An expository survey on epsilon substitution method

Toshiyasu Arai(Kobe, Japan)

Epsilon substitution method is a method proposed by D. Hilbert to prove the consistency of (formal) theories. The idea behind the method is that one could replace consistently transfinite/non-computable objects as a figure of speech by finitary/computable ones as far as transfinite ones are finitely presented as axioms of a theory. In other words, the replacement (epsilon substitution) depends on contexts, i.e., formal proofs in which axioms for the transfinite objects occur. If this attempt would be successfully accomplished, then the (1-)consistency of the theory follows.

For example, for the first order arithmetic PA, replace each existential formula \( \exists x F[x] \) by \( F[\epsilon x.F[x]] \), where the epsilon term \( \epsilon x.F[x] \) intends to denote the least number satisfying \( F[x] \) if such a number exists. Otherwise it denotes an arbitrary object, e.g., zero. Then PA is interpretable in an extended propositional calculus having the epsilon axioms:

\[
(\epsilon) \quad F[t] \rightarrow ex.F[x] \not\geq t \land F[ex.F[x]]
\] (1)

The problem is to find a solving substitution which assigns numerical values to epsilon terms and under which all the epsilon axioms occurring in a given proof are true.

Hilbert’s Ansatz is, starting with the null substitution \( S^0 \) which assigns zero whatever, to approximate a solution by correcting false values step by step, and thereby generate the process \( S^0, S^1, \ldots \) (H-process). The problem is to show that the process terminates.

In this talk we expound basic ideas of the epsilon substitution method à la Ackermann[1], examine its proof-theoretic consequences, and report recent progress on the subject.

1 The first order arithmetic: Ackermann’s proof

In this section we give a sketch of an epsilon substitution method, which is slightly modified from Ackermann[1].

1.1 The H-process

In this subsection we define the H-process.
Let $L_{PA}$ denote a language of the first order arithmetic $PA$. $L_{PA}$ includes some computable function symbols, say $+$ for the addition, $\cdot$ for the multiplication and the predecessor $\hat{-}$, and the relation symbol $<$. Axioms of $PA$ are axioms for computable functions and the well ordering $<$. 

(Lin) The axiom stating that $<$ is a linear ordering.

(WF) Complete induction schema for any formula $F \in L_{PA}$:

$$\forall x[\forall y < x F(y) \rightarrow F(x)] \rightarrow \forall x F(x).$$

$PA$ codifies a transfinite/non-computable concept. Namely the least number satisfying a condition. Now let us introduce a symbolism denoting the concept: the epsilon symbol $\epsilon$, and declare that $\epsilon x.F[\epsilon x]$ is a term ($\epsilon$-term) if $F[\epsilon x]$ is a formula. Observe that, in this symbolism, terms and formulas are generated simultaneously. With the epsilon axiom (1) we don’t need quantifiers $\exists, \forall$ anymore: these are definable from $\epsilon$ by $\exists x F[x] \equiv F[\epsilon x]$ and $\forall x F(x) \equiv F[\epsilon (\epsilon x.F[\epsilon x])]$. Note the fact $\forall F(x) \iff \exists x \neg F[x]$. Moreover the complete induction (WF) follows from (1).

Thus one can interprete $PA$ in a system $PA_{\epsilon}$ with $\epsilon$ but without quantifiers. To be specific, let us describe the system $PA_{\epsilon}$.

The only inference rule of $PA_{\epsilon}$ is modus ponens:

$$\frac{F \quad F \rightarrow G}{G}.$$

Axioms of $PA_{\epsilon}$

Propositional axioms: All tautologies of the language $L_{PA_{\epsilon}}$,

All substitution instances of defining axioms for the computable functions $+, \cdot, \hat{-}$,

Equality axioms: $t = t$ and $s = t \rightarrow (F[s] \rightarrow F[t])$,

Peano axioms for the successor function $S$: $St \neq 0$ and $t \neq 0 \rightarrow t = S(t\hat{-}1)$,

Epsilon axiom (1).

Giving a context, i.e., a (finite) derivation $P$, let us replace the transfinite expressions $ex.F[x]$ occurring in $P$ by finitary objects, i.e., natural numbers. Such a replacement is called an epsilon substitution. In doing so, it should validate the epsilon axioms (1) occurring in $P$.

Specifically an $\epsilon$-substitution $S$ is a finite function assigning values $|ex.F[\epsilon]|_S \in \omega$ of canonical (closed and minimal epsilon) terms $ex.F$. $dom(S)$ denotes its domain.

For a canonical epsilon term $e$ not in $dom(S)$ we set $|e|_S := 0$. $e \leftrightarrow_s e'$ designates that an expression $e$ reduces to another expression $e'$ under $S$ by substituting the values for canonical subexpressions in $e$. Thus any expression $e$ reduces to a unique irreducible form $|e|_S$, $e \leftrightarrow_s |e|_S$. Observe that the value $|e|_S$ of (closed) expressions $e$ under $S$ is computable: $|e|_S$ is a natural number for terms $e$, and $|e|_S \in \{\bot($falsehood$), \top($truth$)\}$ for sentences $e$ in the language of $PA_{\epsilon}$.

Let $Cr = \{C_0, \ldots, C_N\}$ be a fixed finite sequence of closed epsilon axioms. $S$ is solving if $S$ validates any critical formula in $Cr$, $\forall F \in Cr[F \leftrightarrow_s \top]$. Otherwise $S$ is nonsolving.
Lemma 1.1 (The rank lemma)

Thus $0$ is the largest element in $\mathcal{L}$.

Definition 1.2

A substitution $S$ is said to be correct iff $E[m] \land \neg E[0] \iff S \top$ for any $(\epsilon x, m) \in S$. In particular $(\epsilon, 0) \notin S$ for any correct $S$. The reason why $\neg E[0] \iff S \top$ is that if $E[0] \iff S \top$, then $\epsilon x, m$ would suffice to receive the default value $0$, and need not to be in $\text{dom}(S)$.

Obviously $S^0 = \emptyset$ is correct. The $H$-process $S^0, S^1, \ldots$ is defined so that each $S^i$ is correct.

Proposition 1.1.2

Now let $P$ be a proof in PA in which no free variable occurs, and consider the case when $Tr$ is the set of all epsilon axioms occurring as axioms in $P$. Assume $S$ is solving. Then we see that $A \iff S \top$ for any formula $A$ in $P$. Therefore $S$ validates the endformula of $P$. Thus the existence of a solving substitution for any finite set of critical formulas yields the (1-)consistency of PA.

Given the sequence $C r$ we define a sequence $S^0(= \emptyset), S^1, \ldots$ of substitutions, called the H-process (for $C r$) so that the process $S^0, S^1, \ldots$ terminates in a finite number of steps reaching to a solving substitution $S^\omega$.

A substitution $S$ is said to be correct iff $E[m] \land \neg E[0] \iff S \top$ for any $(\epsilon x, m) \in S$. In particular $(\epsilon, 0) \notin S$ for any correct $S$. The reason why $\neg E[0] \iff S \top$ is that if $E[0] \iff S \top$, then $\epsilon x, m$ would suffice to receive the default value $0$, and need not to be in $\text{dom}(S)$.

Obviously $S^0 = \emptyset$ is correct. The H-process $S^0, S^1, \ldots$ is defined so that each $S^i$ is correct.

Proposition 1.1.2

Suppose that a correct substitution $S$ is nonsolving, and let $E[m] \not\iff t \land E(x, m) \iff S \bot$. Putting $m = |t|_S$, this means that $E[m]|_S \iff S \top$ and $E[m]|_S \iff S \top \Rightarrow |x|_S < m$.

First consider the case $\epsilon x, m \in \text{dom}(S)$. Then $|x|_S < m$, and hence $E[m]|_S \iff S \bot$. Second assume $(\epsilon, n) \in S$. By the correctness of $S$, $E[n]|_S \land \neg E[0]|_S \iff S \top \Rightarrow n > m$.

In each case we have $E[n]|_S \land \neg E[0]|_S \iff S \top$. Hence $m \notin S$.

Putting these together, if the value $\epsilon x, m$ is to be corrected, then $m = |t|_S$ should be the new value of $\epsilon x, m$. Then the new value $m$ is smaller than the old one $|x|_S$ in the Ackermann ordering:

$$x <_A y :\Leftrightarrow (x \neq 0 \land y = 0) \lor (x, y \neq 0 \land x < y).$$

Thus $0$ is the largest element in $<_A$. $|x|_A$ denotes the order type of $x$ in the ordering $<_A$.

$$\|x\|_A := \begin{cases} x - 1 & x > 0 \\ \omega & x = 0 \end{cases}$$

Observe that $0 \not\in m <_A \epsilon x, m$ for nonsolving $S$. In particular

$$|\epsilon x, m|_S \notin S \quad (2)$$

The rank $rk(e) < \omega$ of an expression $e$ measures nesting of bound variables in $e$, and is defined so that the following holds:

Lemma 1.1 (The rank lemma)

1. If $\epsilon x, m$ is canonical, then $rk(\epsilon x, m) < rk(\epsilon x, m)$ for each numeral $n$.

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

Definition 1.2 $rk(S) := \max\{rk(e) : e \in \text{dom}(S)\} \cup \{0\}$. For each $e$-substitution $S$ and an ordinal $r \leq \omega$ we set $S_{\leq r} := \{(e, u) \in S : rk(e) \leq r\}$. Analogously we define $S_{< r}, S_{> r}$.

The following lemma is seen from Lemma 1.1.2.
Lemma 1.3 If $S, S'$ are $e$-substitutions with $S_{\leq r} = S'_{\leq r}$, then $|e|_S = |e|_{S'}$ holds for all expressions $e$ of rank $\leq r$.

Now suppose that a correct substitution $S = S'$ is nonsolving. Then $S = S'$ is corrected to define the next substitution $H(S) = S^{i+1}$ in the H-process as follows: Recall that $Cr = \{Cr_I : I \leq N\}$ is a fixed sequence of closed critical formulas. Let $I(S)$ denote the least number $I$ such that $Cr_I \equiv F[t] \rightarrow ex.F[x] \not\equiv t \land F[ex.F[x]] \leftarrow_S \bot$, and let $Cr(S) := Cr_I(S)$.

Putting $e^S := ex[F|_S, r^S = rk(e^S), v^S := |t|_S$, define

$$H(S) := S_{<r} \cup \{(f, u) : rk(f) = r^S \land f \not\equiv e^S\} \cup \{(e^S, v^S)\}.$$ 

Then $rk(S^0) = 0$ and $rk(S^{i+1}) = rk(H(S^i)) = rk(S^i) = rk(S^i)$.

The following lemma is seen from the definition.

Lemma 1.4 Let $S$ and $T$ be substitutions. Suppose $|e|_S = |e|_T$ for any closed $e$-terms occurring in the set $Cr$. Then $Cr(S) = Cr(T)$, and hence $w^S = w^T$ for $w \in \{e, r, v\}$.

The correctness of $H(S)$ follows from Lemmas 1.1.1 and 1.3.

The H-process (for a fixed sequence $Cr$) is defined as follows:

$$S^0 := \emptyset, S^{n+1} := \begin{cases} H(S^n) & \text{if } S^n \text{ is nonsolving} \\ S^n & \text{if } S^n \text{ is solving} \end{cases}$$

The H-process terminates iff there exists an $n$ such that $S^n$ is solving.

Let me explain the reason why we discard the higher-rank part $S_{>r}$: consider the case when $(ey.G[y], m) \in S$ and $e^S \equiv ex.F[x, m]$ such that the variable $y$ occurs in a subterm $ex.F[x, y]$ of $ey.G[y]$. Then $rk(ey.G[y]) > rk(ex.F[x, m])$ and $G[m] \leftarrow_S \top$. However it might be the case $G[m] \leftarrow_T \bot$ for $T = H(S) \cup \{(ey.G[y], m)\}$ since $|ex.F[x, m]|_S = |ex.F[x, m]|_T$.

Moreover note that it might be the case $Cr(S) \equiv F[t] \rightarrow ex.F[x] \not\equiv t \land F[ex.F[x]] \leftarrow_{H(S)} \bot$ since it may happen $|e|_S = |e|_{H(S)}$ for $e \in \{t, F[x]\}$.

### 1.2 Termination proof

In this subsection we show that the H-process terminates. The proof is based on the transfinite induction up to $\varepsilon_0$.

A relation $T \sqsubseteq_A S$ on $e$-substitutions is defined.

**Definition 1.5**

$$T \sqsubseteq_A S \iff \forall (e, u) \in S \exists (e, v) \in T[v \leq_A u]$$

$$\iff |e|_T \leq_A |e|_S \text{ for any canonical } e$$

**Lemma 1.6** $H(S) \sqsubseteq_A S_{\leq r}$ for $r = rk(H(S))$. 

4
We associate an ordinal \( \text{ind}(S) < \omega^\omega \) (index of \( S \)) relative to a fixed sequence \( Cr \) of \( \epsilon \)-axioms.

\( Cl_r(Cr) \) denotes the set of closed \( \epsilon \)-terms occurring in the set \( Cr \). Let \( N_{Cr} := \# Cl_r(Cr) \) (= the cardinality of the set \( Cl_r(Cr) \)). \( N_{Cr} \) is less than or equal to the total number of occurrences of the symbol \( \epsilon \) in the set \( Cr \).

**Definition 1.7**  
1. For an \( e \in Cl_r(Cr) \) put  
   \[
   \varphi(e; S) := \|v\|_A \text{ for } v = |e|_S. 
   \]
2. We arrange the set \( Cl_r(Cr) \) of cardinality \( N_{Cr} \) as follows: \( Cl_r(Cr) = \{e_i : i < N_{Cr}\} \) where  
   \[
   e_j \text{ is a closed subexpression of } e_i \Rightarrow j > i \quad (3)
   \]
3.  
   \[
   \text{ind}(S) = \sum \{(\omega + 1)^i : \varphi(e_i; S) : i < N_{Cr}\}. 
   \]

Let \( S \) and \( T \) be substitutions such that \( T \subseteq_A S \).

Suppose there exists an \( ex.F \in Cl_r(Cr) \) whose values \( |ex.F|_S, |ex.F|_T \) under \( S \) and \( T \) differ. Let \( ex.F \in Cl_r(Cr) \) denote a shortest such expression. Then \( |ex.F|_T < |ex.F|_S \) by \( |F|_T \equiv |F|_S \). Therefore by (3) we have \( \text{ind}(S) > \text{ind}(T) \).

Otherwise we have \( |e|_S = |e|_T \) for any \( e \in Cl_r(Cr) \). Then \( H(T) \subseteq_A H(S) \& rk(H(S)) = rk(H(T)) \) by Lemma 1.4 and, \( \text{ind}(S) = \text{ind}(T) \). Thus we have shown the following lemma.

**Lemma 1.8** Let \( S \) and \( T \) be substitutions such that \( T \subseteq_A S \). Then

1. \( \text{ind}(S) \geq \text{ind}(T) \).
2. \( H(T) \subseteq_A H(S) \& rk(H(S)) = rk(H(T)) \) if \( \text{ind}(S) = \text{ind}(T) \).

Suppose \( rk(S^n) \leq rk(S^{n+1}) \). Then by Lemma 1.6 we have \( S^{n+1} \subseteq_A S^n \). Thus Lemma 1.8.1 yields \( \text{ind}(S^{n+1}) \leq \text{ind}(S^n) \). Assume \( \text{ind}(S^{n+1}) = \text{ind}(S^n) \). Then we would have \( (e^{S^{n+1}}, v^{S^{n+1}}) = (e^{S^n}, v^{S^n}) \in S^{n+1} \). This is not the case by (2). We have shown \( rk(S^n) \leq rk(S^{n+1}) \Rightarrow \text{ind}(S^{n+1}) < \text{ind}(S^n) \). Therefore by transfinite induction up to \( \omega^\omega \) we see that, in the \( H \)-process, any consecutive series which is nondecreasing with respect to ranks, is finite. In particular, if the rank of epsilon terms in the given set \( Cr \) is at most one, then the \( H \)-process terminates.

How about the case when the highest rank is two? Each series of nondecreasing ranks is finite. Moreover we shall see, in Lemma 1.12, that any such series decreases with respect to a lexicographic ordering. In this way the termination of the \( H \)-process is seen when the highest rank is two. Our proof of the termination is completed by nesting this argument.

*Fix a positive integer \( \Lambda \geq 2 \) so that \( \max\{rk(Cr_I) : I = 0, \ldots, N\} < \Lambda \).*

Then for any \( S \) appearing in the \( H \)-process, we have \( rk(S) < \Lambda \).
Definition 1.9 Let \( \vec{S} = \{ S_i : i \leq k \} \) be a consecutive series in the H-process \( S^0, \ldots \). Then
\[
 rk(\vec{S}) := \min(\{ rk(S_i) : 0 < i \leq k \} \cup \{ \Lambda \}) > 0.
\]

Definition 1.10 A consecutive series \( \vec{S} = \{ S_i : i \leq k \} \) in the H-process \( S^0, \ldots \) is a section iff \( rk(S_0) < rk(\vec{S}) \).

Definition 1.11 \( \vec{T} \prec \vec{S} \).
Let \( \vec{S} = S_0, \ldots, S_k, \vec{T} = T_0, \ldots, T_j \) be two sections in the H-process \( S^0, \ldots \). Let \( S_{k+1} = H(S_k) \) and \( T_{j+1} = H(T_j) \).
If \( rk(T_0) < rk(\vec{S}) \) & \( T_0 \subseteq A S_0 \) and one of the following conditions is fulfilled, then we write \( \vec{T} \prec \vec{S} \):

1. There exists a \( p \leq \min\{ k, j \} \) such that \( ind(S_p) > ind(T_p) \) and \( \forall i < p \): \( ind(S_i) = ind(T_i) \).

2. \( j < k \) and \( \forall i \leq j \): \( ind(S_i) = ind(T_i) \).

Lemma 1.12 Let \( \vec{S} = S_0, \ldots, S_k, \vec{T} = T_0, \ldots, T_j \) be two consecutive sections in the H-process \( S^0, \ldots \) such that \( T_0 = H(S_k) \) and \( rk(S_0) \leq rk(T_0) < rk(\vec{S}) \), see the figure H-process. Then

1. \( T_0 \subseteq A S_0 \).

2. \( \vec{T} \prec \vec{S} \).

Proof. 1.12.1 is seen from Lemma 1.6.
1.12.2. By Lemmas 1.12.1 and 1.8 it suffices to show that the following case never happens: \( k \leq j \) and \( \forall i \leq k \): \( ind(S_i) = ind(T_i) \).
This is rejected since we would have \( (e^{S_k}, v^{S_k}) \in T_0 \) and hence \( (e^{T_k}, v^{T_k}) = (e^{S_k}, v^{S_k}) \in T_k \).

In the following figures, let us imagine a substitution, i.e., a finite set of pairs \( (e, v) \) to be arranged in increasing ranks in a vertical line. For example a substitution \( S = \{(e_1, v_1), (e_2, v_2), (e_3, v_3)\} \) with \( rk(e_1) = rk(e_2) < rk(e_3) \) is depicted by the figure:
Lemma 1.12.2 means that each section \( \vec{S} = \{ S_i : i \leq k \} \) codes an ordinal \( o(\vec{S}) < \varepsilon_0 \) in Cantor normal form with base 2: Let \( r := rk(\vec{S}) \). Divide \( \vec{S} \) into substrings which are sections as follows. Put \( \{ k_0 < \cdots < k_l \} = \{ i : i \leq k \land rk(S_i) = r \} \cup \{ 0 \} \), and \( \vec{S} = \vec{S}_0 \cdots \vec{S}_l \) with \( \vec{S}_j = (S_{k_j}, \ldots, S_{k_{j+1}}-1) \) for \( 0 \leq j \leq l \) and \( k_{l+1} = k + 1 \).

The series \( \vec{S}_0, \ldots, \vec{S}_l \) of substrings of \( \vec{S} \) is called the decomposition of \( \vec{S} \). We have \( \forall j < l[S_{j+1} \prec S_j] \).

For ordinals \( a \) and \( \alpha \geq 2 \) and \( k < \omega \), let \( \alpha_0(a) := a \) and \( \alpha_{1+k}(a) := \alpha^{\alpha_k(a)} \).

Also set \( \omega_k := \omega_k(1) \).

**Definition 1.13** \( o(\vec{S}; \xi) \) for \( 0 < \xi \leq rk(\vec{S}) \).

Let \( \xi \) be an ordinal such that \( 0 < \xi \leq r = rk(\vec{S}) \). Then an ordinal \( o(\vec{S}; \xi) \) is defined as follows:
1. \( k = 0 \): Then \( r = \Lambda \). Let \( a := \text{ind}(S_0) \). Set \( o(\vec{S}; \xi) = 2r - \xi(a) \).

2. \( k \geq 0 \): Let \( \vec{S}_0, \ldots, \vec{S}_r \) be the decomposition of \( \vec{S} \). Set

\[
o(\vec{S}; \xi) = 2r - \xi(\sum_{i \leq t} o(\vec{S}_i; r)).
\]

We see

\[
o(\vec{S}; \xi) < \omega\Lambda^2 - \xi.
\]

**Lemma 1.14** Let \( \vec{S} = S_0, \ldots, S_k, \vec{T} = T_0, \ldots, T_j \) be two consecutive sections in the H-process \( S^0, \ldots \) such that \( rk(S_0) \leq rk(T_0) < rk(\vec{S}) \).

Then \( o(\vec{S}; \xi) > o(\vec{T}; \xi) \) for any ordinal \( \xi \leq \min\{rk(\vec{S}), rk(\vec{T})\} \).

This is seen by induction on \( k + j \) using Lemma 1.12.2, \( \vec{T} < \vec{S} \), cf. [2] for a full proof.

**Theorem 1.15** (Transfinite induction up to \( \varepsilon_0 \))

The H-process \( S^0, \ldots \) terminates.

**Proof.** Suppose the H-process \( S^0, \ldots \) is infinite and put \( r_n = rk(S^n) \).

Inductively we define sequences \( \{n_i : i \in \omega\} \) of natural numbers as follows. First set \( n_0 = 0 \). Suppose \( n_i \) has been defined. Then put \( \beta_i = \min\{r_n : n > n_i\} \) and \( n_{i+1} = \min\{n > n_i : r_n = \beta_i\} \).

Then Lemma 1.14 yields an infinite decreasing sequence of ordinals, viz.

\[
\forall i [o(\vec{S}_{i+1}; \xi) < o(\vec{S}_i; \xi) < \omega\Lambda^2 - \xi] \quad \text{for} \quad \xi = \beta_i + 1 \geq 2.
\]

Therefore the H-process \( S^0, \ldots \) for any given sequence \( Cr \) of critical formulas terminates. It provides a closed and solving substitution, which in turn yields the 1-consistency RFN\(_{\Sigma^0_1}\)(PA) of PA stating that any PA-provable \( \Sigma^0_1 \)-sentence is true.

Epsilon substitution methods prove also arithmetical soundness RFN\(_{\Sigma^0_1}\)(PA). For example let PA \( \vdash \exists x \forall y B(x, y) \) with a recursive \( B \), and assume \( \forall x \exists y \neg B(x, y) \). Let \( f \) be the Skolem function \( f(x) := \min\{y \in \omega : \neg B(x, y)\} \). Then PA(\( f \)) \( \vdash \exists x \forall y B(x, f(x)) \) for the extended theory PA(\( f \)) obtained from PA by adding a unary function symbol \( f \). Applying the method, we get a witness \( n \) such that \( B(n, f(n)) \), which contradicts the assumption. Hence \( \exists x \forall y B(x, y) \) must be true. Here we need (a finite amount of) values of the function \( f \) to compute \( |e|_S \), and hence the correctness and the process \( S^0, S^1, \ldots \) are recursive in \( f \). Hence RFN\(_{\Sigma^0_1}\)(PA) follows from transfinite induction up to \( \varepsilon_0 \).

Each H-process for finite sets \( Cr \) of critical formulas provides a solution \( S \) for \( Cr \). Let \( e[x] \) denote an \( \epsilon \)-term obtained from an \( \epsilon \)-term \( e[t] \) occurring in \( Cr \) with a closed term \( t \). Informally \( e[x] \) denotes a number-theoretic function. Under \( S \) the \( \epsilon \)-term \( e[x] \) is interpreted as a function \( e_S : n \mapsto |e[n]|_S \) with a finite support. Namely the set \( \{n \in \omega : e_S(n) \neq 0\} \) is finite. Thus finitary objects are substituted for ideal concepts occurring in (formal) proofs. Note that this does not provide a model of, e.g., Skolemized arithmetic since finitary objects (given by solving substitutions) depend on proofs, i.e., on finite sets of critical formulas occurring in them. In this way epsilon substitution method realizes or is closer to Hilbert’s idea to eliminate ideal concepts as a figure of speech from (formal) proofs of real propositions.
2 Exact bounds

In this section we show that the length of the H-process up to reaching a solution is bounded by an ordinal recursive function. From the bound one can easily read off the bound for the provably recursive functions in PA.

2.1 Ordinal recursive functions

Let us recall the definition and facts on ordinal recursive functions in W. W. Tait[7].

Let $\prec_{\varepsilon_0}$ denote a standard well ordering of type $\varepsilon_0$. Assume that 0 is the least element in $\prec_{\varepsilon_0}$.

For each $\alpha \leq \varepsilon_0$, $\prec_{\alpha}$ denotes the initial segment of $\prec_{\varepsilon_0}$ of type $\alpha$. A number-theoretic function is said to be $\alpha$-recursive iff it is generated from the schemata for primitive recursive functions plus the following schema for introducing a function $f$ in terms of functions $g, h$ and $d$:

\[
 f(y, x) = \begin{cases} 
 g(y, x) & \text{if } d(y, x) \not<_{\alpha} x \\
 h(y, x, f(y, d(y, x))) & \text{if } d(y, x) <_{\alpha} x
\end{cases}
\]

A function is $<_{\alpha}$-recursive iff it is $\beta$-recursive for some $\beta < \alpha$.

W. W. Tait[7], p.163 shows that each class of $\alpha$-recursive functions is closed under the external recursion to introduce a function $f$ in terms of functions $g, h, d$, and $e$:

\[
 f(y, x) = \begin{cases} 
 g(y, x) & \text{if } e(y, d(y, x)) \not<_{\alpha} e(y, x) \\
 h(y, x, f(y, d(y, x))) & \text{if } e(y, d(y, x)) <_{\alpha} e(y, x)
\end{cases}
\]

2.2 Bounds

Given the finite sequence $C_r = \{C_{r_I} : I \leq N \}$ of epsilon axioms, let $\{S^n\}$ denote the H-process for $C_r$. Let $r_n = rk(S^n)$ and $a_n = ind(S^n)$.

Definition 2.1 Define inductively a consecutive series $\vec{S} = \{S^n\}_{m \leq n < k}$ in the H-process to be a $p$-series and a $p$-section as follows:

1. $\vec{S}$ is a 0-series iff $k = m + 1$, i.e., a singleton.
2. A $p$-series is a $p$-section iff it is a section.
3. $\vec{S}$ is a $(p+1)$-series iff there exist substrings $\vec{S}_j = (S^{k_j}, \ldots, S^{k_{j+1}-1}) (j \leq l)$ of $\vec{S}$ enjoying the following conditions:
   
   (a) $S = S_0 * \cdots * S_l (l \geq 0 \& k_{l+1} = k)$,
   
   (b) each $S_j$ is a $p$-section, and
   
   (c) $rk(S^{k_j}) \leq rk(S^{k_{j+1}}) < rk(S^{k_j})$.

Lemma 2.2 1. Each $p$-series is a $(p + 1)$-series.
2. Any \((p+1)\)-series can be decomposed uniquely to a sequence of \(p\)-series so that the sequence enjoys the conditions in Definition 2.1.

Such a sequence of \(p\)-series is said to be the decomposition of the \((p+1)\)-series.

3. Let \(\vec{S} = \{S^n\}_{m \leq n < k}\) be a \(p\)-section, and \(i \in (m, k)\) a number such that \(\{S^n\}_{i \leq n < k}\) is a section. Then \(S^i\) is the first term of an iterated decomposed substrings. Namely \(S^i\) is the first term of the substring of the decomposition of \(\vec{S}\), or the first term of the substring of the last \((p-1)\)-section in the decomposition and so on.

4. Let \(\vec{S}_i = \{S^n: m_i \leq n < k\} (i = 0, 1)\) be two \(p\)-series overlapped, i.e., \([m, k] \cap [m', k'] \neq \emptyset\). Then the union \(\vec{S} = \{S^n\}_{m \leq n < k} (m = \min\{m_0, m_1\}, k = \max\{k_0, k_1\})\) is a \(p\)-series.

5. Let us call a \(p\)-series proper if \(p = 0\) or it is not a \((p-1)\)-series.

(a) If \(\vec{S}\) is a \(p\)-section, then
\[\#rk(\vec{S}) := \#\{rk(S) : S \in \vec{S}\} \geq p + 1.\]

(b) If \(\vec{S}\) is a proper \(p\)-series, then
\[\#rk(\vec{S}) := \#\{rk(S) : S \in \vec{S}\} \geq p.\]

(c) If \(p > 0\) and a proper \(p\)-series \(\vec{S}\) begins with \(S^0 = \emptyset\), then
\[\#rk(\vec{S}) := \#\{rk(S) : S \in \vec{S}\} \geq p + 1.\]

Therefore there is no proper \(\Lambda\)-series beginning with \(S^0\) for \(\Lambda > rk(Cr) := \max\{rk(Cr) : I \leq N\} \).

**Proof.** By induction on \(p\).

2.2.1. A 0-series \(\{S^n\}\) is a 1-series.

2.2.2. Any consecutive series \(\vec{S} = \{S^n\}_{m \leq n < k}\) with \(rk(S^n) \leq rk(\vec{S})\) can be decomposed uniquely to a sequence of sections \(\vec{S}_j = (S^{k_j}, \ldots, S^{k_{j+1}-1})\) with nondecreasing ranks \(rk(S^{k_j})\) of the first terms \(S^{k_j}\).

2.2.3. This is seen by induction on \(p\).

2.2.4. Assume \(p > 0\) and one is not a substring of the other, i.e., \([m_i, k_i] \not\subseteq [m_{i-1}, k_{i-1}]\).

Decompose the \(p\)-series \(\vec{S}\) to the sequence of \((p-1)\)-serieses
\[\vec{S}_i^j = (S^{k_i}, \ldots, S^{k_{i+1}-1}) (j \leq l_i).\] It suffices to show that
\[k_j^i \geq \max\{m_0, m_1\} \text{ or } k_j^i \leq \min\{k_0, k_1\} \Rightarrow j' = k_j^i = k_j^i - j.\]

This is seen from the condition that each decomposition \(\{\vec{S}_i^j : j \leq l_i\}\) is a sequence of sections with nondecreasing ranks of the first terms.

2.2.5. Lemma 2.2.5a is easily seen by induction on \(p\).
Let \( \vec{S} = \vec{S}_0 \ast \cdots \ast \vec{S}_l \) be a proper \((p + 1)\)-series with the displayed decomposition. Then one of \( p \)-sections \( \vec{S}_j \) is proper. Hence Lemma 2.2.5b follows from Lemma 2.2.5a.

Lemma 2.2.5c is seen from the fact \( rk(S^n) > 0 \) for \( n > 0 \).

1. \( o_0(S) = ind(S) \).
2. \( o_{p+1}(\vec{S}) = \sum_{j \leq l} 2^{o_p(\vec{S}_j)} \) for a \((p + 1)\)-series \( \vec{S} = \vec{S}_0 \ast \cdots \ast \vec{S}_l \) with its decomposition \( \{ \vec{S}_j \}_{j \leq l} \).

We see from Lemma 1.12 that \( \vec{S}_{j+1} \prec \vec{S}_j \), and \( o_p(\vec{S}_{j+1}) < o_p(\vec{S}_j) \). Thus \( o_{p+1}(\vec{S}) \) is in Cantor normal form.

Lemma 2.3 Let \( \vec{S} = \{ S^n : m_i \leq n < k_i \} \) \((i = 0, 1)\) be two consecutive \( p \)-series, \( k_0 = m_1 \). Assume \( r_{m_1} \geq r_{m_0} \).

1. The concatenated series \( \vec{S} = \{ S^n : m_0 \leq n < k_1 \} \) is a \((p + 1)\)-series.
2. \( o_p(S^0) > o_p(S^1) \).

Proof. 2.3.1.

Let \( I = \max \{i < m_1 : r_i \leq r_{m_1} \} \). Then \( r_{m_1} < \min \{r_j : I < j < m_1 \} \). This means that \( \{ S^n : I < j < m_1 \} \) is a section. Therefore from Lemma 2.2.3 we see that \( \vec{S}' \) is the first term of an iterated decomposed substring unless \( I = m_0 \).

Decompose \( \vec{S}' \) to the \((p - 1)\)-sections in such a way that \( \vec{S}' \) is the first term of one of these sections, cf. Lemma 2.2.1.

If \( \vec{S}' \) is a section, then append \( \vec{S}' \) to the decomposed sections of \( S^0 \). The result is a desired decomposition of the \((p + 1)\)-series.

Otherwise decompose \( \vec{S}' \) to \((p - 1)\)-sections. Each substring is a \( p \)-section by Lemma 2.2.1. Append the decomposition of \( \vec{S}' \) to the decomposed sections of \( S^0 \). The result is a desired decomposition of the \((p + 1)\)-series.

2.3.2. This is seen from Lemma 1.12 and the proof of Lemma 2.3.1, i.e., the decompositions of \( S^0 \) in the proof.

Note here that we have \( 2^{o_p(\vec{S})} \geq o_{p+1}(\vec{S}) \) for any \( p \)-series \( \vec{S} \). This is seen from the fact \( 2^{\alpha + \beta} \geq 2^\alpha + 2^\beta \) if \( \alpha \geq \beta > 0 \).

Let \( L(0, n) = 1 \) be the length of the longest 0-series starting with \( S^n \).

Let \( L(p + 1, n) \) denote the length of the longest \((p + 1)\)-series starting with \( S^n \):

Case 0 \( S^n \) is solving: \( L(p + 1, n) = 0 \).

In what follows assume that \( S^n \) is nonsolving, and put \( k = L(p, n) \).

Case 1 \( r_{n+k} < r_n \): \( L(p + 1, n) = L(p, n) \).

Case 2 \( r_{n+k} \geq r_n \):

\[
L(p + 1, n) = k + L(p + 1, n + k) = L(p, n) + L(p + 1, n + L(p, n))
\]
Actually the function \( L(p, n) \) depends also on the given sequence \( Cr \) of epsilon axioms. We write \( L(p, n; Cr) \) for \( L(p, n) \) when the parameter \( Cr \) should be mentioned.

**Lemma 2.4**  
1. For each \( p \), the function \((n, Cr) \mapsto L(p, n; Cr)\) is \( \omega_{1+p} \)-recursive.

2. Let \( M(p, n) := n + L(p, n) \). Then \( \{S^i : n \leq i < M(p, n)\} \) is a \( p \)-series.

3. 
\[
  n \leq k < M(p, n) \Rightarrow M(p, k) \leq M(p, n).
\]

4. \( L(p, n) \) is equal to the length of the longest \( p \)-series starting with \( S^n \).

5. Let \( H = H(Cr) := L(\Lambda - 1, 0; Cr) \) for \( rk(Cr) < \Lambda \). Then \( S^H \) is a solution for \( Cr \).

Thus a solution \( S^H(Cr) \) for \( Cr \) can be found by \( \omega_{\Lambda} \)-recursion for \( rk(Cr) < \Lambda \).

**Proof.** By simultaneous induction on \( p \).

2.4.1. The definition of the function \((n, Cr) \mapsto L(p, n; Cr)\) is based on external recursion on \( \omega_{1+p} \). This is seen from Induction Hypothesis, Lemma 2.3.2 and \( \sigma_{p-1}(\bar{S}) < \omega_{1+p} \).

2.4.2. This follows from Lemma 2.3.1.

2.4.3. This follows from Lemmas 2.3.1 and 2.3.2.

2.4.4. This is seen from Lemma 2.4.3.

2.4.5. Assume \( rk(Cr) < \Lambda \). By Lemma 2.2.5c we have \( L(\Lambda, 0; Cr) = L(\Lambda - 1, 0; Cr) = H \). This mean that \( S^H \) is a solution for \( Cr \) by \( r_0 = rk(S^0) = 0 \).

**Lemma 2.5** The function \((p, n, Cr) \mapsto L(p, n; Cr)\) is \( \varepsilon_0 \)-recursive.

**Proof.** The definition of \( L(p, n; Cr) \) is based on nested recursion on the ordinal \( \omega_{\Lambda} \) for \( rk(Cr) < \Lambda \). Therefore it is \( \omega_{1+\Lambda} \)-recursive by a result in W. W. Tait[8].

Next let us give a bound \( B(n) \) for the values \(|e|_{S^n}\), under \( S^n \), of the closed terms \( e \) occurring in \( Cr \) and \( e \in dom(S^n) \). For this, the list of function symbols in the language \( L(PA) \) has to be specified. For example let us assume that the exponential function \( 2^x \) bounds each function: for each function (symbol) \( f \) there exists a constant \( c_0 \) such that \( |f(x)| \leq 2^{x + c_0} \). Let \( c \) be big enough so that for any closed term \( e \) occurring in \( Cr \), \(|e|_S \leq 2_c(x)\) if \( S \) is a substitution such that \( max\{|e_0|_S : e_0 \in dom(S)\} \leq x \).

Pick a bound \( B(0) \) for the given \( Cr \). An inspection on the definition of the \( H \)-process shows that \( B(n) \) bounds the values \(|e|_{S^{n+1}}\) for \( e \in dom(S^{n+1}) \). Therefore \( B(n+1) = 2_c(B(n)) \) is a desired one.

Now suppose that a \( \Pi^0_3 \)-sentence \( \forall x \exists y R(x, y) \) (\( R \): quantifier-free) is provable in \( PA \). For each \( n \), first translate the \( PA \)-proof of \( \exists y R(n, y) \) (\( n \) denotes the \( n \)-th numeral) in \( PA \), and let \( Cr(n) \) denote a sequence of epsilon axioms occurring
in the PA-proof of $R(n, \epsilon.R(\bar{n}, y))$. Then the bound $B(0)$ for $Cr(n)$ is an elementary recursive function, i.e., $B(0) = 2^n(n)$ for a $c$. On the other side, the bound $\Lambda$ for ranks is independent from $n$. Thus $f(n) := B(L(\Lambda - 1, 0; Cr(n)))$ bounds the value of the epsilon term $\epsilon.y.R(\bar{n}, y)$ under the solution for $Cr(n)$. Therefore $\forall n \exists y \leq f(n) R(n, y)$ holds for the $<\omega_0$-recursive function $f$.

### 3 Recent progress

In the final section let us report the recent progress on epsilon substitution method à la Ackermann.

In [2] an epsilon substitution method for theories $(H)_{\alpha_0}$ of absolute jump hierarchies is formulated. The theory $(H)_{\alpha_0}$ is obtained from PA by adding the axiom for absolute jump hierarchies by a quantifier free formula $A(x, y, z, X)$:

$$\alpha \leq \alpha_0 \rightarrow \{z \in H_{\alpha} \leftrightarrow \exists x A(x, \alpha, z, H_{<\alpha})\}.$$  

In [3] the same is done for the theory $ID_1(\Pi^0_1 \lor \Sigma^0_1)$ of non-iterated inductive definitions for disjunctions of simply universal and existential operators. In [4], the method is simplified for the theory $\Pi^0_1$-FIX of non-monotonic $\Pi^0_1$ inductive definitions. It is well known that the theories $ID_1$ of non-iterated inductive definitions for arbitrary positive arithmetical operator, $ID_1(\Pi^0_1 \lor \Sigma^0_1)$, $\Pi^0_1$-FIX and the Kripke-Platek set theory KP$_{\alpha}$ with the axiom of Infinity are mutually equivalent proof-theoretically, and their proof-theoretic ordinal is the Howard ordinal.

[5] gives an epsilon substitution method for the theory $[\Pi^0_1, \Pi^0_1]$-FIX of two steps non-monotonic $\Pi^0_1$ inductive definitions. The theory is proof-theoretically equivalent to the set theory KPM for recursively Mahlo universes.

Recently an epsilon substitution method for the theory $\Pi^0_2$-FIX of non-monotonic $\Pi^0_2$ inductive definitions is formulated in [6]. The theory is proof-theoretically equivalent to the set theory KPH$_{3}$ for $\Pi^0_3$-reflecting universes.

### References


[3] –, Epsilon substitution method for $ID_1(\Pi^0_1 \lor \Sigma^0_1)$, Ann. Pure Appl. Logic 121 (2003), 163-208.


[6] –, Epsilon substitution method for $\Pi^0_2$-FIX, draft.
