldots, s_k\} \hookrightarrow^1_S t' \in \{t_1, \ldots, t_j \mid s_1, \ldots, s_k\}$*
9. *If $s \hookrightarrow^1_S s'$ then $t \in I^s \hookrightarrow^1_S t \in I^{s'}$ and $t \in I^{<s} \hookrightarrow^1_S t \in I^{<s'}$*

*$\hookrightarrow_S$ is the transitive, reflexive closure of $\hookrightarrow^1_S$*

**Lemma 5.1.** *If $e \hookrightarrow^1_S e'$ then $FV(e) = FV(e')$*

*Proof.* By induction on the definition of $\hookrightarrow^1_S$

1. If $e$ is almost canonical then $e'$ is simple, so both are closed
2. Otherwise, the result follows directly from the inductive hypothesis

□

**Lemma 5.2.** *If $e \hookrightarrow^1_S e'$ and $e \hookrightarrow_S e''$ then there is some $u$ such that $e' \hookrightarrow_S u$ and $e'' \hookrightarrow_S u$.*

*Proof.* By induction on the definition of $\hookrightarrow^1_S$

1. If $e$ is almost canonical then there is a unique $u$ such that $e \hookrightarrow^1_S u$, so $e' = e'' = u$.
2. If $e$ has only one immediate subexpression which can be reduced, the result follows directly from IH
3. Otherwise, let $e = f(e_1, \ldots, e_i, \ldots, e_j, \ldots, e_n)$ where $e' = f(e_1, \ldots, e'_i, \ldots, e_j, \ldots, e_n)$ and $e'' = f(e_1, \ldots, e_i, \ldots, e'_j, \ldots, e_n)$. If $i = j$ then the result follows by IH. Otherwise, $u = f(e_1, \ldots, e'_i, \ldots, e'_j, \ldots, e_n)$.

□

**Lemma 5.3.** *If $e \hookrightarrow^1_S e'$ then $d(e') < d(e)$.*

*Proof.* By induction on the definition of $\hookrightarrow^1_S$

1. If $e$ is almost canonical then $e'$ is simple, $d(e) = 1 > 0 = d(e')$
2. Otherwise, the result follows directly from the inductive hypothesis

□

**Lemma 5.4.** *Every expression $e$ has a unique normal form $|e|_S$ such that $e \hookrightarrow_S |e|_S$ and there is no $u$ such that $|e|_S \hookrightarrow^1_S u$.*

*Proof.* By Lemma 5.3, any sequence of reductions must eventually end, and by Lemma 5.2, it must end uniquely. □

**Definition 5.2.** *$e$ is $S$-computable iff $d(|e|_S) = 0$.*

**Definition 5.3.** *If $S$ and $S'$ are $\epsilon$-substitutions then we say $S \trianglelefteq S'$ if for each $(e, u) \in S$, one of the following holds:*

1. $(e, u) \in S'$
2. $e = n \in I^{<\alpha}$, $u = ?$, *and* $(n \in I^{<\beta}, ?) \in S'$ *for some* $\beta > \alpha$
3. $e = n \in I^{<\alpha}$, $u = ?$, *and* $(n \in I^{\beta}, \top) \in S'$ *for some* $\beta \geq \alpha$

$S \trianglelefteq S'$ means that $S$ is the same as $S'$ except that some $(n \in I^{<\alpha}, ?) \in S$ may be improved to some $(e, u) \in S'$ which is stronger, in the sense that $(e, u) \in S'$ implies $(n \in I^{<\alpha}, ?) \in S'^{*}$.

**Lemma 5.5.** *$S \trianglelefteq S'$ implies $S^* \subseteq S'^*$*

*Proof.* Assume $S \trianglelefteq S'$ and $(e, u) \in S^*$; then one of the following holds:

1. $(e, u) \in S$. Then if $(e, u) \in S'$ the result is obvious. Otherwise, $e = n \in I^{<\alpha}$ and either $(n \in I^{<\beta}, ?) \in S'$ for $\beta > \alpha$ or $(n \in I^{\beta}, \top) \in S'$ with $\beta \geq \alpha$. In either case, $(e, u) \in S'^*$.

2. $(e, u) \in S^*$ because there is some appropriate $(n \in I^{\alpha}, v) \in S$. Then $(n \in I^{\alpha}, v) \in S'$, so $(e, u) \in S'^*$.

3. $(e, u) \in S^*$ because there is some appropriate $(n \in I^{<\alpha}, ?) \in S$. If $(n \in I^{<\alpha}, ?) \in S'$ the result is obvious. Otherwise, there is either some $n \in I^{<\beta}$ or $n \in I^{\beta}$ in $S'$. In either case, $(e, u) \in S'^*$ is forced.

□

# 6 Rank

This definition similar to the one used in [Ara03], cf. [Ara02a]. In particular, Lemma 6.1 is essentially Arai's Rank Lemma.

**Definition 6.1.** *Let $\sigma$ be either a variable or $*$. Then if $\sigma \notin FV(e) \cup \{*\}$ then $rk_\sigma(e) = 0$. Otherwise:*

1. $rk_\sigma(\nu) = rk_\sigma(\alpha) = rk_\sigma(\bot) = rk_\sigma(\top) = 0$

2. $rk_\sigma(St) = rk_\sigma(t)$

3. $rk_\sigma(Rt_1 \cdots t_n) = \max\{rk_\sigma(t_1), \ldots, rk_\sigma(t_n)\}$

4. $rk_\sigma(s_1 =^O s_2) = rk_\sigma(s_1 <^O s_2) = \max\{rk_\sigma(s_1), rk_\sigma(s_2)\}$

5. $rk_\sigma(\neg\phi) = rk_\sigma(\phi)$

6. $rk_\sigma(\phi \wedge \psi) = rk_\sigma(\phi \rightarrow \psi) = \max\{rk_\sigma(\phi), rk_\sigma(\psi)\}$

7. $rk_\sigma(\epsilon x < \omega\phi) = \max\{rk_\sigma(\phi), rk_x(\phi) + 1\}$

8. $rk_\sigma(\epsilon\xi < \Omega\phi) = \max\{rk_\sigma(\phi), rk_\xi(\phi) + 1, \Omega + 1\}$

9. $rk_\sigma(\epsilon\xi < s[t \in I^\xi]) = \begin{cases} \max\{rk_\sigma(t), 3\alpha + 1\} & \textit{if } s \equiv \alpha \leq \Omega \\ \max\{rk_\sigma(t), rk_\xi(t) + 1, rk_\sigma(s), \Omega + 2\} & \textit{otherwise} \end{cases}$

10. $rk_\sigma(t \in I^s) = \begin{cases} \max\{rk_\sigma(t), 3\alpha + 2\} & \textit{if } s \in \{\alpha, \alpha\{t\}\} \\ \max\{rk_\sigma(t), rk_\sigma(s), \Omega + 1\} & \textit{otherwise} \end{cases}$

11. $rk_\sigma(t \in I^{<s}) = \begin{cases} \max\{rk_\sigma(t), 3\alpha\} & \text{if } s \equiv \alpha < \Omega \\ \max\{rk_\sigma(t), rk_\sigma(s), \Omega\} & \text{otherwise} \end{cases}$

12. $rk_\sigma(t \in I^{<\Omega}) = \max\{rk_\sigma(t), 3\Omega\}$

$rk(e) = rk_*(e)$ *is the* rank *of* $e$.

Rank is defined to satisfy the following:

**Lemma 6.1.** 1. *If* $e \hookrightarrow^1_S e'$ *then* $rk_\sigma(e') \leq rk_\sigma(e)$

2. *All subexpressions of an expression* $e$ *have ranks* $\leq rk(e)$

3. *If* $\epsilon\nu < s\phi$ *is canonical and* $u \in V^{\iota(\nu)} \cup 0^{\iota(\nu)}$ *then* $rk(\phi(u)) < rk(\epsilon\nu < s\phi)$

4. *If* $n \in I^\alpha$ *is canonical then* $rk(A(I^{<\alpha}, n)) < rk(n \in I^\alpha)$

5. $rk(n \in I^\beta) < rk(\alpha\{n\})$ *whenever* $\beta < \alpha$.

6. $rk(n \in I^\beta) < rk(n \in I^\alpha)$ *whenever* $\beta < \alpha$.

7. *The rank of compound expressions formed by* $R$, $\neg$, $\wedge$, $\rightarrow$, $=^O$, $<^O$, *and* $S$ *is just the maximum of the ranks of the subexpressions.*

We have $rk(e) < \Omega + \omega$ for all $e$.

**Definition 6.2.** *We define* $\mathcal{O}(r)$ *by:*

$$\mathcal{O}(r) = \begin{cases} \alpha & \text{if } r = 3\alpha \text{ or } r = 3\alpha + 1 \\ \alpha + 1 & \text{if } r = 3\alpha + 2 \end{cases}$$

**Definition 6.3.** *If* $S$ *is an* $\epsilon$*-substitution then*

$$S_{\leq r} = \{(e, u) \in S \mid rk(e) \leq r\} \quad \begin{array}{l} \cup\{(n \in I^{<\mathcal{O}(r)}, ?) \mid (n \in I^\beta, \top) \in S, rk(n \in I^\beta) > r\} \\ \cup\{(n \in I^{<\mathcal{O}(r)}, ?) \mid (n \in I^{<\beta}, ?) \in S, rk(n \in I^{<\beta}) > r\} \end{array}$$

This definition differs from the one used in, for instance, [MTB96] with respect to the formula components of $S$. The property we actually maintain is that whenever $rk(e) \leq r$, $|e|_S = |e|_{S_{\leq r}}$. This is necessary to ensure that when a component $(n \in I^{<\beta}, ?)$ is present in $S$ and removed in $S_{\leq r}$, its low-rank consequences (like $(n \in I^{<\mathcal{O}(r)}, ?)$) are kept.

**Lemma 6.2.** 1. *If* $rk(e) \leq r$ *and* $e \hookrightarrow_S u$ *then* $e \hookrightarrow_{S_{\leq r}} u$

2. *If* $(e, u) \in S_{\leq r}$ *and* $rk(e) > r$ *then* $u = ?$, $e = n \in I^{<\beta}$, $rk(e) = r + 1$

3. *If* $(e, u) \in S_{\leq r} \setminus S$ *and* $rk(e) < r$ *then* $u = ?$, $e = n \in I^{<\beta}$, $rk(e) = r - 1$

*Proof.* 1. By straightforward induction on $\hookrightarrow^1_S$.

2. Obvious from the definition

3. Obvious from the definition

□

**Lemma 6.3.** *If* $S_{\leq r} = S'_{\leq r}$ *and* $rk(e) \leq r$ *then* $|e|_S = |e|_{S'}$

*Proof.* By induction on $d(e)$, using Lemma 6.2. □

# 7 $H$-Process

Assume $Cr_0, \ldots, Cr_N$ is a fixed finite sequence of closed critical formulas and that no ordinal constants other than 0 and $\Omega$ appear in any $Cr_I$. This is equivalent to the general case since each constant is definable by a primitive recursive formula.

**Definition 7.1.** *If $|e|_S$ is some true propositional combination of computable formulas we say $e \hookrightarrow_S \top$. If $|e|_S$ is some false propositional combination of computable formulas, we say $e \hookrightarrow_S \bot$.*

*We say $S$ is* solving *if $Cr_I \hookrightarrow_{\overline{S}} \top$ for $I \in \{0, \ldots, N\}$*

## 7.1 $H$-Expressions and $H$-Values

We are going to define a function on $\epsilon$-substitutions which updates them. There will be two types of updates, one for formulas and one for $\epsilon$-terms. When we update a term, we either add $(e, u)$ or we replace some $(e, u)$ with $(e, v)$ where $v < u$ under the appropriate ordering. When we update a formula, we either add $(n \in I^\beta, \top)$ (possibly displacing some $(n \in I^{<\alpha}, ?)$) or replace $(n \in I^\beta, \top)$ with $(n \in I^\gamma, \top)$ where $\gamma <^O \beta$.

**Definition 7.2.** *For each component $(e, u)$, we define $\mathcal{P}(e, u)$ to be the set of pairs which it may displace:*

$$\mathcal{P}(e,u) = \begin{cases} \{(e,v)\}_{u <^O v <^O \alpha} \cup \{(e,?)\} & \text{if } e = \epsilon\xi < \alpha\phi \\ \{(e,v)\}_{u < v < \omega} \cup \{(e,?)\} & \text{if } e = \epsilon x < \omega\phi \\ \{(n \in I^\beta, \top)\}_{\alpha <^O \beta <^O \Omega} \cup \{(n \in I^{<\beta}, ?)\}_{\beta \leq \Omega} & \text{if } e = n \in I^\alpha \end{cases}$$

**Definition 7.3.** *Let $S$ be a nonsolving $\epsilon$-substitution.*

*For $I \leq N$, define $e_I^S$ depending on the type of the formula $Cr_I$:*

1. *If $Cr_I$ is of the form $\neg s = 0 \to s = S\epsilon x < \omega(s = Sx)$ then*

$$e_I^S = \epsilon x < \omega |s|_{\overline{S}} = Sx$$

2. *If $Cr_I$ is of the form $\phi[t] \wedge t < s \to (\epsilon\nu < s\phi) \leq t \wedge \phi[\epsilon\nu < s\phi[\nu]]$ then*

$$e_I^S = \epsilon\nu < |s|_{\overline{S}} |\phi|_{\overline{S}}[\nu]$$

3. *If $Cr_I$ is of the form $t \in I^s \leftrightarrow A[I^{<s}, t]$ then*

$$e_I^S = |t|_{\overline{S}} \in I^{|s|_{\overline{S}}}$$

4. *If $Cr_I$ is of the form $s > 0 \to (t \in I^{<s} \leftrightarrow t \in^{s\{t\}})$ then*

$$e_I^S = \epsilon\nu < |s|_{\overline{S}} |t|_{\overline{S}} \in I^\nu$$

5. *If $Cr_I$ is of the form $A[I^{<\Omega}, t] \to t \in I^{<\Omega}$ then*

$$e_I^S = |t|_{\overline{S}} \in I^\alpha$$

*for suitable $\alpha$.*

Section 7.3 explains what a suitable $\alpha$ is.

**Definition 7.4.** *Define $r_I^S = rk(e_I^S)$ and let $I(S) \leq N$ be the least $I$ such that:*

1. $Cr_I \hookrightarrow_{\overline{S}} \bot$
2. *Whenever $r_J^S < r_I^S$, $Cr_J \hookrightarrow_{\overline{S}} \top$*
3. *If $r_J^S = r_I^S$ and $J < I$, $Cr_J \hookrightarrow_{\overline{S}} \top$*

*Define $Cr(S) = Cr_{I(S)}$, $R(S) = r^S_{I(S)}$, and $e(S) = e^S_{I(S)}$. $e(S)$ is called the $H$-expression of $S$.*

*The $H$-value $v(S)$ is given by:*

1. *If $Cr(S)$ is of the form $\neg s = 0 \to s = S\epsilon x < \omega(s = Sx)$ then $v(S) = |s|_{\overline{S}} - 1$*
2. *If $Cr(S)$ is of the form $\phi[t] \wedge t < s \to (\epsilon\nu < s\phi[\nu]) \leq t \wedge \phi[\epsilon\nu < s\phi[\nu]]$ then $v(S) = |t|_{\overline{S}}$*
3. *If $Cr(S)$ is of the form $t \in I^s \leftrightarrow A[I^{<s}, t]$ then $v(S) = \top$*
4. *If $Cr(S)$ is of the form $s > 0 \to (t \in I^{<s} \leftrightarrow t \in^{s\{t\}})$ then $v(S)$ is the (unique) $\alpha$ such that $(|t|_{\overline{S}} \in I^\alpha, \top) \in S$*
5. *If $Cr(S)$ is of the form $A[I^{<\Omega}, t] \to t \in I^{<\Omega}$ then $v(S) = \top$. In this case, we say $S$ is* at a closure step.

**Definition 7.5.**

$$H^w(S) = S_{\leq R(S)} \setminus \mathcal{P}(e(S), v(S)) \cup \{(e(S), v(S))\}$$

Note that $H^w(S)$ is the same as $H(S)$ in [Ara03].

We can now define the full $H$-step $H(S)$, and the $H$-process (the sequence of substitutions resulting from iteration of the $H$-step) by simultaneous induction. Note that all the definitions in the rest of this section are simultaneous with the definition of the $H$-process.

**Definition 7.6.** *If $S, S'$ are $\epsilon$-substitutions and, for some $n \geq 0$, $H^n(S) = S'$, we say $S \leq S'$. If $n > 0$ we say $S < S'$.*

That is, $S < S'$ exactly if $S'$ comes after $S$ in the $H$-process.

## 7.2 Correction Terms

We need to accomodate the following situation: suppose that, at some closure step, we add $n \in I^\alpha$, and then, at some later step, refute this. Then our would add $(\epsilon x \neg |B(I^{<\alpha}, n, x)|_{\overline{S}}, k)$ for some $k$ which witness this refutation.

Then the immediate next step will be to assign $n \in I^\alpha$ again: while our substitution "knows" that $n \in I^\alpha$ is false, it does not know that $n \in I^{<\Omega}$ is also false. To deal with this, we must copy $(\epsilon x \neg ||B(I^{<\alpha}, n, x)|_{\overline{S}}, k)$ up, by simultaneously adding $(\epsilon x \neg ||B(I^{<\Omega}, n, x)|_{\overline{S}}, k)$.

**Definition 7.7.** *Suppose $S$ is at a closure step and $e^S \equiv n \in I^\alpha$. Then we say $n$ is the* source *of $\alpha$ in $S$. If $S < S'$ in the $H$-process and $n \in I^\alpha$ is in $U$ whenever $S < U \leq S'$ then we also say $n$ is the source of $\alpha$ in $S'$.*

The immediate solution is to say that, whenever $n$ is the source of $\alpha$ and $(e(S), v(S)) = (\epsilon x \neg |B(I^{<\alpha}, n, x)|_{\overline{S}}, k)$ then we add an additional component.

We must also deal with the iterated version of this problem: suppose that, after refuting $n \in I^\alpha$, we add some $m \in I^\beta$ (thereby removing $(\epsilon x \neg |B(I^{<\Omega}, n, x)|_{\overline{S}}, k)$). We then refute $m \in I^\beta$, add an extra component, and then add $n \in I^\alpha$ again, losing the extra component we added to refute $m \in I^{<\Omega}$. To deal with this, when we refute $m \in I^\beta$, we need to add the extra component refuting $n \in I^\alpha$ as well.

We want to copy components of the form $\pi = (\epsilon x \neg |B(I^{<\Omega}, n, x)|_{\overline{U}}, k)$ when $\pi \in U$ where $U \leq S$, but not all such components. When copying from earlier substitutions, we need to ensure that intervening changes have not made the component incorrect. First, we need to require that $I^{<\Omega}$ has the same meaning in both $U$ and $H(S)$. Second, we need to require that all changes made between $U$ and $S$ are irrelevant; it suffices to require that they have rank greater than $\alpha$.

**Definition 7.8.** *$C(S_1, \ldots, S_n)$ is a set of components of the form $(\epsilon x \neg |B(I^{<\Omega}, n, x)|_{\overline{S_i}}, k)$ for some $i \leq n$. $C(S_1, \ldots, S_n) \neq \emptyset$ iff $(e(S_n), v(S_n)) = (\epsilon x \neg |B(I^{<\alpha}, n, x)|_{\overline{S_n}}, k)$ and $n$ is the source of $\alpha$. In this case:*

- $(\epsilon x \neg |B(I^{<\Omega}, n, x)|_{\overline{(S_n)_{\leq r}}}, k) \in C(S_1, \ldots, S_n)$. *This is called the* primary component *of $C(S_1, \ldots, S_n)$*
- *If the following conditions are met then $(\epsilon x \neg |B(I^{<\Omega}, m, x)|_{\overline{S_i}}, l)$:*
  - $i \leq n$
  - $(\epsilon x \neg |B(I^{<\Omega}, m, x)|_{\overline{S_i}}, l) \in S_i$
  - $(n \in I^\xi, \top) \in S_i$ *iff* $\xi < \alpha$ *and* $(n \in I^\xi, \top) \in S_n$
  - *For any $j$ with $i \leq j < n$, $rk(e(S_j)) > 3\alpha$*

*If $C(S_1, \ldots, S_n) \neq \emptyset$ then we say that $S_1, \ldots, S_n$ is at a* corrected H-step.

**Definition 7.9.**

$$H(S_1, \ldots, S_n) = H^w(S_n) \cup C(S_1, \ldots, S_n)$$

Note that $H(S)$ is called $H_\Omega(S)$ in [Ara03].

**Definition 7.10.** *The $H$-process is defined inductively by $H_0 = \emptyset$,*

$$H_{n+1} = \begin{cases} H(H_1, \ldots, H_n) & \text{if } H_n \text{ is not solving} \\ H_n & \text{if } H_n \text{ is solving} \end{cases}$$

*We say the $H$-process (for $Cr_0, \ldots, Cr_n$) terminates if there is some $n$ such that $H_n$ is solving, and therefore that $H_n = H_{n+1}$.*

### 7.3 Ordinal Assignment

We need to select an ordinal to assign when updating closure axioms. The ordinal selected depends on a "height" which we calculate by predicting portions of the tree which will be used to show that the process terminates. Rather than pull all the key definitions out of context, the height will be defined in Section 9.2.

We actually need the part of the process preceeding the substitution at a closure step, so the ordinal has to be defined by simultaneous induction with the $H$-step. Let $ind(S)$ be the operation defined in Section 9.2.

In order to have the properties we will need, we need some information about the ordinals in our substitutions first.

**Definition 7.11.** *1.* $Ord(e) = \{\xi < \Omega \mid \xi$ *is a subterm of* $e\} \cup \{0\}$

*2.*
$$\begin{array}{rcl} Ord(S) & = & \{\xi \mid \exists n[(n \in I^{\xi}) \in \mathrm{dom}(S)]\} \\ & \cup & \{\xi \mid \xi \in \mathrm{rng}(S)\} \\ & \cup & \{0, \Omega\} \end{array}$$

The second clause of $Ord(S)$ may appear unnecessary, and, if $S = H_n$ for some $n$, it is. However in the termination proof we will deal with substitutions which may include some $(e, \alpha)$ when there is no $n \in I^{\alpha}$ in the domain.

**Lemma 7.1 (cf Lemma 9.8 in [Ara03]).** *1. Let $S$ be a substitution. If $e$ is an expression with $Ord(e) \subseteq Ord(S)$ and $e \hookrightarrow_S e'$ then $Ord(e') \subseteq Ord(S)$*

*2. If $S$ is non-solving and $Cr(S)$ is not a closure axiom then $Ord(e(S)) \subseteq Ord(S)$*

*Proof.* 1. By straightforward induction on $\hookrightarrow_S$.

2. By part 1 and the definition of $e(S)$.

□

**Definition 7.12.** *If $O$ is a set of ordinals then*

$$SCl_1(O) = O \cup \{SC(\alpha) \cup \{\gamma\} \mid D(\alpha + \gamma) \in O\}$$

*Then $SCl_{n+1}(O) = SCl_1(SCl_n(O))$, and $SCl(O) = \bigcup_{n<\omega} SCl_n(O)$.*

**Definition 7.13.** *Let $S_0, \ldots, S_m$ be a sequence such that $H(S_0, \ldots, S_i) = S_{i+1}$ and $e(S_i) \geq \Omega$ for $i < m$. Let $\Omega \leq \xi < \Omega + \omega$ be an ordinal and $\Omega \leq r = \min\{rk(e(S_i)) \mid i < m\} < \Omega + \omega$.*

- *If $\xi > r$ then $o(S_0, \ldots, S_m; \xi) = o(S_0, \ldots, S_m; r)$*
- *If $m = 0$ then $o(S_0, \ldots, S_m; \xi) = o(S_0; \xi) = (\Omega + \omega)_{r-\xi}(ind(S_0))$.*
- *If $m > 0$ then let $\{k_1 < \cdots < k_l\} = \{i < m \mid e(S_i) = r\}$, $k_0 = 0$, and set*

$$o(S_0, \ldots, S_m; \xi) = (\Omega + \omega)_{r-\xi} \Sigma_{i<l} o(S_{k_i}, \ldots, S_{k_{i+1}-1}; r)$$

*Set $o(S_0, \ldots, S_m) = o(S_0, \ldots, S_m; \Omega)$.*

**Definition 7.14.** *Suppose $S_0, \dots, S_m = S$ is the maximal sequence such that $H(S_i) = S_{i+1}$ and $e(S_i) \geq \Omega$ for $i < m$. Suppose $SCl(Ord(S)) = \{\xi_0 > \xi_1 > \cdots > \xi_k = 0\}$. Let each $\xi_i = D(\alpha_i + \gamma_i)$ with $\gamma_i < \Omega \leq \alpha_i$.*

*Let $i(S_1, \dots, S_m) = \min\{i < k \mid \alpha_i \geq o(S_0, \dots, S_m)\}$.*

$$k(S_1, \dots, S_m, n) = \begin{cases} \xi_{i(S_1,\dots,S_m)} & \text{if } SC(o(S_1, \dots, S_m)) < \xi_{i(S_1,\dots,S_m)} \\ 0 & \text{otherwise} \end{cases}$$

**Definition 7.15.** *If $Cr_I$ is a closure axiom and $S_0, \dots, S_m = S$ the maximal sequence such that $H(S_0, \dots, S_i) = S_{i+1}$ then $e^I_S = (n \in I^\alpha)$ where*

$$\alpha = D(o(S_0, \dots, S_m) + k(S_0, \dots, S_m))$$

**Lemma 7.2 (Cf. [Ara03], Lemma 9.18).** *Let $O, O'$ be finite sets of ordinals so that $\max O' \leq \max O$, and suppose that $\xi = D(\alpha + \gamma)$ satisfies either:*

- *$\gamma > SC(\alpha)$, and for every $\xi' = D(\alpha' + \gamma') \in SCl(O)$ such that $\alpha \leq \alpha'$, $\xi' \leq \gamma$, or*
- *$\gamma = 0$ and for every $\xi' = D(\alpha' + \gamma') \in SCl(O)$, such that $\alpha \leq \alpha'$, $\xi' \leq SC(\alpha)$*

*Then one of these properties holds for $\xi$ with respect to $O \cup O'$.*

*Proof.* Then we have some $\xi' \in O'$ which violates this property. That is, $\alpha \leq \alpha'$ and $\gamma < \xi'$. Since $\max O' \leq \max O$, we also have some $\xi^* \in O \subseteq SCl(O)$ such that $\xi' \leq \xi^*$. Let $\xi^* = D(\alpha^* + \gamma^*)$ be least such that $\xi^* \in SCl(O)$ and $\xi' \leq \xi^*$.

Suppose $\alpha^* < \alpha'$. Then there is some $\eta \in SC(\alpha^* + \gamma^*)$ such that $\xi' \leq \eta$. But then $\eta \in SCl(O)$ and $\eta < \xi^*$, violating our assumption. Hence $\alpha' \leq \alpha^*$. Then, since $\gamma < \xi^*$, it must be that $\xi^* \leq SC(\alpha)$. But then $\gamma < \xi' \leq SC(\alpha)$, so we have $\gamma = 0$ and the second case remains true. □

**Lemma 7.3.** *If $\xi \in SCl(Ord(S))$ and $Cr_I$ is a closure axiom, and therefore $e^S_I = n \in I^\beta$, then $\xi < \beta$.*

*Proof.* Let $\beta = D(\alpha' + \gamma')$. We proceed by induction. If $\xi \in SCl(Ord(S))$ then $\xi = D(\alpha + \gamma)$, and by IH, $\gamma < \beta$. If $\alpha + \gamma < \alpha' + \gamma'$ then since $SC(\alpha + \gamma) \subseteq SCl(Ord(S))$, by IH we have $SC(\alpha + \gamma) < \beta$, so $\xi < \beta$.

If $\alpha' + \gamma' \leq \alpha + \gamma$ then $\alpha' \leq \alpha$. If $\xi < SC(\alpha)$ then $\xi < \beta$, and otherwise we have $\xi \leq \gamma' < \beta$. □

**Lemma 7.4.** *Let $\xi = D(\alpha + \gamma)$ be some ordinal, and let $\xi' = D(\alpha' + \gamma') < \xi$ be some ordinal such that $\alpha \leq \alpha'$. Then $\xi' \leq \gamma$.*

*Proof.* By induction on $\xi$: let $\xi$ be least that such an $\xi'$ exists.

If $\alpha + \gamma < \alpha' + \gamma'$ then there is some $\eta \in SC(\alpha + \gamma)$ such that $\xi' \leq \eta$. Then either $\eta = \gamma$ or $\eta \in SC(\alpha)$, and therefore $\xi' \leq SC(\alpha)$.

So suppose $\alpha' + \gamma' < \alpha + \gamma$. Then $\alpha' = \alpha$ and $\gamma' < \gamma$. Then $\gamma = D(\alpha^* + \gamma^*)$, and $\alpha \leq \alpha^*$, $SC(\alpha) < \gamma$. Since $SC(\alpha' + \gamma') < \gamma$, it must be that $\alpha^* + \gamma^* < \alpha' + \gamma'$, that is $\alpha^* = \alpha$ and $\gamma^* < \gamma'$.

But then $\gamma^* < \gamma' < \gamma$, so $\gamma' = D(\alpha'' + \gamma'')$, and we must have $\alpha'' \geq \alpha$. But then $\gamma$ and $\gamma'$ provide a smaller example, contradicting the minimality of $\xi$. □

The following lemma is also needed, but we will have to prove it along with the termination proof:

**Lemma 7.5 (Cf. [Ara03], 10.4).** *If $e(S) \equiv n \in I^\alpha$ then there is no $k$ such that $(\epsilon x \neg |B(I^{<\alpha}, n, x)|_{\overline{S}}, k) \in S$.*

Note that the $H$-process we have defined has the following essential property:

**Lemma 7.6.** *If $e(S_k) > r \geq e(S_n)$ then $H(S_1, \dots, S_n) = H(S_{k+1}, \dots, S_n)$.*

*Proof.* Follows from the definition. $\square$

# 8 Correctness

**Definition 8.1.** *We define*

$$F(e,u) = \begin{cases} \phi[u] \wedge \neg\phi[0] & \textit{if } e = \epsilon\nu < s\phi[\nu] \textit{ and } u \neq ? \\ A[I^{<\beta}, n] & \textit{if } e = n \in I^\beta \textit{ and } u \neq ? \\ \top & \textit{if } u = ? \end{cases}$$

$$\mathcal{F}(S) = \{F(e,u) \mid (e,u) \in S \wedge u \neq ?\}$$

Note that $rk(F(e,u)) < rk(e)$ by Lemma 6.1.

**Definition 8.2.** *We say an $\epsilon$ substitution $S$ is correct if $\phi \hookrightarrow_{\overline{S}} \top$ for all $\phi \in \mathcal{F}(S)$.*

**Lemma 8.1.** *If $S$ is correct and nonsolving then $H(S)$ is a correct $\epsilon$-substitution.*

*Proof.* Consider some $(e,u) \in H(S)$. If $(e,u) \in S$ then $rk(F(e,u)) < rk(e(S))$, so the result follows form the correctness of $S$.

On the other hand, suppose $(e,u) \in H(S) \setminus S$. If $u \neq ?$ then either $(e,u) = (e(S), v(S))$ or $(e,u) \in C(S)$. Suppose $(e,u) = (e(S), v(S))$. By Lemma 6.3, it suffices to show that $|F(e,u)|_{\overline{S}} = |F(e,u)|_{\overline{H(S)}}$.

1. If $Cr(S)$ is of the form $\neg s = 0 \to s = S\epsilon x < \omega(s = Sx)$ then $|s|_{\overline{S}} = S(|s|_{\overline{S}} - 1) \hookrightarrow_{\overline{S}} \top$, and therefore $|s|_{\overline{S}} = S(|e|_{\overline{S}} - 1) \hookrightarrow_{\overline{S}} \top$ by Lemma 6.3.

2. If $Cr(S)$ is of the form $\phi[t] \wedge t < s \to (\epsilon\nu < s\phi[\nu]) \leq t \wedge \phi[\epsilon\nu < s\phi[\nu]]$ then $\phi[t] \hookrightarrow_{\overline{S}} \top$, and so $||\phi|_{\overline{S}}[|t|_{\overline{S}}]|_{\overline{S_{\leq rk(e(S))}}} = \top$. By correctness of $S$, we have either $e(S) \notin \mathrm{dom}(S)$, in which case $|\phi[0]|_{\overline{S_{\leq rk(e(S))}}} = \bot$ or $e(S) \in \mathrm{dom}(S)$ in which case, by correctness, the same thing holds. But since $S^*_{\leq rk(e(S))} = H(S)^*_{\leq rk(e(S))}$, we have $|\mathcal{F}(e,u)|_{\overline{H(S)}} = ||\phi|_{\overline{S}}[|t|_{\overline{S}}]|_{\overline{H(S)}} = \top$.

3. If $Cr(S)$ is of the form $t \in I^s \leftrightarrow A[I^{<s}, t]$ then $|A[I^{<|s|_{\overline{S}}}, |t|_{\overline{S}}]|_{\overline{H(S)}} = |A[I^{<s}, t]|_{\overline{S}}$, and since $S$ is correct and $Cr(S) \hookrightarrow_{\overline{S}} \bot$, $A[I^{<s}, t] \hookrightarrow_{\overline{S}} \top$, so $A[I^{<|s|_{\overline{S}}}, |t|_{\overline{S}}] \hookrightarrow_{\overline{H(S)}} \top$.

4. If $Cr(S)$ is of the form $s > 0 \to (t \in I^{<s} \leftrightarrow t \in^{s\{t\}})$ then since $t \in I^{<s} \hookrightarrow_{\overline{S}} \top$, it must be that $u < |s|_{\overline{S}}$. Also, since $rk(n \in I^u) \leq rk(e)$, $(n \in I^u, \top) \in H(S)$.

5. If $Cr(S)$ is of the form $A[I^{<\Omega}, t] \to t \in I^{<\Omega}$ then this follows by Lemma 7.5.

Suppose $(e, u) \in C(S)$. If $(e, u)$ is the primary component of $C(S)$ then let $e(S) \equiv \epsilon x \neg b(I^{<\alpha}, n, x)$. Since $n \in I^{<\xi} \hookrightarrow_{\overline{H(S)}} \top$ iff $n \in I^{<\Omega} \hookrightarrow_{\overline{H(S)}} \top$, and by the argument above, $\neg b(I^{<\alpha}, n, u) \wedge b(I^{<\alpha}, n, 0) \hookrightarrow_{\overline{H(S)}} \top$, the same holds for $\neg b(I^{<\Omega}, n, u) \wedge b(I^{<\Omega}, n, 0)$.

If $(e, u)$ is not the primary component then let $e \equiv \epsilon x \neg b(I^{<\Omega}, m, x)$. We have $\neg b(I^{<\Omega}, m, u) \wedge b(I^{<\Omega}, m, 0) \hookrightarrow_{\overline{U}} \top$ for some $U$ such that $n \in I^{<\xi} \hookrightarrow_{\overline{H(S)}} \top$ iff $n \in I^{<\xi} \hookrightarrow_{\overline{U}} \top$, and $U$ and $H(S)$ agree below the largest ordinal in $H(S)$ (and therefore below the largest ordinal in $U$). But by Lemma 7.1, since $Ord(b) \subseteq \{0\}$, it follows that $|\neg b(I^{<\Omega}, m, u) \wedge b(I^{<\Omega}, m, 0)|_{\overline{H(S)}} = |\neg b(I^{<\Omega}, m, u) \wedge b(I^{<\Omega}, m, 0)|_{\overline{U}} = \top$.

Observe that $H(S)$ is finite, since it is at most one element larger than $S$. It is $I?$-free since $v(S) =?$ is never $?$. It meets the remaining conditions since:

1. No component of $(e, u) \in S_{\leq r}$ has $u = 0^{\iota(e)}$ by the parsimony of $S$, and $v(S) \neq 0^{\iota(e(S))}$
2. If $(e, u) = (\epsilon \nu <^{\iota(\nu)} \alpha.\phi[\nu], u) \in H(S)$ and $u \neq ?$ then either $(e, u) \in S$, in which case $u <^{\iota(\nu)} \alpha$ by the parsimony of $S$, or $e = e(S)$ and $u <^{\iota(\nu)} \alpha$ follows by the definition of the $H$-step.
3. If $(n \in I^{\alpha}, \top), (n \in I^{\beta}, \top) \in H(S)$ then, since $S$ is parsimonious, it must be that (w.l.o.g.) $e(S) = (n \in I^{\alpha})$ and $v(S) = \top$. If $\alpha < \beta$ then $(n \in I^{\beta}, \top) \in \mathcal{P}(e(S), v(S))$, so we could not have $(n \in I^{\beta}, \top) \in H(S)$. If $\beta < \alpha$ then we have $n \in I^{\alpha} \hookrightarrow_{\overline{S}} \top$, so $n \in I^{\alpha}$ could not be the $H$-term of $S$. Therefore $\alpha = \beta$.
4. We cannot have $(n \in I^{<\alpha}, \top) \in S_{\leq r}$ for any $\alpha$ since $S$ is parsimonious, and $e(S) \neq (n \in I^{<\alpha})$.

□

**Definition 8.3.** *$S$ is computationally inconsistent (ci) if $\phi \hookrightarrow_S \perp$ for some $\phi \in \mathcal{F}(S)$ Otherwise it is cc.*

*$S$ is computing iff all formulas $\phi \in \mathcal{F}(S)$ are $S$-computable.*

*$S$ is deciding iff $S$ is computing and the critical formulas $Cr_0, \ldots, Cr_n$ are $S$-computable.*

**Lemma 8.2.** *If $S$ is a correct, nonsolving $\epsilon$-substitution then:*

1. *If $(e, u) \in S$, $u \neq ?$, and $(e, v) \in H(S)$ for some $v \neq u$ then $e = e(S)$, $v = v(S)$, $v \neq ?$ and $v < u$.*
2. *If $(n \in I^{\alpha}, \top) \in S$ and $(n \in I^{\beta}, \top) \in H(S)$ with $\beta \neq \alpha$ then $n \in I^{\alpha}$ is $e(S)$ and $\beta < \alpha$*

*Proof.* 1. Since $e = \epsilon \nu < s.\phi[\nu]$ must be the $H$-term of $S$, we consider which critical formulas could have $e$ as $H$-term. Since $\phi[u] \hookrightarrow_{\overline{S}} \top$ by the correctness of $\overline{S}$, the only axiom which could be false under $\overline{S}$ is the Epsilon Axiom, specifically we must have $(\epsilon \nu.\psi[\nu]) \leq t \hookrightarrow_{\overline{S}} \perp$, and therefore since $u \leq t \hookrightarrow_{\overline{S}} \perp$, we have $v = |t|_{\overline{S}} < u$.

2. $n \in I^\beta$ must be the $H$-term of $S$, so $Cr(S)$ must be an Inductive Definition Axiom, and if we had $\alpha \leq \beta$ then $Cr(S) \hookrightarrow_{\overline{S}} \top$, so we must have $\beta < \alpha$.

$\square$

**Definition 8.4.** *We say the $H$-rule* applies *to $S$ if $S$ is cc, deciding, and nonsolving.*

# 9 Cut Elimination

To prove that the $H$-process terminates, we will create an ad hoc sequent calculus. Our sequents will be $\epsilon$-substitutions augmented with additional information, and the resulting derivations will be similar to inverted $H$-processes, with the empty sequent on the bottom derived from axioms which will include solving $\epsilon$-substitutions. We will then apply a cut-elimination process which will result in a derivation from a solving $\epsilon$-substitution in which each inference corresponds exactly to the $H$-step would be taken (or to certain non-essential operations which do not affect that $H$-process); the well-foundedness of our derivation will then prove that the $H$-process reaches a solving substitution in finitely many steps.

Each pair $(e, u)$ in one of our sequents will be expanded to include a marker, which must be either $t$ (temporary) or $f$ (fixed). This will indicate whether or not that item may be updated in the derivation, that is, whether we allow a step above it in our derivation to represent an $H$-step in which $e$ is the $H$-expression. When a pair is fixed and we would like to update it, we will instead be required to stop at an axiom indicating that we would like to update the pair. (This is a generalization of the distinction between $?$ and $?^\circ$ in [MTB96]; in that paper, only pairs $(e, ?)$ can be updated, while here we must deal with the possibility that any value can be changed by a later $H$-step.)

**Definition 9.1.**

- *A* sequent $\Theta$ *is a set of tuples $(e, u, i)$ satisfying:*
  1. *$\Theta_S = \{(e, u) \mid (e, u, i) \in \Theta\}$ is an $\epsilon$-substitution*
  2. *$i \in \{t, f\}$ for each $(e, u, i) \in \Theta$*
  3. *If $(e, u, i), (e, u, j) \in \Theta$ then $i = j$*
- *A* historical sequent *is a triple $(\Theta, H, A)$ such that:*
  - *$\Theta$ is a sequent*
  - *$H = \langle S_1, \dots, S_n \rangle$ is a finite sequence of $\epsilon$-substitutions*
  - *$A$ is a set of canonical $\epsilon$-terms of rank $\Omega + 1$*
- $\mathrm{dom}(\Theta) = \mathrm{dom}(\Theta) = \mathrm{dom}(\Theta_S)$
- *Suppose $\Theta$ is a sequent, $e$ a canonical expression, $u \in V^{\iota(e)} \cup \{?\}$, and $i \in \{t, f\}$. Then $(e, u, i), \Theta = \Theta \cup \{(e, u, i)\}$ iff $\Theta \cup \{(e, u, i)\}$ is also a sequent; that is, either $e \notin \mathrm{dom}(\Theta)$ or $(e, u, i) \in \Theta$.*
- $\Theta t = \{(e, u, t) \in \Theta\}$
- $\Theta f = \{(e, u, f) \in \Theta\}$

- *We say* $\Theta \trianglelefteq \Sigma$ *if* $\Theta_S \trianglelefteq \Sigma_S$
- *If* $\bowtie \in \{<, \leq, >, \geq, =\}$ *then we say* $\Theta \bowtie r$ *if for every* $e \in \mathrm{dom}(\Theta)$*,* $rk(e) \bowtie r$*.*
- *We say* $\Sigma \lesssim r$ *if there is some* $\Sigma'$ *such that* $\Sigma = \Sigma'_{\leq r}$
- *We say* $r \lesssim \Sigma$ *if, whenever* $(e, u, i) \in \sigma$ *and* $rk(e) < r$ *then* $r = rk(e) + 1$ *and* $e$ *has the form* $n \in I^{<\alpha}$ *for some* $\alpha$
- $Ord(\Theta) = Ord(\Theta_S)$

If $(\Theta, H, A)$ is a historical sequent, we often only mention the sequent $\Theta$, and do not specify that $H$ and $A$ are also present.

## 9.1 $ID_\epsilon$

We introduce a deduction system for historical sequents with three groups of inferences and axioms. We will frequently find it useful to be able to refer to the premises of an inference by parameters, so we introduce notation to make this convenient

**Definition 9.2.** *If* $I$ *is an inference then* $Prem(I, x)$ *refers to the premise of* $I$ *indexed by* $x$*.*

*If* $I$ *as an inference and* $J$ *some instance of an inference occuring in one of the premises of* $I$*, we write* $Param(I, J) = x$ *when* $J$ *occurs in* $Prem(I, x)$*.*

Technically the premise is a deduction, but we will sometimes use $Prem(I, x)$ to refer to the endsequent of the premise; it will be clear from context when we are doing this. In general, if the parameters of $I$ other than ? range over the ordinals below $\alpha$, we equate ? with $\alpha$.

In the definitions below, $I$ always refers to the inference being defined.

### 9.1.1 Generic Axioms

$AxF$ $(\Theta, H, A)$ is an instance of $AxF$ if $\Theta_S$ is ci

$AxS$ $(\Theta, H, A)$ is an instance of $AxS$ is $\Theta_S$ is solving

### 9.1.2 Term Axioms and Inferences

$e$ is an $\epsilon$-term for all axioms and inferences in this subsection, and is called the main expression of the inference or axiom.

$AxH_{e,v}$ $((e, u, f), \Theta, H, A)$ is an instance of $AxH_{e,v}$ if $e$ is the $H$-term and $v$ the $H$-value of $((e, u, f), \Theta)_S$, and the $H$-rule applies

$AxPH_{e,v}$ $((e, ?, t), \Theta, H, A)$ is an instance of $AxH_{e,v}$ if $e$ is the $H$-term and $v$ the $H$-value of $((e, u, t), \Theta)_S$, the $H$-rule applies, and is at a corrected $H$-step, and $(e, v, \alpha) \in A$ for some $\alpha < rk(e)$

$Cut_e$

$$\frac{\{P(I,u) \mid u \in V^{\iota(e)} \cup \{?\}\}}{(\Theta, H, A)}$$

Where the endsequent of $P(I,?)$ is $((e,?,f),\Theta,H,A)$ and the endsequent of $P(I,u)$ for $u \neq ?$ is $((e,u,f),\Theta,H,A)$. We require that $Ord(e) \subseteq Ord(\Theta)$

$CutFr_e$

$$\frac{\{P(I,u) \mid u \in V^{\iota(e)} \cup \{?\}\}}{(\Theta, H, A)}$$

Where the endsequent of $P(I,u)$ is $((e,u,t),\Theta,H,A)$ for $u \neq ?$ and $((e,?,t),\Theta,H,A)$ for $u = ?$. We require that $Ord(e) \subseteq Ord(\Theta)$

$Fr_e$

$$\frac{((e,?,t),\Theta,H,A)}{(\Theta,H,A)}$$

Where $Ord(e) \subseteq Ord(\Theta)$

$H_{e,v}$

$$\frac{((e,v,t),\{(e',v',t) \mid (e',v' \in C(H^\frown\Theta_S \cup \{(e,v)\}))\},\Theta_{\leq rk(e)},H^\frown\Theta_S \cup \{(e,u)\},A)}{((e,u,t),\Theta,A)}$$

If $e$ is the $H$-term and $v$ the $H$-value of $((e,u,t),\Theta)_S$, and the $H$-rule applies to the conclusion.

$CutFr^*_e$

$$\frac{\{P(I,u) \mid u \in V^{\iota(e)} \cup \{?\}\}}{(\Theta, H, A)}$$

Where the endsequent of $P(I,u)$ is $((e,u,t),\Theta,H,A)$ for $u \neq ?$ and $((e,?,t),\Theta,H,A \cup \{e\})$ for $u = ?$ where $\alpha$ is some ordinal, and $\iota(e) = N$. We require that $Ord(e) \subseteq Ord(\Theta)$

### 9.1.3 Formula Axioms and Inferences

$n \in I^{<\alpha}$ (where $\alpha = \Omega$ when appropriate) is the main expression of the axiom or inference except for $FH^2$, $AxFH^2$, and $AxPFH^2$, which have main expression $n \in I^\alpha$. All variants of $FH$ come in two varieties 1 and 2; these are identical except that one applies when the formula being removed is of the form $n \in I^{<\alpha}$ and the other when it is of the form $n \in I^\alpha$. This means that $FH^1$ applies the first time we have an $H$-inference for $n$, and $FH^2$ applies every time we update which $\alpha$ is the first such that $n \in I^\alpha$.

These axioms and inferences are similar to the term axioms and inferences, although somewhat more complicated. We name them be prefixing an $F$ to the name to indicate that they refer to formulas. Axioms $AxPFH$ replace $AxFH$ after partial elimination of cuts (described in detail below).

$AxFH^1_{n,\alpha,\beta}$ $((n \in I^{<\alpha},?,f),\Theta,H,A)$ is an instance of $AxFH^1_{n,\alpha,\beta}$ if $n \in I^\beta$ is the $H$-term of $((n \in I^{<\alpha},?,f),\Theta)_S$ and this is not a closure rule.

$AxFH^2_{n,\alpha,\beta}$ $((n \in I^\alpha, \top, f), \Theta, H, A)$ is an instance of $AxFH^2_{n,\alpha,\beta}$ if $n \in I^\beta$ is the $H$-term of $((n \in I^\alpha, \top, f), \Theta)_S$.

$AxClFH_{n,\beta}$ $((n \in I^{<\Omega}, ?, f), \Theta, H, A)$ is an instance of $AxClFH_{n,\beta}$ if $n \in I^\beta$ is the $H$-term of $((n \in I^{<\Omega}, ?, f), \Theta)_S$ and this substitution is at a closure step.

$AxPFH^1_{n,\alpha,\beta}$ $((n \in I^{<\alpha}, ?, t), \Theta, H, A)$ is an instance of $AxPFH^1_{n,\alpha,\beta}$ if $n \in I^\beta$ is the $H$-term of $((n \in I^{<\alpha}, ?, t), \Theta)_S$ and this is not a closure rule.

$AxPFH^2_{n,\alpha,\beta}$ $((n \in I^\alpha, \top, t), \Theta, H, A)$ is an instance of $AxPFH^2_{n,\alpha,\beta}$ if $n \in I^\beta$ is the $H$-term of $((n \in I^\alpha, \top, t), \Theta)_S$.

$AxPClFH_{n,\beta}$ $((n \in I^{<\Omega}, ?, t), \Theta, H, A)$ is an instance of $AxClFH_{n,\beta}$ if $n \in I^\beta$ is the $H$-term of $((n \in I^{<\Omega}, ?, t), \Theta)_S$ and this substitution is at a closure step.

All these axioms also require that the $H$-rule apply.

$FCut_{n,\alpha,\beta}$

$$\frac{P(I,\alpha) \qquad \{P(I,\gamma) \mid \beta \leq \gamma < \alpha\}}{((n \in I^{<\beta}, ?, f), \Theta, H, A)}$$

Where the endsequent of $P(I,\alpha)$ is $((n \in I^{<\alpha}, ?, f), \Theta, H, A)$ and the endsequent of each $P(I,\gamma)$ is $((n \in I^\gamma, \top, f), \Theta, H, A)$. Also, we require that $\alpha > \beta$ and $\alpha \in Ord(\Theta)$. If $\beta = 0$ then the conclusion is $\Theta$ (and the component $(n \in I^{<\beta}, ?, f)$ is omitted).

$PCut_{n,\alpha,\beta,\delta}$

$$\frac{P(I,\alpha) \qquad \{P(I,\gamma) \mid \beta \leq \gamma < \delta\}}{((n \in I^{<\beta}, ?, f), \Theta, H, A)}$$

Where the endsequent of $P(I,\alpha)$ is $((n \in I^{<\alpha}, ?, t), \Theta, H, A)$ and the endsequent of each $P(I,\gamma)$ is $((n \in I^\gamma, \top, f), \Theta, H, A)$. Also, we require that $\alpha > \delta \geq \beta$ and $\alpha \in Ord(\Theta)$. If $\beta = 0$ then the conclusion is $\Theta$ (and the component $(n \in I^{<\beta}, ?, f)$ is omitted).

$PCutFr_{n,\alpha,\beta,\delta}$

$$\frac{P(I,\alpha) \qquad P(I,\delta) \qquad \{P(I,\gamma) \mid \beta \leq \gamma < \delta\}}{(e(n \in I^{<\beta}, ?, f), \Theta, H, A)}$$

Where the endsequent of $P(I,\alpha)$ is $((n \in I^{<\alpha}, ?, t), \Theta, H, A)$, the endsequent of $P(I,\delta)$ is $((n \in I^\delta, \top, t), \Theta, H, A)$, the endsequent of each $P(I,\gamma)$ is $((n \in I^\gamma, \top, f), \Theta, H, A)$. Also, we require that $\alpha > \delta \geq \beta$ and $\alpha \in Ord(\Theta)$. If $\beta = 0$ then the conclusion is $\Theta$ (and the component $(n \in I^{<\beta}, ?, f)$ is omitted).

$FFr_{n,\alpha,\beta}$

$$\frac{((n \in I^{<\alpha}, ?, t), \Theta, H, A)}{((n \in I^{<\beta}, ?, t), \Theta, H, A)}$$

Provided $\alpha > \beta$. The component $(n \in I^{<\beta}, ?, t)$ is omitted if $\beta = 0$. We require that $\alpha \in Ord(\Theta)$.

$FH^1_{n,\alpha,\beta}$
$$\frac{((n \in I^\alpha, \top, t), \Theta_{\leq rk(n \in I^\alpha)}, H^\frown \Theta_S \cup \{(n \in I^\alpha, ?)\}, A)}{((n \in I^{<\beta}, ?, t), \Theta, H, A)}$$

If $n \in I^\alpha$ is the $H$-term of $((n \in I^{<\beta}, ?, t), \Theta)_S$ and this substitution is not at a closure step, and the $H$-rule applies to the conclusion.

$FH^2_{n,\alpha,\beta}$
$$\frac{((n \in I^\alpha, \top, t), \Theta_{\leq rk(n \in I^\alpha)}, H^\frown \Theta_S \cup \{(n \in I^\alpha, ?)\}, A)}{((n \in I^{\beta}, \top, t), \Theta, H, A)}$$

If $n \in I^\alpha$ is the $H$-term of $((n \in I^{\beta}, \top, t), \Theta)_S$, and the $H$-rule applies to the conclusion.

$ClFH_{n,\alpha}$
$$\frac{((n \in I^\alpha, \top, t), \Theta_{\leq rk(n \in I^\alpha)}, H^\frown \Theta_S \cup \{(n \in I^\alpha, ?)\}, A)}{((n \in I^{<\Omega}, ?, t), \Theta, H, A)}$$

If $n \in I^\alpha$ is the $H$-term of $((n \in I^{<\Omega}, ?, t), \Theta)_S$ and this substitution is at a closure step, and the $H$-rule applies to the conclusion.

While there are a large number of axioms and inferences, most of the differences represent technical variations on the same basic axiom or inference. The $Cut$ and $FCut$ rules are what will be present in the initial derivation. They represent an uninformed guess as to what value to assign some expression we wish to evaluate, a question resolved by considering all possibilities as different branches. Viewed as inference rules, they can be taken to mean that the conclusion is sound precisely when at least one of the input branches is–the inputs represent all possible situations which expand on the conclusion in the necessary way.

The $FCut$, in particular, states that when $\neg n \in I^{<\beta}$ and $\beta < \alpha$, the possibilities are either that $\neg n \in I^{<\alpha}$, or $n \in I^\gamma$ where $\beta \leq \gamma < \alpha$. It represents the formula:

$$\neg n \in I^{<\beta} \to \neg n \in I^{<\alpha} \vee \bigvee_{\beta \leq \gamma < \alpha} n \in I^\gamma \tag{1}$$

Note that we read the inference going upwards.

Also present in the original derivation will be the $AxH$ axiom and its variants, the $AxFH^i$ and $AxClFH$ axioms. These represent attempts to actually apply an $H$-rule, places where the conclusion contains everything needed to pick an $H$-expresion and $H$-value for it, but where the relevant expression is fixed in value.

Our basic operation will be the replacement of $Cut$-type inferences and $AxH$-type axioms with the $Fr$ inference (and its variant the $FFr$ inference) and the $H$ inference (and its variants, the $FH^i$, and $ClFH$ inferences). The $Fr$-type inferences simply mark that we have used some expression–both premise and conclusion will compute every expression exactly the same way, but the premise notes that some default value has been used, which allows us to keep track of where branches belong in our derivation. The $H$-type inferences are the heart of our process, each corresponding to a different type of $H$-step. We could choose to have only one, awkwardly defined, inference for all these cases, but using different inferences makes the case distinctions we will need to make easier.

The remaining inferences are intermediate steps of various kinds, which will appear during cut-elimination, and all eventually be removed. The $PCut$ is an oddity resulting from the fact that, unlike in most cut-elimination arguments, the $FCut$ inference has premises with many different ranks. As a consequence, when we have eliminated cuts above some rank, but not below it, some of the $FCut$'s premises should behave as if the $Cut$ has been eliminated, but some should not. The $PCut$ represents this "partially eliminated" cut–it is essentially an $FFr$ inference with the remains of an $FCut$ added on.

The $CutFr$ and $PCutFr$ inferences are similar, but will exist only in the middle of cut-elimination. They will result when we eliminate a cut of some rank, but are still eliminating cuts of the same rank below that inference, and may need the extra information contained in the other premises. When we have finished eliminating cuts of that rank, we will prune the extra premises, resulting in an $Fr$ or $PCut$ inference.

**Definition 9.3.** *A* deduction *of* $(\Theta, H)$ *in* $ID_\epsilon$ *from a set* $\Sigma$ *of historical sequents is a wellfounded tree according to the rules of inference of* $ID_\epsilon$*. A* derivation *is a deduction from just the axioms of* $ID_\epsilon$*.*

**Definition 9.4.** *We define* $h(d) \leq \alpha$*, the* height *of* $d$ *inductively:*

1. *If* $d$ *is an axiom* $\Theta$ *and* $Ord(\Theta) \setminus \Omega \underline{\ll} \alpha$ *then* $h(d) \leq \alpha$

2. *If* $d$ *ends in an inference* $I$ *with endsequent* $\Theta$ *such that* $Ord(\Theta) \setminus \Omega \underline{\ll} \alpha$ *and for each* $\gamma$ *such that* $Prem(I, \gamma)$ *is defined,* $h(Prem(I, \gamma)) \leq \alpha_\gamma$*, and:*

   (a) $\alpha_\gamma < \alpha$
   
   (b) *If* $\alpha \underline{\ll} \beta$*, and* $\gamma \ll \beta$ *then* $\alpha_\gamma \ll \beta$

   *Then* $h(d) \leq \alpha$*. (This definition is essentially that of [Poh89], Definition 24.27.)*

**Lemma 9.1.** *If* $rk(e) = r$ *then* $r = \alpha + n$ *where* $\alpha$ *is the largest ordinal appearing in* $e$ *or* $\Omega$ *and* $n < \omega$*. Therefore if* $Ord(\{e\}) < \xi$*,* $rk(e) < \Omega$*, and* $\xi$ *is strongly critical then* $rk(e) < \xi$*.*

*Proof.* The first part follows by straightforward induction on the definition of rank. The second follows since $\xi$ is strongly critical, so $\xi > \alpha + \omega > rk(e)$. □

**Lemma 9.2.** *If* $d$ *is a derivation ending in* $\Theta$ *and* $\alpha \in Ord(\Theta)$ *then for every* $\eta$ *such that* $h(d) \leq \eta$*,* $\alpha \underline{\ll} \eta$*.*

*Proof.* By straightforward bottom-up induction on $d$. □

## 9.2 Original Derivation

The construction of the original derivation in this section is the same as in [MTB96] for $\epsilon$-terms $e$, the step described in Lemma 9.4. In Lemma 9.5 we use $FCut$ inferences to decide formulas $n \in I^{<\alpha}$ and $n \in I^{\alpha}$. We do not decide $n \in I^{\alpha}$ by applying a cut directly to this formula; instead we decide $n \in I^{<\alpha+1}$.

Essentially, we attempt to evaluate the critical formulas; if our substitution is correct but non-solving, we will find canonical expressions which have not been assigned values. We apply a cut over some canonical expression appearing in our critical formulas, and repeat the process for every premise of the cut. In order to show that the process halts, we always choose canonical subexpressions of formulas having the maximum possible rank.

We add one essential trick: in order to make sure that ordinals decrease both in terms of $<$ and $\ll$, we have to add ordinals to the height as we add them to sequents. In order to make "room" for this, we have to count the number of places where ordinals could potentially be added.

**Definition 9.5.** *Define $\chi(e)$ by induction:*

1. $\chi(\nu) = \chi(0^{\iota}) = \chi(\alpha) = \chi(\bot) = \chi(\top) = 0$
2. $\chi(St) = \chi(t)$
3. $\chi(pe_1, \ldots, e_n) = \Sigma_{i=1}^{n} \chi(e_i)$
4. $\chi(s_1 =^O s_2) = \chi(s_1 <^O s_2) = \chi(s_1) + \chi(s_2)$
5. $\chi(\neg\phi) = \chi(\phi)$
6. $\chi(\phi \wedge \psi) = \chi(\phi \rightarrow \psi) = \chi(\phi) + \chi(\psi)$
7. $\chi(t \in I^s) = \chi(t \in I^{<s}) = \chi(t) + \chi(s)$
8. $\chi(t \in I^{<\Omega}) = \chi(t) + 1$
9. $\chi(s \in \{t_1, \ldots, t_n \mid s_1, \ldots, s_n\}) = \chi(s)$
10. $\chi(\epsilon\nu < \alpha\phi[\nu]) = \begin{cases} \chi(\phi) + 1 & \textit{if } \alpha = \Omega \\ \chi(\phi) & \textit{otherwise} \end{cases}$

This essentially measures the number of places an unbounded countable ordinal might appear in an expression.

**Lemma 9.3.** *If $e \hookrightarrow^1_S e'$ then $\chi(e') \leq \chi(e)$.*

*Proof.* By the definition of $\hookrightarrow^1_S$. □

**Definition 9.6.** *Let $S$ be an $\epsilon$-substitution and $\Phi = \{A_1, \ldots, A_n\}$ a finite set of closed formulas.*

- $\rho_S(\Phi) = \max\{rk(|A|_S) \mid A \in \Phi, d(|A|_S) > 0\} \cup \{0\}$

- $\chi_S(\Phi, r) = \Sigma_{A\in\Phi, rk(|A|_S)=r}\Omega \cdot \chi(|A|_S) + \#_{\alpha\in Ord(S)}\#_{\beta\in SC(\alpha)}\beta$
- $d_r(F) = \left\{ \begin{array}{cl} 0 & \textit{if } rk(F) < r \\ d(F) & \textit{otherwise} \end{array} \right.$
- $\mu_S(\Phi, r) = \Sigma_{A\in\Phi} d_r(|A|_S)$
- $\nu_S(\Phi) = \Omega^2 \cdot \rho_S(\Phi) + \omega\chi_S(\Phi, \rho_S(\Phi)) + \mu_S(\Phi, \rho_S(\Phi))$

Note $\mu_S(\Phi, r) < \omega$, $\chi_S(\Phi, r) < \omega \cdot \Omega$, $\rho_S(\Phi) < \Omega + \omega$, and therefore $\nu_S(\Phi) < \Omega^3 + \Omega^2$.

**Definition 9.7.**

$$||u||_A = \left\{ \begin{array}{ll} \Omega & \textit{if } u = ? \\ u & \textit{if } u \textit{ is an ordinal or number} \\ 1 & \textit{if } u = \top \end{array} \right\}$$

**Definition 9.8.** *If $S$ is an $\epsilon$-substitution such that the $H$-rule applies to $S$, define an $\epsilon$-substitution $ext_i(S)$ and a set of formulas $\Phi_i(S)$ for $i < \omega$ as follows:*

1. $ext_0(S) = \emptyset$ *and* $\Phi_0(S) = \{Cr_0, \ldots, Cr_N\}$
2. *If the $H$-rule applies to $ext_i(S)$, $ext_{i+1}(S) = ext_i(S)$ and $\Phi_{i+1}(S) = \Phi_i(S)$*
3. *If the $H$-rule does not apply to $ext_i(S)$, let $A_0 \in \Phi_i(S)$ be such that $rk(|A_0|_{ext_i(S)})$ is maximal, and choose some canonical subexpression $e$ of $|A_0|_{ext_i(S)}$. If $e$ is computed by $S$ then $ext_{i+1}(S) = ext_i(S) \cup \{e, |e|_S\}$ and $\Phi_{i+1}(S) = \Phi_i(S) \cup \{F(e, |e|_S)\}$. Otherwise $ext_{i+1}(S) = ext_i(S) \cup \{(e, ?)\}$ and $\Phi_{i+1}(S) = \Phi_i(S)$.*

*$n(S)$ is the leas $n$ such that $ext_n(S) = ext_{n+1}(S)$. $ext(S) = ext_{n(S)}(S)$.*

*Finally if $ext_{i+1} \setminus ext_i(S) = (e, u)$ then $e(i, S) = e$ and $v(i, S) = u$ if $e$ is a term, and $\alpha$ if $e$ is the formula $n \in I^\alpha$ or $n \in I^{<\alpha}$.*

**Definition 9.9.** *Define*

$$ind(S) = \Sigma_{i<n(S)}\Omega^{\omega\nu_{ext_i(S)}(\Phi_i(S))}||v(i, S)||_A + \Omega^{\omega\nu_S(\Phi_{n(S)}(S))+2}$$

The following two lemmas appear complicated, but the concept is simple: we are given an $\epsilon$-substitution and a finite set of formulas. We select a canonical expression $e$ of maximal rank from $\{|A|_S \mid A \in \Phi\}$, assign it a value $u$, and augment $\Phi$ to include a witness to the correctness of $(e, u)$, if necessary. Then we show that the resulting measure by $\nu$ decreases according to both $<$ and $\ll$. We will also show, in particular, that if we are adding expressions in the order used to define $ind(S)$ then also the indices are decreasing.

**Lemma 9.4.** *Let $S$ be an $\epsilon$-substitution and $\Phi$ a finite set of closed formulas such that $\bigcup_{A\in\Phi} Ord(A) \subseteq Ord(S)$. Let $A_0 \in \Phi$ with $rk(|A_0|_S) = \rho_S(\Phi)$, and let $e$ be a canonical $\epsilon$-subterm of $|A_0|_S$. For any $u \in V^{\iota(e)} \cup \{?\}$ let $S^u = S \cup \{(e, u)\}$ and let*

$$\Phi^u = \left\{ \begin{array}{cl} \Phi & \textit{if } u = ? \\ \Phi \cup \{F(e, u)\} & \textit{otherwise} \end{array} \right.$$

*Then for any $u \in V^{\iota(e)} \cup \{?\}$:*

- *$S^u$ is an $\epsilon$-substitution*
- $\rho_{S^u}(\Phi^u) \leq \rho_S(\Phi)$
- $\nu_{S^u}(\Phi^u) < \nu_S(\Phi)$
- *if $\nu_S(\Phi) \underline{\ll} \beta$ and $u \ll \beta$ then $\nu_{S^u}(\Phi^u) \ll \beta$*
- $\bigcup_{A \in \Phi^u} Ord(A) \subseteq Ord(S^u)$
- *if $\Phi = \Phi_{n(S)}(S)$ and $e = e(n(S), S^u)$ then $ind(S^u) < ind(S)$ and if $ind(S) \underline{\ll} \beta$ and $u \ll \beta$ then $ind(S^u) \ll \beta$.*

*Proof.*

- $S^u$ is an $\epsilon$-substitution
  Trivial, since $S$ is an $\epsilon$-substitution
- $\rho_{S^u}(\Phi^u) \leq \rho_S(\Phi)$
  Since $S \subseteq S^u$, we have $||w|_S|_{S^u} = |w|_{S^u}$, and therefore $rk(|w|_{S^u}) \leq rk(|w|_S)$ and $d(|w|_{S^u}) \leq d(|w|_S)$ for each $w$. Also, $rk(|F(e,u)|_{S^u}) \leq rk(F(e,u)) < rk(e) \leq \rho_S(\Phi)$, so $\rho_{S^u}(\Phi^u) \leq \rho_S(\Phi)$.
- $\nu_{S^u}(\Phi^u) < \nu_S(\Phi)$
  If $\rho_{S^u}(\Phi^u) < \rho_S(\Phi)$ then this is obviously the case, so suppose $\rho_{S^u}(\Phi^u) = \rho_S(\Phi)$. If $\max Ord(S^u) > \max Ord(S)$ then it must be that some $\Omega$ has been removed, and therefore $\chi(|A_0|_{S^u}) < \chi(|A_0|_S)$. Therefore $\chi_{S^u}(\Phi^u, \rho_S(\Phi)) \leq \chi_S(\Phi, \rho_S(\Phi))$. If this inequality is strict then we must have $\nu_{S^u}(\Phi^u) < \nu_S(\Phi)$, so assume $\chi_{S^u}(\Phi^u, \rho_S(\Phi)) = \chi_S(\Phi, \rho_S(\Phi))$. For each $A \in \Phi$, $rk(|A|_{S^u}) \leq rk(|A|_S)$ and $d(|A|_{S^u}) \leq d(|A|_S)$. Therefore $\mu_{S^u}(\Phi^u) \leq \mu_S(\Phi)$. But since $d(|A_0|_{S^u}) < d(|A_0|_S)$, this inequality must be strict, so $\nu_{S^u}(\Phi^u) < \nu_S(\Phi)$.
- If $\nu_S(\Phi) \underline{\ll} \beta$ and $u \ll \eta$ then $\nu_{S^u}(\Phi^u) \ll \beta$
  If $\nu_{S^u}(\Phi) \ll \nu_S(\Phi)$ then this follows from transitivity. Otherwise we have $D\nu_S(\Phi) \leq D\nu_{S^u}(\Phi^u)$, and therefore we must have some $\eta \in SC(\nu_{S^u}(\Phi^u))$ such that $D\nu_S(\Phi) \leq \eta$. We must have $\eta \notin SC(\nu_S(\Phi))$. But $SC(\nu_S(\Phi)) \setminus SC(\nu_{S^u}(\Phi^u)) \subseteq SC(u)$. But then we must have $D\nu_S(\Phi) < Du$, so $u \not\ll \nu_S(\Phi)$.
- $\bigcup_{A \in \Phi^u} Ord(A) \subseteq Ord(S^u)$
  Since $Ord(\Phi^u) = Ord(\Phi) \cup Ord(u)$ this follows since $Ord(u) \subseteq Ord(\{(e,u)\})$ and $Ord(\Phi) \subseteq Ord(S)$.
- if $\Phi = \Phi_{n(S)}(S)$ and $e = e(n(S), S^u)$ then $ind(S^u) < ind(S)$ and if $ind(S) \underline{\ll} \beta$ and $u \ll \beta$ then $ind(S^u) \ll \beta$
  Since $ext_{n(S)-1}(S^u) = S$ and $\Phi_{n(S)}(S^u) = \Phi^u$, let

  $$\zeta = \Sigma_{i<n(S)} \Omega^{\omega \nu_{ext_i(S)}(\Phi_i(S))} ||v(i,S)||_A$$

  Then $ind(S) = \zeta + \Omega^{\omega \nu_S(\Phi)+2}$ while $ind(S^u) = \zeta + \Omega^{\omega \nu_S(\Phi)} ||u||_A + \Omega^{\omega \nu_{S^u}(\Phi^u)+2}$. But since $||u||_A \leq \Omega$ and $\nu_{S^u}(\Phi^u) < \nu_S(\Phi)$, $ind(S) < ind(S^u)$.

If $ind(S) \underline{\ll} \beta$ and $u \ll \beta$ then $ind(S^u) \ll \beta$ follows since no strongly critical ordinals appear in $ind(S^u)$ that do not appear in $\nu_{S^u}(\Phi^u)$.

□

The following lemma is essentially the same as the previous one, using a formula instead of a term. One complication arises from the fact that we always need to consider two kinds of formulas: $n \in I^{<\beta}$, and $n \in I^\gamma$ for all $\gamma < \beta$. If the canonical subformula has the form $n \in I^{<\alpha}$ then $\beta = \alpha$, however if it has the form $n \in I^\alpha$, we need to set $\beta = \alpha + 1$, since we need to consider the case where we add $(n \in I^\alpha, \top)$.

**Lemma 9.5.** *Let $S$ be an $\epsilon$-substitution and $\Phi$ a finite set of closed formulas such that $\bigcup_{A \in \Phi} Ord(A) \subseteq Ord(S)$. Let $A_0 \in \Phi$ with $rk(|A_0|_S) = \rho_S(\Phi)$ and let $e'$ be some canonical subformula of $|A_0|_S$. If $e' = I^{<\alpha}$ then set $e = (n \in I^{<\alpha})$ and $\beta = \alpha$, otherwise set $e = (n \in I^{<\alpha+1})$ and $\beta = \alpha + 1$. If there is no component $(n \in I^{<\xi}, ?) \in S$ then let $\delta = 0$, otherwise let $\delta = \xi$. For each $\gamma$ such that $\delta \leq \gamma < \beta$, let $S^\gamma = S \cup \{(n \in I^\gamma, \top)\} \setminus \mathcal{P}(n \in I^\gamma, \top)$ and let $\Phi^\gamma = \Phi \cup \{F(n \in I^\gamma, \top)\}$. Let $S^\beta = S \cup \{(n \in I^{<\beta}, ?)\} \setminus \mathcal{P}(n \in I^{<\beta}, ?)$ and let $\Phi^\beta = \Phi$.*

*Then for every $\delta \leq \gamma \leq \beta$:*

- *$S^\gamma$ is an $\epsilon$-substitution and $g \hookrightarrow_S g'$ implies $g \hookrightarrow_{S^\gamma} g'$*
- *$\rho_{S^\gamma}(\Phi^\gamma) \leq \rho_S(\Phi)$*
- *$\nu_{S^\gamma}(\Phi^\gamma) < \nu_S(\Phi)$*
- *if $\nu_S(\Phi) \underline{\ll} \kappa$ and $\gamma \ll \kappa$ then $\nu_{S^\gamma}(\Phi^\gamma) \ll \kappa$*
- *$\bigcup_{A \in \Phi^\gamma} Ord(A) \subseteq Ord(S^\gamma)$.*
- *if $\Phi = \Phi_{n(S)}(S)$ and $e' = e(n(S), S^u)$ then $ind(S^\gamma) < ind(S)$ and if $ind(S) \underline{\ll} \kappa$ and $\gamma \ll \kappa$ then $ind(S^\gamma) \ll \kappa$*

*Proof.* Similar to Lemma 9.4.

- $S^\gamma$ is an $\epsilon$-substitution and $g \hookrightarrow_S g'$ implies $g \hookrightarrow_{S^\gamma} g'$
  We consider three cases:
  1. There is a component $(n \in I^{<\xi}, ?) \in S$. If $\beta \leq \xi$ then $e'$ is already decided, so we have $\xi < \beta$. Then for $S^\gamma$ with $\xi \leq \gamma \leq \beta$, we have that $S^\gamma$ is an $\epsilon$-substitution since $(n \in I^{<\xi}, ?) \in \mathcal{P}(n \in I^\gamma, \top)$. Also, since $S$ is an $\epsilon$-substitution, $S^\gamma$ meets the remaining conditions.
  2. There is a component $(n \in I^\xi, \top) \in S$. Impossible, since then $e$ would be decided.
  3. There is no such component. Then clearly $S^\gamma$ is an $\epsilon$-substitution since $S$ is.

- $\rho_{S^\gamma}(\Phi^\gamma) \leq \rho_S(\Phi)$
  Since $S \subseteq S^\gamma$, we have $||w|_S|_{S^\gamma} = |w|_{S^\gamma}$, and therefore $rk(|w|_{S^u}) \leq rk(|w|_S)$ and $d(|w|_{S^u}) \leq d(|w|_S)$ for each $w$. Also $rk(|A[I^{<\gamma}, n]|_{S^\gamma}) \leq rk(|A[I^{<\gamma}, n]|_S) < rk(e') \leq \rho_S(\Phi)$.

- $\nu_{S^\gamma}(\Phi^\gamma) < \nu_S(\Phi)$
  As in Lemma 9.4, if we do not have $\rho_{S^\gamma}(\Phi^\gamma) < \rho_S(\Phi)$ then we have $\rho_{S^\gamma}(\Phi^\gamma) = \rho_S(\Phi)$. In this case, if $\max Ord(S^\gamma) > \max Ord(S)$ then we must have $\beta = \Omega$ and $\Omega > \gamma > \max Ord(S)$, and therefore $\chi_{S^u}(\Phi^u, \rho_S(\Phi)) < \chi_S(\Phi, \rho_S(\Phi))$. Otherwise $\chi_{S^u}(\Phi^u, \rho_S(\Phi)) = \chi_S(\Phi, \rho_S(\Phi))$, and in that case, since $e'$ is $S^\gamma$ computable, we have we have $\mu_{S^\gamma}(\Phi^\gamma) < \mu_S(\Phi)$. In either case, $\nu_{S^\gamma}(\Phi^\gamma) < \nu_S(\Phi)$.

- if $\nu_S(\Phi) \underline{\ll} \kappa$ and $\gamma \ll \kappa$ then $\nu_{S^\gamma}(\Phi^\gamma) \ll \kappa$
  If we do not have $\nu_{S^\gamma}(\Phi^\gamma) \ll \nu_S(\Phi)$ then we have $D\nu_S(\Phi) < D\gamma$ and if $\gamma \ll \kappa$ and $\nu_S(\Phi) < \kappa$ then $\nu_{S^\gamma}(\Phi^\gamma)$.

- if $\Phi = \Phi_{n(S)}(S)$ and $e' = e(n(S), S^u)$ then $ind(S^\gamma) < ind(S)$ and if $ind(S) \underline{\ll} \kappa$ and $\gamma \ll \kappa$ then $ind(S^\gamma) \ll \kappa$
  Finally, since $ext_{n(S)-1}(S^\gamma) = S$ and $\Phi_{n(S)}(S^\gamma) = \Phi^\gamma$, let

  $$\zeta = \Sigma_{i<n(S)} \Omega^{\omega \nu_{ext_i(S)}(\Phi_i(S))} v(i, S)$$

  Then $ind(S) = \zeta + \Omega^{\omega\nu_S(\Phi)+2}$ while $ind(S^\gamma) = \zeta + \Omega^{\omega\nu_S(\Phi)}\gamma + \Omega^{\omega\nu_{S^\gamma}(\Phi^\gamma)+2}$. But since $\gamma \leq \Omega$ and $\nu_{S^\gamma}(\Phi^\gamma) < \nu_S(\Phi)$, $ind(S) < ind(S^\gamma)$.

  If $ind(S) \underline{\ll} \kappa$ and $\gamma \ll \beta$ then $ind(S^\gamma) \ll \kappa$ follows since no strongly critical ordinals appear in $ind(S^\gamma)$ that do not appear in $\nu_{S^\gamma}(\Phi^\gamma)$.

□

The Lemma above corresponds to the $FCut$ inference:

$$\frac{(n \in I^{<\beta}, ?), S \qquad (n \in I^\gamma, \top), S, \ldots \qquad \forall\gamma(\delta \leq \gamma < \beta \in V^O)}{(n \in I^{<\delta}, ?), S}$$

The following two lemmata use the previous ones to actually construct a deduction of $\emptyset$. Lemma 9.6 inductively uses Lemma 9.4 and Lemma 9.5 to construct $Cut$ and $FCut$ inferences. Lemma 9.7 applies this to our base case–the empty sequent and the critical formulas we are concerned with–to produce a derivation, and verifies that this derivation is in fact a deduction.

**Lemma 9.6.** *Suppose $\Theta$ is a sequent with $\Theta t = \emptyset$, $L$ a finite set of closed formulas with $\bigcup_{A \in L} Ord(A) \subseteq Ord(\Theta_S)$, and $r = \rho_{\Theta_S}(\mathcal{F}(\Theta_S) \cup L)$. Then there is a derivation $d$ of $(\Theta, \emptyset)$ by $Cuts$ and $FCuts$ of ranks $\leq r$ from computing sequents $\Upsilon$ with $\Upsilon t = \emptyset$ containing $\Theta$ and computing all formulas in $L$. In addition, $h(d) \leq \nu_{\Theta_S}(\mathcal{F}(\Theta_S) \cup L)$.*

*Proof.* By induction on $\nu_{\Theta_S}(\mathcal{F}(\Theta_S) \cup L)$.

Note that we have

$$\max Ord(\Theta) \lll \chi_{\Theta_S}(\mathcal{F}(\Theta_S) \cup L) \lll \nu_{\Theta_S}(\mathcal{F}(\Theta_S) \cup L)$$

Let $\Phi = \mathcal{F}(\Theta_S) \cup L$. If $\Theta$ computes all formulas in $\Phi$ then $\Theta$ satisfies the condition. Otherwise, let $A \in \Phi$ be some formula such that $rk(|A|_{\Theta_S}) = r$ and let $g$ be some canonical subexpression of $|A|_{\Theta_S}$.

If $e$ is a term such that $e \notin \mathrm{dom}(\Theta)$, and each $u \in V^{\iota(e)} \cup \{?\}$, let $\Theta^u = (e, u, f), \Theta$. Then $\Theta^u_S$ satisfies the conditions of Lemma 9.4, so $\nu_{\Theta^u}(\mathcal{F}(\Theta^u_S) \cup L) \ll \nu_\Theta(\Phi)$, and by IH there is a derivation $d_u$ of $\Theta^u$ by $Cut$s and $FCut$s of rank $\leq r$ from appropriate sequents $\Upsilon$ and with $h(d_u) \leq \nu_{\Theta^u}(\mathcal{F}(\Theta^u_S) \cup L)$, and since by Lemma 7.1 $Ord(e) \subseteq Ord(\Theta)$, a $Cut$ with main term $e$ satisfies the theorem.

If $e$ is a formula then if it is $n \in I^\alpha$, let $\beta = \alpha + 1$, and if it is $n \in I^{<\alpha}$ let $\beta = \alpha$. Then for each $\gamma < \beta$ let $\Theta^\gamma = (n \in I^\gamma, \top, f), \Theta$ and let $\Theta^\beta = (n \in I^{<\beta}, ?, f), \Theta$. Then $\Theta^\gamma_S$ satisfies the conditions of Lemma 9.5, so $\nu_{\Theta^\gamma}(\mathcal{F}(\Theta^\gamma) \cup L) \ll \nu_\Theta(\Phi)$, and by IH there is a derivation $d_\gamma$ of $\Theta^\gamma$ by $Cut$s and $FCut$s of rank $\leq r$ from appropriate sequents $\Upsilon$ and with $h(d_\gamma) \leq \nu_{\Theta^\gamma}(\mathcal{F}(\Theta^\gamma) \cup L)$, and since by Lemma 7.1 $Ord(n \in I^\beta) \subseteq Ord(\Theta)$, an $FCut$ with main term $e$ satisfies the theorem. □

**Lemma 9.7.** *There is some $r < \Omega + \omega$ such that there is a derivation $d$ of the empty sequent consisting only of axioms, $Cut$s and $FCut$s with rank $\leq r$, and $h(d) \leq \Omega^3 + \Omega^2$.*

*Proof.* Applying Lemma 9.6 to $\emptyset$, $L = \{Cr_0, \ldots, Cr_N\}$ and $r = \rho_\emptyset(L)$ gives a deduction of $\emptyset$ consisting of only $Cut$s and $FCut$s with rank $\leq \rho_\emptyset(L)$ and axioms.

If some top sequent $\Theta$ of this deduction is not an axiom then $\Theta_S$ must be cc, deciding, and nonsolving. Since the only inferences in the part already constructed are cuts, $\Theta t = \emptyset$. But then $\Theta$ has an $H$-expression $e(\Theta_S)$ appearing in some $Cr_I$. Since $\Theta$ is cc and deciding, $Cr_I$ must be computed, and therefore $e(\Theta_S)$ must be computed, so there must be some $(e, u) \in \mathcal{P}(e(\Theta_S), v(\Theta_S))$ such that $(e, u, f) \in \Theta$. Note that requirements on ordinals of cut terms are satisfied by Lemma 7.1.

Then this must be an $AxH$, $AxFH$, or $AxClFH$ axiom. □

In place of the height bounds given by Lemma 9.7, we will use the height bound given by the function $ind(S)$, which gives a derivation $ind(d) \leq \Omega^{\Omega^3+\Omega^2+1}$. This can obviously be done, using $ind(S)$ in place of $\nu_S(\Phi)$ in Lemma 9.6. While these appears to do nothing but inflate our height bounds, it serves the purpose of synchronizing the height bounds with the ordinal assignment, making a straightforward collapsing argument possible.

## 9.3 Controlling Derivations

A derivation with cuts of rank $r$ and higher eliminated will be called an $r$-derivation. We will define steps which will allows us to convert, for instance, $r + 1$-derivations to $r$-derivations. We will begin at the top of the derivation, and work down to the root (using the well-foundedness of the derivation). As described above, we cannot directly produce an $r$-derivation as we go down, since we may need the additional information retained by $PCutFr$ and $CutFr$ inferences; derivations with this information left in will be called $r^+$-derivations. At the interim stages, we will have $r+1$-derivations below some inference, and $r^+$-derivations above. Once the entire derivation is $r^+$, we will be able to easily prune it into an $r$-derivation.

We will define notations $X(d) \bowtie r$, where $\bowtie$ is some comparison like $<$ or $\geq$ and $X$ is some inference rule or axiom, to indicate that all instances of that rule have the appropriate relation to the rank $r$. $PCutFr$, $FFr$, and the axioms will have slightly modified definitions, and we will need some additional information about $PCut$ inferences, which will be denoted by $PCutF(d) = r$ or $PCutF(d) \approx r$.

**Definition 9.10.** *The* target rank *of an $AxH_{e,v}$ axiom is $rk(e)$.*

*The* target rank *of an $AxFH_{n,\alpha,\beta}$, $AxPFH_{n,\alpha,\beta}$, $AxClFH_{n,\beta}$, or $AxPClFH_{n,\beta}$ is $rk(n \in I^\beta)$.*

**Definition 9.11.** *If the $H$-rule applies to $\Theta$, we say $(\Theta, H)$* conflicts *with $(e, u, i)$ if one of the following holds:*

- $e(\Theta) \equiv n \in X$, $e \equiv n \in Y$, $u = ?$, *and* $rk(n \in Y) \geq rk(n \in X)$
- $e = e(\Theta)$ *and* $u \neq v(\Theta)$
- $(e, v) \in C(H)$ *and* $v \neq u$
- $e \equiv n \in I^\alpha$ *and* $e(\Theta) \equiv \epsilon x \neg |B[I^{<\alpha}, n, x]|_{\overline{S}}$

*We say $\Theta$ conflicts with $(\Sigma, H, A)$ if:*

- *There is some $(e, u, i) \in \Sigma$ such that $\Theta$ conflicts with $(e, u, i)$.*
- *There is some $(e, v) \in C(H)$ such that $e \in A$.*

**Definition 9.12.** *Let $d$ be a derivation.*

- *If*
$$X \in \{Cut, CutFr, CutFr^*, Fr, H, FCut, PCut, FH, ClFH\}$$
*and $\bowtie \in \{<, >, \leq, \geq, =\}$ then we say $X(d) \bowtie r$ if every application of a rule $X$ has main expression with rank $\bowtie r$.*
- *We say $PCutF(d) = r$ if every $PCut_{n,\alpha,\beta,\delta}$ appearing in $d$ satisfies $rk(n \in I^\delta) > r$ and $rk(n \in I^\gamma) < r$ for all $\gamma < \delta$.*
- *We say $PCutF(d) \approx r$ if for every $PCut_{n,\alpha,\beta,\delta}$ appearing, $\delta$ is the least ordinal such that $rk(n \in I^\delta) \geq r$.*
- *We say $PCutFr(d) = r$ if every $PCutFr_{n,\alpha,\beta,\delta}$ appearing in $d$ satisfies $rk(n \in I^\delta) = r$.*
- *If $\bowtie \in \{>, \geq\}$ then we say $FFr(d) \bowtie r$ if every $FFr_{n,\alpha,\beta}$ appearing in $d$ satisfies $rk(n \in I^{<\beta}) \bowtie r$.*
- *We say $Ax(d) \leq r$ if every $AxH$-type axiom $\Theta$ has target rank $\leq r$, and if the target rank is $r$ then $\Theta$ conflicts with the endsequent of $d$.*

- *Both $r$- and $r^+$-derivations are derivations satisfying certain restrictions on the axioms and inferences they include. The restrictions are stated in this table:*

| $r$-derivation | | | $r^+$-derivation | | |
|---|---|---|---|---|---|
| $Cut(d)$ | $<$ | $r$ | $Cut(d)$ | $<$ | $r$ |
| $FCut(d)$ | $<$ | $r$ | $FCut(d)$ | $<$ | $r$ |
| $PCutF(d)$ | $\approx$ | $r$ | $PCutF(d)$ | $=$ | $r$ |
| $H(d)$ | $\geq$ | $r$ | $H(d)$ | $\geq$ | $r$ |
| $FH(d)$ | $\geq$ | $r$ | $FH(d)$ | $\geq$ | $r$ |
| $ClFH(d)$ | $\geq$ | $r$ | $ClFH(d)$ | $\geq$ | $r$ |
| $PCut(d)$ | $\geq$ | $r$ | $PCut(d)$ | $\geq$ | $r$ |
| $Fr(d)$ | $\geq$ | $r$ | $Fr(d)$ | $>$ | $r$ |
| $FFr(d)$ | $\geq$ | $r$ | $FFr(d)$ | $>$ | $r$ |
| $CutFr(d)$ | $<$ | $0$ | $CutFr(d)$ | $=$ | $r$ |
| $CutFr^*(d)$ | $=$ | $\Omega+1$ | $CutFr^*(d)$ | $=$ | $\Omega+1$ |
| $PCutFr(d)$ | $<$ | $0$ | $PCutFr(d)$ | $=$ | $r$ |
| $Ax(d)$ | $\leq$ | $r$ | $Ax(d)$ | $\leq$ | $r$ |

*In addition:*

- *If a $CutFr^*$ inference occurs in an $r$- or $r^+$-derivation then $r \leq \Omega+1$*
- *If $u \neq ?$ and $I$ is a $CutFr^*$ inference then $Prem(I,u)$ is an $\Omega+1$- and an $\Omega+1^+$-derivation (in addition to being an $r$- or $r^+$-derivation)*
- *If $I$ is an $H$-inference of rank $r$ in an $r^+$-derivation and at a corrected $H$-step then the premise of $I$ is an $\Omega+1$- and an $\Omega+1^+$-derivation (in addition to being an $r$- or $r^+$-derivation)*
- *If $(\Theta, H, A)$ is the premise of an $Fr_e$ inference and $rk(e) = \Omega+1$ then either $r \leq o(H^\frown\{(e,?)\} \cup \Theta_S)$ or $e \in A$*

This means that in an $r$ or $r^+$ derivation, all $Cut$ and $FCut$ inferences have rank below $r$, while all $H$, $Fr$, $FH$, and $FFr$ inferences have ranks greater than $r$. Note that we want the conclusion of the $FFr$ inference to have rank at least $r$, not just the premise. $PCut$ inferences are required to 'span' the cut-rank, in the sense that the only premise adding a temporary value adds an expression with rank at least $r$ while the premises adding fixed values add expressions with rank below $r$. $CutFr$ and $PCutFr$ inferences are required to be situated precisely at $r$. For $CutFr$ this means the main term has rank $r$, while for $PCutFr$, this means that the premise $(n \in I^\delta, \top, t)$ which distinguishes it from $PCut$ will have rank $r$.

We wish to measure the height of $r$- and $r^+$-derivations more precisely to ensure that we do not take an overly constrictive upper bound which works before we collapse, but is not generous enough when we try to collapse.

**Definition 9.13.** *If $d$ is an $r$- or $r^+$-derivation and $r \geq \Omega$ then we say $h(d) \leq \eta$ only if the following additional inductive criterion is met: for every $H$-type axiom $(\Sigma, H, A)$ in $d$, $h((\Sigma, H, A)) \leq \eta$ implies $o(H^\frown\Sigma; r)$.*

**Lemma 9.8.** *Suppose $\mathcal{C}$ is some $CutFr_e^*$ inference in an $r$-derivation of $\emptyset$ for some $r < \Omega$, and suppose $(\Theta, H, A)$ is an axiom of type $AxPClFH_{n,\Omega,\alpha}$ in $Prem(\mathcal{C}, u)$, suppose the conclusion of $\mathcal{C}$ is $(\Sigma, H', A')$. Then $o(H) < o((H')^\frown\{(e, ?)\} \cup \Sigma)$.*

*In addition, if $Ord(\Theta) = Ord(\Sigma)$ then*

$$\alpha < D(o((H')^\frown\{(e, ?)\} \cup \Sigma) + k((H')^\frown\{(e, ?)\} \cup \Sigma))$$

*Proof.* We have $H = (H')^\frown\langle S_1, \ldots, S_n$ and there exists $n$ such that $ext_n(S_i) = \Sigma_S$ for some $n$, and $e(n, S_i) = e$, $v(n, S_i) = u$ for each $i \leq n$. But since $||u||_A < ||?||_A$, it follows that $o(H) < o((H')^\frown\{(e, ?)\} \cup \Sigma)$.

The second part follows from the definitions of $o$ and $k$. □

**Definition 9.14.** $(\Theta_0, \ldots, \Theta_n)$ *is a* $r$-prepath *(for $\Theta_n$) if it is a path in some $r$-derivation of $\Theta_0 = \emptyset$. A path is assumed to be given with an analysis of the inference rules constituting the path.*

$(\Theta_0, \ldots, \Theta_n)$ *is an* $r$-path *(for $\Theta_n$) if it is an $r$-prepath and if the inference from $\Theta_{i+1}$ to $\Theta_i$ is a $CutFr^*$ inference then $\Theta_{i+1}$ belongs to $Prem(I, ?)$.*

The key result is Lemma 9.33, which shows that, if we can eliminate cuts, we will prove the termination of the $H$-process.

**Lemma 9.9.** *1. If $\Theta$ is a sequent in an $r + 1$ derivation of $\emptyset$ then $\Theta t \gtrsim r + 1$, $\Theta f \leq r$*

*2. If $\Sigma$ is a sequent in an $r^+$ derivation of $\Theta$ then:*

*(a)* $\Theta_{\leq r} \setminus \Theta t \trianglelefteq \Sigma$

*(b)* $(\Sigma f)_{\geq r} \subseteq \Theta$

*(c)* $\Theta t \geq r \Rightarrow \Sigma t \gtrsim r$

*(d) If $(n \in I^{<\alpha}, ?, t) \in \Theta$ then either there is some $\beta$ such that $(n \in I^{<\beta}, ?, t) \in \Sigma$ or there is some $\beta \geq \mathcal{O}(r)$ such that $(n \in I^{\beta}, \top, t) \in \Sigma$.*

*Proof.* 1. The statement is proved by bottom-up induction on the proof. It obviously holds for $\emptyset$, and in an $r + 1$-derivation viewed bottom up, temporary components are added by $Fr$, $H$, $CutFr^*$, $FFr$, $PCut$, $FH$, and $ClFH$, and these components all have rank at least $r + 1$, unless they belong to $\Upsilon_{\leq r} \setminus \Upsilon$ for some $\Upsilon$, in which case they must have rank $r$. Fixed components are added by $Cut$, $FCut$, and $PCut$, and these components all have rank $< r + 1$, and therefore $\leq r$. The remaining inferences cannot occur.

2. (a) Again by bottom-up induction. The statement obviously holds for $\Theta$ and is trivially preserved by $Fr$, $Cut$, $CutFr$, and $CutFr^*$. Also, since any application of the $H$, $FH$, or $ClFH$ rules is of rank at least $r$, the only term of rank $r$ or less which is removed must be some $(e, v, t)$, which is not in $\Theta_{\leq r} \setminus \Theta t$. Finally, any application of $FFr$, $FCut$, $PCutFr$, or $PCut$ which removes some $(n \in I^{<\beta}, ?, f)$ adds in some $(n \in I^{<\alpha}, ?, i)$ or $(n \in I^{\alpha}, \top, i)$ which satisfies the definition of $\trianglelefteq$.

(b) Going downwards, the only points at which $(e, u, f)$ can vanish are the $Cut$, $FCut$, $PCut$, and $PCutFr$ inferences, and $Cut(d), FCut(d), PCutF(d) < r$ while $PCutFr(d) = r$ so if $rk(e) \geq r$ then $(e, u, f)$ cannot be removed.

(c) Since $Fr(d) > r$,$H(d), FH(d), ClFH(d) \geq r$, $CutFr(d) = r$, $CutFr(d) \geq r$, $PCut(d) > r$, $FFr(d) > r$, and $PCutFr(d) = r$, all components $(e, u, t)$ added going upwards have rank at least $r$, unless they belong to $\Upsilon_{\leq r} \setminus \Upsilon$ for some $\Upsilon$, in which case they have rank $r - 1$.

(d) By bottom-up induction, $(e, u, t)$ cannot be removed by a $Cut$, $CutFr$, $CutFr^*$, $Fr$, or $FFr$ inference. If any other inference removes $e$, it much be replaced by some component which ensures that $n \in I^{<\beta} \hookrightarrow_{\Sigma_S} \perp$.

□

**Lemma 9.10.** *If $d$ is an $r^+$-derivation of $\Theta$ then there is an $r$-derivation $d'$ of $\Theta$.*

*Proof.* If $r \neq \Omega + 1$, prune all $CutFr$ inferences to $Fr$ inferences by deleting all premises except the leftmost one and prune all $PCutFr$ inferences to $PCut$ inferences by deleting the appropriate premise.

If $r = \Omega + 1$, convert all $CutFr$ inferences to $CutFr^*$ inferences. □

**Definition 9.15.** *Let $\Theta$ and $\Sigma$ be two sequents. Then $\Theta$ and $\Sigma$ are* multiplicable *if:*

1. *Whenever $(e, u, i), (e, u', i') \in \Theta \cup \Sigma$, $u' = u$ and $i' = i$*
2. *If $(n \in I^{\alpha}, \top, i), (n \in I^{<\beta}, ?, i') \in \Theta \cup \Sigma$ then $\beta \leq \alpha$*

*We define $R_{\Theta,\Sigma}$ by:*

1. *If $(n \in I^{\alpha}, \top, i), (n \in I^{<\beta}, ?, i') \in \Theta \cup \Sigma$ then $(n \in I^{<\beta}, ?, i') \in R_{\Theta,\Sigma}$*
2. *If $(n \in I^{<\alpha}, ?, i), (n \in I^{<\beta}, ?, i) \in \Theta \cup \Sigma$ and $\beta < \alpha$ then $(n \in I^{<\beta}, ?, i') \in R_{\Theta,\Sigma}$*
3. *If $(n \in I^{\alpha}, \top, i), (n \in I^{\beta}, \top, i') \in \Theta \cup \Sigma$ and $\alpha < \beta$ then $(n \in I^{\beta}, ?, i') \in R_{\Theta,\Sigma}$*

*$\Theta * \Sigma$ is defined and equal to $\Theta \cup \Sigma \setminus R_{\Theta,\Sigma}$ iff $\Theta$ and $\Sigma$ are multiplicable.*

$R_{\Theta,\Sigma}$ is the set of redundant values in $\Theta \cup \Sigma$ which are implied by other values also present, so we remove them to make sure that $\Theta * \Sigma$ is still parsimonious.

**Lemma 9.11.** *If $\Theta$ and $\Sigma$ are multiplicable then:*

1. *If $(e, u, i) \in R_{\Theta,\Sigma}$ then $e \hookrightarrow_{(\Theta*\Sigma)_S} u$*
2. *$\Theta * \Sigma$ is a sequent*

*Proof.* 1. Suppose $(n \in I^{<\beta}, ?) \in (R_{\Theta,\Sigma})_S$ and there is some $(n \in I^{\alpha}, \top) \in (\Theta * \Sigma)_S$. Then, since $\Theta$ and $\Sigma$ are multiplicable, we must have $\beta \leq \alpha$, and therefore, since $(n \in I^{\alpha}, \top) \in (\Theta * \Sigma)_S$, the result follows. The only other possible way there could be some $(n \in I^{<\beta}, ?) \in (R_{\Theta,\Sigma})_S$ is if there is some $(n \in I^{<\alpha}, ?) \in (\Theta * \Sigma)_S$ with $\beta < \alpha$, in which case again the result follows.

If $(n \in I^{\beta}, \top) \in (R_{\Theta,\Sigma})_S$ then we have $(n \in I^{\alpha}, \top) \in (\Theta * \Sigma)_S$, so the result follows.

2. All we need to show is that $(\Theta * \Sigma)_S$ is an $\epsilon$-substitution. This follows directly from the fact that $\Theta_S$ and $\Sigma_S$ are $\epsilon$-substitutions and there are no $n \in I^\alpha, n \in I^\beta \in \mathrm{dom}(\Theta * \Sigma)$ with $\alpha \neq \beta$ since we removed $R_{\Theta,\Sigma}$ from $\Theta * \Sigma$.

$\square$

When we eliminate cuts, we will want to replace certain $AxH$-type axioms with branches from the cut. To do this, it will be necessary to show that we can convert the branch into a derivation of the axiom, since the two will have different components. To do this, we will need to show that, under suitable conditions, the axiom and the conclusion of the cut are multiplicable.

The following lemma will be needed when we wish to eliminate a $Cut_e$ inference. Above that inference is some $AxH_{e,v}$ axiom, which we wish to replace with the suitable branch of the $Cut$ inference. We will replace the axiom with a derivation via an $H$ inference; the premise of this will be $\Sigma$. We will then want to show that the conclusion of the $Cut$ inference, $\Theta$, is sufficiently compatible with $\Sigma$, in the sense that we will be able to convert the relevant branch of the $Cut$ inference into a derivation of $\Sigma$. The situation when we eliminate $FCut$ or $PCut$ inferences is similar, and will make use of this lemma in the same way.

**Lemma 9.12.** *Suppose $d$ is an $r^+$ deduction of $\Theta$ from $\Sigma$, $\Sigma \lesssim r$, and there is an $r+1$ path for $\Theta$. Then $\Sigma$ and $\Theta$ are multiplicable and $\Theta \trianglelefteq \Theta * \Sigma$.*

*Proof.* Suppose $(e, u, i), (e, u', i') \in \Sigma \cup \Theta$. $\Sigma$ and $\Theta$ are sequents, so assume, w.l.o.g., that $(e, u, i) \in \Sigma$ and $(e, u', i') \in \Theta$. We distinguish whether $e$ is a term or a formula; if $e$ is a term then we have $rk(e) \leq r$ since $\Sigma \lesssim r$. But then $i' = f$ by Lemma 9.9(1), and therefore $e \in \Theta_{\leq r} \setminus \Theta t$, so $(e, u', i') \in \Sigma$ by 9.9(2)(a), so $u = u'$ and $i = i'$ since $\Sigma$ is a sequent. If $e$ is a formula then $u = u'$ is determined by $e$ ($u = u' = ?$ if $e$ has the form $n \in I^{<\alpha}$, and $u = u' = \top$ if $e$ has the form $n \in I^\alpha$). Now, if $rk(e) > r$ then we have $i' = t$, and by 9.9(2)(b), we cannot have $i = f$, so $i' = i = t$. If $rk(e) \leq r$ and $i' = f$ then since $(e, u, i) \in \Sigma$ and $\Sigma$ is a sequent, by 9.9(2)(a), $i' = i = f$. In the final case, $rk(e) \leq r$ and $i' = t$. But then by Lemma 9.9(1), it must be that $e$ has the form $n \in I^{<\beta}$ for $\beta = \mathcal{O}(r)$. Then $u = u' = ?$ and $e \in \mathrm{dom}(\Sigma)$, and by Lemma 9.9(2)(d), it must be that $i = t$.

Next suppose $(n \in I^\alpha, \top, i), (n \in I^{<\beta}, ?, i') \in \Sigma \cup \Theta$ where $\alpha < \beta$. Suppose $(n \in I^\alpha, \top, i) \in \Theta$. Then if $rk(n \in I^\alpha) \leq r$ we must have $i = f$ by Lemma 9.9(1), so we must have $n \in I^\alpha \hookrightarrow_{\Sigma_S} \top$ since $\Theta_{\leq r} = \Theta_{\leq r} \setminus \Theta t \trianglelefteq \Sigma$, contradicting the fact that $(n \in I^{<\beta}, ?) \in \Sigma$. On the other hand, if $rk(n \in I^\alpha) > r$ then we cannot have $r < rk(n \in I^\alpha) < rk(n \in I^{<\beta}) \leq r + 1$.

On the other hand, suppose $(n \in I^\alpha, \top, i) \in \Sigma$. Then $rk(n \in I^\alpha) \leq r$. But since $(n \in I^{<\beta}, ?, t) \in \Theta$, by Lemma 9.9(2)(d), it must be that $i = t$ and $\alpha \geq \mathcal{O}(r)$. But $r - 1 \leq 3\mathcal{O}(r)$, so $n \in I^\alpha \geq r + 1$, a contradiction.

Suppose $(e, u, i) \in \Theta$. If $i = t$ then since there is an $r + 1$ path for $\Theta$, by Lemma 9.9(1) either $rk(e) > r$, in which case $(e, u, t) \in \Theta * \Sigma$ since $\Sigma \lesssim r$, or $rk(e) = r$ and $(e, u, t) \in \Theta * \Sigma$ by Lemma 9.9(2)(d). If $i = f$ then $rk(e) \leq r$ by Lemma 9.9(1), so since $\Theta_{\leq r} = \Theta_{\leq r} \setminus \Theta t \trianglelefteq \Sigma$, also $\Theta_{\leq r} \trianglelefteq \Theta * \Sigma$. $\square$

## 9.4 Cut Elimination

A number of technical lemmata are needed for cut-elimination, so a short outline is in order. The core operation is a reduction of one $Cut$ or $FCut$ by replacing the axioms representing $H$ steps with the corresponding inference rule, using another branch, $b$, of the cut to provide a derivation for the premise. Lemmata 9.13 through 9.16 provide operations on derivations which we use to make $b$ fit on top of the $H$-inference even when there are many steps between the cut and the axiom. Lemma 9.19 combines them to show that we can indeed perform a correct $H$-inference from the branch $b$.

Lemma 9.23 uses this to remove a $Cut$, while Lemmata 9.21 and 9.26 remove $FCut$s and $PCut$s respectively. Lemma 9.27 applies these three lemmata inductively to reduce the rank of a derivation of $\emptyset$ from $r+1$ to $r$.

This process is then iterated (Lemma 9.28) to reduce the cut-rank to a limit ordinal. Lemma 9.30 states that we can move from a limit cut-rank to some lower cut-rank in a countable derivation, and Lemma 9.31 lets us collapse $\Omega$ derivations to countable size.

The next two lemmata allows us to replace pairs $(n \in I^{<\alpha}, ?)$ in our derivations with some $(e, u)$ which implies this (that is, either $(n \in I^{<\beta}, ?)$ for $\beta > \alpha$ or $(n \in I^{\beta}, \top)$ for $\beta \geq \alpha$).

Note that the lemmas below largely ignore the history portion of sequents. These portions are changed appropriately as we alter corresponding sequents, but the only situation in which these changes matters is when we change $C(H)$, and this case is dealt with.

**Lemma 9.13 (Persistency).** *Let $d$ be a derivation such that:*

1. *$d$ is an $r^+$ derivation*
2. *The end-sequent of $d$ is of the form $(n \in I^{<\alpha}, ?, f), \Theta$*
3. *There is some $(e, u, t) \in \Theta$ such that $rk(e) = r$ and $u \neq ?$*
4. *$h(d) \leq \eta$*

*Let $\beta \geq \alpha$ and suppose $rk(n \in I^{\beta}) < r$.*
*Then there is a derivation $d_\Sigma$ such that:*

1. *$d_\Sigma$ is an $r^+$ derivation*
2. *The end-sequent of $d_\Sigma$ is $\Sigma = (n \in I^{\beta}, \top, f), \Theta$*
3. *$h(d_\Sigma) \leq \eta$*

*Similarly, if $rk(n \in I^{<\beta}) < r$, there is a derivation $d_{\Sigma'}$ such that:*

1. *$d_{\Sigma'}$ is an $r^+$ derivation*
2. *The end-sequent of $d_{\Sigma'}$ is $\Sigma' = (n \in I^{<\beta}, ?, f), \Theta$*
3. *$h(d_{\Sigma'}) \leq \eta$*

*Proof.* Otherwise, by induction on the last inference of $d$.

1. $Cut$,$CutFr$,$CutFr^*$,$Fr$: The result follows directly from IH.

2. $H$, $FH$: Since $n \in I^{<\alpha} \hookrightarrow_{\Sigma_S} \perp$ and $n \in I^{<\alpha} \hookrightarrow_{\Sigma'_S} \perp$, any expression computed by $(n \in I^{<\alpha}, ?, f), \Theta$ is also computed by $\Sigma$ and $\Sigma'$ and has the same value; so the $H$-expression is unchanged. In addition, while the history may change, these changes will not change $C(H)$

3. $ClFH$: As for an $FH$, with the additional condition that we must make sure the ordinal does not change. But since there is some $e \in \mathrm{dom}(\Theta)$ with $rk(e) = r$, $Ord(\Theta) = Ord(\Sigma)$, Lemma 7.2 applies.

4. $FCut_{m,\gamma,\delta}$: if $m \neq n$ then the result follows directly from IH. If $m = n$ and $\beta \geq \gamma$ then the result follows by applying IH to the subderivation of $(n \in I^{<\gamma}, ?, f), \Theta$. Otherwise $\gamma > \beta \geq \alpha = \delta$, so we trim the $FCut$ to a $FCut_{m,\gamma,\beta}$ to give a derivation of $\Sigma'$, and just take the appropriate subderivation to give $\Sigma$.

   That is, if we start with:

   $$\frac{(n \in I^{<\gamma}, ?, f), \Theta \qquad (n \in I^{\zeta}, \top, f), \Theta, \ldots \qquad \forall \zeta (\delta \leq \zeta < \gamma \in V^O)}{(n \in I^{<\alpha}, ?, f), \Theta} \; FCut_{n,\gamma,\alpha}$$

   we can take the subderivation of $(n \in I^{\beta}, \top, f), \Theta$ for $\Sigma$ and delete extra premises of $FCut_{n,\gamma,\alpha}$ to obtain:

   $$\frac{(n \in I^{<\gamma}, ?, f), \Theta \qquad (n \in I^{\zeta}, \top, f), \Theta, \ldots \qquad \forall \zeta (\beta \leq \zeta < \gamma \in V^O)}{(n \in I^{<\beta}, ?, f), \Theta} \; FCut_{n,\gamma,\beta}$$

   for $\Sigma'$.

5. $PCut_{m,\gamma,\delta,\epsilon}$: if $m \neq n$ then the result follows directly from IH. If $m = n$ then take the appropriate subderivation for $\Sigma$ and truncate to $PCut_{n,\gamma,\beta,\epsilon}$ to give $\Sigma'$.

6. $PCutFr_{m,\gamma,\delta,\epsilon}$: if $m \neq n$ then the result follows directly from IH. If $m = n$ then take the appropriate subderivation for $\Sigma$ and truncate to $PCutFr_{n,\gamma,\beta,\epsilon}$ to give $\Sigma'$.

7. $FFr_{m,\gamma,\delta}$: $m = n$ is impossible by rank considerations, so $m \neq n$ and the result follows directly from IH.

8. $AxPClFH$, $AxClFH$: Since $Ax(d) \leq r$, the main term would have rank at most $r$. But since there is some $(e, u, t) \in \Theta$ with $u \neq ?$ and $rk(e) = r$, Lemma 9.1 requires that the rank of the axiom be greater than $r$. Therefore these axioms do not appear.

9. All axioms other than $AxPClFH$ and $AxClFH$ remain valid

□

**Lemma 9.14 (Persistency).** *Let $d$ be a derivation such that:*

1. *$d$ is an $r^+$ derivation*

2. *The end-sequent of $d$ is of the form $(n \in I^{<\alpha}, ?, i), \Theta$*

3. $h(d) \leq \eta$

*Let $\beta \geq \alpha$ and suppose $rk(n \in I^{\beta}) = r$ and that there is some $(e, u, t) \in \Theta$ such that $rk(e) = r$ and $u \neq ?$. Then there is a derivation $d_{\Sigma}$ such that:*

1. *$d_{\Sigma}$ is an $r^{+}$ derivation*
2. *The end-sequent of $d_{\Sigma}$ is $\Sigma = (n \in I^{\beta}, \top, t), \Theta$*
3. *$h(d_{\Sigma}) \leq \eta$*

*Similarly, if $\beta = \mathcal{O}(r)$ then there is a derivation $d_{\Sigma'}$ such that:*

1. *$d_{\Sigma'}$ is an $r^{+}$ derivation*
2. *The end-sequent of $d_{\Sigma'}$ is $\Sigma' = ((n \in I^{<\beta}, ?, t), \Theta)$*
3. *$h(d_{\Sigma'}) \leq \eta$*

*Proof.* By induction on the last inference of $d$.

1. $Cut$,$CutFr$,$CutFr^{*}$,$Fr$: The result follows directly from IH.
2. $H$, $FH$: Any expression computed by $(n \in I^{<\alpha}, ?, i), \Theta$ is also computed by $\Sigma$ and $\Sigma'$ and has the same value, so the $H$-expression is unchanged and the result follows from IH using the same inference rule. In addition, while the history may change, $C(H)$ will not change.
3. $ClFH$: As for an $FH$, with the additional condition that we must make sure the ordinal does not change. But if this adds an ordinal, we already have some $e \in \mathrm{dom}(\Theta)$ with $rk(e) = r$, so $Ord(\Theta) = Ord(\Sigma)$.
4. $FCut_{m,\gamma,\delta}$: if $m \neq n$ then the result follows directly from IH. If $m = n$ then since $rk(n \in I^{<\gamma}) < r$, the result follows by applying IH to the subderivation of $(n \in I^{<\gamma}, ?, f), \Theta$.
5. $PCut_{m,\gamma,\delta,\epsilon}$: if $m \neq n$ then the result follows directly from IH. If $m = n$ and $\alpha = \gamma$ then we have an appropriate subderivation, otherwise replace the inference with an $FFr$ inference.
6. $PCutFr_{m,\gamma,\delta,\epsilon}$: if $m \neq n$ then the result follows directly from IH. If $m = n$ then it must be that $\alpha = \delta$ and $\Sigma$ is a conclusion of a subderivation.
7. $FFr_{m,\gamma,\delta}$: if $m \neq n$ the result follows directly from IH. If $m = n$ then $\delta = \beta$, so we are done.
8. $AxFH_{n,\alpha,\gamma}$: replace with an $AxPFH_{n,\beta,\gamma}$.
9. $AxPClFH$, $AxClFH$: In the $\Sigma$ case, since $Ax(d) \leq r$, the main term would have rank at most $r$. But since there is some $(e, u, t) \in \Theta$ with $u \neq ?$ and $rk(e) = r$, Lemma 9.1 requires that the rank of the axiom be greater than $r$. Therefore these axioms do not appear.

   In the $\Sigma'$ case, these axioms remain valid except that some $AxClFH$ axioms might have to be replaced with $AxPClFH$ axioms.

10. Other axioms are unchanged

The same functions that measure $d$ can measure $d_\Sigma$ and $d_{\Sigma'}$. $\square$

**Lemma 9.15 (Weakening).** *Let $d$ be a derivation of $\Theta$ and $\Sigma$ a sequent such that:*

1. *$d$ is an $r^+$ derivation*
2. *There is some $(e, u, t) \in \Theta$ with $u \neq ?$ and $rk(e) = r$*
3. *$\Sigma \lesssim r$*
4. *$(\Sigma f)_{\geq r} \subseteq \Theta$*
5. *$\Sigma t \gtrsim r$*
6. *$\Theta_{\leq r} \trianglelefteq \Sigma$*
7. *$\Theta \trianglelefteq \Theta * \Sigma$*
8. *$h(d) \leq \eta$*

*Then there is an derivation $d * \Sigma$ of $\Theta * \Sigma$ such that:*

1. *$d * \Sigma$ is an $r^+$ derivation*
2. *$h(d') \leq \eta$*

This lemma is one of the core operations we will use when eliminating cuts. The complicated statement hides the basic situation we are dealing with: $\Sigma$ is the premise of an inference which results from replacing some $AxH$-type axiom with the corresponding $H$-type inference. $\Theta$ is a premise of the $Cut$-type inference which first introduced the main expression of that inference. Our goal here is to use the derivation $d$ which we can place on top of our new $H$-type inference to make this deduction into a derivation (the conclusion of $d'$ is not quite $\Sigma$–the next lemma will resolve this).

*Proof.* By induction on the last inference of $d$.

1. $Cut$: Let the main term be $e$. Then $rk(e) < r$ and either:
   (a) There is some $u$ such that $(e, u, t) \in \Sigma$: not possible, since $\Sigma t \gtrsim r$
   (b) There is some $u$ such that $(e, u, f) \in \Sigma$: then $((e, u, f), \Theta) * \Sigma = \Theta * \Sigma$, and therefore $d * \Sigma = Prem(I, u) * \Sigma$
   (c) There is no such $u$: then by I.H. for each $u$, $Prem(I, u) * \Sigma$ is defined, and $d * \Sigma$ just applies the $Cut$ rule to $(Prem(I, u) * \Sigma)_{u \in \mathbb{N} \cup \{?\}}$
2. $CutFr$, $CutFr^*$: Let the main term be $e$. Then $rk(e) \geq r$ and either:
   (a) There is some $u$ such that $(e, u, t) \in \Sigma$: then $((e, u, t), \Theta) * \Sigma = \Theta * \Sigma$, and therefore $d * \Sigma = Prem(I, u) * \Sigma$

(b) There is $u$ such that $(e, u, f) \in \Sigma$: not possible, since $(\Sigma f)_{\geq r} \subseteq \Theta$ and $(e, u, f) \notin \Theta$.

(c) There is no such $u$: then by I.H. for each $u$, $Prem(I, u) * \Sigma$ is defined, and $d * \Sigma$ just applies the $CutFr$ rule to $(Prem(I, u) * \Sigma)_{u \in \mathbb{N} \cup \{?\}}$

3. $Fr$: Let the main term be $e$. Then $((e, ?, t), \Theta) * \Sigma$ is defined, since $\Sigma \lesssim r < rk(e)$, and $d * \Sigma$ is just the $Fr$ rule applied to the derivation of $\Theta * \Sigma$.

4. $H$: Let the main term be $e$. Then $rk(e) \geq r$ and $\Theta = (e, u, i), \Upsilon$ is derived from $\Theta' = (e, v, t), \Upsilon_{\leq rk(e)}$. Since $\Sigma \lesssim r \leq rk(e)$, $\Sigma' = \{(e', u', i') \in \Sigma \mid e' \neq e\}$ is also correct, and $(\Sigma' f)_{\geq r} \subseteq \Theta'$.

   Since $\Sigma' \subseteq \Sigma$ and $\Theta' \subseteq \Theta$, $\Theta'$ and $\Sigma'$ are multiplicable and $\Theta' * \Sigma' = (e, v, t), \Upsilon_{\leq rk(e)} * \Sigma' = (e, v, t), (\Upsilon * \Sigma')_{\leq rk(e)}$, while $\Theta * \Sigma = (e, u, i), \Upsilon * \Sigma'$. Therefore $d * \Sigma$ is obtained by applying the same inference to the derivation given by I.H..

   Now, suppose $C(H)$ changes. Then this must be at a corrected $H$-step, and $rk(e) = r$. But then the derivation above is also an $\Omega + 1^+$ derivation, so, by IH, we have a derivation of $d * \Sigma \cup \{(e, u, t) \mid (e, u) \in C(H)\}$.

5. $FCut_{n,\alpha,\beta}$,$PCut_{n,\alpha,\beta,\delta}$,$PCutFr_{n,\alpha,\beta,\delta}$: Call the inference $I$ and let $\Theta = (n \in I^{<\beta}, ?, f), \Theta'$.

   (a) If there are no $n \in I^{\gamma}$ or $n \in I^{<\gamma}$ for any $\gamma$ in $\mathrm{dom}(\Sigma)$ then just apply IH to each subderivation and end the derivation with the same inference applied to the new subderivations.

   (b) Suppose we have $\Sigma = (n \in I^{\gamma}, \top, i), \Sigma'$. Then by IH, if $Prem(I, u)$ is one of the immediate subderivations then $Prem(I, u) * \Sigma'$ is defined. If $i = t$ then $rk(n \in I^{\gamma}) = r$, so $R$ is not a $PCut$ inference. If $I$ is a $PCutFr$ inference, there is a subderivation ending in $(n \in I^{\delta}, \top, t), \Theta'$, and since $\delta = \gamma$, by IH there is a derivation of $((n \in I^{\gamma}, \top, t), \Theta') * \Sigma = \Theta * \Sigma$. If $R$ is an $FCut$ inference, we have $\gamma > \alpha$, so the result follows from Lemma 9.14 followed by IH.

   If $i = f$ then we cannot have $rk(n \in I^{\gamma}) \geq r$, since $(\Sigma f)_{\geq r} \subseteq \Theta$ and $n \in I^{\gamma} \notin \mathrm{dom}(\Theta)$. Since $rk(n \in I^{\gamma}) < r$ either we have some subderivation ending in $(n \in I^{\gamma}, \top, f), \Theta$, and we apply IH to that, or we apply IH to the result of Lemma 9.13.

   (c) Suppose we have $\Sigma = (n \in I^{<\gamma}, ?, i), \Sigma'$. Then by IH, if $Prem(I, u)$ is some subderivation then $Prem(I, u) * \Sigma'$ is defined. If $i = t$ then $r + 1 \geq rk(n \in I^{<\gamma}) \geq r - 1$. If $I$ is a $PCut$ inference then there is a subderivation ending in $(n \in I^{<\alpha}, ?, t), \Theta'$. If $\gamma = \alpha$ then we apply IH to the subderivation of $(n \in I^{\gamma}, \top, t), \Theta'$, otherwise we truncate the $PCut$ to an $FFr$ inference. If $R$ is an $FCUt$ inference, we have $\gamma > \alpha$, so the result follows from Lemma 9.14 and IH. If $R$ is a $PCutFr$ inference then we truncate to an $FFr$ inference and apply IH to $Prem(I, \alpha)$.

   If $i = f$ then we cannot have $rk(n \in I^{<\gamma}) \geq r$, since $(\Sigma f)_{\geq r} \subseteq \Theta$ and $n \in I^{<\gamma} \notin \mathrm{dom}(\Theta)$. If $I$ is an $FCut$ inference with $\alpha < \gamma$, we apply Lemma 9.13 followed by IH, otherwise we prune the inference by replacing $\beta$ with $\gamma$.

6. $FFr_{n,\alpha,\beta}$: There can be no $n \in I^{\gamma}$ in $\mathrm{dom}(\Sigma)$, and if $\Sigma = (n \in I^{<\gamma}, ?, i), \Sigma'$ then $i = t$ since the sequents are multiplicable, and $\gamma \leq \beta$, so $\Theta * \Sigma = \Theta * \Sigma'$, and we just apply IH to the premise of the $FFr$ inference, and add the same $FFr$ inference to give the required derivation.

7. $FH_{n,\alpha,\beta}$:

   (a) If there are no $n \in I^{\gamma}$ or $n \in I^{<\gamma}$ for any $\gamma$ in $\mathrm{dom}(\Sigma)$ then just apply IH to the subderivation and end the derivation with the same inference applied to the new subderivation.

   (b) We cannot have $\Sigma = (n \in I^{\gamma}, \top, i), \Sigma'$, since either $n \in I^{\alpha} \in \mathrm{dom}(\Theta)$ or $n \in I^{<\alpha} \in \mathrm{dom}(\Theta)$ and $\alpha > \gamma$.

   (c) Suppose $\Sigma = (n \in I^{<\gamma}, ?, i), \Sigma'$. We have $\Theta * \Sigma = \Theta * \Sigma'$, since $\alpha > \gamma$ and the inference applied to the result of the inductive hypothesis remains valid.

8. $ClFH_{n,\xi}$: As with an $FH$ inference, and note that since there is some $(e, u, t) \in \Theta$ with $u \neq ?$ and $rk(e) = r$, it follows that $r < rk(n \in I^{\beta})$. Then since $\Sigma \lesssim r$, it follows by 7.2 that the ordinal remains unchanged.

9. $AxPClFH$, $AxClFH$: Since $Ax(d) \leq r$, the main term would have rank at most $r$. But since there is some $(e, u, t) \in \Theta$ with $u \neq ?$ and $rk(e) = r$, Lemma 9.1 requirs that the rank of the axiom be greater than $r$ if $r < \Omega$. Therefore these axioms do not appear if $r < \Omega$.

10. Axioms: Otherwise, if $\Theta$ is an axiom then $\Theta * \Sigma$ is an axiom of the same kind.

□

**Lemma 9.16 (Repetition).** *Let $p = (\Theta_0, \ldots, \Theta_n)$ be a $r+1$ path for $\Theta = \Theta_n$. Let $\Sigma \lesssim r$ be a correct sequent such that $\Theta_{\leq r} \trianglelefteq \Sigma$, $\Theta * \Sigma$ is defined, and $\Theta \trianglelefteq \Theta * \Sigma$.*

*Then there is a derivation $d'$ of $\Sigma$ from $\Theta * \Sigma$ consisting only of $Fr$, $H$, $FFr$, $FH$, and $ClFH$ inferences of ranks $> r$ copied from $p$ and in the same order.*

This lemma completes the work of the previous one, providing a series of inferences we can use to place the derivation given by the previous one on top of the $H$-type inference we have created.

*Proof.* By induction on $n$, and trivial if $n = 0$. Suppose $n > 0$ and let $\Theta' = \Theta_{n-1}$. Since $\Theta$ is on an $r+1$ path, it follows that $\Theta'_{\leq r} \trianglelefteq \Theta_{\leq r}$ (since no inference in an $r+1$ path will remove elements of rank $\leq r$, nor elements of rank greater than $r$ which might be in $\Theta'_{\leq r}$, without replacing them by something appropriate) and $\Theta' * \Sigma$ is defined, so by IH there is a derivation of $\Sigma$ from $\Theta' * \Sigma$. Consider the inference from $\Theta$ to $\Theta'$.

1. $Cut$: We have $\Theta = (e, u, f), \Theta'$ and $rk(e) \leq r$. Therefore $\Theta * \Sigma = \Theta' * \Sigma$

2. $CutFr$,$PCutFr$: impossible in an $r+1$-path

3. $CutFr^*$: This must be the $?$ branch, so an $Fr$ inference applies. Since $p$ is a path for $\Theta$, $e \in A$.

4. $Fr$: $\Theta = (e, ?, t), \Theta'$. Then $rk(e) > r$, so $\Theta * \Sigma = (e, ?, t), \Theta' * \Sigma$ and the same $Fr$ inference applies.

5. $H$: $\Theta = (e, v, t), \Upsilon_{\leq rk(e)}$ and $\Theta' = (e, u, t), \Upsilon$. Then since $rk(e) > r$, $\Theta * \Sigma = (e, v, t), (\Upsilon * \Sigma)_{\leq rk(e)}$. And $\Theta' * \Sigma = (e, u, t), \Upsilon * \Sigma$, so the $H$-rule applies to $\Theta' * \Sigma$ and has the same $H$-value.

6. $FCut_{n,\alpha,\beta}$: We have $\Theta = (e, u, f), \Theta^-$ and $\Theta' = (e', u', f), \Theta^-$. Also, $\{(e, u, f)\} * \Theta' = \Theta$, and since $\Theta_{\leq r} \trianglelefteq \Sigma$ and $rk(e) \leq r$, we have $\Theta * \Sigma = \Theta' * \Sigma$.

7. $PCut_{n,\alpha,\beta,\delta}$: We have $\Theta = (e, u, i), \Theta^-$ and $\Theta' = (e', u', i'), \Theta^-$. If $i = f$ then $rk(e) < r$, so $\Sigma * \Theta = \Sigma * \Theta'$. Otherwise, $e = n \in I^{<\alpha}$ and $n \in I^{<\mathcal{O}(r)} \in \mathrm{dom}(\Sigma)$, so an $FFr$ inference applies or $\Sigma * \Theta = \Sigma * \Theta'$.

8. $FFr_{n,\alpha,\beta}$: since $rk(n \in I^{<\alpha}) > rk(n \in I^{<\beta}) > r$, we have $(n \in I^{<\beta}, ?, t) \in \mathrm{dom}(\Theta' * \Sigma)$, so $FFr_{n,\alpha,\beta}$ is an inference from $\Theta * \Sigma$ to $\Theta' * \Sigma$.

9. $FH$: $\Theta = (e, v, t), \Upsilon_{\leq rk(e)}$ and $\Theta' = (e', \top, t), \Upsilon$. Since $rk(e) > rk(e') \geq r + 1 > r$, $\Theta * \Sigma = (e, v, t), (\Upsilon * \Sigma)_{\leq rk(e)}$. And $\Theta' * \Sigma = (e', \top, t), \Upsilon * \Sigma$, so the $FH$-rule applies to $\Theta' * \Sigma$ and has the same $H$-value.

10. $ClFH$: As for an $FH$. The rank of the ordinal is greater than any ordinal in $Ord(\Sigma)$ by Lemma 7.3, and unchanged by Lemma 7.4.

□

**Lemma 9.17.** *Let $d$ be an $r^+$-derivation of $(\Theta, H, A)$ with $r \geq \Omega$ and $H'$ a sequence of $\epsilon$-substitutions, and $S$ a substitution with $rk(e(S)) = r$. Suppose $h(d) \leq \eta$. Then there is an $r^+$-derivation $d'$ of $(\Theta, (H')^\frown S^\frown H, A)$ with $h(d') \leq o((H')^\frown S; r) + \eta$.*

*Proof.* By induction on $d$. The only case we need to check is if $d$ is an axiom. Then we had $o(H; r) \leq \eta$, and therefore

$$o((H')^\frown S^\frown H; r) = o((H')^\frown S; r) + o(H; r)$$

□

**Lemma 9.18.** *Suppose that there is an $r^+$ derivation $d$ of $\Theta$, $\Sigma$ is a correct sequent such that:*

1. $\Sigma \lesssim r$
2. $\Theta * \Sigma$ *is defined*
3. $(\Sigma f)_{\leq r} \subseteq \Theta$
4. $\Sigma t \gtrsim r$

*5.* $\Theta_{\leq r} \trianglelefteq \Sigma$

*6.* $\Theta \trianglelefteq \Theta * \Sigma$

*7. There is an* $r+1$ *path* $p$ *for* $\Theta$

*8.* $h(d) \leq \eta$

*9. There is some* $(e, u, t) \in \Theta$ *such that* $rk(e) = r$

*Then there is an* $r^+$*-derivation* $d'$ *of* $\Sigma$ *with* $h(d') \leq \eta \# n$ *for some* $n$*.*

This lemma combines the previous two into a single operation: given the premise of an $H$-type inference, $\Sigma$, and the appropriate subderivation of a $Cut$-type inference, this lemma produces a derivation of $\Sigma$.

*Proof.* By Lemma 9.15 there is a derivation $d^*$ of $(\Theta * \Sigma, a \cup a' \upharpoonright Ord(\Theta * \Sigma))$ with $h(d^*) \leq \eta$, and by Lemma 9.16 there is a deduction of $(\Sigma, a')$ from $(\Theta * \Sigma, a \cup a' \upharpoonright Ord(\Theta * \Sigma))$ consisting only of $Fr$, $H$, $FFr$, $FH$, and $ClFH$ inferences of rank $> r$.

To see that the height bound holds, note that the 'tail' attached by Lemma 9.16 consists of finitely many inferences. By [Poh89], Lemma 24.16(iv), $\alpha \ll \alpha \# 1$. If we define $d_0 = d^*$ and $d_i$ to be $d^*$ with the first $i$ inferences from the tail, by induction, $h(d_i) \leq \eta \# i$, since $\eta \# i \ll \eta \# i \# 1$, and therefore $h(d') \leq \eta \# n$ for some $n$. $\Box$

**Lemma 9.19.** *Let* $((e, u, i), \Upsilon, H, A)$ *be some instance of* $AxH$*,* $AxPFH$ *or* $AxPClFH$ *with main expression* $e$*. Let* $S = ((e, u, i), \Upsilon)_S$*,* $e' = e(S)$*,* $u' = v(S)$*, and* $r = rk(e')$*.*

*Suppose* $d$ *is an* $r^+$ *derivation of* $(e, u, i), \Theta$ *which contains one or more instances of* $(e, u, i), \Upsilon$*, and suppose there exists an* $r+1$ *path* $p$ *for* $(e, u, i), \Theta$ *and an* $r^+$ *derivation* $d'$ *of* $(e', u', t), \Theta$*.*

*Then there is an* $r^+$ *derivation* $d^*$ *of* $(e, u, t), \Theta$ *in which the axiom* $(e, u, i), \Upsilon$ *is not present, and if* $h(d) \leq \eta$ *and* $h(d') \leq \zeta$ *then* $h(d^*) \leq \zeta \# \omega \# \eta$*.*

This is the core lemma which we will actually apply in eliminating cuts. It applies the previous ones to replace all occurances of an axiom, setting the stage for elimination of a $Cut$-type inference. The complexity is necessary to deal with the various kinds of axioms which are all handled by this lemma.

*Proof.* First, observe that $(e, u, i), \Upsilon$ cannot be an axiom in the derivation $d'$ of $(e', u', t), \Theta$. We must have $(e, u) \in \mathcal{P}(e', u')$, and therefore $(e, u, i)$ and $(e', u', t)$ cannot be present in the same sequent. We also have $u' \neq ?$, so by bottom-up induction, the only place $(e', u', t)$ could disappear (going upwards) in an $r^+$ derivation is at an $H$ inference in which it is replaced by $(e', v)$ with $v < u'$. But either $u = ?$ or $u' < u$, so $(e, u, i)$ does not occur in any sequent in $d'$.

Define $\Sigma = (e', u', t), \Upsilon_{\leq r}$.

By Lemma 9.12, we have that $\Sigma$ and $\Theta$ are multiplicable and $\Theta \trianglelefteq \Theta * \Sigma$. Since $(e', u', t) \in \Sigma$, we have $((e', u', t), \Theta) * \Sigma = \Theta * \Sigma$, so by Lemma 9.9(2)(b), we have $(((e, u, i), \Upsilon)f)_{\leq r} \subseteq (e, u, i), \Theta$, and by Lemma 9.9(2)(c) we have $((e, u, i), \Theta)t \geq r \Rightarrow ((e, u, i), \Upsilon)t \gtrsim r$. Since $\Theta$ has an $r+1$ path, $\Theta t \gtrsim r+1$, so $(\Sigma f)_{\geq r} \subseteq (e, u, f), \Theta$ and $\Sigma t \gtrsim r$. We have $\Theta_{\leq r} \setminus \Theta t \trianglelefteq \Sigma$ by Lemma 9.9(2)(a).

By Lemma 9.18, there is a derivation $d^h$ of $\Sigma$ with $h(d^h) \leq \zeta \# n$ for some $n$.

Now, we construct $d^*$ by induction on the last inference of $d$. If $d$ is the axiom $(e, u, i), \Upsilon$, replace it with an $H$, $FH$, or $ClFH$ inference as appropriate from $\Sigma$ to $(e, u, t), \Upsilon$. If $r \geq \Omega$, we apply Lemma 9.17 to $d^h$ (the lemma applies since $rk(e') = r$).

Let $\delta = u'$ if $e$ is an $O$-term and $u' \notin Ord(\Theta)$, let $\delta = \beta$ if the axiom is $AxPFH_{n,\alpha,\beta}$ and $\beta \notin Ord(\Theta)$, and 0 otherwise. Since $\delta \in Ord(\Upsilon)$, $\delta \leq \zeta$. Also, if $r \geq \Omega$, $o(H; r) \leq \zeta$. Therefore $h(d^*) \leq \zeta \# \omega \# \eta$ (this is trivial if $\delta = 0$; otherwise it follows from the fact that $\delta \ll \eta$).

If $d$ is some other axiom then $d^* = d$.

If $d$ ends in some inference $I$, replace each $Prem(I, u)$ with $Prem^*(I, u)$ using IH, and let $d^*$ be the result of applying $I$ to the $Prem^*(I, u)$, changing $(e, u, f)$ to $(e, u, t)$ in the conclusion of $I$ if necessary (this can be done since a fixed expression will not be removed going up). Let $h(Prem(I, u))) \leq \eta_u$; then since each $h(Prem^*(I, u)) \leq \zeta \# \omega \# \eta_u$, $h(Prem^*(I, u)) \leq \zeta \# \omega \# \eta$.

$$
\begin{array}{ll}
 & \vdots \\
d' * \Sigma & \Theta * \Sigma \\
 & \vdots \\
H_{e,v} & \dfrac{(e', u', t), \Upsilon_{\leq r}}{(e, u, t), \Upsilon} \\
d & \\
 & \vdots \\
 & (e, u, t), \Theta
\end{array}
$$

□

**Lemma 9.20.** *Let $d$ be an $r^+$ derivation ending in $(n \in I^{<\alpha}, ?, f), \Theta$ with $rk(n \in I^{<\alpha}) = r$. Then there is an $r^+$ derivation $d'$ ending in $(n \in I^{<\alpha}, ?, t), \Theta$ and if $h(d) \leq \eta$ then $h(d') \leq \eta$.*

*Let $d$ be an $r^+$ derivation ending in $(n \in I^{\alpha}, \top, f), \Theta$ with $rk(n \in I^{\alpha}) = r$. Then there is an $r^+$ derivation $d'$ ending in $(n \in I^{\alpha}, \top, t), \Theta$ and if $h(d) \leq \eta$ then $h(d') \leq \eta$.*

This lemma, simple in concept, if not statement, lets us transform the branches above an $FCut$ or $PCut$, which add fixed main expressions, into derivations which end with a temporary main expression, which we will need for cut-elimination to proceed below that inference.

*Proof.* Replace any $AxFH_{n,\alpha,\beta}$ axioms with $AxPFH_{n,\alpha,\beta}$ axioms and $AxClFH_{n,\beta}$ axioms with $AxPClFH_{n,\beta}$ axioms. Then all inferences remain valid. □

**Lemma 9.21.** *Let $d$ be an derivation ending in an $FCut_{n,\alpha,\beta}$ with $rk(n \in I^{<\alpha}) = r$ such that the immediate subderivations are $r^+$ derivations and $p$ is an $r + 1$ path for the end-sequent $(n \in I^{<\beta}, ?, f), \Theta$. Then there is an $r^+$ derivation $d'$ of $(n \in I^{<\beta}, ?, f), \Theta$ and if $h(d) \leq \eta$ then $h(d') \leq \eta$.*

This lemma replaces an $FCut$ with a $PCut$.

*Proof.* Replace the $FCut$ inference with a $PCut$ inference and apply Lemma 9.20. $\square$

**Lemma 9.22.** *Let $d$ be an derivation*

$$\dfrac{\begin{array}{c} d_? \\ \vdots \\ (e,?,f),\Theta \end{array} \qquad \begin{array}{c} d_u \\ \vdots \\ (e,u,f),\Theta \end{array} \qquad u \in V^{\iota(e)}}{\Theta}\ Cut_e$$

*such that $rk(e) = r$, the immediate sub-derivations of $d$ are $r^+$ derivations, and there exists an $r+1$ path $p$ for the end-sequent $\Theta$ of $d$. Then for each $u \in V^{\iota(e)} \cup \{?\}$ there is an $r^+$ derivation $d'_u$ of $(e,u,t),\Theta$. Also if $h(d) \leq \eta$ then*

$$h(d'_u) \leq (\omega \# \eta) \times ||u||_{\prec_\iota}$$

*where*

$$||u||_{\prec_\iota} = \begin{cases} u & \text{If } u \in \mathbb{N} \text{ or } u \text{ is an ordinal} \\ \omega & \text{If } u = ? \text{ and } \iota = N \\ \iota(e) & \text{If } u = ? \text{ and } \iota(e) \text{ is an ordinal} \end{cases}$$

This lemma constructs the new branch we will need when we eliminate a $Cut$ inference.

*Proof.* By transfinite induction on $u$.

Suppose we have already constructed $Prem'(I,v)$ for all $v < u$ (or $v <^O u$), or for all $v \in V^{\iota(e)}$ if $u = ?$. Then for each $AxH_{e,v}$ appearing in $d$ which conflicts with $(e,u,t)$ but not with $\Theta$, we already have an $r^+$ derivation of $Prem'(I,v)$, since when $u \neq ?$, $v < u$ (or $v <^O u$). This satisfies the conditions of Lemma 9.19, so we apply this to each $AxH_{e,v}$ to get $Prem'(i,u)$.

If $AxH_{e,v}$ conflicts with $\Theta$, replace it with $AxPH_{e,v}$.

Note that if $e$ is an $O$-term and $u \notin Ord(\Theta)$ then $h(Prem(I,u)) \leq \eta + u$, and $h(Prem(I,u)) \leq \eta$ otherwise.

If $h(Prem'(I,v)) \leq \zeta$ then $h(Prem'(I,u)) \leq \zeta \# \omega \# \eta$, so by IH:

$$h(Prem'(I,u)) \leq (\omega \# \eta) \times ||u||_{\prec_\iota} \# \omega \# \eta = (\omega \# \eta) \times (||u||_{\prec_\iota} + 1)$$

$\square$

**Lemma 9.23.** *Let $d$ be a derivation ending with a $Cut$ inference $C$ of rank $r$ such that the immediate sub-derivations of $d$ are $r^+$-derivations and $p$ is an $r+1$-path for the end-sequent $\Theta$ of $d$. Then there is an $r^+$-derivation $d'$ of $\Theta$ and if $h(d) \leq \eta$ then $h(d') \leq (\omega \# \eta) \times (\max\{r, \omega\} + 1) + 1$.*

This lemma shows that it is possible to eliminate a single $Cut$.

*Proof.* By Lemma 9.22, for each premise $d_u : (e,u,f),\Theta$, there are $r^+$ derivations $d'_u$ of $(e,u,t),\Theta$, so we apply a $CutFr$ to these to give the desired derivation.

The height bound follows since $||?||_{\prec_{\iota(e)}} \leq \max\{r, \omega\}$: $||?||_{\prec_N} = \omega$, while $||?||_{\prec_\alpha} = \alpha$. $\square$

**Lemma 9.24.** *Let $d$ be a derivation ending with a $CutFr^*_e$ inference $C$ such that the immediate sub-derivations of $d$ are $r^+$-derivations and $p$ is an $r+1$-path for the end-sequent $\Theta$ of $d$. Then there is an $r^+$-derivation $d'$ of $\Theta$ and if $h(d) \leq \eta$ then $h(d') \leq \eta + \omega + \eta$.*

*Proof.* Let $(\Sigma, H)$ be some $AxH_{e',v}$ in $d$ which conflicts with $(e, ?, t)$, but does not conflict with $\Theta$. Then $(e, u) \in H(H)$ for some $u$.

Set $\Upsilon = \{(e, u, t) \mid (e, u) \in C(H)\} \cup \Theta_{\leq r}$. Observe that $\Upsilon \lesssim r$, $\Upsilon f \subseteq \Theta$, $\Upsilon t = \Omega + 1$, $\Theta_{\leq r} \trianglelefteq \Upsilon$, and $\Theta \trianglelefteq \Theta * \Sigma$. Then by Lemma 9.15, since $d_u$ is an $\Omega + 1^+$ derivation, we have a derivation of $d_u * \Upsilon$. Now we can apply Lemma 9.19 to give an $r^+$-derivation of $\Theta$. $\square$

**Lemma 9.25 (Path Weakening).** *Suppose $p$ is an $s$-path for $\Theta$ and there is some $r < s$ such that $Cut(p), FCut(p) \leq r$. Then there is an $r+1$-path $p'$ for $\Theta$ that is obtained by changing the subscripts of some of the $PCut_{n,\alpha,\beta,\delta}$ inferences in $p$.*

In general, an $s$-path will not be an $r$-path for $r \neq s$; however, if the $Cut$ and $FCut$ inferences all have ranks at most $r$, the only problem is excess branches on $PCut$ inferences. These can be pruned to given an $r+1$-path. We will need to do this in order to make our lemmata general enough to handle cut-elimination past a limit ordinal.

*Proof.* By induction on the length of $p$. Let $p = (\Theta_0, \ldots, \Theta', \Theta)$. If the inference from $\Theta$ to $\Theta'$ is anything other than some $PCut_{n,\alpha,\beta,\delta}$ then the result follows from IH and the fact that $Cut(p), FCut(p) \leq r$.

If the inference is $PCut_{n,\alpha,\beta,\delta}$ then we prune it to $PCut_{n,\alpha,\beta,\gamma}$ where $\gamma$ is the least ordinal such that $rk(n \in I^\gamma) \geq r$. $\square$

**Lemma 9.26.** *Let $d$ be an derivation ending in a $PCut_{n,\alpha,\beta,\delta}$ with $rk(n \in I^{<\alpha}) > r$, such that the immediate subderivations are $r^+$ derivations, and there is an $s$ path $p$ with $s > r$ for the end-sequent $(n \in I^{<\beta}, ?, f), \Theta$ such that $Cut(p), FCut(p) \leq r$. Then there is an $r^+$ derivation $d'$ for $(n \in I^{<\beta}, ?, f), \Theta$ such that every $AxPFH_{n,\alpha,\gamma}$ or $AxPClFH_{n,\gamma}$ in $d'$ satisfies $rk(n \in I^\gamma) < r$ and if $h(d) \leq \eta$ then $h(d') \leq \eta \# \omega \# \eta$.*

This lemma takes a $PCut$ which would be unacceptable in an $r^+$-derivation and converts it to an inference which is allowed in an $r^+$-derivation. The resulting inference will be a $PCutFr$ if $r = rk(n \in I^\gamma)$ for some $\gamma$, and a $PCut$ inference otherwise. Note that this lemma allows for the possibility of jumping multiple ranks, when $s > r+1$.

*Proof.* If $r = rk(n \in I^\gamma)$ for a suitable $\gamma$ then we apply Lemma 9.20 to the subderivations $Prem(I, \gamma)$ to give $r^+$ derivations $Prem'(I, \gamma)$ of $(n \in I^\gamma, \top, t), \Theta$ and prune the inference to $PCutFr_{n,\alpha,\beta,\gamma}$. If there is no $\gamma$ such that $r = rk(n \in I^\gamma)$ then we just prune to a $PCut_{n,\alpha,\beta,\gamma}$ inference.

If any $AxPFH_{n,\alpha,\gamma}$ axioms with $rk(n \in I^\gamma) \geq r$ appear in this derivation, apply Lemma 9.19 to them and $Prem'(I, \gamma)$, using Lemma 9.25 to get an $s'$ path for $s \geq$

$s' > r$. The height bound holds since $h(Prem'(I,\gamma)) \leq \eta$ if $\gamma$ is in the ordinals of the end-sequent, and $h(Prem'(I,\gamma)) \leq \eta + \gamma$ otherwise.

If any $AxPClFH_{n,\alpha,\gamma}$ axioms with $rk(n \in I^\gamma) \geq r$ appear in this derivation, we do the same thing, but must ensure that the height bound holds. But, by the definition of $\gamma$, we must have $\gamma \underset{-}{\ll} \eta$, so $h(Prem'(I,\gamma)) \leq \eta$. □

**Lemma 9.27.** *If $d$ is an $r+1$ derivation of $\Theta$ and $\Theta$ has an $r+1$ path $p$ then there is an $r^+$ derivation $d'$ of $\Theta$ and if $h(d) \leq \eta$ then*

$$h(d') \leq (\max\{r,\omega\} + \omega)^{\omega \,\#\, \eta \,\#\, \eta+2}$$

*Recall that we may use Definition 2.2 to shorten this to $h(d') \leq (\alpha+\omega)_1(\eta)$.*

*Proof.* By induction on $d$.

If $d$ is an axiom and $r < \Omega$ then the result is trivial.

If $d$ is an $H$-type axiom $(\Sigma, H, A)$, $r \geq \Omega$, we already have $o(H^\frown\Sigma_S; r+1) \leq \eta$. Since this exists in a derivation of $\emptyset$, for each substitution $S$ in this sequence, $rk(e(S)) > r$. Therefore

$$o(H^\frown\Sigma_S; r) = (\Omega+\omega)^{o(H^\frown\Sigma_S; r+1)} \leq (\Omega+\omega)^{\omega \,\#\, \eta \,\#\, \eta+2}$$

Otherwise, let $\mathcal{I}$ be the last inference of $d$, let $\{Prem(I,u)\}$ be the family of immediate sub-derivations of $d$, and let $\Theta_u$ be the end-sequent of $Prem(I,u)$. Then by I.H., for each $u$ there is an $r^+$ derivation $Prem'(I,u)$ of $\Theta_u$ with $h(Prem'(I,u)) \leq (\max\{r,\omega\}+\omega)^{\omega \,\#\, h(Prem(I,u)) \,\#\, h(Prem(I,u))+2}$. Let $d^+$ be the derivation of $\Theta$ by $\mathcal{I}$ from $\{Prem'(I,u)\}$.

If $\mathcal{I}$ is a $Cut$ of rank $r$, we apply Lemma 9.23 to $d^+$ to get an $r^+$ derivation $d'$ such that

$$\begin{aligned} h(d') &\leq (\omega \,\#\, h(d^+)) \times (\max\{r,\omega\}+1) + 1 \\ &= (\omega \,\#(\max\{r,\omega\}+\omega)^{\omega \,\#\, h(d) \,\#\, h(d)}) \times (\max\{r,\omega\}+1) + 1 \\ &\leq (\omega \times (\max\{r,\omega\}+1)) \,\#(\max\{r,\omega\}+\omega)^{\omega \,\#\, h(d) \,\#\, h(d)+1} + 1 \\ &\leq (\max\{r,\omega\}+\omega)^{\omega \,\#\, h(d) \,\#\, h(d)+2} \end{aligned}$$

(Note that we use standard ordinal exponentiation rather than the $\alpha^{\underline{\beta}}$ operation which corresponds to the iteration of $\times$. This is justified since $\max\{r,\omega\}+\omega$ is a limit, so the two operations agree in all cases we are interested in [Bac55](§23.1).)

If $\mathcal{I}$ is an appropriate $PCut$ inference not allowable in an $r^+$ derivation, apply Lemma 9.26, and if $\mathcal{I}$ is an appropriate $FCut$ inference not allowable in an $r^+$ derivation, apply Lemma 9.21.

If $\mathcal{I}$ is a $CutFr^*$ inference, we deal with two cases. If there is some $u \neq ?$ such that $Prem(I,u)$ is not an $r^+$-derivation (note that we cannot apply IH to these premises, because there is not a valid path) then trim this to a $Fr$ inference. This can only happen when $r \leq o(H^\frown\{(e,?\} \cup \Theta_S))$, so the derivation remains valid. Otherwise, apply Lemma 9.24.

Otherwise $d^+$ is an $r^+$ derivation of $\Theta$.

We must check that $Ax(d) \leq r$ is met. Let $\Sigma$ be an axiom appearing in $d$ with target rank $\geq r$. If $\Sigma$ is an $AxH$, $AxFH$, or $AxClFH$ axiom then by Lemma 9.9(2)(b),

$(\Sigma f)_{\geq r} \subseteq \Theta$, so it follows that $e \in \mathrm{dom}(\Theta)$. But $\Theta f \leq r$ by Lemma 9.9(1). The target rank must be $r$, since $d$ together with $p$ gives an $r+1$-derivation of $\emptyset$.

If $\Sigma$ is an $AxPH$ which conflicts with the premise of $\mathcal{I}$, but not the conclusion, then $\Sigma$ is at a corrected $H$-step, so $r = 3\alpha + 1$ for some $\alpha$. Consider the element in the premise of $\mathcal{I}$ which $\Sigma$ conflicts with. If it is $(e(\Sigma), ?, t)$ then $\mathcal{I}$ is a $Cut_e$, and since $\alpha \in Ord(e)$ and there is an $r+1$-path for $\Theta$, $n \in I^\alpha \in \mathrm{dom}(\Theta)$. But this conflicts with $\Sigma$.

If the element in the premise is $n \in I^\alpha$ then $\mathcal{I}$ must be a $ClFH$ inference, so the $H$-rule applies to $\Theta$, and therefore

$$(\epsilon x \neg |B[I^{<\Omega}, n, x]|_{\overline{\Sigma_{\leq r}}}, ?) = (\epsilon x \neg |B[I^{<\Omega}, n, x]|_{\overline{\Theta_{\leq r}}}, ?) \in \Theta$$

Finally, if the element in the premise is $(\epsilon x \neg |B[I^{<\Omega}, n, x]|_{\overline{\Sigma_{\leq r}}}, ?)$ then $\mathcal{I}$ is either an $Fr$ or a $CutFr^*$ inference. But if this is an $Fr$ inference then we must have $r > h(\Theta)$ since $(\epsilon x \neg |B[I^{<\Omega}, n, x]|_{\overline{\Sigma_{\leq r}}}, ?) \in \Sigma_S$. But then $e \in A$, so $\Sigma$ conflicts with $(\Theta, H, A)$.

If $\mathcal{I}$ is a $CutFr^*$ inference then since $(e, ?, t) \in \Sigma$, it follows from Lemma 9.8 than $o(H^\frown\{(e, ?)\} \cup \Theta_S) < r$, so this axiom was eliminated when we applied Lemma 9.24.

Suppose $\Sigma$ is either an $AxPClFH_{n,\beta}$ or $AxPFH_{n,\beta}$ axiom which conflicts with a premise of $\mathcal{I}$ but not the conclusion. Let $e(\Sigma) \equiv n \in I^\alpha$. Then it must conflict with some $n \in Y$ with $rk(n \in Y) \geq rk(n \in X)$. But then $\mathcal{I}$ must be a $PCut$ inference (it cannot be an $FFr$ inference, since then the conflicting component would be replaced by another conflicting component). But then we eliminated $\Sigma$ when we applied Lemma 9.26. $\square$

**Lemma 9.28.** *If $d$ is an $\alpha + r$-derivation of $\Theta$ with $r < \omega$ such that $h(d) \leq \eta$ and $\Theta$ has an $\alpha + r$-path $p$ in which all $Cut$ and $FCut$ inferences have rank $\leq \alpha$ then there is an $\alpha$ derivation $d'$ of $\Theta$ such that $h(d') \leq (\alpha + \omega)_r(\eta)$.*

*Proof.* Note that for all $n$ such that $0 \leq n \leq r$, we have an $\alpha + n$ path $p_n$ by Lemma 9.25.

By induction on $r$. If $r = 0$ then we are done; otherwise, let $r = s + 1$. Then by Lemmata 9.27, there is an $s^+$ derivation $d^*$ of $\Theta$ with $h(d^*) \leq (\alpha + \omega)_1(\eta)$. By Lemma 9.10, there is an $s$ derivation of $\Theta$. $\square$

**Lemma 9.29.** *Let $\lambda_0 < \lambda$, and define $(\varphi\lambda_0)^n\eta$ to be the result of iterating $\varphi\lambda_0$ $n$-times on $\eta$. Then $(\varphi\lambda_0)^n(\varphi\lambda\eta + 1) \leq \varphi\lambda(\eta + 1)$.*

*Proof.* Clearly $\varphi\lambda\eta + 1 < \varphi\lambda(\eta + 1)$, and therefore

$$(\varphi\lambda_0)^n(\varphi\lambda\eta + 1) \leq (\varphi\lambda_0)^n(\varphi\lambda(\eta + 1))$$

By straightforward induction on $n$, $(\varphi\lambda_0)^n(\varphi\lambda(\eta + 1)) = \varphi\lambda(\eta + 1)$. $\square$

**Lemma 9.30.** *If $d$ is an $r + \omega^\lambda$ derivation of $\Theta$ with $\lambda$ a limit, $h(d) \leq \eta$, and $\Theta$ has an $r + \omega^\lambda$-path $p$ such that $Cut(p), FCut(p) \leq r$, then there is an $r^+$ derivation $d'$ of $\Theta$ with $h(d') \leq \varphi(\lambda + 1)(r + \omega^\lambda + \eta + 1)$.*

*Proof.* By main induction on $\lambda$ and side induction on $h(d)$. Assume that the lemma holds for all limit ordinals $< \lambda$.

Consider the last inference of $d$:

1. A $Cut$ or $FCut$, of rank $\sigma \geq r$. If $\lambda = 0$ then by Lemma 9.28, there is a $d'$ with $h(d') \leq \varphi(\lambda+1)(r+\omega^{\lambda}+\eta+1)$. If $\lambda > 0$ then there are $n$ and $\lambda_0$ such that $\sigma < r + n\omega^{\lambda_0}$. Let $Prem(I,u)$ be the $\lambda$-subderivations of the premise of $I$ with $h(Prem(I,u)) \leq \eta_u$; then by the side induction, we have $(r+n\omega^{\lambda_0})^+$ derivations $Prem'(I,u)$ with $h(Prem'(I,u)) \leq \phi(\lambda+1)(r+\omega^{\lambda}+\eta)$. Let the derivation $c$ be the result of applying $Cut$ or $FCut$ to the $Prem'(I,u)$.

   Applying the main IH $n$ times gives an $r^+$ derivation $d'$ of $\Theta$ with

   $$h(d') \leq (\varphi(\lambda_0+1))^n(\varphi(\lambda+1)(r+\omega^{\lambda}+\eta)+1) \leq \varphi(\lambda+1)(r+\omega^{\lambda}+\eta+1)$$

   by Lemma 9.29.

2. A $PCut_{n,\alpha,\beta,\delta}$. Let $Prem(I,\gamma)$ be the $\lambda$-subderivations of its premises. By the side induction, each $Prem(I,\gamma)$ can be tranformed into an $r^+$ derivation $Prem'(I,\gamma)$. Let $c$ be the result of applying $PCut$ to the $Prem'(I,\gamma)$.

   Applying Lemma 9.26 gives an $r^+$ derivation $d'$ of $\Theta$. The height bound is trivial.

3. Otherwise, using IH, replace each subderivation $Prem(I,u)$ with an $r^+$ derivation $Prem'(I,u)$, and the resulting derivation, $d'$, is an $r^+$ derivation and the height bound is trivial.

$\square$

**Lemma 9.31 (Collapsing).** *Suppose $d$ is an $\Omega$-derivation $\Theta$ such that $h(d) \leq \eta$. Then for every $r \geq Dh(d)$, there is an $r$-derivation $d'$ of $\Theta$ with $h(d') \leq D\eta$.*

*Proof.* By induction on $d$. If $d$ is an axiom, the result is trivial, since if $0 \ll \eta$, also $0 \ll D\eta$. Otherwise, apply IH to each premise of the final inference $I$ of $d$. If $I$ is an $H, FH, Fr$, or $FFr$ inference, the result follows directly from IH, since the only premise $d'$ has index 0, so $0 \ll \eta$ and $h(d') \ll h(d)$ implies $Dh(d') \ll D\eta$.

If $I$ is a $Cut$, $CutFr^*$, or $FCut$ inference then we may apply $I$ to the result of applying IH to each premise. Then each premise is indexed by an ordinal $\leq \gamma < \Omega$, where $\gamma \ll \eta$, and therefore $\gamma \ll D\eta$, so by definition, if some premise has height $\alpha$ then $D\alpha \ll D\eta$, and we are done.

Now suppose $I$ is a $PCut_{n,\Omega,\beta}$. If $(\Sigma, H, A)$ is some $AxPClFH_{n,\Omega,\alpha}$ axiom above this, we have $\alpha = D(o(H^\frown\Sigma;\Omega)+k(H^\frown\Sigma))$. But since this is an $\Omega$-derivation, $h((\Sigma,H,A)) \leq \eta$ implies $o(H^\frown\Sigma;\Omega) \leq \eta$, and using Lemma 9.2, $\alpha \ll D\eta$, so we may trim this to a $PCut_{n,r,\beta}$ inference with height $\leq D\eta$ for any $r \geq D\eta$. $\square$

**Lemma 9.32.** *If $d$ is an $\Omega + r$-derivation of $\emptyset$ with $r < \omega$ then there is a $0$-derivation $d'$ of $\emptyset$, and if $h(d) < \Omega^{\Omega^3+\Omega^2+1}$ then $h(d') < D\epsilon_{\Omega+1}$.*

*Proof.* By Lemma 9.28 there is an $\Omega$-derivation $d_1$ of $\emptyset$ with $h(d_1) \leq (\Omega+\omega)_r(\Omega^{\Omega^3+\Omega^2+1})$. By Lemma 9.31, there is a $Dh(d_1)$-derivation $d_2$ of $\emptyset$ with $h(d_2) \leq D(\Omega+\omega)_r(\Omega^{\Omega^3+\Omega^2+1})$. Finally, by Lemma 9.30 there is a $0$-derivation $d'$ of $\emptyset$ with

$$h(d') \leq \varphi(D(\Omega+\omega)_r(\Omega^{\Omega^3+\Omega^2+1})+1)(2D(\Omega+\omega)_r(\Omega^{\Omega^3+\Omega^2+1})+1) < D\epsilon_{\Omega+1}$$

□

**Lemma 9.33.** *In a $0$-derivation $d$ of $\emptyset$ all sequents are correct, the top sequent is an $AxS$, and all other inferences are either $Fr$, $H$, $FFr$, $FH$, or $ClFH$.*

*Proof.* Since $d$ is a $0$-derivation, there are no $Cut$, $CutFr$, $FCut$, $PCut$, or $PCutFr$ inferences, so $d$ is linear. By bottom-up induction, all sequents in $d$ are correct and and $\Theta f = \emptyset$ for every sequent $\Theta$ in $d$. Since $d$ is well-founded, there is a top-sequent $\Upsilon$ which must be an axiom, and since it must be correct, $\Upsilon f = \emptyset$, and since $Ax(d) \leq 0$ and the endsequent is $\emptyset$, it must be an $AxS$. □

**Lemma 9.34.** *If there is a $0$-derivation $d$ of $\emptyset$ then the $H$-process terminates.*

*Proof.* By Lemma 9.33, all inferences in $d$ are $Fr$, $H$, $FFr$, $FH$, or $ClFH$. Since the derivation is wellfounded, it corresponds to a finite sequence $\Theta_0, \ldots, \Theta_n$. Define $S_i = \{(e,u) \in (\Theta_i)_S \mid u \neq ?\}$. Since $Fr$ and $FFr$ inferences only add or remove expressions with default values, we have a sequence of $\epsilon$-substitutions and if $S_i \neq S_{i+1}$ then $S_{i+1} = H(S_i)$ since the inference must be one of $H$, $FH$, or $ClFH$. Since $\Theta_n$ is an instance of $AxS$, it follows that the $H$-process terminates. □

**Theorem 1.** *The $H$-process terminates.*

*Proof.* By Lemma 9.7 there is an $r$ derivation of $\emptyset$ for some $r < \Omega+\omega$. If $r \geq \Omega$, apply Lemma 9.32, otherwise just apply Lemma 9.30. The result is a supported $0$-derivation of $\emptyset$, so by Lemma 9.34, the $H$-process terminates. □

# References


[Ack40] Wilhelm Ackermann, *Zur Widerspruchsfreiheit der Zahlentheorie*, Math. Ann. **117** (1940), 162–194.

[Ara] Toshiyasu Arai, *Ideas in the epsilon substitution method for* $\Pi^0_1 - FIX$.

[Ara02a] ______, *Epsilon substitution method for* $ID_1(\Pi^0_1 \vee \Sigma^0_1)$, preprint (2002).

[Ara02b] ______, *Epsilon substitution method for theories of jump hierarchies*, Archive for Mathematical Logic **41** (2002), 123–153.

[Ara03] ______, *Epsilon substitution method for* $ID_1(\Pi^0_1 \vee \Sigma^0_1)$, Annals of Pure and Applied Logic **121** (2003), no. 2, 163–208.

[Bac55] Heinz Bachmann, *Transfinite zahlen*, Springer-Verlag, 1955.

[BFPS81] W. Buchholz, S. Feferman, W. Pohlers, and W. Seig, *Iterated inductive definitions and subsystems of analysis: Recent proof-theoretical studies*, Lecture Notes in Mathematics, no. 897, Springer-Verlag, 1981.

[Fef70] Soloman Feferman, *Formal theories for transfinite iterations of generalized inductive definitions and some subsystems of analysis*, Intuitionism and Proof Theory (R.E. Vesley A Kino, J. Myhil, ed.), 1970, Proceedings of the summer conference at Buffalo, N.Y. 1968, pp. 303–326.

[Gen34] Gerhard Gentzen, *Untersuchungen über das logische schließen i, ii*, Mathematische Zeischrift **39** (1934), 176–210, 405–431.

[Gen36] ______, *Die widerspruchsfreiheir der reinen zahlentheorie*, Mathematische Annalen **112** (1936), 494–565.

[HB70] David Hilbert and Paul Bernays, *Grundlagen der Mathematik*, vol. 2, Springer, 1970.

[Kre63] G. Kreisel, *Generalized inductive definitions*, Stanford Report (1963), 115–139.

[Min94] Grigori Mints, *Gentzen-type systems and Hilbert's epsilon substitution method*, Logic, Method and Philosophy of Science (Amsterdam), vol. IX, Elsevier, 1994, pp. 91–122.

[Min03] ______, *Extension of epsilon substitution method to ID1, streamlined version*, preprint (2003).

[Mos74] Y. N. Moschovakis, *Elementary induction on abstract structures*, North Holland, 1974.

[MT99] Grigori Mints and Sergei Tupailo, *Epsilon-substitution for the ramified language and $\Delta^1_1$-comprehension rule*, Logic and Foundations of Mathematics (Andrea Cantinia et. al., ed.), Kluwer Academic Publishers, 1999, pp. 107–130.

[MTB96] Grigori Mints, Sergei Tupailo, and Wilfried Buchholz, *Epsilon substitution method for elementary analysis*, Archive for Mathematical Logic **35** (1996), 103–130.

[Poh89] Wolfram Pohlers, *Proof theory: An introduction*, Lecture Notes in Mathematics, no. 1407, Springer-Verlag, 1989.

[Tak67] Gaisi Takeuti, *Consistency proofs of subsystems of classical analysis*, Annals of Mathematics **86** (1967), no. 299-348.

[Tow03] Henry Towsner, *Epsilon substitution for transfinite induction*, unpublished (2003).