Epsilon Substitution Method for Elementary Analysis

Grigori Mints
Department of Philosophy, Stanford University, Stanford, CA 94305
Sergei Tupailo
Institute of Cybernetics, Estonian Academy of Sciences, Tallinn, EE0026, Estonia
Wilfried Buchholz
Mathematisches Institut der Universität München,
Theresienstrasse 39, D-80333 München, Bundesrepublik Deutschland

July 25, 1997

Abstract

We formulate epsilon substitution method for elementary analysis EA (second order arithmetic with comprehension for arithmetical formulas with predicate parameters). Two proofs of its termination are presented. One uses embedding into ramified system of level one and cutelimination for this system. The second proof uses non-effective continuity argument.

Introduction

The epsilon substitution method is based on the language introduced by Hilbert [7] (and used later by Bourbaki [3]). The main non-boolean construction of this language is $\epsilon x F[x]$, read as "an x satisfying the condition $F[x]$". In number-theoretic contexts it is often interpreted as least x satisfying $F[x]$". Existential and universal quantifiers become explicitly definable by

$$\exists x F[x] = F[\epsilon x F[x]]; \qquad \forall x F[x] = F[\epsilon x \neg F[x]] \tag{1}$$

The main axioms of the corresponding formalism are *critical formulas*

$$F[t] \to F[\epsilon x F[x]] \tag{2}$$

Hilbert's approach (Ansatz) to transforming arbitrary (non-finitistic) number-theoretic proofs into finitistic (combinatorial) proofs by means of the substitution method is described in [7]. Cf. also the short and lively presentation by Hermann Weyl in [19]. The approach is as follows.

Take all critical formulas (2) occurring in a given proof P. There is only a finite number of them, so one always deals with a finite system E of critical formulas. Consider any substitution S of numerals for constant epsilon-terms. If all critical formulas (2) are true under S, it is called a solving substitution for the system E. Hilbert proposed a specific plan for finding a solving substitution by a series of successive approximations, described below. If it succeeds and if the last formula of the proof P, i.e. the formula proved by P, is a constant combinatorial identity such as $1 + 2 + \ldots + 10 = 55$, replacing all free variables by any numerals and then each epsilon-term t by $S(t)$ immediately yields a variable-free (finitistic, combinatorial) proof of the same identity. Moreover, it was noted by Ackermann and stressed later by Kreisel that the same device allows one to extract the numerical content of existential proofs, i.e. proofs of existential formulas $\exists x F[x]$ with combinatorial (free variable) $F[x]$. Indeed, $\exists x F[x]$ is translated as $F[\epsilon x F[x]]$. If S is a solving substitution for the proof P of such a formula, and $N = S(\epsilon x F[x])$ then $S(P)$ is a proof of $F[N]$. So N is a numerical realization of the existential quantifier in $\exists x F[x]$.

Hilbert's suggestion for finding a solving substitution by a successive approximation method is based on the following idea of generating substitutions of numerals for closed epsilon-terms. The initial approximation S_0 is identically 0. At every stage only a finite number of epsilon-terms are assigned non-zero values. If approximations $S_0, \ldots, S_i$ are already generated, and S_i is not yet a solving substitution, then S_{i+1} is found as follows. Fix appropriate ordering of the critical formulas (2) and take the first formula in this ordering which is false under S_i, i.e. for which

$$S_i(F[t]) = true, S_i(F[\epsilon x F]) = false$$

This means that the value $S_i(\epsilon x F)$ is incorrect: this value does not satisfy F, while $F[S_i(t)]$ is true under S_i. Then the value of $\epsilon x F$ is corrected by putting

$$S_{i+1}(\epsilon x F) = (\textit{the least } N \leq t)\ (S_i(F[n]) = true)$$

The problem stated by Hilbert was to prove termination of the sequence $S_0, S_1, S_2, \ldots$ after a finite number of steps for any system of critical formulas (2). After von Neumann's [15] attack on this (see below) Hilbert [6] stated further problems:

find a proof of termination for pure number theory, for analysis, and for analysis with the axiom of choice, when each of these systems is suitably reformulated in the epsilon-calculus.

The first attempt by Ackermann [1] to prove termination for analysis was shown by von Neumann [15] to contain a serious defect connected with the treatment of equality. Von Neumann introduced a device allowing to avoid this defect, gave an exact formulation based on Hilbert's Ansatz and presented a termination proof for the case when the terms $\epsilon x F[x]$ in critical formulas involve only free-variable (combinatorial) formulas $F[x]$. This corresponds to number theory with the induction axiom for quantifier free formulas.

After Gentzen's proof [4] of cut-elimination and consequently of consistency for arithmetic (with respect to closed equations) revealed the role of the ordinal ϵ_0, Ackermann [2] was able to find a final formulation and to give a termination proof for full first order arithmetic (pure number theory). His formulation for first order arithmetic, described also by Hilbert and Bernays [7], and used in all the subsequent research, consists of Hilbert's Ansatz modified in accordance with von Neumann's remarks plus the following stipulation: after the value of $\epsilon x F$ was corrected, the values of all terms of greater complexity (rank) are set to 0.

Ackermann's proof [2] is rather involved. On the other hand, there exists much easier non-effective termination proof of the substitution method for the first order arithmetic. It uses a very simple non-effective proof of the existence of solving substitution (cf. [18], section 5.1) and continuous dependence of a solution of a system of critical formulas from function parameters (cf. [18], end of the section 5.3). These ideas were elaborated by Mints in [10] for the first order arithmetic and in [11] for the theory of hereditarily finite sets into a non-effective termination proof which uses continuity in Baire topology in the space of number-theoretic functions. A possibility to simplify a system of critical formulas by substituting computable values of subordinate ϵ-matrices and a need for some restrictions on the functionals used was pointed out in [8], p. 259 (ii), 260.
The existence of a solving substitution for the case of analysis (second order arithmetic) is proved as simply as in the first order case, but nothing similar to the simple termination proof is known for the second order case. G. Kreisel [8, 9] further investigated the substitution method for the first order case and made it a basis for the no-counterexample interpretation, the first published functional interpretation producing non-trivial constructive results for non-constructive proofs. W. Tait [18] presented a finer analysis of the rate of termination in terms of relevant parameters based on his analysis [17] of recursion schemata.

The next step was made in [12] where a Gentzen-type system $PA\epsilon$ in the epsilon language for first order arithmetic was proposed. The axioms of this calculus depend on a particular system E of critical formulas. Normalization (cut-elimination) steps for this calculus were defined and the following statements proved.
Theorem 1. The sequent expressing existence of a solving substitution for the system E of critical formulas is derivable in $PA\epsilon$ by a derivation of special form, called f-derivation.
Theorem 2. f-derivations can be normalized, i.e. the cut rule can be eliminated by a series of normalization steps.
Theorem 3. A cut-free derivation of the sequent in Theorem 1 encodes a finite sequence

$$S_0, S_1, \ldots, S_n \tag{3}$$

of successive substitutions (in the sense of Ackermann) terminating in a solving substitution for the original system E of critical formulas.

Hence the rate of convergence (i.e. the value of n in (3) as a function of the system E) is determined by theorems 3 and 2, and is measured by ϵ_0 . The derivation mentioned in Theorem 1 in fact formalizes the non-effective proof of the existence of a solving substitution mentioned earlier. In this paper we apply the proof strategy expressed by Theorems 1-3 to the subsystem of the second order arithmetic based on the comprehension schema for arithmetical formulas with predicate parameters. This system is called EA in [16], and is obviously predicative: it can be easily embedded into ramified analysis of the level 1, and its proof-theoretic ordinal is proved in [16] to be ϵ_{ϵ_0}. This opens the way to the extension of our approach to ramified systems where bound predicate variables have levels and range over predicate terms of lower levels.

Present approach to the definition of the epsilon substitution for the second order epsilon-terms has two important new features. The values are closed arithmetical abstracts, and the passage to the new substitution is done according to the rank of some intermediate object (canonical form, cf. Section 5). These definitions have been found in the framework of the termination proof via cut-elimination which is presented in the section 6. After this it became possible to give a shorter non-effective proof which is presented in the section 7.

This research was supported by Center for the Study of Language and Information, Stanford University, and by the NSF grant DMS-9206976, and its results with proofs appeared in the preprint [13] by the first two authors. Chapters 2 and 4 of the preprint were written by S. Tupailo, Chapter 3 was due to G. Mints. W. Buchholz wrote a new version of the core parts of the paper with several technical improvements and simplifications.

1 The System $EA\epsilon$

The formulation here is very close to [7].

1.1 The language $L2\epsilon$

Let us describe in detail the language of ϵ-calculus.

Basic Symbols
0-variables (denoted by $x, y, z, \ldots$);
1-variables (denoted by $X, Y, Z, \ldots$);
the 0-ary function constant 0 (zero), and the unary function constant S (successor);
predicate constants for n-ary computable predicates ($n \geq 1$) including = (equality), *add* (addition) and *prod* (multiplication);
the propositional logical connectives $\neg$, $\wedge$, $\rightarrow$;
the epsilon symbol ϵ and the application symbol App.

Definition 1 (Terms and formulas)
1. Each ι-variable is a ι-term ($\iota = 0, 1$).
2. The constant 0 is a 0-term.
3. If t is a 0-term, then $\mathsf{S}t$ is a 0-term.
4. If t_1,...,t_n are 0-terms and p is an n-ary predicate constant, then $pt_1 \ldots t_n$ is a formula.
5. If t is a 0-term and P is a 1-term then $\mathsf{App}Pt$ is a formula.
6. If A, B are formulas then $\neg A$, $\wedge AB$, $\rightarrow AB$ are formulas.
7. If F is a formula and ξ is a ι-variable then $\epsilon\xi F$ is a ι-term (ι=0,1).

To increase readability, we sometimes use infix notation for binary logical connectives, insert parentheses and use standard abbreviations like $A \vee B = \neg A \rightarrow B$.

Definition 2 (λ-terms)
If G is a formula and z is a 0-variable occurring free in G, then $\lambda z G$ is a λ-term.

ι-terms ($\iota = 0, 1$), formulas and λ-terms are called *expressions.*
Terms of the form $\epsilon\xi F$ are called *ϵ-terms.*
The 0-terms $0, \mathsf{S}0, \mathsf{SS}0, \ldots$ are called *numerals.*

Var denotes the set of all variables, $\mathbb{N}$ denotes the set of all numerals.
We define $0^0 := 0, \quad 0^1 := \lambda z(z = 0)$.
For each term u we set
$$\iota(u) := \begin{cases} 0 & \text{if } u \text{ is a 0-term} \\ 1 & \text{if } u \text{ is a 1-term} \end{cases}$$
.

Note 1. Instead of usual computable functions we use computable predicates (their graphs) to simplify technical details. Computable functions can be introduced in a familiar way via their graphs [14].
Note 2. Lambda-symbols are not allowed to occur inside other expressions. The result of their substitution into other expressions is understood via λ-conversion (cf. Definitions 6,7 below), which allows to treat them together with 1-terms.

Syntactic variables:
e, u, v, w for expressions,
ξ, η for variables,
p for any predicate constant and the symbols $\mathsf{S}, \neg, \wedge, \rightarrow, \mathsf{App}$,
s, t for 0-terms,
n for a numeral $\mathsf{S}^n 0$,
P for 1-terms,
T for 1-terms and λ-terms,
A, B, F, G for formulas.
The set $FV(e)$ of free variables of an expression e is defined in the standard way.

Definition 3
$FV(0) := \emptyset$ and $FV(\xi) = \{\xi\}$ for each variable ξ,
$FV(pe_1 \dots e_n) = FV(e_1) \cup \dots \cup FV(e_n)$,
$FV(\epsilon\xi F) := FV(F) \setminus \{\xi\}$,
$FV(\lambda z G) := FV(G) \setminus \{z\}$.
e is called closed *iff $FV(e) = \emptyset$*

Substitution
We identify expressions which are equivalent modulo renaming of bound variables. If u is a $\iota(\xi)$-term then $e[\xi/u]$ denotes the result of substituting u for each free occurrence of ξ in e, where bound variables in e are renamed if necessary. If ξ is known from the context we write $e[u]$ for $e[\xi/u]$.

The next series of definitions is needed mainly in the section 3 below to determine computations ¿from inside replacing closed ϵ-terms by their values. The depth $d(e)$ is a measure of nesting of closed ϵ-subterms, taking into account that 1-ϵ-subterms can be substituted by λ-terms, and this can increase nesting by an arbitrary finite amount. It uses the natural sum function $\#$ on ordinal numbers, which is a commutative associative analog of the ordinal sum $+$:

$$\omega^\alpha \# \omega^\beta = \omega^{max(\alpha,\beta)} + \omega^{min(\alpha,\beta)}$$

Definition 4
1. $d(\xi) := d(0) := 0$
2. $d(pe_1 \dots e_n) := d(e_1)\# \dots \#d(e_n)$, *for* $p \neq \mathsf{App}$
3. $d(\mathsf{App}Pt) := \omega \cdot (d(P)\#d(t))$
4. $d(\epsilon\xi F) := \begin{cases} 1 + d(F) & \textit{if } \epsilon\xi F \textit{ is closed} \\ d(F) & \textit{otherwise} \end{cases}$
5. $d(\lambda z G) := d(G)$

Note 1. $d(e) < \omega^\omega$.
Note 2. $d(e) = 0$ iff e does not contain closed ϵ-subterms.

Definition 5
1. An ϵ-term $e = \epsilon\xi F$ is called canonical *if it is closed and $d(e) = 1$ (i.e. $d(F) = 0$).*
2. A λ-term $\lambda z G$ is called canonical *if $FV(G) = \{z\}$ and $d(G) = 0$.*

3. e is called simple *if $d(e)=0$ and e is closed.*
4. e is called arithmetical *if e contains no 1-variable.*
5. TRUE (FALSE) *denotes the set of all true (false) simple formulas. [Note that a simple formula contains no variables and is constructed from computable atomic formulas by boolean connectives. Every simple term is a numeral].*
6. $\mathbb{B}_0 := \mathbb{N}$*;* $\mathbb{B}_1$*:= the set of all canonical arithmetical λ-terms.*

The objects to be immediately evaluated are canonical ϵ-terms, and the values of ι-terms will be elements of B_ι.
To define substitution $e[Y/T]$ for expressions e and λ-terms T, we extend the 'operation' App:

Definition 6
$\mathsf{App}(T,t) := \begin{cases} G[t] & \text{if } T=\lambda zG \\ \mathsf{App}Tt & \text{otherwise} \end{cases}$

For $p \neq \mathsf{App}$ we set $p(e_1,\ldots,e_n) := pe_1\ldots e_n$.

Definition 7
1. If $e \in \mathrm{Var}\cup\{0\}$ *then* $e[Y/T] := \begin{cases} T & \text{if } e=Y \\ e & \text{otherwise} \end{cases}$
2. $(pe_1\ldots e_n)[Y/T] := p(e_1[Y/t],\ldots,e_n[Y/t])$
3. $(\epsilon\xi F)[Y/T] := \begin{cases} \epsilon\xi F & \text{if } \xi=Y \\ \epsilon\xi F[Y/T] & \text{otherwise} \end{cases}$ *where it is assumed without loss of generality that* $\xi \notin FV(T)$.
4. $(\lambda zF)[Y/T] := \lambda zF[Y/T]$ *assuming that* $z \notin FV(T)$.

The next definition concerns a restriction on the second order critical formulas which arise in the process of translating EA-derivations (Section 2 below) into $EA\epsilon$-derivations.

Definition 8 (Regular λ-terms)
A λ-term λzG is called regular *if it can be obtained from a λ-term without bound 1-variables by substitution of ϵ-terms for free variables and $z \in FV(G)$.*

1.2 Axioms and inference rules of $EA\epsilon$

The language of $EA\epsilon$ is $L2\epsilon$.
The only **inference rule** of $EA\epsilon$ is *modus ponens*: $\frac{F \quad F\to G}{G}$.
Axioms of $EA\epsilon$
Propositional axioms: all propositional tautologies of the language $L2\epsilon$,
All substitution instances of defining axioms for the predicate constants, including the predicates of addition and multiplication:
$add(s,0,s)$ and $add(s,t,r) \to add(s,\mathsf{S}t,\mathsf{S}r)$
$prod(s,0,0)$ and $prod(s,t,r)\wedge add(r,s,r_1) \to prod(s,\mathsf{S}t,r_1)$
Equality axioms: $t=t$ and $s=t\to(F[s]\to F[t])$,
Peano axioms for S: $\mathsf{S}t\neq 0$ and $\mathsf{S}s=\mathsf{S}t\to s=t$,
Minimality axioms: $\epsilon xF[x]=\mathsf{S}t\to\neg F[t]$,
Critical formulas:
$F[t]\to F[\epsilon xF[x]]$,
$s\neq 0\to F[\epsilon xF[x]]$ with $F:=(s=\mathsf{S}x)$,
$F[T]\to F[\epsilon XF[X]]$ with T being a 1-term or a regular λ-term.
Comment. Critical formulas of the second kind are not present in [7]. They are needed here to interprete Robinson's axiom $s\neq 0\to\exists x(s=\mathsf{S}x)$.
This concludes the description of $EA\epsilon$. Note that the formulas

$$s=t\to(\epsilon xF[x,s]=\epsilon xF[x,t])$$

are consequences of the equality axioms of $EA\epsilon$.
Note: $EA\epsilon$ is closed under the *substitution rule*: if F is derivable and u is a $\iota(\xi)$-term, then $F[\xi/u]$ is derivable. Indeed, all axioms and inference rules of $EA\epsilon$ are closed under substitution.

2 Embedding EA into $EA\epsilon$

EA is the usual system of *elementary analysis* (i.e. second order arithmetic with the axiom scheme of arithmetical comprehension). The language $L2$ of EA is obtained from $L2\epsilon$ by dropping the epsilon symbol ϵ and adding the existential quantifier $\exists$. In the definition of terms and formulas case 7 is replaced by:
7. If F is a formula and ξ is a ι-variable ($\iota = 0, 1$) then $\exists\xi F$ is a formula.

Note that the only 1-terms of $L2$ are 1-variables.

The universal quantifier is defined via $\exists$: $\forall\xi F := \neg\exists\xi\neg F$.

Axioms of EA

(1) Propositional axioms: all propositional tautologies of the language $L2$,
(2) Defining axioms for the predicate constants, including the predicates of addition and multiplication:
$add(x, 0, x)$ and $add(x, y, z) \to add(x, \mathsf{S}y, \mathsf{S}z)$
$prod(x, 0, 0)$ and $prod(x, y, z) \wedge add(z, x, z_1) \to prod(x, \mathsf{S}y, z_1)$
(3) Equality axioms: $x = x$ and $x = y \to (F[x] \to F[y])$,
(4) Peano axioms for S: $\mathsf{S}x \neq 0$ and $\mathsf{S}x = \mathsf{S}y \to x = y$,
(5) Induction axioms: $F[0] \to \forall x(F[x] \to F[\mathsf{S}x]) \to \forall xF[x]$,
(6) First order existential axioms: $F[t] \to \exists xF[x]$,
(7) Second order existential axioms: $F[T] \to \exists XF[X]$,
where T is a 1-variable or a λ-term containing no bound 1-variables.

Inference rules of EA

$$\text{modus ponens}\quad \frac{F \quad F \to G}{G}$$

$$\exists \to \quad \frac{F[\eta] \to G}{\exists\xi F[\xi] \to G}$$

where ξ, η are both 0-variables or 1-variables and the standard proviso is satisfied: the *eigenvariable* η does not occur free in the conclusion $\exists\xi F[\xi] \to G$.

Definition 9
For any formula F of EA define inductively an $L2\epsilon$-formula F^*:
$F^* := F$ for atomic F
$(\neg F)^* := \neg F^*$ and $(F \odot G)^* := (F^* \odot G^*)$ for $\odot = \wedge, \to$
$(\exists\xi F[\xi])^* := F^*[\epsilon\xi F[\xi]^*]$
$(\lambda zG)^* := \lambda zG^*$

Theorem 2.1
a) If $EA \vdash F$ then $EA\epsilon \vdash F^*$.
b) If $EA \vdash F$ and F is closed then there exists an $EA\epsilon$-derivation of F^* in which all formulas are closed.

Proof: a) Let EA' be the set of all $L2$-formulas F with $EA\epsilon \vdash F^*$.
EA' is closed under modus ponens and the existential rule $\exists\to$. (Note that $(\exists\xi F[\xi] \to G)^* = F^*[\epsilon\xi F^*[\xi]] \to G^*$, and therefore closure of EA' under $\exists\to$ follows from the fact that $EA\epsilon$ is closed under substitution.)
The *-translation of any EA-axiom of kind (1)–(4) is an $EA\epsilon$-axiom of the same kind.
The *-translation of any existential axiom of EA is a critical formula.
Below we show that the *-translation of every induction axiom of EA is derivable in $EA\epsilon$. This will finish the proof of a).
Observe that b) is an immediate consequence of a): Given an $EA\epsilon$-derivation d of a closed formula F one simply replaces every free ι-variable in d by 0^ι.

Now we consider the *-translation A of an induction axiom.
Obviously A is of the form $G[0] \to \overline{G}[u] \to G[e]$ with $\overline{G}[x] := G[x] \to G[\mathsf{S}x]$, $u := \epsilon x\neg\overline{G}[x]$, $e := \epsilon x\neg G[x]$.
Let $e^- := \epsilon y(e = \mathsf{S}y)$.
We derive $G[0] \to \overline{G}[u] \to G[e]$ by case distinction $e = 0$, $e \neq 0$. The first case is settled by the equality axiom $0 = e \to G[0] \to G[e]$. For the second case observe that $e \neq 0 \to \overline{G}[u] \to G[e]$ follows by propositional logic ¿from the minimality axiom $e = \mathsf{S}e^- \to \neg\neg G[e^-]$, the critical formulas $e \neq 0 \to e = \mathsf{S}e^-$, $\neg\overline{G}[e^-] \to \neg\overline{G}[u]$ and the equality axiom $e = \mathsf{S}e^- \to G[\mathsf{S}e^-] \to G[e]$. □

3 Computations with the ϵ-Substitutions

Definition 10
An ϵ-substitution *is a function* S *such that*
$dom(S)$ *(domain of* S*) is a set of canonical* ϵ*-terms,*
if $e \in dom(S)$ *then* $S(e) \in \mathbb{B}^{\iota(e)} \cup \{?\}$.
An ϵ*-substitution* S *is called* total *if* $dom(S)$ *is the set of all canonical* ϵ*-terms.*
$\overline{S} := S \cup \{(e, ?) : e$ *is a canonical* ϵ*-term* $\notin dom(S)\}$ *is called the* standard extension of S.

Comment. We consider a function as a set of ordered pairs. So an ϵ-substitution is a set of pairs (e, u) where e is a canonical ϵ-term and $u \in \mathbb{B}^{\iota(e)} \cup \{?\}$. Hilbert's finitist position allowed only for finite ϵ-substitutions S. All (canonical) ϵ-terms $\epsilon\xi F$ not mentioned in S explicitly, have default value $0^{\iota(\xi)}$. In fact the possible value ? for a term indicates exactly this default value. The ϵ-substitution process is defined entirely in terms of finite substitutions: the standard extension $\overline{S}$ of S is used only to simplify notation for computations with default values. Essentially infinite substitutions appear only in section 7 below.

3.1 Computation Steps

Let S be an arbitrary ϵ-substitution.

Definition 11 (Inductive definition of $e \hookrightarrow^1_S e'$)
1.1. If $(e, u) \in S$ *and* $u \neq ?$ *then* $e \hookrightarrow^1_S u$
1.2. If $(e, ?) \in S$ *then* $e \hookrightarrow^1_S 0^{\iota(e)}$
2. If $1 \le i \le n$, $e_i \hookrightarrow^1_S e'_i$ *and* e'_i *is not a* λ*-term then* $pe_1 \dots e_n \hookrightarrow^1_S pe_1 \dots e_{i-1}e'_ie_{i+1} \dots e_n$
3. If $(P, \lambda zG) \in S$ *then* $\mathsf{App}Pt \hookrightarrow^1_S G[t]$
4. If $F \hookrightarrow^1_S F'$ *then* $\epsilon\xi F \hookrightarrow^1_S \epsilon\xi F'$
5. If $G \hookrightarrow^1_S G'$ *then* $\lambda zG \hookrightarrow^1_S \lambda zG'$.

Comment. Since S is defined only for canonical terms which do not contain any proper closed ϵ-subterms, computations proceed from inside, and since canonical terms are closed, no term containing variables is immediately computed according to S: only its closed subterms can be replaced.

Note. If $e \hookrightarrow^1_S \lambda zG$ and e is not a λ-term then either $(e, \lambda zG) \in S$ or $(e, ?) \in S$ and $\lambda zG = 0^1$.

Definition 12
1. e *is called* S-reducible *if there exists an* e' *with* $e \hookrightarrow^1_S e'$. *Otherwise* e *is called* S-irreducible *or* in S-normal form.
2. $\hookrightarrow_S$ *denotes the transitive and reflexive closure of* $\hookrightarrow^1_S$.

Note. e is S-reducible iff e contains canonical ϵ-subterms in the domain of S. In particular no default computations are allowed unless S contains pairs $(e, ?)$.

We are going to prove that $\hookrightarrow_S$ is well-founded, i.e. every sequence of computation steps is terminating, by showing that each step decreases the depth $d(e)$ (Definition 4).

Lemma 3.1 *If* $e \hookrightarrow^1_S e'$ *then* $FV(e') = FV(e)$.

Proof. Consider the cases in the definition of $\hookrightarrow^1_S$. If the computation is by 1.1 or 1.2, then $FV(e) = \emptyset = FV(e')$. In the cases 4,5 use the induction hypothesis. The same works in case 2 when $p \neq \mathsf{App}$ or $p = \mathsf{App}$ but the second argument is changed.
Let now $e = \mathsf{App}Pt$ and $e' = \mathsf{App}(P', t)$ with $P \hookrightarrow^1_S P'$ (recall the definition 6). Then P is an ϵ-term, since 1-variables are not reducible.
1. P is canonical. Then $P' = \lambda zG \in \mathbb{B}_1$, $FV(G) = \{z\}$ and $e' = G[t]$. Hence $FV(e') = FV(t) = FV(e)$.
2. P is not canonical. Then $e' = \mathsf{App}P't$ and $FV(e') = FV(P') \cup FV(t) \stackrel{IH}{=} FV(P) \cup FV(t) = FV(e)$. $\square$

Lemma 3.2 *For each arithmetical expression* e *one has:*
a) $d(e) < \omega$,
b) $d(e[y/t]) < (d(t) + 1) \cdot \omega$.

Proof. a) is obvious, since App does not occur in e and all clauses of the definition 4 except 3 add only finite amount.
b) Induction on e. Let $y \in FV(e)$, since otherwise $d(e[y/t]) = d(e) < \omega$. Set $\alpha := (d(t)+1)\cdot\omega$.
1. If $e = y$ then $d(e[y/t]) = d(t) < d(t) + 1 < \alpha$.
2. If $e = pe_1 \ldots e_n$ with $p \neq \mathsf{App}$ then $d(e[t]) = d(pe_1[t] \ldots e_n[t]) = d(e_1[t])\# \ldots \# d(e_n[t]) < \alpha$, since by the induction hypothesis (IH) $d(e_i[t]) < \alpha$ and α is closed under $\#$.
3. If $e = \epsilon x F$ then we can assume $x \notin FV(t)$ and $d(e[t]) = d(\epsilon x F[t]) \leq 1 + d(F[t]) \overset{IH}{<} 1 + \alpha = \alpha$.
4. If $e = \lambda z F$, apply IH. □

Lemma 3.3 *If $e \hookrightarrow^1_S e'$ then $d(e') < d(e)$.*

Proof by induction on the definition of $\hookrightarrow^1_S$.
1.1, 1.2. If e is a canonical ϵ-term then $d(e') = 0 < 1 = d(e)$.
2. If $e = pe_1 \ldots e_n,\ e' = pe_1 \ldots e'_i \ldots e_n$ and $e_i \hookrightarrow^1_S e'_i$, then:
2.1. if $p \neq \mathsf{App}$ then $d(e') = d(e_1)\# \ldots \# d(e'_i)\# \ldots \# d(e_n) \overset{IH}{<} d(e_1)\# \ldots \# d(e_n) = d(e)$
2.2. if $e = \mathsf{App}Pt$ and $e' = \mathsf{App}P't$, then $d(e') = \omega \cdot (d(P')\# d(t)) \overset{IH}{<} \omega \cdot (d(P)\# d(t)) = d(e)$.
2.3. if $e = \mathsf{App}Pt$ and $e' = \mathsf{App}Pt'$, one proceeds as in the case 2.2.
3. Let $e = \mathsf{App}Pt$ and $e' = G[t]$. Note that $\forall \alpha < \omega^\omega (\alpha \cdot \omega \leq \omega \cdot \alpha)$.
Then by Lemma 3.2 $d(e') < (d(t)+1)\cdot\omega \leq \omega \cdot (d(t)+1) \leq \omega \cdot (d(P)\# d(t)) = d(e)$.
4, 5. If $e = \epsilon\xi F, e' = \epsilon\xi F'$ or $e = \lambda z F, e' = \lambda z F'$ with $F \hookrightarrow^1_S F'$ then, since $FV(F) = FV(F')$, we have $d(e') = j + d(F') \overset{IH}{<} j + d(F) = d(e)$ for some $j = 0, 1$. □

We are going to prove that $\hookrightarrow^1_S$ is locally confluent.

Lemma 3.4 *If e is arithmetical and $t \hookrightarrow^1_S t'$ then $e[t] \hookrightarrow_S e[t']$.*

Proof is routine. □

Lemma 3.5 *If $e \hookrightarrow^1_S e'$ and $e \hookrightarrow^1_S e''$ then there is an expression u such that $e' \hookrightarrow_S u$ and $e'' \hookrightarrow_S u$.*

Proof by induction on e.
1. If $e \in dom(S)$ then $e' = e''$.
2. Let $e = pe_1 \ldots e_i \ldots e_n,\ e' = pe_1 \ldots e'_i \ldots e_n$ and $e_i \hookrightarrow^1_S e'_i$.
2.1. $e'' = pe_1 \ldots e''_i \ldots e_n$ with $e_i \hookrightarrow^1_S e''_i$. Then by the IH one has $e'_i \hookrightarrow_S u_i$ and $e''_i \hookrightarrow_S u_i$ for some u_i. Take $u := pe_1 \ldots u_i \ldots e_n$.
2.2.$e'' = pe_1 \ldots e'_j \ldots e_n$ with $e_j \hookrightarrow^1_S e'_j$ and $i \neq j$. Take $u := pe_1 \ldots e'_i \ldots e'_j \ldots e_n$.
3. $e = \mathsf{App}Pt,\ e' = G[t],\ e'' = \mathsf{App}Pt'$ with $P \hookrightarrow^1_S \lambda z G$ and $t \hookrightarrow^1_S t'$. Then $e'' \hookrightarrow_S G[t']$ and, by Lemma 3.4 $e' \hookrightarrow_S G[t']$.
The remaining cases are easy: apply I.H. □

Theorem 3.6 (Church-Rosser Property)
For each expression e there exists a unique S-irreducible expression e^* with $e \hookrightarrow_S e^*$.

Proof. By Lemma 3.3 computations terminate, and together with Lemma 3.5 this implies uniqueness. □

Definition 13 (Normal form $|e|_S$ for expression e)
The unique expression e^* in the previous theorem is called the S-normalform of e *and denoted by* $|e|_S$.

Definition 14 *An expression e is* S-computable *iff* $d(|e|_S) = 0$

Note.
1. e is S-computable iff $|e|_S = |e|_{\overline{S}}$.
2. If S is total then every expression is S-computable.

We prove next that $\hookrightarrow_S$ is preserved under substitution. The main problem will be with λ-terms.

Lemma 3.7
a) $e \hookrightarrow^1_S e'$ *implies* $e[\eta/u] \hookrightarrow_S e'[\eta/u]$.
b) $u \hookrightarrow^1_S u'$ *implies* $e[\eta/u] \hookrightarrow_S e[\eta/u']$.
c) $e[\eta/u] \hookrightarrow_S |e|_S[\eta/|u|_S]$.

Proof. We first prove a) and b) under the assumption that u is not a λ-term.
a) We use induction on the definition of $\hookrightarrow^1_S$.
The only non-trivial case is $e = \mathsf{App}Pt,\ e' = G[z/t],\ P \hookrightarrow^1_S \lambda zG$.
Then since P is closed and $FV(G) = \{z\}$, one has $e[u] = \mathsf{App}Pt[u],\ e'[u] = G[z/t[u]]$. Hence by the definition of $\hookrightarrow^1_S$ one has $e[u] \hookrightarrow^1_S G[z/t[u]] = e'[u]$.
b) We use induction on e assuming $\eta \in FV(e)$. Again the only non-trivial case is $e = \mathsf{App}Pt,\ P[u'] = \lambda zG$. Then $e[u] = \mathsf{App}P[u]t[u],\ e[u'] = G[z/t[u']]$ and by IH $P[u] \hookrightarrow_S \lambda zG,\ t[u] \hookrightarrow_S t[u']$. Since u is not a λ-term, we have $\lambda zG \in \mathbb{B}_1$. Hence there exists a canonical ϵ-term Q such that $P[u] \hookrightarrow_S Q \hookrightarrow^1_S \lambda zG$. Hence $e[u] \hookrightarrow_S \mathsf{App}Qt[u'] \hookrightarrow^1_S G[z/t[u']] = e[u']$.
Now we assume that u is a λ-term. We have to consider two additional cases.
a') $e = \mathsf{App}Pt$ and $P[u] = \lambda zG$. Then $P = \eta$ (otherwise P contains λ) and thus $u = \lambda zG$, $e' = \mathsf{App}Pt'$ with $t \hookrightarrow^1_S t'$. Hence $e[u] = G[z/t[u]]$, $e'[u] = G[z/t'[u]]$ and $t[u] \hookrightarrow_S t'[u]$ by IH. Since $t[u]$ is not a λ-term, by b) we obtain $e[u] \hookrightarrow_S e'[u]$.
b') $e = \mathsf{App}Pt$ and $P[u] = \lambda zG$. Then $P = \eta$ (otherwise P contains λ) and thus $u = \lambda zG, u' = \lambda zG'$ with $G \hookrightarrow_S G'$. Hence $e[u] = G[z/t[u]]$ and $e[u'] = G'[z/t[u']]$. By IH $t[u] \hookrightarrow_S t[u']$. Since $t[u]$ is not a λ-term, by a), b) we obtain $G[z/t[u]] \hookrightarrow_S G'[z/t[u]] \hookrightarrow_S G'[z/t[u']]$.
c) By induction on the number of computation steps, we obtain versions of a),b) with $\hookrightarrow^1_S$ replaced by $\hookrightarrow_S$. These imply c): $e[u] \hookrightarrow_S |e|_S[|u|] \hookrightarrow_S |e|_S[|u|_S]$. □

The next proposition clarifies the structure of normal forms for partial substitutions S, where some canonical ϵ-terms are irreducible. It is needed in section 6.3.

Lemma 3.8
a) If v is an S-irreducible subterm of e, $d(v) > 0$ and $e \hookrightarrow^1_S e'$ then v is also a subterm of e'.
b) If $e[\eta/u]$ is S-computable and $\eta \in FV(e)$ then u is S-computable.
c) If e is S-computable and $e \hookrightarrow_{\overline{S}} e'$ then $e \hookrightarrow_S e'$.

Proof. a) We use induction on the definition of $\hookrightarrow^1_S$. Since e is reducible, v has to be a proper subterm of e and thus e cannot be a canonical ϵ-term.
1. Let $e = pe_1 \dots e_i \dots e_n,\ e' = pe_1 \dots e'_i \dots e_n$ and $e_i \hookrightarrow^1_S e'_i$. Then the claim follows immediately from the IH.
2. $e = \mathsf{App}Pt,\ e' = G[z/t]$ with $P \hookrightarrow^1_S \lambda zG$. Then P is a canonical ϵ-term $\neq v$, and therefore v is not a subterm of P. Hence v is a subterm of t and thus of e' since z occurs free in G.
3. The cases $e = \epsilon\xi F, e' = \epsilon\xi F'$ or $e = \lambda zF, e' = \lambda zF'$with $F \hookrightarrow^1_S F'$ follow immediately from IH.
b) Assume for contradiction that $d(|u|_S) > 0$. Then $|u|_S$ and hence $|e|_S[|u|_S]$ contains an S-irreducible, closed ϵ-subterm v. By Lemma 3.7c and Theorem 3.6 we have $|e|_S[|u|_S] \hookrightarrow_S |e[u]|_S$, and therefore [by a)] v is a subterm of $|e[u]|_S$. This contradicts the S-computability of $e[u]$.
c) It suffices to prove the Lemma for one step reductions. Assume that e is S-computable and $e \hookrightarrow^1_{\overline{S}} e'$ holds. By b) every canonical ϵ-subterm v of e is S-computable and therefore in $dom(S)$. Hence $e \hookrightarrow^1_S e'$. □

Let us establish some properties of regular λ-terms (Definition 8) needed in section 5.

Lemma 3.9
a) If T is regular and $T \hookrightarrow^1_S T'$ then T' is regular.
b) If T is a closed regular λ-term with $d(T) = 0$ then $T \in \mathbb{B}_1$.
c) If T is a closed regular λ-term then $|T|_{\overline{S}} \in \mathbb{B}_1$.

Proof. We have $T = e[\eta_1/u_1, ..., \eta_n/u_n]$ where $u_1, ..., u_n$ are ϵ-terms and no bound 1-variable occurs in e.
a) Without loss of generality we may assume that either $T' = e[u'_1, u_2, ..., u_n]$ with $u_1 \hookrightarrow^1_S u'_1$ or $T' = e'[u_1, ..., u_n]$ with $e \hookrightarrow^1_S e'$. In both cases one easily sees that T' is again regular. (Only if $u'_1 \in \mathbb{B}_1$ a short argument is needed.)

b) For contradiction let us assume that $\eta_i \in FV(e)$. Then, since T is closed, u_i is a closed ϵ-term and thus $0 < d(u_i) \leq d(T)$. Hence none of $\eta_1, ..., \eta_n$ occurs free in e and we have $T = e$ which means that T is a closed λ-term containing no 1-variables. Hence $T \in \mathbb{B}_1$, since $d(T) = 0$.
c) follows from a) and b). □

4 The rank function

The rank will measure nesting of bound variables. We extend to the ϵ-language a definition known for Ramified Analysis [16].
Set $o(x) := 0$, $o(X) := \omega$, $o(\mathsf{App}) := \omega$, $o(p) := o(0) := 0$, for $p \neq \mathsf{App}$.
In the following σ denotes elements from $\mathrm{Var} \cup \{*\}$.

Definition 15 Definition of $\mathrm{rk}_\sigma(e)$
1. If $\sigma \notin \mathrm{FV}(e) \cup \{*\}$ *then* $\mathrm{rk}_\sigma(e) := 0$.
2. For $\sigma \in \mathrm{FV}(e) \cup \{*\}$ *we define:*

$$\mathrm{rk}_\sigma(e) := \begin{cases} o(e) & \text{if } e \in \mathrm{Var} \cup \{0\} \\ \max\{o(p), \mathrm{rk}_\sigma(e_1), ..., \mathrm{rk}_\sigma(e_n)\} & \text{if } e = pe_1...e_n \\ \max\{o(\xi), \mathrm{rk}_\sigma(F), \mathrm{rk}_\xi(F) + 1\} & \text{if } e = \epsilon\xi F \\ \mathrm{rk}_\sigma(G) & \text{if } e = \lambda z G \end{cases}.$$

Note 1. $\mathrm{rk}_\sigma(e) < \omega \cdot 2$.
Note 2. $\mathrm{rk}_\sigma(e)$ is a measure of nesting of bound variables in subterms of e containing free variable σ , and rk_* takes account of all ϵ-subterms. More precisely,

Lemma 4.1

$$\mathrm{rk}_*(e) = \sup\{o(\eta), \mathrm{rk}_\sigma(F)+1 \ : \eta, \epsilon\sigma F \ \textit{occur in } e\} \tag{4}$$

Proof by induction on e. Define

$$o(e) := \sup\{o(\eta) : \ \eta \text{ occurs in } e\}, \qquad \mathrm{rk}_1(e) := \sup\{\mathrm{rk}_\sigma(F)+1 \ : \epsilon\sigma F \text{ occurs in } e\}$$

Note that $o(pe_1...e_n) = \max\{o(e_1), ..., o(e_n)\}$.
1. $e \in \mathrm{Var} \cup \{0\}$. Then $\mathrm{rk}_*(e) = o(e)$, $\mathrm{rk}_1(e) = 0$
2. $e = pe_1...e_n$. Then
$\mathrm{rk}_*(e) \overset{IH}{=} \max\{o(p), \max_i \max\{o(e_i), \mathrm{rk}_1(e_i)\}\} =$
$= \max\{o(p), \max_i o(e_i), \max_i \mathrm{rk}_1(e_i)\} = \max\{o(e), \mathrm{rk}_1(e)\}$.
3. $e = \epsilon\xi F$. Then
$\mathrm{rk}_*(e) = \max\{o(\xi), \mathrm{rk}_*(F), \mathrm{rk}_\xi(F)+1\} \overset{IH}{=} \max\{o(\xi), o(F), \mathrm{rk}_1(F), \mathrm{rk}_\xi(F)+1\} = \max\{o(e), \mathrm{rk}_1(e)\}$
4.$e = \lambda z G$. Then $\mathrm{rk}_*(e) = \mathrm{rk}_*(G) \overset{IH}{=} \max\{o(G), \mathrm{rk}_1(G)\} = \max\{o(e), \mathrm{rk}_1(e)\}$. □
Next Lemmas establish properties of rank.

Lemma 4.2
a) e arithmetical $\implies \mathrm{rk}_\sigma(e) < \omega$
b) $X \in \mathrm{FV}(e) \implies \omega \leq \mathrm{rk}_X(e)$.
c) $\omega \leq \mathrm{rk}_*(P)$*, for each 1-term P.*

Proof. Easy induction on e, P.

Lemma 4.3
If $d(e) = 0$ and e is not a λ-term then $\mathrm{rk}_*(e) \leq \sup\{\mathrm{rk}_\sigma(e) : \sigma \in \mathrm{FV}(e)\}$.

Proof:
1. $e = 0$: Then $\mathrm{rk}_*(e) = 0$. $e = \xi \in \mathrm{Var}$: Then $\mathrm{rk}_*(e) = o(\xi) = \mathrm{rk}_\xi(e)$.
2. $e = \epsilon\xi F$: Then, since $d(e) = 0$, we have $d(F) = 0$ and $\mathrm{FV}(e) \neq \emptyset$. Therefore

$\max\{o(\xi), \mathrm{rk}_\xi(F){+}1\} \leq \sup\{\mathrm{rk}_\sigma(e) : \sigma \in \mathrm{FV}(e)\}$ and, by I.H., $\mathrm{rk}_*(F) \leq \sup\{\mathrm{rk}_\sigma(F) : \sigma \in \mathrm{FV}(F)\}$.
Hence $\mathrm{rk}_*(e) = \max\{o(\xi), \mathrm{rk}_*(F), \mathrm{rk}_\xi(F){+}1\} \leq \sup\{\mathrm{rk}_\sigma(e) : \sigma \in \mathrm{FV}(e)\}$.
3. $e = pe_1...e_n$. Since $\omega \leq \mathrm{rk}_*(P)$ for each 1-term P, we have
$\mathrm{rk}_*(e) = \max\{\mathrm{rk}_*(e_1), ..., \mathrm{rk}_*(e_n)\} \overset{\mathrm{IH}}{\leq} \sup\{\mathrm{rk}_\sigma(e_i) : \sigma \in \mathrm{FV}(e_i) \,\&\, 1{\leq}i{\leq}n\} \leq \sup\{\mathrm{rk}_\sigma(e) : \sigma \in \mathrm{FV}(e)\}$. □

Lemma 4.4
If $\epsilon\xi F$ *is canonical then* $\mathrm{rk}_*(F) \leq \mathrm{rk}_\xi(F)$ *and* $\mathrm{rk}_*(\epsilon\xi F) = \max\{o(\xi), \mathrm{rk}_\xi(F){+}1\}$.

Proof:
Since $\epsilon\xi F$ is canonical, we have $d(F) = 0$. By 4.3 this yields
$\mathrm{rk}_*(F) \leq \sup\{\mathrm{rk}_\sigma(F) : \sigma \in \mathrm{FV}(F)\}$ and thus (since $\mathrm{FV}(F) \subseteq \{\xi\}$) $\mathrm{rk}_*(F) \leq \mathrm{rk}_\xi(F)$.
Hence $\mathrm{rk}_*(\epsilon\xi F) = \max\{o(\xi), \mathrm{rk}_*(F), \mathrm{rk}_\xi(F) + 1\} = \max\{o(\xi), \mathrm{rk}_\xi(F) + 1\}$. □

The next two lemmas show that the rank does not increase during computation.

Lemma 4.5
$\mathrm{rk}_\sigma(e[y/t]) < \max\{\omega, \mathrm{rk}_\sigma(t){+}1\}$, *if* e *is arithmetical.*

Proof: Let $y \in \mathrm{FV}(e)$ and $\sigma \in \mathrm{FV}(e[t]) \cup \{*\}$. (Otherwise $\mathrm{rk}_\sigma(e[t]) = \mathrm{rk}_\sigma(e) < \omega$ or $\mathrm{rk}_\sigma(e[t]) = 0$.)
1. $e = y$: $\mathrm{rk}_\sigma(e[t]) = \mathrm{rk}_\sigma(t)$.
2. $e = pe_1...e_n$: Then $\mathrm{rk}_\sigma(e[t]) = \max\{o(p), \mathrm{rk}_\sigma(e_1[t]), ..., \mathrm{rk}_\sigma(e_n[t])\}$, and the assertion follows by I.H..
3. $e = \epsilon xF$. Then $o(x) = 0$, $\mathrm{rk}_\sigma(e[t]) = \max\{\mathrm{rk}_\sigma(F[t]), \mathrm{rk}_x(F[t]) + 1\}$ and, by I.H.,
$\mathrm{rk}_\sigma(F[t]) < \max\{\omega, \mathrm{rk}_\sigma(t) + 1\}$ and $\mathrm{rk}_x(F[t]) < \max\{\omega, \mathrm{rk}_x(t) + 1\} = \omega$.
The last equation holds by clause 1 in the definition of rank, since $x \notin \mathrm{FV}(t) \cup \{*\}$.
4. $e = \lambda zG$. The assertion follows by I.H. □

Lemma 4.6
If $e \hookrightarrow^1_S e'$ *then* $\mathrm{rk}_\sigma(e') \leq \mathrm{rk}_\sigma(e)$.

Proof by induction on the definition of $\hookrightarrow^1_S$*:*
Let $\sigma \in \mathrm{FV}(e') \cup \{*\}$. (Otherwise $\mathrm{rk}_\sigma(e') = 0$.)
1.1. If e is a canonical 0-ϵ-term then $\mathrm{rk}_\sigma(e') = 0$, since $e' \in \mathbb{N}$.
1.2. If e is a canonical 1-ϵ-term then $\mathrm{rk}_\sigma(e') < \omega \leq \mathrm{rk}_\sigma(e)$, since e' is arithmetical and $\sigma = *$.
2. $e = pe_1...e_n$ or $e = \lambda yF$:
2.1. $e = \mathsf{App}Pt$ and $e' = G[z/t]$ with $P \hookrightarrow^1_S \lambda zG$. Then
$\mathrm{rk}_\sigma(e') \overset{4.5}{\leq} \max\{\omega, \mathrm{rk}_\sigma(t)\} \leq \mathrm{rk}_\sigma(e) = \max\{\omega, \mathrm{rk}_\sigma(P), \mathrm{rk}_\sigma(t)\}$.
2.2. otherwise: immediate from I.H.
3. $e = \epsilon\xi F$ and $e' = \epsilon\xi F'$ with $F \hookrightarrow^1_S F'$:
$\mathrm{rk}_\sigma(e') = \max\{o(\xi), \mathrm{rk}_\sigma(F'), \mathrm{rk}_\xi(F'){+}1\} \overset{\mathrm{IH}}{\leq} \max\{o(\xi), \mathrm{rk}_\sigma(F), \mathrm{rk}_\xi(F){+}1\} = \mathrm{rk}_\sigma(e)$. □
Substitution of a variable by an appropriate canonical value also does not increase the rank.

Lemma 4.7
$\mathrm{rk}_\sigma(e[\eta/u]) \leq \mathrm{rk}_\sigma(e)$, *for each* $u \in \mathbb{B}_{\iota(\eta)}$.

Proof by induction on e*:* Let $e' := e[\eta/u]$ and assume that $\eta \in \mathrm{FV}(e)$ and $\sigma \in \mathrm{FV}(e') \cup \{*\}$.
Note that $\mathrm{FV}(e') = \mathrm{FV}(e) \setminus \{\eta\}$.
1. $e = \eta$: Then $\sigma = *$ and $\mathrm{rk}_*(e') = 0$ (if $\iota(\eta) = 0$) or $\mathrm{rk}_*(e') < \omega = \mathrm{rk}_*(e)$ (if $\iota(\eta) = 1$).
2. $e = \mathsf{App}\eta t$ and $u = \lambda zG$: Then $e' = G[z/t[u]]$ and thus
$\mathrm{rk}_\sigma(e') \overset{4.5}{\leq} \max\{\omega, \mathrm{rk}_\sigma(t[u])\} \overset{\mathrm{IH}}{\leq} \max\{\omega, \mathrm{rk}_\sigma(t)\} = \mathrm{rk}_\sigma(e)$.
3. $e = \epsilon\xi F$. Then
$\mathrm{rk}_\sigma(e') = \mathrm{rk}_\sigma(\epsilon\xi F') = \max\{o(\xi), \mathrm{rk}_\sigma(F'), \mathrm{rk}_\xi(F') + 1\} \overset{\mathrm{IH}}{\leq} \max\{o(\xi), \mathrm{rk}_\sigma(F), \mathrm{rk}_\xi(F) + 1\} = \mathrm{rk}_\sigma(e)$.
4. In all other cases the assertion follows immediately from IH. □

The next statement shows that our definition of rank is suitable: the rank decreases when the 'body' of a canonical ϵ-term is substituted by a canonical value.

Lemma 4.8
If $\epsilon\xi F$ is canonical then $\mathrm{rk}_*(F[u]) < \mathrm{rk}_*(\epsilon\xi F)$, *for each* $u \in \mathbb{B}_{\iota(\xi)}$.

Proof: $\mathrm{rk}_*(F[u]) \overset{4.7}{\leq} \mathrm{rk}_*(F) \overset{4.4}{\leq} \mathrm{rk}_\xi(F) < \mathrm{rk}_*(\epsilon\xi F)$. □

Definition 16 $\mathrm{rk}(e) := \mathrm{rk}_*(e)$ *is called the* rank *of* e.

Definition 17 (Truncation to a given rank)
For each ϵ-substitution S and $r \in On$ we set $S_{\leq r} := \{(e,u) \in S : \mathrm{rk}(e) \leq r\}$.
Analogously we define $S_{\geq r}$, $S_{<r}$, $S_{>r}$.

Lemma 4.9
If S, S' are ϵ-substitutions with $S_{\leq r} = S'_{\leq r}$ then $|e|_S = |e|_{S'}$ holds for all expressions e of rank $\leq r$.

Proof:
Since all subterms of an expression e have ranks $\leq \mathrm{rk}(e)$, we have:
$\mathrm{rk}(e) \leq r \Rightarrow \forall e'(\, e \hookrightarrow_S e' \Leftrightarrow e \hookrightarrow_{S'} e'\,)$.
Together with Lemma 4.6 this yields the assertion by induction on $d(e)$. □

5 The H-process

Let us recall that critical formulas are formulas of three types:
$F[t] \to F[\epsilon x F[x]]$,
$s \neq 0 \to F[\epsilon x F[x]]$ with $F := (s = \mathsf{S}x)$,
$F[T] \to F[\epsilon X F[X]]$ with T a 1-ϵ-term or regular lambda-term.
We assume that $Cr_0, ..., Cr_N$ (with $N \in \mathbb{N}$) **is a fixed sequence of closed critical formulas**.

In this section we define a successive approximation process for finding a solution of this system according to Hilbert's approach. It is useful to recall here a Comment from Section 3. The H-process will be arranged in such a way that all non-default values of ϵ-terms are *correct*: $S(\epsilon x F[x])$ is the least n satisfying $F[n]$, and $S(\epsilon X F[X])$ is an arithmetical λ-term T satisfying $F[T]$.

Definition 18
$F[\![x/n]\!] := F[x/n] \wedge \neg F[x/n{-}1] \wedge \ldots \wedge \neg F[x/0]$,
$F[\![X/T]\!] := F[X/T]$.

Definition 19
Let S be an ϵ-substitution:
$e \hookrightarrow_S \mathsf{TRUE}\,(\mathsf{FALSE}) \;:\Longleftrightarrow\; |e|_S \in \mathsf{TRUE}\,(\mathsf{FALSE})$.
$\mathcal{F}(S) := \{F[\![\xi/u]\!] : (\epsilon\xi F, u) \in S \,\&\, u \neq ?\,\}$
S is correct *iff $A \hookrightarrow_S \mathsf{TRUE}$ for all $A \in \mathcal{F}(S)$.*
S is solving *iff $Cr_I \hookrightarrow_S \mathsf{TRUE}$ for $I = 0, ..., N$. Otherwise S is* nonsolving.
$\overline{S} := S \cup \{(e, ?) : e$ *canonical ϵ-term* $\notin \mathrm{dom}(S)\}$ *is called the* standard extension of S, *cf. Section 3.*

The definition of a term whose value is to be corrected (H-term) to a new value (H-value) and of the ϵ-substitution $\mathrm{H}(S)$ to which S is reduced by this change of value, is given in terms of the standard extension $\overline{S}$. In fact the first term of minimal rank is chosen, and all values of higher rank are cancelled. Since ϵ-substitutions are defined only for canonical terms, all these operations are preceded by transforming arbitrary ϵ-term $\epsilon\xi F$ into its *canonical form* $\epsilon\xi|F|_S$.

Definition 20
Let S be an ϵ-substitution such that $\overline{S}$ is nonsolving. (Then $|Cr_I|_{\overline{S}} \in \mathsf{FALSE}$ for some $I \leq N$.)
Set $r_I := \mathrm{rk}(\epsilon\xi|F|_{\overline{S}})$, where $Cr_I = F_0 \to F[\epsilon\xi F]$.
$\mathrm{Cr}(S) := Cr_I$, *where $I \leq N$ is such that*
$$|Cr_I|_{\overline{S}} \in \mathsf{FALSE} \,\&\, \forall J \leq N[\,|Cr_J|_{\overline{S}} \in \mathsf{FALSE} \Rightarrow r_I < r_J \vee (r_I = r_J \wedge I \leq J)].$$

Let $\mathrm{Cr}(S) = F_0 \to F[\epsilon\xi F]$*:*
$\epsilon\xi|F|_{\overline{S}}$ *is called the* H-term of S.
The H-value v *of* S *is defined as follows*
a) *if* $\iota(\xi) = 1$ *and* $F_0 = F[T]$ *then* $v := |T|_{\overline{S}}$,
b) *if* $\iota(\xi) = 0$, $F_0 = (s \neq 0)$, *and* $F = (s = \mathsf{S}x)$ *then* $v := |s|_{\overline{S}} - 1$,
c) *if* $\iota(\xi) = 0$ *and* $F_0 = F[t]$ *then* $v :=$ *the unique* $n \in \mathbb{N}$ *with* $|F|_{\overline{S}}[\![n]\!] \hookrightarrow_{\overline{S}} \mathsf{TRUE}$.

Remark: If e is the H-term and v the H-value of S, then $v \in \mathbb{B}_{\iota(e)}$. (For $\iota(e) = 1$ cf. Lemma 3.9c.)
The next definition is central for the substitution method.

Definition 21 *If* $\overline{S}$ *is nonsolving then*
$\mathrm{H}(S) := (S \setminus \{(e, ?)\})_{\leq \mathrm{rk}(e)} \cup \{(e, v)\}$, *where* e *is the H-term and* v *the H-value of* S.

Let us prove that $\overline{\mathrm{H}(S)}$ is indeed a correct ϵ-substitution if $\overline{S}$ is correct and nonsolving.

Lemma 5.1
Let S *be an* ϵ*-substitution such that* $\overline{S}$ *is correct and nonsolving,*
and let e *be the H-term,* v *the H-value of* S. *Then the following holds:*
a) $(e, ?) \in \overline{S}$,
b) $|e|_{\overline{\mathrm{H}(S)}} = v \neq 0^{\iota(e)}$,
c) $\overline{\mathrm{H}(S)}$ *is correct.*

Proof:
Let $Cr(S) = F_0 \to F[\epsilon\xi F]$. Then $e = \epsilon\xi|F|_{\overline{S}}$.
a) *Assumption:* $(e, w) \in S$ and $w \neq ?$. Then, since $\overline{S}$ is correct, by Lemma 3.7c $F[\epsilon\xi F] \hookrightarrow_{\overline{S}} |F|_{\overline{S}}[w] \hookrightarrow_{\overline{S}}$ TRUE.
On the other hand, since $\mathrm{Cr}(S) \hookrightarrow_{\overline{S}} \mathsf{FALSE}$, we have $F[\epsilon\xi F] \hookrightarrow_{\overline{S}} \mathsf{FALSE}$. *Contradiction.*
b) By Lemma 3.7b we have $F[e] \hookrightarrow_{\overline{S}} \mathsf{FALSE}$ and $F[v] \hookrightarrow_{\overline{S}} \mathsf{TRUE}$. Hence, by 3.7b, $v \neq |e|_{\overline{S}} = 0^{\iota(e)}$.
The equation $|e|_{\overline{\mathrm{H}(S)}} = v$ holds, since $(e, v) \in \mathrm{H}(S)$.
c) Let $(\epsilon\xi A, w) \in \mathrm{H}(S)$ with $w \neq ?$. Then $(\epsilon\xi A, w) \in S$ or $\epsilon\xi A$ is e and w is v. One has $|A[\![w]\!]|_{\overline{S}} \in \mathsf{TRUE}$: in the first case since $\overline{S}$ is correct, and in the second case by definition of $\mathrm{H}(S)$ and Lemma 3.7c. By Lemma 4.8 one has $\mathrm{rk}(A[\![w]\!]) < \mathrm{rk}(e)$, and by Lemma 4.9 $|A[\![w]\!]|_{\overline{\mathrm{H}(S)}} = |A[\![w]\!]|_{\overline{S}}$. □

Definition 22
The H-process *(for* $Cr_0, ..., Cr_N$*) is defined as follows:*
$$S_0 := \emptyset, \quad S_{n+1} := \begin{cases} \mathrm{H}(S_n) & \text{if } \overline{S_n} \text{ is nonsolving} \\ \emptyset & \text{otherwise} \end{cases}.$$
The H-process terminates *iff there exists an* $n \in \mathbb{N}$ *such that* $\overline{S_n}$ *is solving.*

The next definition determines when the substitution S itself, (and not $\overline{S}$) contains sufficient information to compute all necessary values. An important instance is the provable incorrectness of S.

Definition 23
Let S *be an* ϵ*-substitution.*
S *is* computationally inconsistent (ci) *iff* $A \hookrightarrow_S \mathsf{FALSE}$ *for some* $A \in \mathcal{F}(S)$.
Otherwise S *is* computationally consistent (cc).
An expression e *is* S-computable *iff* $d(|e|_S) = 0$.
S *is* computing *iff all formulas* $A \in \mathcal{F}(S)$ *are* S*-computable.*
S *is* deciding *iff* S *is computing and the critical formulas* $Cr_0, ..., Cr_N$ *are* S*-computable.*

Remark
1. A closed formula A is S-computable iff $A \hookrightarrow_S \mathsf{TRUE}$ or $A \hookrightarrow_S \mathsf{FALSE}$.
2. e is S-computable iff $|e|_S = |e|_{\overline{S}}$.
3. If S is computing and cc then S is correct.

The next definition requires that all steps for computing $\mathrm{H}(S)$ are possible in terms of S itself.

Definition 24
Let S be an ϵ substitution. We say that the H-rule applies to S *iff*
(1) S is cc, deciding, nonsolving, and
(2) if $\mathrm{Cr}(S) = F[t] \to F[\epsilon x F]$ *then there exists an* $n \in \mathbb{N}$ *with* $|F|_S[\![n]\!] \hookrightarrow_S \mathsf{TRUE}$.

Lemma 5.2
Let S be cc, deciding, nonsolving, and $\mathrm{Cr}(S) = F[t] \to F[\epsilon x F]$*. Then*
a) $|t|_S \in \mathbb{N}$ *and* $F[|t|_S] \hookrightarrow_S \mathsf{TRUE}$.
b) If the formulas $F[i]$ ($i < |t|_S$) are S-computable then the H-rule applies to S.

Proof:
a) Since $\mathrm{Cr}(S) \hookrightarrow_{\overline{S}} \mathsf{FALSE}$ and $\mathrm{Cr}(S)$ is S-computable, we have $F[t] \hookrightarrow_S \mathsf{TRUE}$. Now the assertion follows by the Lemmas 3.7b and 3.8b.
b) Let $m := |t|_S$. Since $F[m] \hookrightarrow_S \mathsf{TRUE}$ and since $F[0], ..., F[m-1]$ are S-computable, for
$n := \min\{k : F[k] \hookrightarrow_S \mathsf{TRUE}\}$ we have $n \leq m$ and $|F|_S[\![n]\!] \hookrightarrow_S \mathsf{TRUE}$. □

Lemma 5.3
If the H-rule applies to S then $\mathrm{H}(S)$ *is correct.*

Proof: (cf. proof of 5.1)
Let $(\epsilon\xi A, w) \in \mathrm{H}(S)$ with $w \neq ?$. We have to prove $|A[\![w]\!]|_{\mathrm{H}(S)} \in \mathsf{TRUE}$. But $\mathrm{rk}(A[\![w]\!]) < \mathrm{rk}(e)$ (with e the H-term of S) and therefore $|A[\![w]\!]|_{\mathrm{H}(S)} = |A[\![w]\!]|_S$. If $(\epsilon\xi A, w) \in S$ then $|A[\![w]\!]|_S \in \mathsf{TRUE}$, since S is correct. Otherwise $\epsilon\xi A = e$ and $w =$ H-value of S. We treat the case where $\mathrm{Cr}(S) = F[T] \to F[\epsilon X F[X]]$ and leave the two other cases to the reader. Then $\epsilon\xi A = e = \epsilon\xi|F|_{\overline{S}}$ and $w = |T|_{\overline{S}}$. Since S is deciding, $F[T]$ is S-computable and therefore (by Lemmas 3.7c and 3.8c) $|F[T]|_S = |\,|F|_{\overline{S}}[\,|T|_{\overline{S}}\,]\,|_S = |A[\![w]\!]|_S$. Since $|\mathrm{Cr}(S)|_{\overline{S}} \in \mathsf{FALSE}$ and S is deciding, we have $|F[T]|_S = |F[T]|_{\overline{S}} \in \mathsf{TRUE}$. □

6 Termination proof via Cut-elimination

In this section we apply the general schema from [12] for reducing the termination problem for the H-process to the cut-elimination problem in some specially devised sequent calculus with a kind of ω-rule. The termination proof is constructive and uses induction up to ϵ_{ϵ_0}. In the sequel we use r as syntactic variable for ordinals ($< \epsilon_{\epsilon_0}$).

6.1 The system ϵEA

The original infinite derivation constructed for a given system E of critical formulas represents the tree of all finite ϵ-substitutions: empty substitution is put at the bottom, and all one-component extensions of S are placed above S (the rule Cut below). The branch of this tree is terminated (cf. Axioms below) when a provably incorrect substitution, solution, or a substitution admitting H-rule is reached. This 'bottom-up' view of the inference rules is helpful below.

The Cut-elimination process introduces steps of the H-process into this tree in the form of the rules Fr, H. The rule Fr temporarily 'freezes' the default value of e, which may be changed by the rule H into non-default values.

Definition 25
A sequent *is a function Θ such that* $\mathrm{dom}(\Theta)$ *is a set of canonical ϵ-terms, and* $\Theta(e) \in \mathbb{B}_{\iota(e)} \cup \{?, ?^\circ\}$ *for each* $e \in \mathrm{dom}(\Theta)$.

So a sequent is almost the same as an ϵ-substitution. A component $(e, ?^0)$ of a sequent S indicates that the default value for e is fixed and will not be changed in the extensions of S to be considered. By identifying $?$ and $?^\circ$ we associate with every sequent Θ an ϵ-substitution which is also denoted by Θ. A sequent Θ is called *correct (cc, deciding, computing)* if the associated ϵ-substitution has the respective property. An analoguous convention is followed with respect to notations like $|e|_\Theta$, $\hookrightarrow_\Theta$, etc.

Abbreviation:
$(e, u), \Theta := \{(e, u)\} \cup \Theta$, if $e \notin \mathrm{dom}(\Theta)$.
$\Theta \leq r \;:\iff\; \forall (e, u) \in \Theta (\mathrm{rk}(e) \leq r) \quad (\iff \Theta = \Theta_{\leq r})$. In the same way we define "$\Theta \geq r$".

Definition 26
$\Theta f := \{(e, u) \in \Theta : u = ?^\circ\}$, $\Theta t := \{(e, u) \in \Theta : u = ?\}$*: fixed and temporary part of a sequent.*

The system ϵEA
Rules of inference:

$$(\mathsf{Cut}) \quad \frac{(e, ?^\circ), \Theta \quad \ldots (e, u), \Theta \ldots (u \in \mathbb{B}_{\iota(e)})}{\Theta} \qquad (\mathsf{CutFr}) \quad \frac{(e, ?), \Theta \quad \ldots (e, u), \Theta \ldots (u \in \mathbb{B}_{\iota(e)})}{\Theta}$$

$$(\mathsf{Fr}) \quad \frac{(e, ?), \Theta}{\Theta} \qquad (\mathsf{H}) \quad \frac{(e, v), \Theta_{\leq \mathrm{rk}(e)}}{(e, ?), \Theta},$$

if the H-rule applies to $(e, ?), \Theta$, and
e is the H-term, v the H-value of $(e, ?), \Theta$.

Axioms:
(AxF) Θ, if Θ is ci
(AxS) Θ, if Θ is cc, deciding, and solving
(AxH) $(e, ?^\circ), \Theta$, if the H-rule applies to $(e, ?), \Theta$ and e is the H-term of $(e, ?), \Theta$.

In the above rules and axioms e always denotes a canonical ϵ-term not in $\mathrm{dom}(\Theta)$.
We call e the *main term* of the respective inference.

Definition 27 *By a* deduction in ϵEA *we mean a deduction (i.e. wellfounded tree) according to the rules of inference of ϵEA from axioms of ϵEA and additional sequents. By a* derivation in ϵEA *we mean a deduction in ϵEA from axioms of ϵEA only.*

By $\mathrm{h}(d)$ we denote the *height* of the deduction d, i.e. $\mathrm{h}(d) := \sup\{\mathrm{h}(d_u) + 1 : u \in I\}$ where $(d_u)_{u \in I}$ is the family of immediate subdeductions of d.
If $\mathcal{I}$ is an inference then $\mathrm{rk}(\mathcal{I})$ denotes the rank of its main term.
If d is a deduction, and X is one of the symbols Cut, CutFr, Fr, H, and $\bowtie$ is one of the symbols $<, \leq, >, \geq, =$ then
$\mathsf{X}(d) \bowtie r \;:\iff\; \mathrm{rk}(\mathcal{I}) \bowtie r$ for every X-inference $\mathcal{I}$ in d.
Hence "$\mathsf{Cut}(d) < r$" means that all cuts in d have rank $< r$, and "$\mathsf{X}(d) < 0$" means that there are no X-inferences in d.

6.2 Cut-elimination

Cuts will be eliminated in the usual way beginning with the maximal rank r. Eliminated cuts will be replaced by CutFr and H with the same main term, i.e. with the same rank. More precisely, a cut will be replaced by CutFr and then moved (permuted) up the derivation until one encounters AxS with a main term e traceable to the main formula $(e, ?)$ of CutFr. Then the AxS is replaced by the rule H, and the derivation of the corresponding right premise of the cut is placed over the rule H. After all cuts of rank r are eliminated, these CutFr will be pruned to Fr. So finally cuts of rank r will be replaced by Fr of rank r. This motivates the following

Definition 28 *Let d be a deduction.*
*d is an r-*deduction *iff* $\mathsf{Cut}(d) < r$ & $\mathsf{CutFr}(d) < 0$ & $\mathsf{Fr}(d) \geq r$ & $\mathsf{H}(d) \geq r$.
*d is an r^+-*deduction *iff* $\mathsf{Cut}(d) < r$ & $\mathsf{CutFr}(d) = r$ & $\mathsf{Fr}(d) > r$ & $\mathsf{H}(d) \geq r$.

Lemma 6.1
Every r^+-derivation d of Θ can be transformed into an r-derivation d' of Θ with $\mathrm{h}(d') \leq \mathrm{h}(d)$ by pruning each CutFr *to* Fr.

Proof is obvious. □

Lemma 6.2 (Properties of fixed and temporary parts of a sequent)
a) If Θ is a sequent in an $r{+}1$-deduction of $\emptyset$ then $\Theta t > r$ and $\Theta f \le r$.
b) If Σ is a sequent in an r^+-deduction of a sequent Θ then
(1) $\Theta_{\le r} \setminus \Theta t \subseteq \Sigma$,
(2) $(\Sigma f)_{\ge r} \subseteq \Theta$,
(3) $\Theta t \ge r \Rightarrow \Sigma t \ge r$.

Proof.
a) is proved by the bottom-up induction on the given deduction. This statement is obviously true for $\emptyset$, and is inherited from conclusion to the premise of a rule: temporary part Θt is increased only by Fr which has rank $\ge r{+}1 > r$, and Θf is increased by Cut which has rank $< r+1$ (i.e. $\le r$).
b)(1) The only nontrivial case is (H). Let $\Theta = (e, ?), \Upsilon$ and $\Sigma = (e, v), \Upsilon_{\le \mathrm{rk}(e)}$. Then $\Theta_{\le r} \setminus \Theta t \subseteq \Upsilon_{\le r} \subseteq \Sigma$, since $r \le \mathrm{rk}(e)$.
(2) Going from Σ down to Θ the only points where some $(e, ?^\circ)$ could vanish are Cut-inferences. But each Cut in an r^+-derivation has rank $< r$.
(3) holds, since $\mathsf{Fr}(d) > r$ and $\mathsf{CutFr}(d) = r$, so that formulas added to Θt are of the rank $\ge r$. □

Applicability of the standard cut-elimination transformations in our case is seriously restricted by non-admissibility (in general) of the weakening rule: adding components to a sequent can fail to produce a sequent.

Definition 29 *Two sequents Θ, Σ are* multiplicable *if $\Theta \cup \Sigma$ is a function. In this case we write $\Theta * \Sigma$ for $\Theta \cup \Sigma$, and say that $\Theta * \Sigma$ is* defined.

Lemma 6.3
Let d be an r^+-derivation of Θ. Let $\Sigma \le r$ be a correct sequent such that $\Theta * \Sigma$ is defined and $(\Sigma f)_{\ge r} \subseteq \Theta$, $\Sigma t \ge r$. Then there exists an r^+-derivation $d * \Sigma$ of $\Theta * \Sigma$ with $\mathrm{h}(d * \Sigma) \le \mathrm{h}(d)$.

Proof by induction on d:
We distinguish cases according to the last inference of d.
1. Cut with main term e: Then $\mathrm{rk}(e) < r$.
(a) $e \notin \mathrm{dom}(\Sigma)$: $d * \Sigma$ is obtained from $(d_u * \Sigma)_{u \in \mathbb{B}_{\iota(e)} \cup \{?^\circ\}}$ by the same Cut.
(b) $(e, ?) \in \Sigma$: This cannot happen, since $\Sigma t \ge r$ and $\mathrm{rk}(e) < r$.
(c) $(e, u) \in \Sigma$ with $u \in \mathbb{B}_{\iota(e)} \cup \{?^\circ\}$:
Then $((e, u), \Theta) * \Sigma = \Theta * \Sigma$, and therefore the Cut is pruned: the derivation $d * \Sigma := d_u * \Sigma$ has endsequent $\Theta * \Sigma$.
2. CutFr with main term e. Then $\mathrm{rk}(e) = r$ and $(e, ?^\circ) \notin \Theta$. Hence $(e, ?^\circ) \notin \Sigma$, since $(\Sigma f)_{\ge r} \subseteq \Theta$.
(a) $e \notin \mathrm{dom}(\Sigma)$: As (a) above.
(b) $(e, u) \in \Sigma$ with $u \in \mathbb{B}_{\iota(e)} \cup \{?\}$: as (c) above.
3. Fr with main term e: Then $((e, ?), \Theta) * \Sigma$ is defined, since $\Sigma \le r < \mathrm{rk}(e)$.
4. H with main term e: Then $\mathrm{rk}(e) \ge r$, and $\Theta = (e, ?), \Upsilon$ is derived from $\Theta' := (e, v), \Upsilon_{\le \mathrm{rk}(e)}$.
¿From $\Sigma \le r \le \mathrm{rk}(e)$ it follows that $\Sigma' := \Sigma \setminus \{(e, ...)\}$ is still correct, and that $(\Sigma' f)_{\ge r} \subseteq \Theta'$.
¿From $e \notin \mathrm{dom}(\Sigma')$ and $\Theta' \subseteq (e, v), \Upsilon$ it follows that $\Theta' * \Sigma'$ is defined.
Since $\Sigma' \le r \le \mathrm{rk}(e)$, we also get $\Theta' * \Sigma' = (e, v), \Upsilon_{\le \mathrm{rk}(e)} * \Sigma' = (e, v), (\Upsilon * \Sigma')_{\le \mathrm{rk}(e)}$ and $\Theta * \Sigma = (e, ?), \Upsilon * \Sigma'$.
So $d * \Sigma$ is obtained from $d_0 * \Sigma'$ by an H-inference of the same kind.
5. Axioms: If Θ is an axiom then $\Theta * \Sigma$ is an axiom of the same kind. □

We continue to investigate admissibility of weakening .

Definition 30 $(\Theta_0, ..., \Theta_n)$ *is an* r-path (for Θ_n) *if it is a path in some r-deduction of $\Theta_0 = \emptyset$.*

Lemma 6.4
Let $(\Theta_0, ..., \Theta_n)$ be an $r{+}1$-path for $\Theta := \Theta_n$. Let $\Sigma \le r$ be a correct sequent such that $\Theta_{\le r} \subseteq \Sigma$.
Then $\Theta * \Sigma$ is defined, and there exists a deduction of Σ from $\Theta * \Sigma$ consisting only of Fr- and H-inferences of ranks $> r$.

Proof by induction on n: Let $n > 0$ and $\Theta' := \Theta_{n-1}$.
$\Theta * \Sigma$ is defined, since $\Theta_{\leq r} \subseteq \Sigma = \Sigma_{\leq r}$. Note that $\Theta'_{\leq r} \subseteq \Theta_{\leq r}$, and therefore by I.H. there is a deduction of Σ from $\Theta' * \Sigma$. We now show that either $\Theta' * \Sigma = \Theta * \Sigma$ or $\Theta' * \Sigma$ derives from $\Theta * \Sigma$ by a Fr- or H-inference of rank $> r$. For this we distinguish cases according to the topmost inference in $(\Theta_0, ..., \Theta_n)$, i.e. the inference from Θ to Θ'.
1. Cut : We have $\Theta = (e,u), \Theta'$ and $\mathrm{rk}(e) \leq r$. Then the Cut is pruned. Indeed, $(e,u) \in \Sigma$ and thus $\Theta * \Sigma = \Theta' * \Sigma$.
2. CutFr : This cannot happen.
3. Fr with main term e, and $\Theta = (e,?), \Theta'$: Then $\mathrm{rk}(e) > r$ and $\Theta * \Sigma = (e,?), \Theta' * \Sigma$.
4. H with main term e, and $\Theta = (e,v), \Upsilon_{\leq \mathrm{rk}(e)},\ \Theta' = (e,?), \Upsilon$:
Then $\Theta' * \Sigma$ is derived from $\Theta * \Sigma$ by H. Indeed $\mathrm{rk}(e) > r,\ \Theta * \Sigma = (e,v), (\Upsilon * \Sigma)_{\leq \mathrm{rk}(e)}$, and $\Theta' * \Sigma = (e,?), \Upsilon * \Sigma$. The H-rule applies to $\Theta' * \Sigma$ (cf. Definition 24), since the H-rule applies to Θ', and Σ is correct. $\square$

Applicability of the cut-reduction transformation is restricted more or less to subderivations of 'good' derivations of an empty sequent $\emptyset$.

Lemma 6.5
Let d be a derivation ending with a cut $\mathcal{C}$ of rank r such that the immediate subderivations of d are r^+-derivations, and there exists an $(r{+}1)$-path for the endsequent Θ of d.
Then there exists an r^+-derivation d' of Θ with $\mathrm{h}(d') \leq \mathrm{h}(d) + \omega + \mathrm{h}(d)$.

Proof. Let e be the main term of $\mathcal{C}$, and d_u the immediate subderivations of d.
We transform d as follows:
(1) The cut $\mathcal{C}$ is turned into an inference CutFr by changing every sequent $(e,?^\circ), \Theta'$ of $d_{?^\circ}$ into $(e,?), \Theta'$. The only rules that are damaged by this transformation are axioms AxH of the form $(e,?^\circ), \Upsilon$ which now become $(e,?), \Upsilon$. At each of these points we extend the deduction by the corresponding H-inference, obtaining thereby the new top sequent $(e,v), \Upsilon_{\leq r}$.
(2) We consider now one such top sequent $\Sigma := (e,v), \Upsilon_{\leq r}$.
By Lemma 5.3 $\Sigma = \Sigma_{\leq r}$ is correct, and by 6.2a,b(1) we get $\Theta_{\leq r} = \Theta_{\leq r} \setminus \Theta t \subseteq \Sigma$.
(3) By Lemma 6.4 $\Theta * \Sigma$ is defined, and there exists a deduction of Σ from $\Theta * \Sigma$ consisting only of Fr- and H-inferences of ranks $> r$.
(4) To derive top sequents $\Theta * \Sigma$ in (3) we consider now the subderivation d_v of $\Theta' := (e,v), \Theta$ in the original derivation d.
$\Theta' * \Sigma$ is defined, since $\Theta * \Sigma$ is defined and $(e,v) \in \Sigma$. By Lemma 6.2b(2),(3) (applied to $d_{?^\circ}$) we have $((e,?^\circ), \Upsilon) f_{\geq r} \subseteq (e,?^\circ), \Theta$ and $(\, ((e,?^\circ), \Theta) t \geq r \Rightarrow ((e,?^\circ), \Upsilon) t \geq r\,)$. By 6.2a (applied to the $(r{+}1)$-path for Θ) we have $\Theta t > r$. Hence $(\Sigma f)_{\geq r} \subseteq \Theta'$ and $\Sigma t \geq r$. Now by Lemma 6.3 there is an r^+-derivation $d_v * \Sigma$ of $\Theta' * \Sigma$, i.e. of $\Theta * \Sigma$.
(5) The structure of d' is the following: to some tops of the deduction $d'_{?^\circ}$ of $(e,?), \Theta$ (which is an r^+-deduction) one-branch deductions from Lemmma 6.4 are added, and the tops of the latter deductions are r^+-derived by Lemma 6.3. Hence the entire derivation is an r^+-derivation. $\square$

Now cut reduction is iterated in the standard way.

Lemma 6.6
If d is an $r{+}1$-derivation of Θ, and Θ has an $r{+}1$-path then there exists an r^+-derivation d' of Θ with $\mathrm{h}(d') \leq \omega^{\mathrm{h}(d)+1}$.

Proof by induction on $\mathrm{h}(d)$:
If $h(d) = 0$, i.e. d consists of an axiom, the assertion is obvious. Assume $h(d) > 0$.
Let $\mathcal{I}$ be the last inference of d. Let $(d_u)_{u \in I}$ be the family of immediate subderivations of d, and Θ_u the endsequent of d_u. Then by I.H. for each $u \in I$ we have an r^+-derivation d'_u of Θ_u with $\mathrm{h}(d'_u) \leq \omega^{\mathrm{h}(d_u)+1} \leq \omega^{\mathrm{h}(d)}$. Let d^+ be the derivation of Θ which is obtained from the family $(d'_u)_{u \in I}$ by $\mathcal{I}$.
Obviously $\mathrm{h}(d^+) \leq \omega^{\mathrm{h}(d)} + 1$.
1. $\mathcal{I}$ is not a cut of rank r: Then d^+ is an r^+-derivation of Θ.
2. $\mathcal{I}$ is a cut of rank r: In that case we apply Lemma 6.5 to d^+, and obtain an r^+-derivation d' of Θ with $\mathrm{h}(d') \leq \mathrm{h}(d^+) + \omega + \mathrm{h}(d^+) < \omega^{\mathrm{h}(d)+1}$. $\square$

Lemma 6.7
If d is an ω-derivation of Θ, and Θ has an ω-path f in which all cuts are of rank $\leq r$ (with $r < \omega$) then there exists an r^+-derivation d' of Θ with $\mathrm{h}(d') \leq \epsilon_{\mathrm{h}(d)+1}$.

Proof by induction on $\mathrm{h}(d)$*:*
1. Suppose that d ends in a cut of rank $r+n$. Let d_u be the ω-subderivations of its premises. By I.H. each d_u can be transformed into an $(r+n)^+$-derivation d'_u. By replacement of d_u by d'_u for all u the derivation d is turned into a derivation c with $\mathrm{h}(c) = \sup_u(\mathrm{h}(d'_u)+1) \leq \sup_u(\epsilon_{\mathrm{h}(d_u)+1}+1) \leq \epsilon_{\mathrm{h}(d)}+1$.
To the derivation c we apply Lemma 6.5 and obtain an $(r+n)^+$-derivation c' of Θ with $\mathrm{h}(c') \leq \mathrm{h}(c)+\omega+\mathrm{h}(c) < \epsilon_{\mathrm{h}(d)+1}$.
Now n applications of Lemmata 6.1,6.6 yield the claim.
Note that for each $i \in \mathbb{N}$ f is an $(r+i+1)$-path, and $(\alpha < \epsilon_{\mathrm{h}(d)+1} \Rightarrow \omega^{\alpha+1} < \epsilon_{\mathrm{h}(d)+1})$.
2. Suppose that d ends in some other inference $\mathcal{I}$. Again by I.H. each immediate subderivation d_u transforms into an r^+-derivation d'_u. By replacement of d_u by d'_u for all u the derivation d is turned into an r^+-derivation d' with $\mathrm{h}(d') = \sup_u(\mathrm{h}(d'_u)+1) \leq \sup_u(\epsilon_{\mathrm{h}(d_u)+1}+1) \leq \epsilon_{\mathrm{h}(d)+1}$. $\square$

Lemma 6.8 (Cut-elimination)
Let d be an r-derivation of $\emptyset$ with $r < \omega+\omega$ and $\mathrm{h}(d) < \epsilon_0$.
Then there exists a 0-derivation d' of $\emptyset$ with $\mathrm{h}(d') < \epsilon_{\epsilon_0}$.

Proof: w.l.o.g. $r = \omega+n$. Then n applications of 6.6,6.1 yield an ω-derivation d° of $\emptyset$ with $\mathrm{h}(d^\circ) < \epsilon_0$. To d° we apply 6.7, 6.1 and obtain a 0-derivation d' of $\emptyset$ with $\mathrm{h}(d') \leq \epsilon_{\mathrm{h}(d^\circ)+1} < \epsilon_{\epsilon_0}$. $\square$

The next Lemma says in fact that a cut-free derivation of the empty sequent is a protocol of a terminating H-process.

Lemma 6.9
A 0-derivation d of $\emptyset$ consists of exactly one branch and the following holds:
(a) all sequents in d are correct;
(b) the top sequent of d is an axiom AxS*, and all other inferences in d are of the kind* Fr *or* H.

Proof.
(a) Proof by bottom-up induction: H : cf. Lemma 5.3. Fr : If Θ is correct and $e \notin \mathrm{dom}(\Theta)$ then $(e,?),\Theta$ is correct.
(b) Since d is a 0-derivation, there are no Cut- or CutFr-inferences in d. Hence d is linear. By bottom-up induction we obtain $\Theta f = \emptyset$ for each sequent Θ in d. Since d is wellfounded, there exists a top sequent Υ. This has to be an axiom. By (a) Υ is not ci. Since $\Theta f = \emptyset$, Υ cannot be an axiom AxH. $\square$

6.3 Construction of the original derivation

Here we construct the tree of finite ϵ-substitutions mentioned at the beginning of the section 6. General idea here is the same as in [12]. At each stage leaves of the tree are extended (by the bottom-up application of Cut) to make them “more computed” till the axioms are reached. Subterms of the non-computed (but needed) ϵ-terms of maximum rank are computed till these maximum ϵ-terms can be reduced to a canonical form and then computed. Note that the following definitions are stated for a given substitution S, and not for its completion $\overline{S}$.

Definition 31 *Let S be an ϵ-substitution and Φ a finite set of closed formulas.*
$\rho_S(\Phi) := \max\{\mathrm{rk}(|A|_S) : A \in \Phi,\ d(|A|_S) > 0\} \cup \{0\}$
$\nu_S(\Phi) := \omega^\omega \cdot r + \#_S(\Phi, r)$ *where* $r := \rho_S(\Phi)$,
$\#_S(\Phi, r) := d_r(|A_1|_S)\# \ldots \# d_r(|A_n|_S)$, *where* $\Phi = \{A_1, ..., A_n\}$ *without repetitions,*
$$d_r(F) := \begin{cases} 0 & \textit{if } \mathrm{rk}(F) < r \\ d(F) & \textit{otherwise} \end{cases}.$$

Note that $\rho_S(\Phi) < \omega+\omega$, $\#_S(\Phi, r) < \omega^\omega$, $\nu_S(\Phi) < \omega^{\omega+1}\cdot 3$.

Lemma 6.10 (One step of extension)
Let S be an ϵ-substitution and Φ a finite set of closed formulas.
Let $e = \epsilon\xi F \notin \mathrm{dom}(S)$ be a canonical ϵ-subterm of a formula $|A_0|_S$ with $A_0 \in \Phi$, $\mathrm{rk}(|A_0|_S) = \rho_S(\Phi)$.
Let $u \in \mathbb{B}_{\iota(e)} \cup \{?\}$, $S' := S \cup \{(e,u)\}$, and $\Phi' :=$ if $u = ?$ then Φ else $\Phi \cup \{F[\![u]\!]\}$.
Then $\rho_{S'}(\Phi') \leq \rho_S(\Phi)$ and $\nu_{S'}(\Phi') < \nu_S(\Phi)$.

Proof:
(a) Since $S \subseteq S'$, we have $||w|_S|_{S'} = |w|_{S'}$ and thus $\mathrm{rk}(|w|_{S'}) \leq \mathrm{rk}(|w|_S)$, $d(|w|_{S'}) \leq d(|w|_S)$ for each w.
Let $r := \rho_S(\Phi)$ and $r' := \rho_{S'}(\Phi')$.
One easily sees that $r' \leq r$. Indeed, $\mathrm{rk}(|F[\![u]\!]|_{S'}) \leq \mathrm{rk}(F[\![u]\!]) < \mathrm{rk}(e) \leq r$, and for remaining terms w occurring in Φ cf (a) .
Let $r' = r$ (for $r' < r$ the claim is trivial: $\nu_{S'}(\Phi') < \omega^\omega \cdot (r'+1) \leq \omega^\omega \cdot r$).
For each $A \in \Phi$ we have $\mathrm{rk}(|A|_{S'}) \leq \mathrm{rk}(|A|_S)$, $d(|A|_{S'}) \leq d(|A|_S)$ and thus $d_r(|A|_{S'}) \leq d_r(|A|_S)$.
Moreover $|A_0|_S$ is S'-reducible and $\mathrm{rk}(|A_0|_S) = r$. Hence $d_r(|A_0|_{S'}) < d_r(|A_0|_S)$ (Lemma 3.3).
Finally $d_r(|F[\![u]\!]|_{S'}) = 0$, since $\mathrm{rk}(|F[\![u]\!]|_{S'}) < r$. □

Lemma 6.11 (Rank reduction)
Let Θ be a sequent, L a finite set of closed formulas, and $r := \rho_\Theta(\mathcal{F}(\Theta) \cup L)$ (cf. Definition 19).
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. Moreover $\mathrm{h}(d) \leq \nu_\Theta(\mathcal{F}(\Theta) \cup L)$.

Proof by induction on $\nu_\Theta(\mathcal{F}(\Theta) \cup L)$:
Let $\Phi := \mathcal{F}(\Theta) \cup L$. If Θ computes all formulas in Φ we are done.
Otherwise there exists a canonical ϵ-subterm $e = \epsilon\xi F$ of a formula $|A_0|_\Theta$ with $A_0 \in \Phi$, $\mathrm{rk}(|A_0|_\Theta) = r$.
Let $u \in \mathbb{B}_{\iota(e)} \cup \{?^\circ\}$ and $\Theta' := (e,u), \Theta$. Then $\mathcal{F}(\Theta') =$ if $u = ?^\circ$ then $\mathcal{F}(\Theta)$ else $\mathcal{F}(\Theta) \cup \{F[\![u]\!]\}$.
By the Lemma 6.10 $r' := \rho_{\Theta'}(\mathcal{F}(\Theta') \cup L) \leq r$ and $\nu_{\Theta'}(\mathcal{F}(\Theta') \cup L) < \nu_\Theta(\mathcal{F}(\Theta) \cup L)$.
Hence (by I.H.) there exists a deduction d_u of Θ' by cuts of ranks $\leq r'$ from computing sequents Υ containing Θ' and computing all formulas in L, and $\mathrm{h}(d_u) \leq \nu_{\Theta'}(\mathcal{F}(\Theta') \cup L)$.
A cut with main term e yields the desired deduction d. □

Lemma 6.12
There exists an $r < \omega + \omega$ and an $r{+}1$-derivation d of the empty sequent containing only axioms and cuts. In addition we have $\mathrm{h}(d) < \omega^{\omega+2}$.

Proof:
First apply Lemma 6.11 to the empty sequent and the set $L := \{Cr_0, ..., Cr_N\}$. Let $r := \rho_\emptyset(L)$, and consider an arbitrary top sequent Θ of the resulting $r{+}1$-deduction d° which is not an axiom. Then Θ is cc, deciding and nonsolving. Since the only inferences in d° are cuts of rank $\leq r$, we have $\Theta t = \emptyset$ and $\mathrm{rk}(e) \leq r$ for each $e \in \mathrm{dom}(\Theta)$. Let e be the H-term of Θ. Since Θ is deciding, e is Θ-computable (cf. Lemmas 3.7c and 3.8b). Together with Lemma 5.1a and $\Theta t = \emptyset$ this implies $(e, ?^\circ) \in \Theta$. Since Θ is not an axiom AxH, it follows that $\mathrm{Cr}(\Theta)$ is of the form $F[t] \to F[\epsilon x F]$ and $e = \epsilon x A$ with $A := |F|_\Theta$. By Lemma 5.2 $A[n] \hookrightarrow_\Theta$ TRUE for $n := |t|_\Theta$.
Now let $L' := \{A[n{-}1], \ldots, A[0]\}$. Then $\rho_\Theta(\mathcal{F}(\Theta) \cup L') \leq r$. Apply Lemma 6.11 to Θ, L' and consider any cc top sequent Υ of the resulting $r{+}1$-deduction. Υ contains Θ and computes all formulas in L'. Now Lemma 5.2b yields that Υ is an axiom AxH with main term e. □

Theorem 6.13 . *The H-process terminates.*

Proof. Combine Lemmata 6.12, 6.8, 6.9. Cf. [12].

6.4 Producing a substitution in terms of finite predicates

6.4.1 Soundness

Lemma 6.14 *Let S be a correct and total ϵ-substitution. Then all closed axioms of $EA\epsilon$ except maybe critical formulas are satisfied by S. Modus ponens rule preserves truth under S.*

Proof. All instances of propositional tautologies and defining axioms for predicate constants are satisfied by S by the Lemma 3.7. Modus ponens preserves truth under S, since values of composite formulas are calculated accordingly to standard boolean rules. Equality axioms are satisfied, since by Lemma 3.7c: if $|t|_S = |u|_S$ for 0-terms t, u then $|e[t]|_S = |e[u]|_S$ for any expression e.
Consider a minimality axiom $\epsilon x F[x] = \mathsf{S}t \to \neg F[t]$: Assume $|\epsilon x F[x] = \mathsf{S}t|_S \in \mathsf{TRUE}$ and let $n := |t|_S$. Then $(\epsilon x |F|_S, n{+}1) \in S$ and by correctness and Lemma 3.7 we have $|F[\![n{+}1]\!]|_S = |\,|F|_S[\![n{+}1]\!]\,|_S \in \mathsf{TRUE}$, in particular $|\neg F[t]|_S = |\neg F[n]|_S \in \mathsf{TRUE}$. $\square$

6.4.2 ϵ-free derivation

Let d^* be a closed derivation in $EA\epsilon$, and let S be a correct, total, and solving ϵ-substitution for the system $Cr_0, ..., Cr_N$ of critical formulas of d^*. (For example $S = \overline{S_n}$, where S_n is produced by the H-process for $Cr_0, ..., Cr_N$, cf. Definition 22.) Since S is correct and total, all axioms of d^* except critical formulas are satisfied by S and modus ponens rule preserves truth under S. Since S is solving, critical formulas are satisfied, too. Hence all formulas in d^* are true under S.

Closed formulas are constructed by propositional connectives ¿from atomic formulas of the form $pe_1 \dots e_n$ and $\mathsf{App}(\epsilon X F)e_1$ where e_i are numerals or closed 0-ϵ-terms possibly preceded by several S, and $\epsilon X F$ is a closed 1-term. Let M be the maximum of all numerals, including all values $|u|_S$ of all 0-terms u mentioned in all computations above needed to verify the axioms of the derivation d^*.

Replace exterior occurrences of 1-epsilon-terms $\epsilon X F$ in d^* by finite predicates

$$\{n \le M \mid |\mathsf{App}(\epsilon X F)n|_S \in \mathsf{TRUE}\}$$

and exterior occurrences of 0-epsilon-terms u by their S-values. We obtain an ϵ-free derivation from formulas which are true under the standard interpretation of predicate constants and boolean connectives, and the rule

$$(\,\mathsf{App}\{n_1, \dots, n_k\}(n) \text{ is true }) \overset{def}{\iff} (n \text{ occurs among } n_1, \dots, n_k)\,.$$

Note that thereby an $EA\epsilon$ derivation d^* of a formula $F[\epsilon x F]$ $(= (\exists x F)^*)$ or $F[\epsilon X F]$ $(= (\exists X F)^*)$ with F quantifier- and ϵ-free is transformed into a derivation of $F[n]$ or $F[P]$ for some numeral n or finite predicate P.

7 Non-constructive proof of termination

In this section $S, S_n, ...$ always denote ϵ-substitutions with $\{e \in \mathrm{dom}(S) : S(e) = ?\} = \emptyset$. For each pair (e, u) we set $\mathrm{rk}((e, u)) := \mathrm{rk}(e)$.

Definition 32 *Let S be an ϵ-substitution such that $\overline{S}$ is correct and nonsolving.*
Let e be the H-term and v the H-value of S.
We set $\mathrm{rk}(S) := \mathrm{rk}(e)$ and $\pi(S) := (e, v)$.

Note that if $\overline{S}$ is correct and nonsolving, then according to our general assumption on S we have $e \notin \mathrm{dom}(S)$ and $\mathrm{H}(S) = S_{\le \mathrm{rk}(e)} \cup \{(e, v)\}$.

Definition 33 *An ϵ-substitution S is called* r-substitution *($r \in On$) iff $\overline{S}$ is correct and $\mathrm{rk}(\pi) < r$ for all pairs $\pi \in S$, i.e. $S = S_{<r}$.*

Definition 34 *Let $r \in On$. An* r-process *is a sequence $(S_i)_{i<\nu}$ such that:*
- *$0 < \nu \le \omega$,*
- *S_0 is an r-substitution,*
- *if $i < \nu$ and $\overline{S_i}$ is nonsolving and $\mathrm{rk}(S_i) \ge r$ then $i + 1 < \nu$ and $S_{i+1} = \mathrm{H}(S_i)$,*
- *if $i < \nu$ and ($\overline{S_i}$ is solving or $\mathrm{rk}(S_i) < r$) then $\nu = i + 1$.*
  (In this case S_i is the last substitution, and $S_\nu = S_{i+1}$ is not defined.)

Remark

1. For each r-substitution S there is a unique r-process $(S_i)_{i<\nu}$ with $S = S_0$.
2. If $(S_i)_{i<\nu}$ is an r-process then
   a) for all $i < \nu$, $\overline{S_i}$ is correct and $(S_i)_{<r} = S_0$,
   b) for all i with $i{+}1 < \nu$, $\overline{S_i}$ is nonsolving and $\mathrm{rk}(S_i) \geq r$.

The values of rank r are preserved in an r-process.

Lemma 7.1
If $(S_i)_{i\in\nu}$ is an r-process and $i \leq j < \nu$ then $(S_i)_{\leq r} \subseteq S_j$.

Proof: If $i{+}1 < \nu$ then $S_{i+1} = (S_i)_{\leq \mathrm{rk}(e)} \cup \{(e,v)\}$ with $\mathrm{rk}(e) \geq r$, and therefore $(S_i)_{\leq r} \subseteq S_{i+1}$. ¿From this the claim follows by induction on j. $\Box$

An infinite r-process if it existed, would necessarily introduce values of rank $> r$.

Lemma 7.2 *If $(S_i)_{i\in\omega}$ is an r-process then $\forall k \exists i \geq k(\,\mathrm{rk}(S_i) > r\,)$.*

Proof:
Assumption: $k \in \mathbb{N}\;\&\;\forall i \geq k(\,\mathrm{rk}(S_i) = r\,)$. — We write $|e|_n$ for $|e|_{\overline{S_n}}$.
(1) For each canonical ϵ-term e there is an n such that $\forall i \geq n(\,|e|_n = |e|_i\,)$.
Proof:
1. $\mathrm{rk}(e) = r$ and $e \notin \bigcup_{i\in\omega} \mathrm{dom}(S_i)$: Then $|e|_i = 0^{\iota(e)}$ for all i.
2. $\mathrm{rk}(e) = r$ and $e \in \mathrm{dom}(S_n)$: By Lemma 7.1 it follows that $|e|_i = |e|_n$ for all $i \geq n$.
3. $\mathrm{rk}(e) < r$: Then $|e|_i = |e|_0$ for all i.
4. $\mathrm{rk}(e) > r$: By assumption we have $\forall i > k(e \notin \mathrm{dom}(S_i))$ and therefore $|e|_i = 0^{\iota(e)}$ for all $i > k$. $\Box$
(2) For each expression e there is an n such that $\forall i \geq n(|e|_i = |e|_n)$.
Proof by induction on $d(e)$. If $d(e) = 0$ then $|e|_i = e = |e|_n$. Assume now that $d(e) > 0$ and let u be some canonical ϵ-subterm of e. By (1) there is an m such that $|u|_i = |u|_m$ for all $i \geq m$. Let e' result from e by 'contraction' of u. Then $e \hookrightarrow^{1}_{\overline{S_i}} e'$ for all $i \geq m$, and by I.H. there is an $n \geq m$ such that $|e'|_i = |e'|_n$ for all $i \geq n$. Hence $|e|_i = |e'|_i {=} |e'|_n = |e|_n$ for all $i \geq n$. $\Box$ (2).
By (2) there is an $n \geq k$ such that $|F|_{n+1} = |F|_n$ and $|\epsilon\xi F|_{n+1} = |\epsilon\xi F|_n$ for each critical formula $F_0 {\to} F[\epsilon\xi F]$ from the list $Cr_0, ..., Cr_N$. So especially for the H-term $e = \epsilon\xi|F|_n$ of S_n we have $|e|_{n+1} = |\epsilon\xi|F|_{n+1}|_{n+1} = |\epsilon\xi F|_{n+1} = |\epsilon\xi F|_n = |e|_n$. (Note that $|\epsilon\xi F|_S = |\,\epsilon\xi|F|_S\,|_S$). But, since $S_{n+1} = \mathrm{H}(S_n)$, this is a contradiction to Lemma 5.1b. $\Box$

Each infinite r-process can be extended to an infinite $r+1$-process.

Theorem 7.3
Let $(S_i)_{i\in\omega}$ be an r-process. Then (by adding pairs of rank r) S_0 can be extended to an $(r{+}1)$-substitution $S{+}$ such that the $(r{+}1)$-process beginning with $S{+}$ is infinite.

Proof: Again we set $|e|_n := |e|_{\overline{S_n}}$.
$S{+} := \bigcup_{i\in\omega} (S_i)_{\leq r}$. By Lemma 7.1 $S{+}$ is a function. We show that the $r+1$-process beginning with $S{+}$ is an accelerated version of the given infinite r-process $(S_i)_{i\in\omega}$. (cf. (2) below).
(1) $\overline{S{+}}$ is correct.
Proof: Let $(\epsilon\xi A, u) \in (S_i)_{\leq r}$. Then for each $v \in \mathbb{B}_{\iota(\xi)}$, $\mathrm{rk}(A[v]) < \mathrm{rk}(\epsilon\xi A) \leq r$ and therefore $|A[v]|_i = |A[v]|_0 = |A[v]|_{\overline{S+}}$. Hence $|A[\![u]\!]|_i = |A[\![u]\!]|_{\overline{S+}}$. Since S_i is correct, we have $|A[\![u]\!]|_i \in \mathsf{TRUE}$.
Let $(S{+}_i)_{i<\nu}$ be the $(r{+}1)$-process starting with $S{+}$. Abbreviation: $|e|^{+}_{n} := |e|_{\overline{S+_n}}$.
The following proposition (2a) together with (1) yields the theorem.
(2) For every $i \in \omega$ holds
  (a) $i < \nu$,
  (b) there exists $j \in \omega$ such that $S{+}_i$ is an extension of S_j by some pairs of rank r.
Proof by induction on i:
I. $i = 0$: (a) trivial. (b) Take $j := 0$.
II. $i \to i{+}1$:

By I.H. $i < \nu$, and there is a j such that $S{+}_i = S_j \uplus \Psi$ with $\forall\pi \in \Psi(\,\mathrm{rk}(\pi) = r)$.
(The symbol $\uplus$ indicates disjoint union).
Let Δ be the finite set of all pairs $\pi \in S+$ of rank r which are used in the computation of $|Cr_I|^+_0, ..., |Cr_I|^+_i$ $(I = 0, ..., N)$ as well as $\pi(S{+}_0), \ldots, \pi(S{+}_i)$.
Let $k := \min\{l \geq j : \Delta \subseteq S_l \,\&\, \mathrm{rk}(S_l) > r\}$ (cf. Lemma 7.2). We have $k \geq j \,\&\, \Delta \subseteq S_k \,\&\, \mathrm{rk}(S_k) > r$
CASE A: $k = j$.
(a) Then $S{+}_i = S_k \uplus \Psi$, and Ψ is not used in the computation of $|Cr_I|^+_i$ since S_k already contains Δ. Hence $|Cr_I|^+_i = |Cr_I|_k$ for $I = 0, ..., N$. Since $\overline{S_k}$ is nonsolving and $\mathrm{rk}(S_k) > r$, it follows that $\overline{S{+}_i}$ is nonsolving, $\pi(S{+}_i) = \pi(S_k)$, and $\mathrm{rk}(S{+}_i) \geq r + 1$. Hence $i + 1 < \nu$.
(b) $S{+}_{i+1} = (S{+}_i)_{\leq \mathrm{rk}(S+_i)} \uplus \{\pi(S{+}_i)\} = (S_k \uplus \Psi)_{\leq \mathrm{rk}(S_k)} \uplus \{\pi(S_k)\} = (S_k)_{\leq \mathrm{rk}(S_k)} \uplus \{\pi(S_k)\} \uplus \Psi = S_{k+1} \uplus \Psi$.
CASE B: $j < k$. Then $\mathrm{rk}(S_{k-1}) = r$. Otherwise $(S_k)_{\leq r} = (S_{k-1})_{\leq r}$ and $\mathrm{rk}(S_{k-1}) > r$, so that k is not minimal.
Hence in the step from S_{k-1} to S_k all pairs of rank $> r$ are removed. It follows that $S{+}_0 = S_k \uplus \Phi$ where Φ consists of pairs of rank r which are not used in the computation of $|Cr_I|^+_0, ..., |Cr_I|^+_i$ and $\pi(S{+}_i)$.
Proposition. $S{+}_l = S_{k+l} \uplus \Phi$, for $l = 0, ..., \min\{i + 1, \nu - 1\}$ $\quad(*)$.
Proof by induction on l: Let $l < \min\{i + 1, \nu - 1\}$ and $S{+}_l = S_{k+l} \uplus \Phi$.
By the definition of Φ, all pairs of rank r which are used in the computation of $|Cr_I|^+_0, ..., |Cr_I|^+_i$ and $\pi(S{+}_i)$ are contained in $\Delta \subseteq S_k$. Since $l \leq i$, we have $\pi(S{+}_l) = \pi(S_{k+l})$ and $\mathrm{rk}(S{+}_l) = \mathrm{rk}(S_{k+l})$. Since $l{+}1 < \nu$, the substitution $S{+}_l$ is not terminal, and we have $\mathrm{rk}(S{+}_l) \geq r{+}1$. Hence $S{+}_{l+1} = (S{+}_l)_{\leq \mathrm{rk}(S+_l)} \cup \{\pi(S{+}_l)\} = (S_{k+l} \uplus \Phi)_{\leq \mathrm{rk}(S_{k+l})} \uplus \{\pi(S_{k+l})\} = (S_{k+l})_{\leq \mathrm{rk}(S_{k+l})} \uplus \{\pi(S_{k+l})\} \uplus \Phi = S_{k+l+1} \uplus \Phi$.
(a) The above proposition $(*)$ yields $|Cr_I|^+_i = |Cr_I|_{k+i}$ since Ψ is not used in the computation of $|Cr_I|^+_i$ .Hence $\overline{S{+}_i}$ is nonsolving, since $\overline{S_{k+i}}$ is nonsolving. Now by $(*)$ we also have $\pi(S{+}_i) = \pi(S_{k+i})$ and $\mathrm{rk}(S{+}_i) = \mathrm{rk}(S_{k+i})$. Assume $\mathrm{rk}(S_{k+i}) = r$; then $\pi(S_{k+i}) \in S+ \subseteq S{+}_i$ and thus $\pi(S_{k+i}) \neq \pi(S{+}_i)$, since $\pi(S{+}_i) \notin S{+}_i$. Contradiction.
Hence $\mathrm{rk}(S{+}_i) = \mathrm{rk}(S_{k+i}) \geq r{+}1$ and thus $i + 1 < \nu$.
(b) As in the proof of $(*)$ we get $S{+}_{i+1} = S_{k+i+1} \uplus \Phi$. $\quad\square$

Theorem 7.4
The 0-process Π *beginning with the empty substitution* $\emptyset$ *terminates in a solving substitution.*

Proof: Obviously it suffices to prove that Π is finite (terminates).
For contradiction we assume that Π is infinite.
Below we define substitutions S^r for all $r \in On$ such that:
(1) $S^0 = \emptyset$,
(2) S^r is an r-substitution,
(3) the r-process starting with S^r is infinite,
(4) $S^q \subseteq S^r$, for all $q < r$.

Let $R := \max\{\mathrm{rk}(Cr_I) : I \leq N\} + 1$. Then $\mathrm{rk}(S) < R$ for each substitution S. But on the other hand $\mathrm{rk}(S^R) \geq R$, since by (3) the R-process starting with S^R is infinite. *Contradiction.*

Definition of S^r by transfinite recursion on r:
$S^0 := \emptyset$,
$S^{r+1} := (S^r)+$ (cf. Theorem 7.3),
$S^r := \bigcup_{q<r} S^q$, if $r \in Lim$.

In parallel with that definition we prove by transfinite induction on r that S^r satisfies the above conditions (2),(3),(4). The successor step is settled by theorem 7.3. Now assume $r \in Lim$.
Then (4) is trivial, and (2) follows ¿from I.H.(2),(4), since each particular value (e, u) in S^r is verified by computation in some S^q, $q < r$.
For (3) assume that the r-process for S^r is finite. Then it uses information only from S^q for finitely many $q < r$, hence there is a $q < r$ such that S^q is finite which contradicts I.H.(3). $\quad\square$

References

[1] W.Ackermann, Begründung des Tertium non datur mittels der Hilbertschen Theorie der Widerspruchsfreiheit, Math.Ann. 1925, 93, 1-36

[2] W.Ackermann, Zur Widerspruchsfreiheit der Zahlentheorie, Math.Ann. 1940, 117, 162-194

[3] N. Bourbaki, Theorie des ensembles, Hermann, 1958

[4] G. Gentzen, Die Widerspruchsfreiheit der reinen Zahlentheorie, Math Ann., 1936, 112, N4, 493-565

[5] J.-Y.Girard, Une extension de l'interpretation de Gödel a l'analyse et la application a l'elimination des coupures dans l'analyse et la theorie des types, Proc. 2-nd Scand.Logic Symp., North-Holland, 1972, 63-92

[6] D.Hilbert, Probleme der Grundlegung der Mathematik, Math.Ann. 1929, 102, 1-9

[7] D.Hilbert, P.Bernays, Grundlagen der Mathematik, Bd.2, Springer, 1970

[8] G.Kreisel, On the Interpretation of Non-finitist proofs I, J. Symbolic Logic 1951, 16, 241-267

[9] G.Kreisel On the Interpretation of Non-finitist proofs II, J. Symbolic Logic 1952, 17, 43-58

[10] G.Mints, Simplified Consistency Proof for Arithmetic (Russian), Proc. Estonian Acad. of Sci. Fiz.-Math.1982, 31 N4, 376-382

[11] G. Mints, Epsilon Substitution Method for the Theory of Hereditarily Finite Sets (Russian), Proc. Eston. Acad. of Sci. Fiz.-Math. 1989 N2, 154-164

[12] G.Mints, Gentzen-type Systems and Hilbert's Epsilon Substitution Method. I. In: Logic, Method. and Philos. of Sci. IX, Elsevier, 1994, 91-122

[13] G. Mints, S. Tupailo, Epsilon Substitution Method for Elementary Analysis, Report No. CSLI-93-175, 1993, CSLI, Stanford University, February 1993

[14] Kleene S.C., Introduction to Methamathematics, Van Nostrand, 1952

[15] J.von Neumann, Zur Hilbertschen Beweistheorie, Math. Zeitschrift 26, 1927, 1-46

[16] K. Schutte, Proof Theory, Springer, 1977

[17] W.Tait, Functionals Defined by Transfinite Recursion, J. Symbolic Logic 1965, 30 N2, 155-174

[18] W.Tait, The Substitution Method, J. Symbolic Logic 1965, 30, N2, 175-192

[19] H.Weyl, David Hilbert and His Mathemtical Work, Bull. Amer. Math. Society, 1944, 50, 612-654