Epsilon Substitution for Transfinite Induction

Henry Towsner

May 20, 2005

Abstract

We apply Mints' technique for proving the termination of the epsilon substitution method via cut-elimination to the system of Peano Arithmetic with Transfinite Induction given by Arai.

1   Introduction

Hilbert introduced the epsilon calculus in [Hilbert, 1970] as a method for proving the consistency of arithmetic and analysis. In place of the usual quantifiers, a symbol ϵ 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 we consider the (finite) list of critical formulas which appear in a proof. Then we define a series of functions, called ϵ-substitutions, 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 it 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 same formula. Typically, we start with a proof of $\exists x \phi[x]$ with ϕ quantifier free. At the conclusion of the H-process, we have $\phi[n]$ and a proof that n does, in fact, satisfy ϕ.

After several attempts, Ackermann [Ackermann, 1940] proved that the method terminates for first order arithmetic, and therefore that a substitution of numerical values for all infinitary terms can be found finitistically. While no proof currently exists for the termination of the method for full analysis (indeed, it is not clear how to formulate the process for full analysis), proofs have been given for several fragments, including elementary analysis [Mints et al., 1996], ramified analysis [Mints and Tupailo, 1999], the hyperarithmetical hierarchy [Arai, 2002], and, most recently, the impredicative system ID_1 [Arai, 2003], [Mints, 2003]. Some of these proofs are done by an extension of Ackermann's technique, while others are done using newer techniques.

While Arai's proof of the termination of the epsilon substitution method for ID_1 is based on Ackermann's method, it requires using the complete induction axiom

$$\forall x((\forall y < x \phi[y]) \rightarrow \phi[x]) \rightarrow \forall x \phi[x]$$

in place of the usual induction axiom

$$\forall x(\phi[x] \rightarrow \phi[Sx]) \rightarrow \forall x$$

As a preliminary step, he proves termination for first order arithmetic with the complete induction axiom using ordinal assignments in the style of [Ackermann, 1940]. This has the added advantage of allowing the consideration of more powerful systems in which the usual ordering of the natural numbers $<$ is replaced by some transfinite well-ordering.

Here we give a different proof that the epsilon substitution method terminates for this alternative formulation of arithmetic, based on a technique developed by Mints [Mints, 1994]. We define a Gentzen type proof system for epsilon substitutions (where the rules of the proof system depend on the particular set of critical formulas being considered). We then show that there is a derivation (including cuts) of the empty sequent, that cuts can be eliminated in a series of standard steps, and that a cut free derivation of the empty sequent is just an encoding of the H-process, terminating after a finite number of steps.

We modify Ackermann's original H-process to follow [Arai, 2003], which differs in that when $\phi[t]$ is true, we simply take the value of t under the current substitution to be the new value for $\epsilon x.\phi[x]$, rather than selecting the least number satisfying t. Something like this seems necessary, since when $\prec$ is transfinite, we cannot select a least n effectively. 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.

Our sequent system requires modifications in order to accommodate this. [Mints, 1994] distinguishes between fixed and temporary default values, using them to keep track of which cuts have yet to be eliminated below. Here we

need to distinguish the fixed and temporary cases regardless of the value of the substitution, since we can directly substitute one numerical value for another.

The method in this paper can also be extended to a system with bounded ϵ-terms of the form $\epsilon x \leq t\phi$.

2 $L\epsilon$

Our theory will use a language $L\epsilon$, essentially the usual language of first order arithmetic with ϵ in place of quantifiers. We assume that a recursive well-ordering relation $\prec$ is fixed.

The symbols of $L\epsilon$ are:

1. Variables $x, y, z, \ldots$

2. The 0-ary function constant 0 and the unary function constant S

3. Predicate constants for n-ary computable predicates, $n \geq 1$, including $\prec$, $=$, addition, and multiplication

4. Propositional logical connectives $\wedge$, $\neg$, and $\rightarrow$

5. The epsilon symbol ϵ

We let $\phi \vee \psi$ abbreviate $\neg(\neg\phi \vee \neg\psi)$.

The terms and formulas of $L\epsilon$ are defined simultaneously by:

1. Every variable is a term

2. 0 is a term

3. If t is a term then St is a term

4. If t_i is a term for $i \leq n$ and p is an n-ary predicate constant then $pt_1 \cdots t_n$ is a formula

5. If ϕ and ψ are formulas then $\neg\phi$, $\wedge\phi\psi$, and $\rightarrow \phi\psi$ are formulas

6. If ϕ is a formula and x a variable occurring free in ϕ then $\epsilon x\phi$ is a term.

Definition 1.
- *Terms and formulas are collectively called* expressions.

- *A term is a* numeral *if it is either* 0 *or* Sn *for some numeral* n. *The set of numerals is denoted* $\mathbb{N}$.

- $FV(e)$ *is the set of* free variables *of an expression* e, *defined in the usual way, with* ϵx *binding* x.

- *An expression is* closed *if* $FV(e) = \emptyset$.

- *If* u *is a term then* $e[x/u]$ *denotes the result of substituting* u *for each free occurrence of* x *in* e, *with free variables in* e *renamed if necessary. Often we write* $e[u]$ *for* $e[x/u]$ *when* x *is clear from the context.*

3

Definition 2.   • *The* depth $d(e)$ *of an expression e is the number of closed ϵ-subterms of e.*

• *An ϵ-term $e = \epsilon x.\phi[x]$ is* canonical *if it is closed and $d(e) = 1$ (that is, it has no closed proper ϵ-subterms)*

• *TRUE (FALSE) is the set of all true (false) closed, variable-free formulas.*

3 $PA\epsilon$

We introduce a system which is the translation of Peano Arithmetic with transfinite $\prec$-induction into $L\epsilon$.

The only inference rule of $PA\epsilon$ is modus ponens: $\dfrac{\phi \qquad \phi \rightarrow \psi}{\psi}$

Let α denote the order type of the well-ordering $\prec$. If $\prec = <$, the usual ordering of the natural numbers, then this system will be equivalent to PA, but if $\prec$ is the standard ordering of an ordertype $\epsilon_0 \preceq \alpha$, this system will be stronger. In order to simply the calculation of bounds later, we assume $\alpha \geq \omega^2$.

We let $s \preceq t$ be an abbreviation for $s \prec t \vee s = t$.

The axioms are:

1. All propositional tautologies of the language $L\epsilon$

2. All substitution instances of the defining axioms for predicate constants

3. Equality axioms $t = t$ and $s = t \rightarrow \phi[s] \rightarrow \phi[t]$

4. Peano axioms $\neg St = 0$ and $Ss = St \rightarrow s = t$

5. Ordering axioms

    (a) $s \prec t \vee t \prec s \vee s = t$
    (b) $\neg t \prec t$
    (c) $s \prec t \wedge t \prec u \rightarrow s \prec u$
    (d) $\neg t \prec 0$
    (e) $\neg(s \prec t \wedge t \prec s)$

6. Critical formulas:

$$\phi[t] \rightarrow \phi[\epsilon x.\phi[x]] \wedge (\epsilon x.\phi[x]) \preceq t$$

$$\neg s = 0 \rightarrow s = S\epsilon xs = Sx$$

4

4 ϵ-Substitutions

We will define ϵ-substitutions which tell us how to replace (some) ϵ-terms with numeric values. We will have a default value, ?, which indicates that, while a term will be replaced by 0, 0 may not actually be a number which satisfies the corresponding formula. In particular, this value for $\epsilon x\phi[x]$ is compatible with the possibility that $\forall x\neg\phi[x]$. Whenever we assign a value n other than ? to $\epsilon x\phi[x]$, we intend that $\phi[n]$ holds.

Definition 3.

An ϵ-substitution is a function S such that:

- *The domain of S is a set of canonical ϵ-terms*

- *If $e \in \mathrm{dom}(S)$ then $S(e) \in \mathbb{N} \cup \{?\}$*

  An ϵ-substitution is total *if its domain is the set of all canonical ϵ-terms*

We will also consider the standard extension of an ϵ-substitution. This makes a substitution total by simply assigning the default value to any term not already assigned.

Definition 4. *The* standard extension $\overline{S}$ *of an ϵ-substitution S is*

$$\overline{S} \quad \cup \quad \{(e,?) \mid \neg\exists v(e,v) \in S\}$$

5 Computations with ϵ-Substitutions

Given an expression e, we can define the reduction of a single occurrence of an ϵ-term inductively.

Definition 5.   *1. If $(e,u) \in S$ and $u \neq ?$ then $e \hookrightarrow^1_S u$*

2. *If $(e,?) \in S$ then $e \hookrightarrow^1_S 0$*

3. *If $1 \leq i \leq n, e_i \hookrightarrow^1_S e'_i$ then $pe_1\cdots e_i\cdots e_n \hookrightarrow^1_S pe_1\cdots e'_i\cdots e_n$*

4. *If $\phi \hookrightarrow^1_S \phi'$ then $\epsilon x.\phi[x] \hookrightarrow^1_S \epsilon x.\phi'[x]$*

Let $\hookrightarrow_S$ denote the reflexive, transitive closure of $\hookrightarrow^1_S$. Say e is S-irreducible if there is no e' such that $e \hookrightarrow^1_S e'$.

Lemma 1.   *1. $e \hookrightarrow_S e' \to FV(e') = FV(e)$*

2. *$e \hookrightarrow^1_S e' \to d(e') < d(e)$*

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

4. *Every expression e has a unique S-irreducible normal form $|e|_S$ under $\hookrightarrow_S$*

Proof. The first three parts are proved by induction on the structure of e.

1. If e is canonical then $FV(e) = \emptyset = FV(e')$. Otherwise, the statement follows by I.H..

2. If e is canonical then $d(e') = 0 < 1 = d(e)$. Otherwise the statement follows from I.H..

3. If e is canonical then $e' = e''$. If $e = pe_1 \ldots e_n$ then $e' = pe_1 \ldots e_i' \ldots e_n$ and either $e'' = pe_1 \ldots e_i'' \ldots e_n$, in which case the result follows by I.H., or $e'' = pe_1 \ldots e_j' \ldots e_n$, in which case both e' and e'' reduce to $pe_1 \ldots e_i' \ldots e_j' \ldots e_n$. In the remaining cases, just apply I.H..

4. By part two, a series of reductions must terminate, and by part 3, the result is unique.

$\square$

Definition 6.
- $e \hookrightarrow_S TRUE(FALSE)$ *iff* $|e|_S \in TRUE(FALSE)$

- e *is* S-computable *if* $d(|e|_S) = 0$

- $\mathcal{F}(S) = \{\phi[v] \wedge \neg\phi(0) \mid (\epsilon x.\phi[x], v) \in S \wedge v \neq ?\}$

- S *is* correct *if for any* $f \in \mathcal{F}(S)$, $f \hookrightarrow_{\overline{S}} TRUE$

- S *is* computationally inconsistent *(ci) if there is some* $\phi \in \mathcal{F}(S)$ *such that* $\phi \hookrightarrow_S FALSE$

- S *is* deciding *if every critical formula* Cr_I *is* S-computable, that is, computes to a formula in either $TRUE$ or $FALSE$.

$\mathcal{F}(S)$ is the set of formulas which S asserts to be true.

Lemma 2. *If* $s = t \hookrightarrow_{\overline{S}} TRUE$ *and* $u[x]$ *is some term then* $|u[s]|_{\overline{S}} = |u[t]|_{\overline{S}}$.

Proof. Note that since $s = t \hookrightarrow_{\overline{S}} TRUE$, we have $|s|_{\overline{S}} = |t|_{\overline{S}}$.
By induction on u:

1. If $u = x$ or 0 then this is trivial

2. If $u = Su'$ then this follows directly from IH

3. If $u = \epsilon y\phi[x, y]$ then $\phi[|s|_{\overline{S}}, y]$ is (syntactically) the same as $\phi[|t|_{\overline{S}}, y]$, so $|\epsilon y\phi[s, y]|_{\overline{S}} = |\epsilon y\phi[t, y]|_{\overline{S}}$.

$\square$

Lemma 3. *If* $Cr = \{Cr_1, \ldots, Cr_N\}$ *is the set of critical formulas appearing in the proof of some formula* ϕ *and each* $Cr_i \hookrightarrow_{\overline{S}} TRUE$ *then* $\phi \hookrightarrow_{\overline{S}} \top$.

Proof. Each $Cr_I \hookrightarrow_{\overline{S}} TRUE$. If ψ is some axiom other than a critical formula or $s = t \to \psi'[s] \to \psi'[t]$ then obviously $\psi \hookrightarrow_{\overline{S}} TRUE$. Modus ponens preserves truth under S, since $\psi_0 \hookrightarrow_{\overline{S}} TRUE$ and $(\psi_0 \to \psi_1) \hookrightarrow_{\overline{S}} TRUE$ then $\psi_1 \hookrightarrow_{\overline{S}} TRUE$.

Finally, suppose ψ is $s = t \to \psi'[s] \to \psi'[t]$. Assume $s = t \hookrightarrow_{\overline{S}} TRUE$ and $\psi'[s] \hookrightarrow_{\overline{S}} TRUE$, and proceed by induction on ψ:

1. If $\psi[x] = p t_1[x] \cdots t_n[x]$ then $|t_i[s]|_{\overline{S}} = |t_i[t]|_{\overline{S}}$, so $\psi[t] \hookrightarrow_{\overline{S}} TRUE$.

2. Otherwise $\psi[t] \hookrightarrow_{\overline{S}} TRUE$ follows straightforwardly from the inductive hypothesis.

$\square$

6   Rank

The rank of a formula is the level of nesting present in its ϵ-subterms when the main variable of a term appears in a subterm. The definition here is the same as [Mints et al., 1996]. So $\epsilon x.(\epsilon y. y = 2) = x$ has rank 1, while $\epsilon x.(\epsilon y. y = x) = x$ has rank 2.

Definition 7.   • $\sigma \notin FV(e) \cup \{*\} \to rk_\sigma(e) = 0$, *otherwise:*

- $rk_\sigma(pe_1 \cdots e_n) = \max\{rk_\sigma(e_1), \ldots, rk_\sigma(e_n)\}$

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

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

Definition 8. *If S is an ϵ-substitution and $r < \omega$ then $S_{\leq r} = \{(e, n, u) \in S \mid rk(e) \leq r\}$*

Lemma 4. *If S and S' are ϵ-substitutions with $S_{\leq r} = S'_{\leq r}$ and $rk(e) \leq r$ then $|e|_S = |e|_{S'}$*

Proof. All subterms of e have ranks $\leq rk(e)$, and therefore $e \hookrightarrow^1_S e' \Leftrightarrow e \hookrightarrow^1_{S'} e'$, so $e \hookrightarrow_S e' \Leftrightarrow e \hookrightarrow_{S'} e'$. $\square$

Lemma 5. *If $e \hookrightarrow^1_S e'$ then $rk_\sigma(e') \leq rk_\sigma(e)$*

Proof. By induction on the definition of $\hookrightarrow^1_S$. If $\sigma \notin FV(e) \cup \{*\}$ then $\sigma \notin FV(e') \cup \{*\}$ since $FV(e) = FV(e')$. If $e = pe_1 \cdots e_n$ or e is a non-canonical ϵ-term then this follows from IH. If e is a canonical ϵ-term then $rk(e') = 0 < rk(e)$. $\square$

7 H-Process

This is essentially identical to that given in [Arai, 2003].

Let $Cr = \{Cr_0, \ldots, Cr_N\}$ be a fixed sequence of closed critical formulas.

Definition 9. *We say S is* solving *if for each $I \leq N$, $Cr_I \hookrightarrow_{\overline{S}} TRUE$.*

Let S be a finite, correct, nonsolving ϵ-substitution. We will consider the critical formulas made false by $\overline{S}$ and select the first one of minimal rank to be fixed. For $I = 0, \ldots, N$, we define parameters needed for the H-process.

Definition 10. *If the I-th critical formula has the form*

$$\phi[t] \rightarrow \phi[\epsilon x.\phi[x]] \wedge (\epsilon x.\phi[x]) \preceq t$$

we define:

- $e_I^S = \epsilon x.|\phi[x]|_{\overline{S}}$

- $r_I^S = rk(e_I^S)$

- $v_I^S = |t|_{\overline{S}}$

If the I-th critical formula has the form

$$\neg s = 0 \rightarrow s = S(\epsilon x s = Sx)$$

we define:

- $e_I^S = \epsilon x.|s|_{\overline{S}} = Sx$

- $r_I^S = rk(e_I^S)$

- $v_I^S = |s|_{\overline{S}} - 1$

Next, we define:

- $I(S) = I \Leftrightarrow$
  $Cr_I \hookrightarrow_{\overline{S}} \bot \wedge \forall J \leq N[Cr_J \hookrightarrow_{\overline{S}} \bot \rightarrow r_I^S < r_J^S \vee (r_I^S = r_J^S \wedge I \leq J)]$
  That is, $I(S)$ is the index of a critical formula which is false under S such that all critical formulas of smaller rank and all formulas of equal rank which appear earlier are true under S.

- $Cr(S) = Cr_{I(S)}$

- $r^S = r_{I(S)}^S$

- $e^S = e_{I(S)}^S$

- $v^S = v_{I(S)}^S$ *is called the H-value of S*

- $H(S) = S_{<r^S} \cup \{(e, u) \in S \mid rk(e) = r^S \wedge e \neq e^S\} \cup (e^S, v^S)$

8

Lemma 6. *If S is a finite, correct, nonsolving ϵ-substitution then $v^S \neq 0$*

Proof. Suppose $v_I^S = 0$ for some I. Then the conclusion of Cr_I is of one of the forms

$$\phi(\epsilon x \phi(x)) \wedge (\epsilon x \phi(x)) \preceq t$$

$$s = S(\epsilon x s = S x)$$

In this first case, suppose $|\epsilon x \phi(x)|_{\overline{S}} \neq 0$. Then, by correctness, $\neg\phi(0) \hookrightarrow_{\overline{S}} TRUE$. On the other hand, if $|\epsilon x \phi(x)|_{\overline{S}}$ then, since obviously $0 \preceq t \hookrightarrow_{\overline{S}} TRUE$ then we have $\phi(0) \hookrightarrow_{\overline{S}} TRUE$. In either case, $Cr_I \hookrightarrow_S TRUE$, so $I \neq I(S)$.

In the second case, if $v_I^S = 0$ then $|s|_{\overline{S}} = S0$, and therefore, by correctness, $|\epsilon x s = S x|_{\overline{S}} = 0$, so $s = S(\epsilon x s = S x) \hookrightarrow_{\overline{S}} TRUE$. Again, $Cr_I \hookrightarrow_S TRUE$, so $I \neq I(S)$. $\square$

Lemma 7. *If S is a finite, correct, nonsolving ϵ-substitution then $H(S)$ is finite and correct.*

Proof. Suppose $(\epsilon x.\psi[x], v) \in H(S)$. If $(\epsilon x.\psi[x], v) \in S$ then it follows that $rk(\epsilon x.\psi[x]) \leq rk(e^S)$, and therefore that $|\psi[\epsilon x.\psi[x]]|_{\overline{H(S)}} = |\psi[\epsilon x.\psi[x]]|_{\overline{S}} = TRUE$ since S is correct.

On the other hand, suppose $(\epsilon x.\psi[x], v) \in H(S) \setminus S$ and

$$Cr(S) = \phi(t) \rightarrow \phi[\epsilon x.\phi[x]] \wedge (\epsilon x.\phi[x]) \preceq t$$

Then since $S_{<rk(e^S)} = H(S)_{<rk(e^S)}$, we have $|\phi[e^S]|_{\overline{H(S)}} = |\phi[v^S]|_{\overline{S}} = TRUE$. If $\phi[0] \hookrightarrow_{\overline{H(S)}} TRUE$ then we would have $\phi[0] \hookrightarrow_{\overline{S}} TRUE$. But then either $|\epsilon x \phi[x]|_{\overline{S}} = 0$, in which case we would have $Cr(S) \hookrightarrow_{\overline{S}} TRUE$, or $|\epsilon x \phi[x]|_{\overline{S}} > 0$, in which case $\phi[0] \hookrightarrow_{\overline{S}} FALSE$ by the correctness of S.

Finally, if $(\epsilon x.\psi[x], v) \in H(S) \setminus S$ and

$$Cr(S) = \neg s = 0 \rightarrow s = S(\epsilon x s = S x)$$

then we have $\psi[x] = |s|_{\overline{S}} = Sx$. Since $v = |s|_{\overline{S}} - 1$ and $|s|_{\overline{S}} - 1$, we have $|s|_{\overline{S}} = S|s|_{\overline{S}} \in TRUE$.

Hence, $H(S)$ is correct. $\square$

Lemma 8. *If S is a finite, correct, nonsolving ϵ-substitution such that $(e, u) \in S$, $u \neq ?$, and $(e, v) \in H(S)$ for some $v \neq u$ then $v \neq ?$ and $v \prec u$.*

Proof. Since $e = \epsilon x.\phi[x]$ must be the H-term of S, $Cr(S)$ is of the form:

$$\psi[t] \rightarrow \psi[\epsilon x.\psi[x]] \wedge (\epsilon x.\psi[x]) \preceq t$$

where $|\psi|_{\overline{S}} = \phi$, $|t|_{\overline{S}} = v$, and (e, v) is placed into $H(S)$ by the H-step. ($Cr(S)$ cannot be of the form $\neg s = 0 \rightarrow s = S(\epsilon x s = S x)$, since such an axiom must be made true by $\overline{S}$ if $u \neq ?$). Since $Cr(S) \hookrightarrow_{\overline{S}} \bot$, the left side of the implication reduces to $\top$ while one of the conjuncts on the right must reduce to $\bot$. But if $u = |\epsilon x \psi(x)|_{\overline{S}} \preceq |t|_{\overline{S}} = v$ then, since S is correct, the right side is true under S. So $v \prec u$. $\square$

9

Definition 11. *The H-process for a fixed set Cr of critical formulas is defined by:*

- $S_0 = \emptyset$

- $S_{n+1} = \begin{cases} H(S_n) & \text{if } \overline{S_n} \text{ is nonsolving} \\ \emptyset & \text{otherwise} \end{cases}$

The process terminates *or* halts *if there is some n such that $\overline{S_n}$ is solving.*

8 ϵPA

In order to prove that the H-process eventually halts, we expand substitutions to sequents, which are substitutions with additional information. We then define a Gentzen-style deduction system, ϵPA, for these sequents.

The additional information required in a sequent is a marker $i \in \{t, f\}$, where t stands for temporary and f for fixed. When $(e, ?, i) \in S$ and e is the H-term of S, we derive the corresponding sequent in different ways depending on whether $i = t$ or $i = f$. In the former case, we apply an H-rule, deriving the sequent from one corresponding to $H(S)$. However if $(e, ?, f)$ is in the sequent, it is an axiom. This proves necessary in order to distinguish which elements of the sequent can serve as principal formulas in a cut rule.

Definition 12.   • *A sequent Θ is a set of tuples (e, u, i) where $\Theta_S = \{(e, u) \mid (e, u, i) \in \Theta\}$ is a finite ϵ-substitution with $i \in \{t, f\}$, and at most one of $(e, u, t), (e, u, f)$ is in Θ for any e and u.*

- $\text{dom}(\Theta) = \{e \mid (e, u, i) \in \Theta\}$

- *Suppose Θ is a sequent, e a canonical ϵ-term, $n \in \mathbb{N}$, $u \in \mathbb{N} \cup \{?\}$, 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, $e \notin \text{dom}(\Theta)$ or $(e, u, i) \in \Theta$.*

- $\Theta t = \{(e, u, t) \in \Theta\}$

- $\Theta f = \{(e, u, f) \in \Theta\}$

We introduce a deduction system for sequents. The axioms are:

- Θ is an instance of AxF if Θ_S is ci

- Θ is an instance of AxS is Θ_S is solving

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

The inference rules are:

$$\frac{(e, ?, f), \Theta \ldots \qquad (e, u, f), \Theta, \ldots \qquad (u \in \mathbb{N})}{\Theta} \; Cut_e$$

$$\frac{(e,?,t),\Theta\ldots \qquad (e,u,t),\Theta,\ldots \qquad (u\in\mathbb{N})}{\Theta}\ CutFr_e$$

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

$$\frac{(e,v,t),\Theta_{\leq rk(e)}}{(e,u,t),\Theta}\ H_{e,v}\ \text{If } e \text{ is the } H\text{-term and } v \text{ the } H\text{-value of } ((e,u,t),\Theta)_S.$$

8.1 Original Derivation

Our initial derivation of the empty sequent will be constructed solely out of cuts. It assigns arbitrary values to canonical subterms of ϵ-terms, starting with terms of maximal rank, until all of them are reduced to a canonical form and then computed. We use the construction given in [Mints et al., 1996], slightly simplified since all formulas have finite depth in the present system.

Definition 13. *Let S be an ϵ-substitution and $\Phi = \{A_1, \ldots, A_n\}$ a finite set of closed formulas.*

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

- $d_r(F) = \begin{cases} 0 & \text{if } rk(F) < r \\ d(F) & \text{otherwise} \end{cases}$

- $\#_S(\Phi, r) = \Sigma_{A\in\Phi} d_r(|A|_S)$

- $\nu_S(\Phi) = \omega \cdot \rho_S(\Phi) + \#_S(\Phi, \rho_S(\Phi))$

Note that $\#_S(\Phi, r), \rho_S(\Phi) < \omega$, and therefore $\nu_S(\Phi) < \omega^2$.

Lemma 9. *Let S be an ϵ-substitution and Φ a finite set of closed formulas. Pick some $A_0 \in \Phi$ with $rk(|A_0|_S) = \rho_S(\Phi)$, and let e be a canonical ϵ-subterm of $|A_0|_S$. For any $u \in \mathbb{N} \cup \{?\}$ let $S^u = S \cup \{(e,u)\}$ and let*

$$\Phi^u = \begin{cases} \Phi & \text{if } u =? \\ \Phi \cup \{\phi[u]\} & \text{otherwise} \end{cases}$$

Then for any $u \in \mathbb{N} \cup \{?\}$, $\rho_{S^u}(\Phi^u) \leq \rho_S(\Phi)$ and $\nu_{S^u}(\Phi^u) < \nu_S(\Phi)$.

Proof. 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 formula or term w. Also, $rk(|\phi[u]|_{S^u}) \leq rk(\phi[u]) < rk(e) \leq \rho_S(\Phi)$, so $\rho_{S^u}(\Phi^u) \leq \rho_S(\Phi)$.

If $\rho_{S^u}(\Phi^u) < \rho_S(\Phi)$ then obviously $\nu_{S^u}(\Phi^u) < \nu_S(\Phi)$, so suppose $\rho_{S^u}(\Phi^u) = \rho_S(\Phi)$. Since for each $A \in \Phi$, $rk(|A|_{S^u}) \leq rk(|A|_S)$ and $d(|A|_{S^u}) \leq d(|A|_S)$, so $d_{\rho_S(\Phi)}(|A|_{S^u}) \leq d_{\rho_S(\Phi)}(|A|_S)$. Since there is some e such that $|A_0|_S \hookrightarrow_{S^u} e$ and $rk(|A_0|_S) = \rho_S(\Phi)$, we have $d_{\rho_S(\Phi)}(|A_0|_{S^u}) < d_{\rho_S(\Phi)}(|A_0|_S)$. Also, $d_{\rho_S(\Phi)}(|\phi[u]|_{S^u}) = 0$, since $rk(|\phi[u]|_{S^u}) < r$. $\square$

Lemma 10. *Suppose Θ is a sequent, L a finite set of closed formulas, and $r = \rho_{\Theta_S}(\mathcal{F}(\Theta_S) \cup L)$. Then there is a deduction d of Θ by cuts of ranks $\leq r$ from computing sequents Υ containing Θ and computing all formulas in L. In addition, $h(d) \leq \nu_{\Theta_S}(\mathcal{F}(\Theta_S) \cup L)$.*

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

Let $\Phi = \mathcal{F}(\Theta_S) \cup L$. If Θ computes all formulas in Φ then Θ satisfied the condition. Otherwise, let $A \in \Phi$ be some formula such that $rk(|A|_{\Theta_S}) = r$ and let e be some canonical ϵ-subterm of $|A|_{\Theta_S}$. Then for each $u \in \mathbb{N} \cup \{?\}$, let $\Theta^u = (e, u, f), \Theta$.

Then Θ_S^u satisfies the conditions of the previous lemma, so $\nu_{\Theta^u}(\Phi^u) < \nu_{\Theta}(\Phi)$, and therefore by the inductive hypothesis, there is a deduction d_u of Θ^u by cuts of rank $\leq r$ from appropriate sequents Υ and with $h(d_u) \leq \nu_{\Theta^u}(\Phi^u)$.

Adding a cut with main term e gives the required deduction. $\square$

Lemma 11. *There is some $r < \omega$ such that there is a deduction d of the empty sequent consisting only of axioms and cuts with rank $\leq r$, and $h(d) < \omega^2$.*

Proof. Take $r = \rho_\Theta(L)$. Applying Lemma 10 to $\emptyset$ and $L = \{Cr_0, \ldots, Cr_N\}$ gives a deduction of $\emptyset$ consisting of only cuts with rank $\leq \rho_\emptyset(L)$ and axioms.

If some top sequent Θ of this deduction is not AxF or AxS then Θ_S must be cc, deciding, and nonsolving. Since the only inferences in the part already constructed are cuts, $\Theta t = \emptyset$. But then Θ has an H-term e, and since Θ is cc and deciding, $(e, ?, f) \in \Theta$. Then there is a critical formula

$$Cr_I = \psi[t] \to \psi[\epsilon x.\psi[x]] \wedge (\epsilon x.\psi[x]) \preceq t$$

such that $e = \epsilon x.|\psi[x]|_{\overline{\Theta_S}}$ and $|t|_{\overline{\Theta_S}} = u$ and $Cr_I \hookrightarrow_{\overline{\Theta_S}} \bot$.

Hence Θ is an instance of AxH since $(e, ?, f) \in \Theta$. $\square$

8.2 Cut-elimination

In cut-elimination, we proceed by first eliminating cuts of maximal rank, then of the next rank, and so on until all cuts are eliminated. We eliminate a cut by replacing Cut with $CutFr$, then replacing the corresponding AxH by an H rule. The premise of the H rule is found by taking the corresponding branch of the $CutFr$ derivation, modifying it to match the premise of the H rule, and placing it on top. The $CutFr$ rules are then pruned to Fr rules, once all cuts of the current rank have been eliminated.

The technique is essentially that used by Mints, with appropriate changes corresponding to the fact that non-? values will sometimes need to be updated. The substantial changes are found in Lemma 17; the rest of the lemmas are identical to those of [Mints et al., 1996] except where changes were necessary to accommodate the differences in the definition of sequents here.

A deduction with cuts of rank r and higher eliminated will be called an r deduction, and during the interim stages, when the cuts have been eliminated but the $CutFr$'s have not been pruned, we will have an r^+ deduction.

Definition 14.
- *If d is a deduction, $X \in \{Cut, CutFr, Fr, H\}$, and $\bowtie \in \{<, >, \leq, \geq, =\}$ then we say $X(d) \bowtie r$ if every application of a rule X in d has main term with rank $\bowtie r$.*

- *d is an r-deduction if $Cut(d) < r \wedge CutFr(d) < 0 \wedge Fr(d) \geq r \wedge H(d) \geq r$*

- *d is an r^+-deduction if $Cut(d) < r \wedge CutFr(d) = r \wedge Fr(d) > r \wedge H(d) \geq r$*

Lemma 12. *Every r^+-derivation can be transformed into an r-derivation.*

Proof. Just prune all $CutFr$ inferences to Fr inferences by deleting all premises except the leftmost one. $\qquad\square$

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

Lemma 13. *A 0-derivation d of $\emptyset$ consists of exactly one branch such that all sequents in d are correct, the top sequent is an axiom AxS, and all other inferences are either Fr or H.*

Proof. Since d is a 0-derivation, there are no Cut or $CutFr$ inferences in d, so d is linear. The rule H, viewed bottom up, preserves correctness by a previous lemma, and Fr preserves correctness since if $e \notin \mathrm{dom}(\Theta)$ then $\{(e, ?)\} \cup \Theta_S$ is correct, so every sequent in d is correct. By bottom-up induction, $f(\Theta) = \emptyset$ for any Θ in d, and since d is well founded, there is some top sequent Σ. Since it is correct and does not contain any fixed values, it must be AxS. $\qquad\square$

Lemma 14.
1. *If Θ is a sequent in an $r+1$ deduction of $\emptyset$ then $\Theta t > r$ and $\Theta f \leq r$*

2. *If Σ is a sequent in an r^+ deduction of Θ then:*
   - *(a) $\Theta_{\leq r} \setminus \Theta t \subseteq \Sigma$*
   - *(b) $(\Sigma f)_{\geq r} \subseteq \Theta$*
   - *(c) $\Theta t \geq r \Rightarrow \Sigma t \geq r$*

Proof.
1. The statement is proved by bottom-up induction on the proof. It obviously holds for $\emptyset$, and in an $r + 1$-deduction viewed bottom up, temporary elements are added by Fr and H, which have rank at least $r+1$, while fixed elements are added only by Cut, which has rank $< r+1$, that is, $\leq r$.

2. (a) Again by bottom-up induction. The statement obviously holds for Θ and is trivially preserved by Fr, Cut, and $CutFr$. Also, since any application of the H rule 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 $\Theta_{\leq r} \setminus \Theta t$.

   (b) Going downwards, the only point at which (e, u, f) can vanish is the Cut inference, and $Cut(d) < r$ so if $rk(e) \geq r$ then it cannot be removed.

13

(c) Since $Fr(d), H(d) > r$ and $CutFr(d) = r$, all (e, u, t) added going upwards have rank at least r.

$\square$

Definition 15. *Let Θ and Σ be two sequents. Then Θ and Σ are* multiplicable *if $\Theta \cup \Sigma$ is a sequent. $\Theta * \Sigma$ is defined and equal to $\Theta \cup \Sigma$ iff Θ and Σ are multiplicable.*

In particular, since $(\Theta \cup \Sigma)_S$ must be a finite ϵ-substitution, if $e \in \operatorname{dom}(\Theta) \cap \operatorname{dom}(\Sigma)$ then there are u and i such that $(e, u, i) \in \Theta$ and $(e, u, i) \in \Sigma$.

Lemma 15. *Let d be an r^+ derivation of Θ and let $\Sigma \leq r$ be a correct sequent such that Θ and Σ are multiplicable and $(\Sigma f)_{\geq r} \subseteq \Theta$, $\Sigma t \geq r$. Then there exists an r^+ derivation $d * \Sigma$ of $\Theta * \Sigma$ with $h(d * \Sigma) \leq h(d)$.*

Proof. By induction on the last inference of d.

1. Cut_e: Then $rk(e) < r$ and either:

   (a) There is some u such that $(e, u, t) \in \Sigma$: not possible, since $\Sigma t \geq r$

   (b) There is some u such that $(e, u, f) \in \Sigma$: then $((e, u, f), \Theta) * \Sigma = \Theta * \Sigma$, so the cut is pruned and $d * \Sigma = d_u * \Sigma$

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

2. $CutFr_e$: Then $rk(e) = r$ and either:

   (a) There is u such that $(e, u, t) \in \Sigma$: then $((e, u, t), \Theta) * \Sigma = \Theta * \Sigma$, and therefore the $CutFr$ is pruned and $d * \Sigma = d_u * \Sigma$

   (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 every u, $d_u * \Sigma$ is properly defined, and $d * \Sigma$ just applies the $CutFr_e$ rule to $(d_u * \Sigma)_{u \in \mathbb{N} \cup \{?\}}$.

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

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

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

5. Axioms: If Θ is an axiom then $\Theta * \Sigma$ is an axiom of the same kind.

$\square$

Definition 16. $(\Theta_0, \ldots, \Theta_n)$ *is an* r-path *(for Θ_n) if it is a path in some* r-*deduction of* $\Theta_0 = \emptyset$

Lemma 16. *Let* $(\Theta_0, \ldots, \Theta_n)$ *be an* $r + 1$ *path for* $\Theta = \Theta_n$. *Let* $\Sigma \leq r$ *be a correct sequent such that* $\Theta_{\leq r} \subseteq \Sigma$.

Then Θ *and* Σ *are multiplicable and there is a deduction of* Σ *from* $\Theta * \Sigma$ *consisting only of* Fr *and* H *inferences of ranks* $> r$.

Proof. By induction on n. Suppose $n > 0$ and let $\Theta' = \Theta_{n-1}$. Then $\Theta * \Sigma$ is defined since, by Lemma 14(2a) $\Theta_{\leq r} \subseteq \Sigma = \Sigma_{\leq r}$. If some $(e, u, i) \in \Theta'$ is not present in Θ then the step from Θ' to Θ is an application of the H-rule, and since the path $(\Theta_1, \ldots, \Theta_n)$ is taken from an $r + 1$ derivation, it follows that $\Theta'_{\leq r} \subseteq \Theta_{\leq r}$. Then by I.H. there is a deduction of Σ from $\Theta' * \Sigma$. Consider the inference from Θ to Θ'.

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

2. *CutFr*: impossible in an $r + 1$ derivation

3. *Fr*: $\Theta = (e, ?, t), \Theta'$. Then $rk(e) > r$, so $\Theta * \Sigma = (e, ?, t), \Theta' * \Sigma$ and is obtained by Fr_e.

4. *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 H_e applies $\Theta' * \Sigma$.

$\square$

Lemma 17. *Let d be a derivation ending with a cut C of rank r such that the immediate sub-derivations of d are r^+ derivations and there exists an $r + 1$ path for the end-sequent Θ of d. Then for each $u \in \mathbb{N} \cup \{?\}$ there is an r^+ derivation d'_u of $(e, u, t), \Theta$ such that*

$$h(d'_u) \leq (\omega + h(d))||u||_{\prec} + \omega + h(d_u) < (\omega + h(d))(||u||_{\prec} + 1)$$

where
$$||u||_{\prec} = \begin{cases} \text{The order type of } u \text{ under } \prec & \text{If } u \in \mathbb{N} \\ \alpha & \text{If } u = ? \end{cases}$$

Proof. By transfinite induction on u. Let e be the main term and d_u the immediate sub-derivations.

Suppose we have already constructed d'_v for all $v \prec u$ and $v \in \mathbb{N}$, or for all $v \in \mathbb{N}$ if $u = ?$. First replace all occurrences of (e, u, f) in sequents in d_u with (e, u, t). All inference rules will still hold, but an AxH of the form $(e, u, f), \Upsilon$ ceases to be an axiom. We instead derive this from

$$\Sigma = (e, v, t), \Upsilon_{\leq r}$$

where v is the H-value of the AxH. Note that when $u \neq ?$, $v \prec u$ by Lemma 8.

15

We have $\Theta_{\leq r} \setminus \Theta t \subseteq \Sigma$ by Lemma 14(2a), and since $\Theta t > r$ and Σ is correct, by Lemma 16 there is a deduction of Σ from $\Theta * \Sigma$ consisting only of Fr and H inferences of ranks $> r$.

Since $v \prec u$, we have a derivation d'_v of $(e, v, t), \Theta$. Since $(e, v, t) \in \Sigma$, we have $((e, v, t), \Theta) * \Sigma = \Theta * \Sigma$. By Lemma 14(2)(b), we have $(((e, u, f), \Upsilon)f)_{\leq r} \subseteq (e, u, f), \Theta$ and by Lemma 14(2)(c) we have $((e, u, f), \Theta)t \geq r \Rightarrow ((e, u, f), \Upsilon)t \geq r$. Since Θ has an $r+1$ path, $\Theta t > r$, so $(\Sigma f)_{\geq r} \subseteq (e, u, f), \Theta$ and $\Sigma t \geq r$. Then by Lemma 15 there is an r^+ derivation $d'_v * \Sigma$ of $\Theta * \Sigma$.

$$
\begin{array}{l}
d'_v * \Sigma \quad \vdots \\
\qquad\quad \Theta * \Sigma \\[1em]
\qquad\quad \vdots \\[0.5em]
H_{e,v} \dfrac{(e, v, t), \Upsilon}{(e, u, t), \Upsilon} \\
\quad d_u \\[1em]
\qquad\quad \vdots \\[0.5em]
\qquad (e, u, t), \Theta
\end{array}
$$

The resulting derivation consists of an r^+ derivation, $d'_v * \Sigma$, a finite sequence of Fr and H inferences of rank $> r$, and another r^+ derivation, so the entire derivation is r^+. If $u \in \mathbb{N}$, we have

$$
h(d'_u) \leq \left(\sup_{v \prec u} h(d'_v) \right) + \omega + h(d_u)
$$

By the inductive hypothesis, this implies that

$$
h(d'_u) \leq (\omega + h(d)) ||u||_\prec + \omega + h(d_u) < (\omega + h(d))(||u||_\prec + 1)
$$

$\square$

Lemma 18. *Let d be a derivation ending with a cut C of rank r such that the immediate sub-derivations of d are r^+ derivations and there exists an $r+1$ path for the end-sequent Θ of d. Then there is an r^+ derivation d' of Θ with $h(d') \leq (\omega + h(d))(\alpha + 1) + 1$.*

Proof. Let e be the main term and d_u the immediate sub-derivations. By Lemma 17, for each $u \in \mathbb{N} \cup \{?\}$ there is an r^+ derivation d'_u of $(e, u, t), \Theta$, so construct an r^+ derivation d' of Θ by applying the $CutFr$ rule to $\{d'_u\}_{u \in \mathbb{N} \cup \{?\}}$.

$$
h(d') \leq (\omega + h(d))(\alpha + 1) + 1
$$

$\square$

Lemma 19. *If d is an $r+1$ derivation of Θ and Θ has an $r+1$ path then there is an r^+ derivation d' of Θ with $h(d') \leq (\alpha + \omega)^{h(d)+1}$.*

Proof. By induction on $h(d)$. If $h(d) = 0$, then d is an axiom and the result is trivial. Assume $h(d) > 0$.

16

Let $\mathcal{I}$ be the last inference of d, let $(d_u)_{u \in I}$ be the family of immediate sub-derivations of d, and let Θ_u be the end-sequent of d_u. Then by I.H., for each $u \in I$ there is an r^+ derivation d'_u of Θ_u with $h(d'_u) \leq (\alpha + \omega)^{h(d_u)+1} \leq (\alpha + \omega)^{h(d)}$. Let d^+ be the derivation of Θ by $\mathcal{I}$ from $(d'_u)_{u \in I}$.

If $\mathcal{I}$ is not a cut of rank r then d^+ is an r^+ derivation of Θ. If $\mathcal{I}$ is a cut of rank r, we apply Lemma 18 to d^+ to get an r^+ derivation d' such that

$$h(d') \leq (\omega + h(d^+))(\alpha + 1) + 1 \leq (\alpha + \omega)^{h(d)}(\alpha + 1) + 1 < (\alpha + \omega)^{h(d)+1}$$

$\square$

Lemma 20. *If d is an r-derivation of $\emptyset$ with $r < \omega$ and $h(d), \alpha < \beta$ for some ϵ-number β then there is a 0-derivation d' of $\emptyset$ with $h(d') < \beta$.*

Proof. By induction on r. If $r = 0$ then we are done; otherwise, let $r = s + 1$. Then by Lemma 19, there is an s^+ derivation d^* of $\emptyset$ with $h(d^*) \leq (\alpha + \omega)^{h(d)+1}$. By Label 12, there is an s derivation of $\emptyset$. $\square$

References

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

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

[Arai, 2003] Arai, T. (2003). Epsilon substitution method for $ID_1(\Pi_1^0 \vee \Sigma_1^0)$. *Annals of Pure and Applied Logic*, 121(2):163–208.

[Hilbert, 1970] Hilbert, D. (1970). *Grundlagen der Mathematik*, volume 2. Springer.

[Mints, 1994] Mints, G. (1994). Gentzen-type systems and Hilbert's epsilon substitution method. In *Logic, Method and Philosophy of Science*, volume IX, pages 91–122, Amsterdam. Elsevier.

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

[Mints and Tupailo, 1999] Mints, G. and Tupailo, S. (1999). Epsilon-substitution for the ramified language and Δ_1^1-comprehension rule. In et. al., A. C., editor, *Logic and Foundations of Mathematics*, pages 107–130. Kluwer Academic Publishers.

[Mints et al., 1996] Mints, G., Tupailo, S., and Buchholz, W. (1996). Epsilon substitution method for elementary analysis. *Archive for Mathematical Logic*, 35:103–130.