The proof-theoretic analysis of the Suslin operator in applicative theories

Gerhard Jäger* Thomas Strahm*

Abstract

In this article we introduce the Suslin quantification functional E_1 into the framework of Feferman's explicit mathematics and analyze it from the point of view of proof theory. More precisely, we work in the first order part of explicit mathematics augmented by appropriate axioms for E_1. Then we establish the exact proof-theoretic relationship between these applicative theories and (subsystems of) the second order theory $(\Delta^1_2\text{-}\mathsf{CA})$, depending on the induction principles permitted.

1 Introduction

During the last decades a series of interesting axiomatic frameworks for constructive mathematics have been proposed. Depending on the philosophical position of their inventors, numerous routes have been pursued, some of which are quite different in nature. Comprehensive surveys of many of these approaches are given in the textbooks by Beeson [1] and Troelstra and Van Dalen [23, 24].

In the seventies Feferman developed his own approach towards constructive mathematics and was particularly interested in laying the logical foundations of Bishop-style constructive mathematics. In his seminal paper Feferman [4] he introduces the famous theory T_0 and more generally the program of *explicit mathematics*. Feferman [8] discusses the relationships between systems of explicit mathematics and alternative approaches towards constructivism on the one hand side and to subsystems of second order arithmetic on the other hand. It soon turned out that T_0 and its subsystems play an important role in reductive proof theory and, in particular, in reducing classical systems to constructively justifiable ones.

*Research partly supported by the Swiss National Science Foundation.

Feferman's development of explicit mathematics is strongly influenced by his work on generalized recursion theory and also motivated by the idea of abstract computations. It is one of the attractive features of explicit mathematics that it possesses models in the sense of classical set theory as well as models which are constructive/computational in flavor. As a consequence, we can directly relate the classical meaning of a sentence provable in explicit mathematics to its computational interpretation. This approach is brought up and studied in Feferman [7].

A further philosophically interesting approach towards mathematics goes back to work of Poincaré and Weyl and is generally referred to as the *arithmetical* foundations of mathematics. Feferman's logical analysis of this approach in Feferman [5, 9] reveals the crucial role of the so-called non-constructive minimum operator μ in this context.

Later it turned out that this work also has its natural place in explicit mathematics. Since the operations of explicit mathematics can be regarded as abstract computations, functionals of higher types can be added in a direct and perspicuous way. In explicit mathematics we basically deal with first and second order entities: *operations* which can be understood as abstract computations and *classifications* or *types* which are collections of operations. For the proof-theoretic analysis of functionals of higher types in this framework already the first-order applicative core of explicit mathematics is of significant interest. The corresponding axiomatic first order systems are often referred to as *applicative theories*. For a survey of this area see Jäger, Kahle and Strahm [18].

A first important step in the proof-theoretic treatment of functionals of higher types in the framework of applicative theories and explicit mathematics was the analysis of the non-constructive μ operator over the basic theory BON of operations and numbers in Feferman and Jäger [11]. In the subsequent articles Feferman and Jäger [12] and Glaß and Strahm [13] type structures which are closed under elementary comprehension and join have been added to $\mathsf{BON}(\mu)$.

The non-constructive μ operator is also very interesting from the point of view of type two recursion theory and corresponds to the number quantifier functional E_0. The 1-section of μ and of E_1 consists exactly of the hyperarithmetic sets and, hence, provides the least standard model of Δ^1_1 comprehension,

$$(\mathbb{N}, \mathsf{1\text{-}sec}(\mu), \dots) \models (\Delta^1_1\text{-}\mathsf{CA}).$$

Moreover, the least ordinal which is not recursive in E_0 is the ordinal ω_1^{ck}, the first admissible ordinal greater than ω.

It is common practice in applicative theories to represent subsets of the natural numbers as total function from natural numbers to $\{0,1\}$. From the recursion-theoretic results we can therefore immediately read off that the sets of $\mathsf{BON}(\mu)$ in its least standard model coincide with the hyperarithmetic sets.

Proof-theoretically, however, $\mathsf{BON}(\mu)$ is comparatively weak provided that induction on the natural numbers is restricted to sets: $\mathsf{BON}(\mu)$ plus set induction $(\mathsf{S}\text{-}\mathsf{I}_\mathsf{N})$ is a conservative extension of Peano arithmetic PA and so proof-theoretically equivalent to the theory $(\Delta^1_1\text{-}\mathsf{CA})\!\restriction$. If the schema $(\mathrm{L}\text{-}\mathsf{I}_\mathsf{N})$ of complete induction for all formulas of the underlying language L is permitted then $\mathsf{BON}(\mu)$ becomes equivalent to $(\Delta^1_1\text{-}\mathsf{CA})$:

$$\mathsf{BON}(\mu) + (\mathsf{S}\text{-}\mathsf{I}_\mathsf{N}) \;\equiv\; (\Delta^1_1\text{-}\mathsf{CA})\!\restriction, \qquad \mathsf{BON}(\mu) + (\mathrm{L}\text{-}\mathsf{I}_\mathsf{N}) \;\equiv\; (\Delta^1_1\text{-}\mathsf{CA}).$$

According to the articles by Feferman, Jäger, Glaß and Strahm cited above, even elementary comprehension and join may be added to BON plus set induction without going beyond PA with respect to proof-theoretic strength.

$\mathsf{BON}(\mu)$ plus set induction and with or without elementary comprehension and join provides a solid logical basis for the Poincare-Weyl approach. It is also an interesting system in connection with a series of other formalisms introduced and studied in Feferman [6], where the problem of designing adequate higher type formalisms for mathematical practice is addressed. In this article Feferman also points out the relevance of the Suslin quantifier functional E_1.

This type two functional tests for wellfoundedness of binary relations and thus has at least the power of Π^1_1 comprehension. The 1-section of E_1 coincides with the sets of natural numbers in the constructible hierarchy up the the first recursively inaccessible ordinal ι_0. This ordinal is also the least ordinal not recursive in E_1. According to a result of Gandy, one can also show that the 1-section of E_1 builds the least standard model of Δ^1_2 comprehension,

$$(\mathbb{N}, \text{1-}\mathsf{sec}(\mathsf{E}_1), \dots) \models (\Delta^1_2\text{-}\mathsf{CA}).$$

In this article we will analyze the proof-theoretic strength of E_1 in the context of applicative theories. The theory $\mathsf{BON}(\mu)$ with additional axioms for E_1 is baptized SUS in the following. Then we show the proof-theoretic equivalences

$$\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I}_\mathsf{N}) \;\equiv\; (\Delta^1_2\text{-}\mathsf{CA})\!\restriction, \qquad \mathsf{SUS} + (\mathrm{L}\text{-}\mathsf{I}_\mathsf{N}) \;\equiv\; (\Delta^1_2\text{-}\mathsf{CA}).$$

We also obtain results concerning intermediate forms of induction and their relationship to systems like $(\Delta^1_2\text{-}\mathsf{CR})$. Obviously, there is a striking analogy concerning the relationship between proof-theoretic and recursion-theoretic results for μ and E_1.

The methods of proof, however, are rather different. For establishing the upper bounds of, say, $\mathsf{SUS} + (\mathsf{S\text{-}I_N})$ we make use of a very specific positive Δ^1_2 inductive definition. We interpret the application operation by a Σ definable fixed point of this inductive definition and have to exploit a delicate interplay between proper set-theoretic functions and functions defined in terms of our application relation, cf. Theorem 16 below.

The plan of the paper is a follows. In the next section we introduce the basic applicative framework, the axiomatizations of μ and E_1 and several forms of complete induction on the natural numbers. In Section 3 we establish proof-theoretic lower bounds of the various applicative theories by embedding subsystems of analysis with (iterated) Π^1_1 comprehension. Section 4 constitutes the central part of the paper. We obtain proof-theoretic upper bounds for SUS plus various forms of induction in appropriate theories for iterated admissible sets.

2 The applicative framework and E_1

It is the purpose of this section to introduce the basic applicative framework as well as the precise axiomatizations of the non-constructive μ operator and the Suslin operator E_1. Further, we will distinguish three forms of complete induction on the natural numbers which will be relevant in the sequel.

The language of our applicative theories is a first order language L of partial terms with individual variables $a, b, c, f, g, h, u, v, w, x, y, z \dots$ (possibly with subscripts). L includes individual constants k, s (combinators), $\mathsf{p}, \mathsf{p_0}, \mathsf{p_1}$ (pairing and unpairing), 0 (zero), $\mathsf{s_N}$ (numerical successor), $\mathsf{p_N}$ (numerical predecessor), $\mathsf{d_N}$ (definition by numerical cases), $\mathsf{r_N}$ (primitive recursion), μ (non-constructive μ operator), and E_1 (Suslin operator). Further, L has a binary function symbol $\cdot$ for (partial) term application, unary relation symbols $\downarrow$ (defined) and N (natural numbers), as well as a binary relation symbol $=$ (equality).

The *individual terms* $(r, s, t, r_1, s_1, t_1, \dots)$ of L are inductively generated as follows:

1. The individual variables and individual constants are individual terms.

2. If s and t are individual terms, then so also is $(s \cdot t)$.

In the following we often abbreviate $(s \cdot t)$ simply as (st), st or sometimes also $s(t)$; the context will always ensure that no confusion arises. We further adopt the convention of association to the left so that $s_1 s_2 \dots s_n$ stands for $(\dots(s_1 s_2)\dots s_n)$. Further, we put $t' := \mathsf{s_N} t$ and $1 := 0'$. We define general n-tupling by induction on $n \geq 2$ as follows:

$$(s_1, s_2) := \mathsf{p} s_1 s_2, \qquad (s_1, \dots, s_{n+1}) := ((s_1, \dots, s_n), s_{n+1}).$$

Finally, we also use quite frequently the vector notation $\vec{\mathcal{Z}}$ for a finite string of objects $\mathcal{Z}_1, \dots, \mathcal{Z}_n$ of the same sort. Whenever we write $\vec{\mathcal{Z}}$ the length of this string is either irrelevant or given by the context.

The *formulas* $(A, B, C, A_1, B_1, C_1, \dots)$ of L are inductively generated as follows:

1. Each atomic formula $\mathsf{N}(t)$, $t\downarrow$, and $(s = t)$ is a formula.
2. If A and B are formulas, then so also are $\neg A$, $(A \vee B)$, $(A \wedge B)$, and $(A \rightarrow B)$.
3. If A is a formula, then so also are $(\exists x)A$ and $(\forall x)A$.

Our applicative theories are based on *partial* term application. Hence, it is not guaranteed that terms have a value, and $t\downarrow$ is read as "t is defined" or "t has a value". Accordingly, the *partial equality relation* $\simeq$ is introduced by

$$s \simeq t \ := \ (s\downarrow \vee\, t\downarrow) \rightarrow (s = t).$$

In addition, we write $(s \neq t)$ for $(s\downarrow \wedge\, t\downarrow \wedge\, \neg(s = t))$. Finally, we use the following abbreviations concerning the predicate N:

$$\begin{aligned}
t \in \mathsf{N} \ &:= \ \mathsf{N}(t), \\
(\exists x \in \mathsf{N})A \ &:= \ (\exists x)(x \in \mathsf{N} \wedge A), \\
(\forall x \in \mathsf{N})A \ &:= \ (\forall x)(x \in \mathsf{N} \rightarrow A), \\
t \in (\mathsf{N} \rightarrow \mathsf{N}) \ &:= \ (\forall x \in \mathsf{N})(tx \in \mathsf{N}), \\
t \in (\mathsf{N}^{m+1} \rightarrow \mathsf{N}) \ &:= \ (\forall x \in \mathsf{N})(tx \in (\mathsf{N}^m \rightarrow \mathsf{N})).
\end{aligned}$$

Now we are going to recall the basic theory BON of operations and numbers which has been introduced in Feferman and Jäger [11]. Its underlying logic is the *classical logic of partial terms* due to Beeson [1]; it is also described in Feferman [10] and corresponds to E^+ logic with strictness and equality of Troelstra and Van Dalen [23]. The non-logical axioms of BON are divided into the following five groups.

I. Partial combinatory algebra.

(1) $\mathsf{k}ab = a$,

(2) $\mathsf{s}ab\downarrow \wedge \mathsf{s}abc \simeq ac(bc)$.

II. Pairing and projection.

(3) $\mathsf{p}_0(a,b) = a \wedge \mathsf{p}_1(a,b) = b$.

III. Natural numbers.

(4) $0 \in \mathsf{N} \wedge (\forall x \in \mathsf{N})(x' \in \mathsf{N})$,

(5) $(\forall x \in \mathsf{N})(x' \neq 0 \wedge \mathsf{p_N} x' = x)$,

(6) $(\forall x \in \mathsf{N})(x \neq 0 \rightarrow \mathsf{p_N} x \in \mathsf{N} \wedge (\mathsf{p_N} x)' = x)$.

IV. Definition by numerical cases.

(7) $u \in \mathsf{N} \wedge v \in \mathsf{N} \wedge u = v \rightarrow \mathsf{d_N} abuv = a$,

(8) $u \in \mathsf{N} \wedge v \in \mathsf{N} \wedge u \neq v \rightarrow \mathsf{d_N} abuv = b$.

V. Primitive recursion on N.

(9) $f \in (\mathsf{N} \rightarrow \mathsf{N}) \wedge g \in (\mathsf{N}^3 \rightarrow \mathsf{N}) \rightarrow \mathsf{r_N} fg \in (\mathsf{N}^2 \rightarrow \mathsf{N})$,

(10) $f \in (\mathsf{N} \rightarrow \mathsf{N}) \wedge g \in (\mathsf{N}^3 \rightarrow \mathsf{N}) \wedge a \in \mathsf{N} \wedge b \in \mathsf{N} \wedge h = \mathsf{r_N} fg \rightarrow$
$ha0 = fa \wedge ha(b') = gab(hab)$.

As usual the axioms of a partial combinatory algebra allow one to define λ abstraction and to prove a recursion or fixed point theorem. For proofs of these standard results the reader is referred to [1, 4]. For the second assertion of the following lemma, which is a slight extension of the usual λ abstraction, one makes in addition use of pairing and projections.

Lemma 1 *1. For each* L *term t and all variables x there exists an* L *term $(\lambda x.t)$ whose variables are those of t, excluding x, so that* BON *proves*

$$(\lambda x.t)\downarrow \wedge (\lambda x.t)x \simeq t.$$

2. For each L *term t and all variables $x_1, \ldots, x_n (n \geq 2)$ there exists an* L *term s whose variables are those of t, excluding $x_1, \ldots, x_n$, so that* BON *proves*

$$s\downarrow \wedge s(x_1, \ldots, x_n) \simeq t.$$

3. There exists a closed L *term* rec *so that* BON *proves*

$$\mathsf{rec} f\downarrow \wedge\ \mathsf{rec} f x \simeq f(\mathsf{rec} f)x.$$

Let us now turn to the two type 2 functionals which will be relevant in the sequel. The non-constructive or unbounded μ operator is characterized by the following two axioms.

The non-constructive μ operator

(μ.1) $\quad f \in (\mathsf{N} \to \mathsf{N}) \leftrightarrow \mu f \in \mathsf{N}$,

(μ.2) $\quad f \in (\mathsf{N} \to \mathsf{N}) \wedge (\exists x \in \mathsf{N})(fx = 0) \to f(\mu f) = 0$.

A much stronger functional is the Suslin operator E_1, which tests for the wellfoundedness of a binary relation on N (given as a total operation from N^2 to N).

The Suslin operator E_1

(E_1.1) $\quad f \in (\mathsf{N}^2 \to \mathsf{N}) \leftrightarrow \mathsf{E}_1 f \in \mathsf{N}$,

(E_1.2) $\quad f \in (\mathsf{N}^2 \to \mathsf{N}) \to$
$\qquad [(\exists g \in \mathsf{N} \to \mathsf{N})(\forall x \in \mathsf{N})(f(gx')(gx) = 0) \leftrightarrow \mathsf{E}_1 f = 0]$.

Since we want to study E_1 also in the presence of very weak induction principles, we include the non-constructive μ operator in our basic axiomatic framework for the Suslin operator. Accordingly, we let SUS denote the L theory which extends BON by the axioms about μ and E_1,

$$\mathsf{SUS} := \mathsf{BON} + (\mu.1) + (\mu.2) + (\mathsf{E}_1.1) + (\mathsf{E}_1.2).$$

In the sequel we will be interested in three forms of complete induction on the natural numbers N, namely *set induction,* N *induction* and *formula induction.*

Let us first recall the notion of a *subset of* N from [5, 11]. Sets of natural numbers are represented via their characteristic functions which are total on N. Accordingly, we define

$$f \in \mathcal{P}(\mathsf{N}) := (\forall x \in \mathsf{N})(fx = 0 \vee fx = 1)$$

with the intention that an object x belongs to the set $f \in \mathcal{P}(\mathsf{N})$ if and only if $(fx = 0)$. The three relevant induction principles are now given as follows.

Set induction on N $\quad$ (S-I_N).

$$f \in \mathcal{P}(\mathsf{N}) \wedge f0 = 0 \wedge (\forall x \in \mathsf{N})(fx = 0 \to fx' = 0) \to (\forall x \in \mathsf{N})(fx = 0).$$

N induction on N (N-$\mathsf{I_N}$).

$$f0 \in \mathsf{N} \wedge (\forall x \in \mathsf{N})(fx \in \mathsf{N} \to fx' \in \mathsf{N}) \to (\forall x \in \mathsf{N})(fx \in \mathsf{N}).$$

Formula induction on N (L-$\mathsf{I_N}$). For all formulas $A(x)$ of L:

$$A(0) \wedge (\forall x \in \mathsf{N})(A(x) \to A(x')) \to (\forall x \in \mathsf{N})A(x).$$

In the rest of this paper we will be concerned with the proof-theoretic analysis of the three systems

$$\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I_N}), \qquad \mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N}), \qquad \mathsf{SUS} + (\mathrm{L}\text{-}\mathsf{I_N}).$$

In the next section we will establish lower proof-theoretic bounds by embedding suitable subsystems of analysis into these applicative systems.

3 Lower bounds

Lower bounds for the theory SUS plus various forms of induction on the natural numbers will be established in this section by embedding suitable systems of second order arithmetic. Actually, theories with Π^1_1 comprehension and transfinitely iterated forms thereof will play a crucial role. More precisely, we will show that $\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I_N})$ contains $(\Pi^1_1\text{-}\mathsf{CA})\!\restriction$, $\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N})$ contains $(\Pi^1_1\text{-}\mathsf{CA})_{<\omega^\omega}\!\restriction$, and $\mathsf{SUS} + (\mathrm{L}\text{-}\mathsf{I_N})$ contains $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}\!\restriction$.

3.1 Second order arithmetic with Π^1_1 comprehension

In the following we introduce the notions and results of second order arithmetic which will be important for our embeddings. We deviate from the usual conventions in working with a form of second order arithmetic with set and function variables. The reason for this choice is that the following Π^1_1 normal form theorem is most suitable for our purposes.

Let $\mathcal{L}_2$ denote a language of second order arithmetic with *number variables* $a, b, c, f, g, h, u, v, w, x, y, z, \dots$, *set variables* $U, V, W, X, Y, Z, \dots$. and *function variables* $F, G, H, \dots$ (all possibly with subscripts). In addition, $\mathcal{L}_2$ includes a constant 0 as well as function and relation symbols for all primitive recursive functions and relations. The *number terms* $(r, s, t, r_1, s_1, t_1, \dots)$ of $\mathcal{L}_2$ and the *formulas* $(A, B, C, A_1, B_1, C_1, \dots)$ of $\mathcal{L}_2$ are defined as usual.

An $\mathcal{L}_2$ formula is called *arithmetic*, if it does not contain bound set or function variables; let Π^1_0 denote the class of arithmetic $\mathcal{L}_2$ formulas. The Π^1_1 [Σ^1_1]

formulas of $\mathcal{L}_2$ are obtained from the arithmetic formulas by closing under universal [existential] set and function quantification.

In the following we make use of the usual primitive recursive coding machinery in $\mathcal{L}_2$: $\langle \ldots \rangle$ is a standard primitive recursive function for forming n-tuples $\langle t_1, \ldots, t_n \rangle$; Seq is the primitive recursive set of sequence numbers; $lh(t)$ denotes the length of (the sequence number coded by) t; $(t)_i$ is the ith component of (the sequence coded by) t if $i < lh(t)$, i.e. $t = \langle (t)_0, \ldots, (t)_{lh(t) \dot{-} 1} \rangle$ if t is a sequence number. In addition, we write $s \in (U)_t$ for $\langle s, t \rangle \in U$.

If $\mathcal{F}$ is a collection of $\mathcal{L}_2$ formulas, then $\mathcal{F}$ comprehension ($\mathcal{F}$-CA) is the schema

$$(\exists X)(\forall x)(x \in X \leftrightarrow A(x))$$

for all formulas $A(u)$ in the collection $\mathcal{F}$. The relationship between sets and functions is given by the so-called graph principle ($\mathcal{GP}$),

$$(\forall X)[(\forall x)(\exists! y)\langle x, y \rangle \in X \;\rightarrow\; (\exists F)(\forall x)\langle x, F(x) \rangle \in X].$$

Moreover, $\mathcal{L}_2$ induction on the natural numbers ($\mathcal{L}_2$-$\mathsf{I_N}$) comprises

$$A(0) \wedge (\forall x)(A(x) \rightarrow A(x')) \;\rightarrow\; (\forall x)A(x)$$

for all formulas $A(u)$ of $\mathcal{L}_2$. Set induction (S-$\mathsf{I_N}$) with respect to $\mathcal{L}_2$, on the other hand, is the axiom

$$(\forall X)(0 \in X \wedge (\forall x)(x \in X \rightarrow x' \in X) \;\rightarrow\; (\forall x)(x \in X)).$$

(Π^1_0-CA) is the $\mathcal{L}_2$ theory which contains the usual axioms of Peano arithmetic PA, all instances of Π^1_0 comprehension, the graph principle ($\mathcal{GP}$) as well as formula induction ($\mathcal{L}_2$-$\mathsf{I_N}$). The restricted version (Π^1_0-CA)$\restriction$ of (Π^1_0-CA) is obtained if we replace ($\mathcal{L}_2$-$\mathsf{I_N}$) by (S-$\mathsf{I_N}$). The theories (Π^1_1-CA) and (Π^1_1-CA)$\restriction$ are defined accordingly.

We now state a version of the normal form theorem for Π^1_1 formulas tailored for our later purposes. Its proof is more or less folklore and can be found at many places (for example in Simpson [22]).

Theorem 2 (Π^1_1 normal forms) *For every Π^1_1 formula A there exists an arithmetic formula $B_A(u, v)$ which contains the free variables of A plus two fresh variables u and v so that (Π^1_0-CA)$\restriction$ proves*

$$A \;\leftrightarrow\; \neg(\exists F)(\forall x)B_A(F(x{+}1), F(x)).$$

Now let $\prec$ be a standard wellordering of ordertype ε_0. In addition we assume that 0 is the least element with respect to $\prec$ and the field of $\prec$ is the set of natural numbers. Furthermore, if n is a natural number, then we write $\prec_n$ for the restriction of $\prec$ to the numbers $m \prec n$.

For each Π^1_1 formula $A(U, v)$ with all its free variables in U, v, the *A-hyperjump hierarchy along $\prec_n$ starting with U* is defined by the following transfinite recursion,

$$(W)_0 := U \quad \text{and} \quad (W)_i := \{\langle m, j\rangle : j \prec i \wedge A((W)_j, m)\}$$

for all $0 \prec i \prec n$. We write $\mathsf{Hier}_A(U, W, n)$ for the $\mathcal{L}_2$ formula which formalizes this definition.

If α is an ordinal less than ε_0, then $(\Pi^1_1\text{-}\mathsf{CA})_{<\alpha}{\restriction}$ is the $\mathcal{L}_2$ theory which extends $(\Pi^1_0\text{-}\mathsf{CA}){\restriction}$ by the axioms

$$(\forall X)(\exists Y)\mathsf{Hier}_A(X, Y, n)$$

for each natural number n so that $\prec_n$ has order type less than α and each Π^1_1 formula $A(U, v)$ with all its free variables in U, v. Moreover, $(\Pi^1_1\text{-}\mathsf{CA})_{<\alpha}$ denotes $(\Pi^1_1\text{-}\mathsf{CA})_{<\alpha}{\restriction}$ plus $(\mathcal{L}_2\text{-}\mathsf{I}_\mathsf{N})$.

Lemma 3 *The theories $(\Pi^1_1\text{-}\mathsf{CA})_{<\omega^\omega}$ and $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}$ are proof-theoretically equivalent to $(\Pi^1_1\text{-}\mathsf{CA})_{<\omega^\omega}{\restriction}$ and $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}{\restriction}$, respectively.*

The proof of this lemma makes use of standard recursion-theoretic arguments. Actually, the lemma holds for ordinals of the form ω^λ with λ limit.

3.2 Modeling hyperjumps via E_1

Recursion-theoretic arguments clearly indicate that E_1 can be used to deal with hyperjumps in an applicative framework. For this purpose, we work with the natural embedding of $\mathcal{L}_2$ into L so that (i) the number variables of $\mathcal{L}_2$ are interpreted as ranging over N, (ii) the set variables of $\mathcal{L}_2$ as ranging over $\mathcal{P}(\mathsf{N})$, and (iii) the function variables as ranging over $(\mathsf{N} \to \mathsf{N})$.

In the following we assume that we have a translation of the number, set and function variables of $\mathcal{L}_2$ into the variables of L so that no conflicts arise. For convenience we often simply write, for example, a, x, f for the translations of the number, set and function variable a, X, F, respectively. Furthermore, we can use the recursion operator $\mathsf{r_N}$ to associate a suitable L term to each symbol for a primitive recursive function on the natural numbers and prove the corresponding recursion equations in $\mathsf{BON} + (\mathsf{S}\text{-}\mathsf{I_N})$. Thus, every $\mathcal{L}_2$ term

t has a canonical translation t^{N} in L. Similarly, each symbol for a primitive recursive relation on N can be represented by an L term which represents its characteristic function in the sense above.

Now let R be a symbol for an n-ary primitive recursive relation and t_R the corresponding L term. If $s, t_1, \dots, t_n$ are terms of $\mathcal{L}_2$, then the atomic formulas of $\mathcal{L}_2$ are translated into L formulas as follows:

$$(s \in U)^{\mathsf{N}} := ((us^{\mathsf{N}}) = 0); \qquad (t_1 = t_2)^{\mathsf{N}} := (t_1^{\mathsf{N}} = t_2^{\mathsf{N}});$$
$$R(t_1, \dots, t_n)^{\mathsf{N}} := (t_R t_1^{\mathsf{N}} \dots t_n^{\mathsf{N}} = 0).$$

We extend this translation in the usual way and associate to each $\mathcal{L}_2$ formula $A(\vec{U}, \vec{F}, \vec{v})$ an L formula $A^{\mathsf{N}}(\vec{u}, \vec{f}, \vec{v})$ such that

$$\begin{aligned} ((\exists X)A(X))^{\mathsf{N}} &= (\exists x \in \mathcal{P}(\mathsf{N}))A^{\mathsf{N}}(x), \\ ((\exists F)A(F))^{\mathsf{N}} &= (\exists f \in \mathsf{N} \to \mathsf{N})A^{\mathsf{N}}(f), \\ ((\exists y)A(y))^{\mathsf{N}} &= (\exists y \in \mathsf{N})A^{\mathsf{N}}(y) \end{aligned}$$

and similarly for universal quantifiers. A further convention is that we often identify $\mathcal{L}_2$ terms and arithmetic $\mathcal{L}_2$ formulas with their translation in L as long as no conflict arises.

Usual arguments as for example in Feferman [5] and Feferman and Jäger [11] show that with help of the unbounded minimum operator μ each arithmetic formula of $\mathcal{L}_2$ can be coded up by an L term in the sense of the following lemma.

Lemma 4 *For every arithmetic formula $A(\vec{U}, \vec{F}, \vec{v})$ of $\mathcal{L}_2$ with all its free variables in $\vec{U}, \vec{F}, \vec{v}$ there exists a closed individual term t_A of* L *so that* $\mathsf{BON}(\mu) + (\mathsf{S\text{-}I_N})$ *proves*

1. $(\forall \vec{x} \in \mathcal{P}(\mathsf{N}))(\forall \vec{f} \in \mathsf{N} \to \mathsf{N})(\forall \vec{y} \in \mathsf{N})(t_A(\vec{x}, \vec{f}, \vec{y}) = 0 \lor t_A(\vec{x}, \vec{f}, \vec{y}) = 1)$,
2. $(\forall \vec{x} \in \mathcal{P}(\mathsf{N}))(\forall \vec{f} \in \mathsf{N} \to \mathsf{N})(\forall \vec{y} \in \mathsf{N})(A(\vec{x}, \vec{f}, \vec{y}) \leftrightarrow t_A(\vec{x}, \vec{f}, \vec{y}) = 0)$.

This lemma is the crucial step for the interpretation of $(\Pi^1_0\text{-}\mathsf{CA})\!\restriction$ modulo the embedding described above. It is even the case that $(\Pi^1_0\text{-}\mathsf{CA})\!\restriction$ and $\mathsf{BON}(\mu) + (\mathsf{S\text{-}I_N})$ have the same proof-theoretic strength; cf. Feferman and Jäger [11]. For the following, however, we only need that $(\Pi^1_0\text{-}\mathsf{CA})\!\restriction$ is contained in $\mathsf{BON}(\mu) + (\mathsf{S\text{-}I_N})$.

Lemma 5 *Let $A(\vec{U}, \vec{F}, \vec{v})$ be an $\mathcal{L}_2$ formula with all its free variables in $\vec{U}, \vec{F}, \vec{v}$ and assume that $(\Pi^1_0\text{-}\mathsf{CA})\!\restriction$ proves $A(\vec{U}, \vec{F}, \vec{v})$. Then we have*

$$\mathsf{BON}(\mu) + (\mathsf{S\text{-}I_N}) \vdash \vec{u} \in \mathcal{P}(\mathsf{N}) \land \vec{f} \in (\mathsf{N} \to \mathsf{N}) \land \vec{v} \in \mathsf{N} \to A^{\mathsf{N}}(\vec{u}, \vec{f}, \vec{v}).$$

Proof All axioms of Peano arithmetic PA without the schema of complete induction obviously create no problems. Set induction of $(\Pi^1_0\text{-}\mathsf{CA})\!\upharpoonright$ directly translates into set induction in the applicative context. The schema of arithmetic comprehension follows from the lemma above. So there only remains the graph principle $(\mathcal{GP})$. Hence, let a be an element of $\mathcal{P}(\mathsf{N})$ so that

$$(\forall x \in \mathsf{N})(\exists y \in \mathsf{N})(a(\langle x, y\rangle) = 0).$$

Then we define f to be the term $(\lambda x.\mu(\lambda y.a(\langle x, y\rangle)))$. As a consequence, we have $(\lambda y.a(\langle x, y\rangle)) \in (\mathsf{N} \to \mathsf{N})$ and therefore $\mu(\lambda y.a(\langle x, y\rangle)) \in \mathsf{N}$ for all $x \in \mathsf{N}$. This implies $f \in (\mathsf{N} \to \mathsf{N})$ and, moreover,

$$(\forall x \in \mathsf{N})(a(\langle x, fx\rangle) = 0).$$

Hence, also the graph principle has been established. Straightforward induction on the length of the proof in $(\Pi^1_0\text{-}\mathsf{CA})\!\upharpoonright$ now yields our result. □

One important consequence of this lemma is that our Π^1_1 normal form theorem is available in $\mathsf{BON}(\mu) + (\mathsf{S}\text{-}\mathsf{I}_\mathsf{N})$. Based on these normal forms, we can now employ the Suslin operator E_1 in order to lift Lemma 4 from arithmetic to Π^1_1 formulas of $\mathcal{L}_2$.

Lemma 6 *For every Π^1_1 formula $A(\vec{U}, \vec{F}, \vec{v})$ of $\mathcal{L}_2$ with all its free variables in $\vec{U}, \vec{F}, \vec{v}$ there exists a closed individual term t_A of* L *so that* $\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I}_\mathsf{N})$ *proves*

1. $(\forall \vec{x} \in \mathcal{P}(\mathsf{N}))(\forall \vec{f} \in \mathsf{N} \to \mathsf{N})(\forall \vec{y} \in \mathsf{N})(t_A(\vec{x}, \vec{f}, \vec{y}) = 0 \vee t_A(\vec{x}, \vec{f}, \vec{y}) = 1)$,
2. $(\forall \vec{x} \in \mathcal{P}(\mathsf{N}))(\forall \vec{f} \in \mathsf{N} \to \mathsf{N})(\forall \vec{y} \in \mathsf{N})(A^\mathsf{N}(\vec{x}, \vec{f}, \vec{y}) \leftrightarrow t_A(\vec{x}, \vec{f}, \vec{y}) = 0)$.

Proof Let $A(\vec{U}, \vec{F}, \vec{v})$ be a Π^1_1 formula of $\mathcal{L}_2$ with its free variables as indicated. In view of Theorem 2 there exists an arithmetic formula B_A with the free variables of A plus the two new variables a and b so that

$$A^\mathsf{N}(\vec{u}, \vec{f}, \vec{v}) \leftrightarrow \neg(\exists g \in \mathsf{N} \to \mathsf{N})(\forall z \in \mathsf{N})B_A(\vec{u}, \vec{f}, \vec{v}, g(z+1), g(z))$$

for all $\vec{u} \in \mathcal{P}(\mathsf{N})$, all $\vec{f} \in (\mathsf{N} \to \mathsf{N})$ and all $\vec{v} \in \mathsf{N}$. Due to Lemma 4 there exists a closed L term t representing B_A. Then the term

$$(\lambda a, b.t(\vec{u}, \vec{f}, \vec{v}, a, b))$$

is a total operation from N^2 to N for all $\vec{u} \in \mathcal{P}(\mathsf{N})$, all $\vec{f} \in (\mathsf{N} \to \mathsf{N})$ and all $\vec{v} \in \mathsf{N}$. If we now define s to be the term

$$1 \dot{-} \mathsf{E}_1(\lambda a, b.t(\vec{u}, \vec{f}, \vec{v}, a, b)),$$

then we can finally take t_A to be the closed L term which is associated to s according to Lemma 1(2). It is easy to check that t_A has the properties stated in the lemma. □

This lemma together with Lemma 5 immediately implies our first result, namely that $(\Pi^1_1\text{-}\mathsf{CA})\!\restriction$ is contained in $\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I_N})$.

Theorem 7 *Let $A(\vec{U}, \vec{F}, \vec{v})$ be an $\mathcal{L}_2$ formula with all its free variables in $\vec{U}, \vec{F}, \vec{v}$ and assume that $(\Pi^1_1\text{-}\mathsf{CA})\!\restriction$ proves $A(\vec{U}, \vec{F}, \vec{v})$. Then we have*

$$\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I_N}) \vdash \vec{u} \in \mathcal{P}(\mathsf{N}) \wedge \vec{f} \in (\mathsf{N} \to \mathsf{N}) \wedge \vec{v} \in \mathsf{N} \;\to\; A^{\mathsf{N}}(\vec{u}, \vec{f}, \vec{v}).$$

Now we turn to transfinite hyperjump hierarchies. We first consider the system $\mathsf{SUS} + (\text{L-}\mathsf{I_N})$ and show that it contains $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}$. We simply have to follow the argument in Feferman and Jäger [11] and replace ordinary jump hierarchies by hyperjump hierarchies, which is possible since E_1 is available in SUS.

Lemma 8 *Let $A(U, v)$ be a Π^1_1 formula of $\mathcal{L}_2$ with all its free variables in U, v and assume that n is an arbitrary natural number. Then there exists a closed* L *term h_A so that $\mathsf{SUS} + (\text{L-}\mathsf{I_N})$ proves:*

1. $u \in \mathcal{P}(\mathsf{N}) \;\to\; h_A u \in \mathcal{P}(\mathsf{N})$,
2. $u \in \mathcal{P}(\mathsf{N}) \;\to\; \mathsf{Hier}_A(u, h_A u, n)$.

Proof Utilizing Lemma 6 we choose closed L terms t_A and t_B for representing the Π^1_1 formula $A(U, v)$ and the arithmetic formula $B(u, v)$,

$$B(u, v) \;:=\; (u = \langle (u)_0, (u)_1 \rangle \wedge (u)_1 \prec v \wedge v \prec n).$$

Then we apply the recursion theorem in order to obtain a closed term s with the property

$$suvw \;\simeq\; \begin{cases} uw, & \text{if } v = 0, \\ t_A(su(w)_1, (w)_0), & \text{if } t_B wv = 0, \\ 1, & \text{otherwise} \end{cases}$$

for all $u \in \mathcal{P}(\mathsf{N})$ and $v, w \in \mathsf{N}$. Then the following three implications are provable in $\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I_N})$:

(1) $u \in \mathcal{P}(\mathsf{N}) \wedge w \in \mathsf{N} \;\to\; su0w = uw$,

(2) $u \in \mathcal{P}(\mathsf{N}) \wedge v \in \mathsf{N} \wedge w \in \mathsf{N} \wedge t_B wv = 0 \;\to\; suvw \simeq t_A(su(w)_1, (w)_0)$,

(3) $u \in \mathcal{P}(\mathsf{N}) \wedge v \in \mathsf{N} \wedge w \in \mathsf{N} \wedge 0 \prec v \wedge t_B wv = 1 \;\to\; fuvw = 1$.

Since full formula induction is available in $\mathsf{SUS} + (\mathsf{L}\text{-}\mathsf{I_N})$ and the order type of $\prec_n$ is less than ε_0, we can show by transfinite induction along $\prec_n$ that

$$(4) \qquad v \in \mathsf{N} \wedge v \prec n \;\rightarrow\; suv \in \mathcal{P}(\mathsf{N})$$

for all $u \in \mathcal{P}(\mathsf{N})$. Finally, we let h_A be the term $(\lambda u, v.su(v)_1(v)_0)$. It is now easy to verify that h_A has all the desired properties. □

This lemma basically states that arbitrary hyperjumps can be iterated in $\mathsf{SUS} + (\mathsf{L}\text{-}\mathsf{I_N})$ along all ordinals less than ε_0. In view of Lemma 5 we can therefore conclude that $\mathsf{SUS} + (\mathsf{L}\text{-}\mathsf{I_N})$ contains $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}$.

Theorem 9 *Let $A(\vec{U}, \vec{F}, \vec{v})$ be an $\mathcal{L}_2$ formula with all its free variables in $\vec{U}, \vec{F}, \vec{v}$ and assume that $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}$ proves $A(\vec{U}, \vec{F}, \vec{v})$. Then we have*

$$\mathsf{SUS} + (\mathsf{L}\text{-}\mathsf{I_N}) \vdash \vec{u} \in \mathcal{P}(\mathsf{N}) \wedge \vec{f} \in (\mathsf{N} \rightarrow \mathsf{N}) \wedge \vec{v} \in \mathsf{N} \;\rightarrow\; A^{\mathsf{N}}(\vec{u}, \vec{f}, \vec{v}).$$

This embedding of $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}$ into $\mathsf{SUS} + (\mathsf{L}\text{-}\mathsf{I_N})$ is rather standard. The situation is more complicated, however, in the case of the system $\mathsf{SUS}+(\mathsf{N}\text{-}\mathsf{I_N})$, i.e. if full induction is restricted to $(\mathsf{N}\text{-}\mathsf{I_N})$.

In $\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N})$ one can show transfinite induction with respect to formulas of the form $ta \in (\mathsf{N} \rightarrow \mathsf{N})$ up to ω^k for all k less than ω. This argument requires some technical effort, but all details are carried through in Jäger and Strahm [20], even for the theory $\mathsf{BON}(\mu) + (\mathsf{N}\text{-}\mathsf{I_N})$.

As soon as these transfinite inductions are made available, we can step back and follow the pattern of the proof of Lemma 8 and show that hyperjump hierarchies up to each ordinal ω^k for k less than ω can be constructed in $\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N})$. Therefore, we have the following embedding theorem.

Theorem 10 *Let $A(\vec{U}, \vec{F}, \vec{v})$ be an $\mathcal{L}_2$ formula with all its free variables in $\vec{U}, \vec{F}, \vec{v}$ and assume that $(\Pi^1_1\text{-}\mathsf{CA})_{<\omega^\omega}\!\restriction$ proves $A(\vec{U}, \vec{F}, \vec{v})$. Then we have*

$$\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N}) \vdash \vec{u} \in \mathcal{P}(\mathsf{N}) \wedge \vec{f} \in (\mathsf{N} \rightarrow \mathsf{N}) \wedge \vec{v} \in \mathsf{N} \;\rightarrow\; A^{\mathsf{N}}(\vec{u}, \vec{f}, \vec{v}).$$

4 Upper bounds

A natural way for computing the upper proof-theoretic bounds of SUS with set, N and formula induction is to embed them in suitable set theories. For dealing with μ and especially E_1 we need an admissible universe which is in addition a limit of admissibles. Hence, the theory KPi or more precisely some of its subsystems, all described for example in Jäger [17], are appropriate for our purpose. By referring to known results about the relationship between these set theories and the subsystems of second order arithmetic of the previous section we will obtain exact proof-theoretic characterizations.

4.1 Set theories for a recursively inaccessible universe

Let $\mathcal{L}_1$ be the first order part of our language $\mathcal{L}_2$ from above. The theory KPi is formulated in the extension $\mathcal{L}^* = \mathcal{L}_1(\in, \mathsf{N}, \mathsf{S}, \mathsf{Ad})$ of $\mathcal{L}_1$ by the membership relation symbol $\in$, the set constant N for the set of natural numbers and the unary relation symbols S and Ad for sets and admissibles, respectively.

The terms $(\mathfrak{r}, \mathfrak{s}, \mathfrak{t}, \mathfrak{r}_1, \mathfrak{s}_1, \mathfrak{t}_1, \dots)$ of $\mathcal{L}^*$ are the terms of $\mathcal{L}_1$ plus the set constant N. The formulas $(A, B, C, A_1, B_1, C_1, \dots)$ of $\mathcal{L}^*$ as well as the Δ_0, Σ, Π, Σ_n and Π_n formulas of $\mathcal{L}^*$ are defined as usual. Equality between objects is not represented by a primitive symbol but defined by

$$(\mathfrak{s} = \mathfrak{t}) \;:=\; \begin{cases} (\mathfrak{s} \in \mathsf{N} \,\wedge\, \mathfrak{t} \in \mathsf{N} \,\wedge\, (\mathfrak{s} =_{\mathsf{N}} \mathfrak{t})) \;\vee \\ (\mathsf{S}(\mathfrak{s}) \,\wedge\, \mathsf{S}(\mathfrak{t}) \,\wedge\, (\forall x \in \mathfrak{s})(x \in \mathfrak{t}) \,\wedge\, (\forall x \in \mathfrak{t})(x \in \mathfrak{s})) \end{cases}$$

where $=_{\mathsf{N}}$ is the symbol for the primitive recursive equality on the natural numbers. The formula $A^{\mathfrak{s}}$ is the result of replacing each unrestricted quantifier $(\exists x)(\dots)$ and $(\forall x)(\dots)$ in A by $(\exists x \in \mathfrak{s})(\dots)$ and $(\forall x \in \mathfrak{s})(\dots)$, respectively. In addition, we freely make use of all standard set-theoretic notations and write, for example $(\mathfrak{s} : \mathsf{N} \to \mathsf{N})$ to express that $\mathfrak{s}$ is a function from N to N, $\mathsf{Tran}(\mathfrak{s})$ for the Δ_0 formula saying that $\mathfrak{s}$ is a transitive set and $\mathsf{Ord}(\mathfrak{s})$ for Δ_0 formula stating that $\mathfrak{s}$ is an ordinal; in the following small Greek letters range over ordinals.

Let $\mathcal{F}$ be a collection of $\mathcal{L}^*$ formulas. Induction on the natural numbers with respect to $\mathcal{F}$ consist of all formulas

$$(\mathcal{F}\text{-}\mathsf{I}_{\mathsf{N}}) \qquad A(0) \wedge (\forall x \in \mathsf{N})(A(x) \to A(x')) \;\to\; (\forall x \in \mathsf{N})A(x),$$

so that $A(a)$ belongs to $\mathcal{F}$. Accordingly, $\in$ induction with respect to $\mathcal{F}$ comprises for all $A(a)$ from $\mathcal{F}$ the formulas

$$(\mathcal{F}\text{-}\mathsf{I}_{\in}) \qquad (\forall x)[(\forall y \in x)A(y) \to A(x)] \;\to\; (\forall x)A(x).$$

Now we introduce the theory KPi for a recursively inaccessible universe. Its logical axioms comprise the usual axioms of classical first order logic with equality. The non-logical axioms of KPi can be divided into the following five groups.

I. Ontological axioms. We have for all terms $\mathfrak{r}$, $\vec{\mathfrak{s}}$ and $\mathfrak{t}$ of $\mathcal{L}^*$, all function symbols $\mathcal{H}$ and relation symbols R of $\mathcal{L}_1$ and all axioms $A(\vec{a})$ of group III whose free variables belong to the list $\vec{a}$:

(1) $\mathfrak{r} \in \mathsf{N} \leftrightarrow \neg\mathsf{S}(\mathfrak{r})$.

(2) $\vec{\mathfrak{s}} \in \mathsf{N} \;\rightarrow\; \mathcal{H}(\vec{\mathfrak{s}}) \in \mathsf{N}$.

(3) $R(\vec{\mathfrak{s}}) \;\rightarrow\; \vec{\mathfrak{s}} \in \mathsf{N}$.

(4) $\mathfrak{r} \in \mathfrak{t} \;\rightarrow\; \mathsf{S}(\mathfrak{t})$.

(5) $\mathsf{Ad}(\mathfrak{t}) \;\rightarrow\; (\mathsf{N} \in \mathfrak{t} \wedge \mathsf{Tran}(\mathfrak{t}))$.

(6) $\mathsf{Ad}(\mathfrak{t}) \;\rightarrow\; (\forall \vec{x} \in \mathfrak{t}) A^{\mathfrak{t}}(\vec{x})$.

(7) $\mathsf{Ad}(\mathfrak{r}) \wedge \mathsf{Ad}(\mathfrak{t}) \;\rightarrow\; \mathfrak{r} \in \mathfrak{t} \vee \mathfrak{r} = \mathfrak{t} \vee \mathfrak{t} \in \mathfrak{r}$.

II. Number-theoretic axioms. We have for all axioms $A(\vec{a})$ of Peano arithmetic PA which are not instances of the schema of complete induction and whose free variables belong to the list $\vec{a}$:

(Number theory) $(\forall \vec{x} \in \mathsf{N}) A^{\mathsf{N}}(\vec{x})$.

III. Kripke Platek axioms. We have for all terms $\mathfrak{s}$ and $\mathfrak{t}$ and all Δ_0 formulas $A(a)$ and $B(a, b)$ of $\mathcal{L}^*$:

(Pair) $\exists x(\mathfrak{s} \in x \;\wedge\; \mathfrak{t} \in x)$.

(Tran) $\exists x(\mathfrak{s} \subset x \;\wedge\; \mathsf{Tran}(x))$.

(Δ_0-Sep) $\exists y(\mathsf{S}(y) \;\wedge\; y = \{x \in \mathfrak{s} : A(x)\})$.

(Δ_0-Coll) $(\forall x \in \mathfrak{s}) \exists y B(x, y) \;\rightarrow\; \exists z (\forall x \in \mathfrak{s})(\exists y \in z) B(x, y)$.

IV. Limit axiom. It is used to formalize that each set is element of an admissible set, hence we claim:

(Lim) $(\forall x)(\exists y)(x \in y \;\wedge\; \mathsf{Ad}(y))$.

V. Induction axioms. These consist of the schemas ($\mathcal{L}^*$-I_{N}) of complete induction on the natural numbers and ($\mathcal{L}^*$-$\mathsf{I}_{\in}$) of $\in$ induction, both for arbitrary $\mathcal{L}^*$ formulas.

The theory KPi without the limit axiom corresponds to the usual system of Kripke-Platek set theory above the natural numbers as urelements; its typical models are the structures $\mathsf{L}(\alpha)_{\mathbb{N}}$ with α an admissible ordinal greater than ω. Hence, the least standard model of KPi is the structure $\mathsf{L}(\iota)_{\mathbb{N}}$ for ι being the least recursively inaccessible ordinal.

For modeling SUS with various forms of induction, however, we do not need the full strength of KPi. So let $\mathsf{KPi}^{\mathsf{r}}$ be the subsystem of KPi in which complete induction on the natural numbers and $\in$ induction are restricted to

Δ_0 formulas, i.e. ($\mathcal{L}^*$-$\mathsf{I_N}$) and ($\mathcal{L}^*$-$\mathsf{I}_\in$) are replaced by (Δ_0-$\mathsf{I_N}$) and (Δ_0-$\mathsf{I}_\in$), respectively. Moreover, $\mathsf{KPi^w}$ is defined to be $\mathsf{KPi^r}$ plus ($\mathcal{L}^*$-$\mathsf{I_N}$).

The proof-theoretic analysis of these set theories has been established a long time ago and their relationship to subsystems of second order arithmetic is well-known. In the following theorem (Δ^1_2-CA) and (Δ^1_2-CR) are the usual subsystems of analysis with Δ^1_2 comprehension axiom and rule, respectively; (BI) denotes the standard schema of bar induction.

Theorem 11 *We have the following proof-theoretic equivalences:*

1. $(\Pi^1_1\text{-}\mathsf{CA})\!\restriction\; \equiv\; (\Delta^1_2\text{-}\mathsf{CA})\!\restriction\; \equiv\; \mathsf{KPi^r}$,
2. $(\Pi^1_1\text{-}\mathsf{CA})_{<\omega^\omega}\; \equiv\; (\Delta^1_2\text{-}\mathsf{CR})\; \equiv\; \mathsf{KPi^r} + (\Sigma_1\text{-}\mathsf{I_N})$,
3. $(\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0}\; \equiv\; (\Delta^1_2\text{-}\mathsf{CA})\; \equiv\; \mathsf{KPi^w}$,
4. $(\Delta^1_2\text{-}\mathsf{CA}) + (\mathsf{BI})\; \equiv\; \mathsf{KPi}$.

For the proof of assertions (1), (3) and (4) of this theorem consult e.g. Buchholz et. al. [2], Jäger [14, 16] and Jäger and Pohlers [19]. The second assertion can be obtained by making use of similar techniques.

4.2 Modeling SUS in $\mathsf{KPi^r}$

Now we turn to the central part of this paper, namely to construct a model of SUS in $\mathsf{KPi^r}$. The crucial idea is to make use of a very specific positive Δ^1_2 inductive definition for interpreting the application relation ($rs \simeq t$) of SUS. The comparative strength of this inductive definition is needed in order to handle E_1; additional difficulties arise since everything has to be carried through within $\mathsf{KPi^r}$. However, the syntactic structure of our Δ^1_2 inductive definition makes this possible.

An additional technical complication is created by representing the primitive recursion operator $\mathsf{r_N}$ as well as the non-constructive μ operator. Both have been treated in detail in Feferman and Jäger [11] so that we confine ourselves to repeating some basic notations.

Let $A(a,b,c)$ be an $\mathcal{L}^*$ formula with at most a,b,c free and n a natural number greater than 0. Then we define for each natural number n greater than 0 and each vector $\vec{b} = b_1,\dots,b_n$ an $\mathcal{L}^*$ formula $\mathsf{Ap}^n_A(a,\vec{b},c)$ by induction on n, and from that, an $\mathcal{L}^*$ formula $\mathsf{Tot}^n_A(a)$:

$$\begin{aligned}
\mathsf{Ap}^1_A(a,b_1,c) &:= a,b_1,c \in \mathsf{N} \land A(a,b_1,c),\\
\mathsf{Ap}^{n+1}_A(a,\vec{b},b_{n+1},c) &:= (\exists x \in \mathsf{N})(\mathsf{Ap}^n_A(a,\vec{b},x) \land A(x,b_{n+1},c)),\\
\mathsf{Tot}^n_A(a) &:= (\forall \vec{x} \in \mathsf{N})(\exists y \in \mathsf{N})\mathsf{Ap}^n_A(a,\vec{x},y).
\end{aligned}$$

Later we will choose an appropriate formula $A(a,b,c)$ as interpretation of $(ab \simeq c)$. Then $\mathsf{Ap}^n_A(a, b_1, \dots, b_n, c)$ translates the formula $(ab_1 \dots b_n \simeq c)$ of SUS. In this sense, $\mathsf{Tot}^n_A(a)$ means that a is (the code of) a n-ary total operation on N. Observe that $\mathsf{Ap}^n_A(a, \vec{b}, c)$ and $\mathsf{Tot}^n_A(a)$ are Σ formulas provided that $A(a,b,c)$ is a Σ formula.

Further, if a is (a code of) a unary function on N in the sense of A and b (a code of) a ternary function on N in the sense of A, then the following formula $\mathsf{Rc}_A(a,b,u,v,w)$ describes the graph of the function which is defined from a and b by primitive recursion in the sense of A:

$$\mathsf{Rc}_A(a,b,u,v,w) \;:=\; \begin{cases} (\exists x \in \mathsf{N})[Seq(x) \land lh(x) = v{+}1 \land A(a,u,(x)_0) \\ \quad \land (\forall y \in \mathsf{N})(y < v \to \mathsf{Ap}^3_A(b,u,y,(x)_y,(x)_{y+1})) \\ \quad \land w = (x)_v]. \end{cases}$$

Recall that Seq is the primitive recursive predicate for sequence numbers. We also set for all natural numbers n

$$Seq_n(\mathsf{t}) := Seq(\mathsf{t}) \land lh(\mathsf{t}) = n$$

and assume that our coding of sequences is so that $\neg(Seq_m(\mathsf{t}) \land Seq_n(\mathsf{t}))$ if the natural numbers m and n are different.

Finally, we are approaching the interpretation of L in $\mathcal{L}^*$. First we choose pairwise different numerals $\hat{\mathsf{k}}, \hat{\mathsf{s}}, \hat{\mathsf{p}}, \hat{\mathsf{p}}_0, \hat{\mathsf{p}}_1, \hat{\mathsf{s}}_\mathsf{N}, \hat{\mathsf{p}}_\mathsf{N}, \hat{\mathsf{d}}_\mathsf{N}, \hat{\mathsf{r}}_\mathsf{N}, \hat{\mu}$ and $\hat{\mathsf{E}}_1$ (the values of) which do not belong to the set $\{0\} \cup \{x \in \mathbb{N} : Seq(x)\}$; they will serve as codes of the corresponding constants of L. The treatment of compound L terms will be described below.

For the purpose of the following inductive definition we choose a fresh ternary relation symbol R and define $\mathfrak{A}(R,a,b,c)$ to be the $\mathcal{L}^*(R)$ formula which is the disjunction of the following formulas (1)–(24):

(1) $a = \hat{\mathsf{k}} \land c = \langle \hat{\mathsf{k}}, b \rangle$,

(2) $Seq_2(a) \land (a)_0 = \hat{\mathsf{k}} \land c = (a)_1$,

(3) $a = \hat{\mathsf{s}} \land c = \langle \hat{\mathsf{s}}, b \rangle$,

(4) $Seq_2(a) \land (a)_0 = \hat{\mathsf{s}} \land c = \langle \hat{\mathsf{s}}, (a)_1, b \rangle$,

(5) $Seq_3(a) \land (a)_0 = \hat{\mathsf{s}} \land$
$(\exists x, y \in \mathsf{N})[R((a)_1, b, x) \land R((a)_2, b, y) \land R(x, y, c)]$,

(6) $a = \hat{\mathsf{p}} \land c = \langle \hat{\mathsf{p}}, b \rangle$,

(7) $Seq_2(a) \wedge (a)_0 = \hat{\mathsf{p}} \wedge c = \langle (a)_1, b \rangle$,

(8) $a = \hat{\mathsf{p}}_0 \wedge (\exists x \in \mathsf{N})(b = \langle c, x \rangle)$,

(9) $a = \hat{\mathsf{p}}_1 \wedge (\exists x \in \mathsf{N})(b = \langle x, c \rangle)$,

(10) $a = \hat{\mathsf{s}}_{\mathsf{N}} \wedge c = b+1$,

(11) $a = \hat{\mathsf{p}}_{\mathsf{N}} \wedge b = c+1$,

(12) $a = \hat{\mathsf{d}}_{\mathsf{N}} \wedge c = \langle \hat{\mathsf{d}}_{\mathsf{N}}, b \rangle$,

(13) $Seq_2(a) \wedge (a)_0 = \hat{\mathsf{d}}_{\mathsf{N}} \wedge c = \langle \hat{\mathsf{d}}_{\mathsf{N}}, (a)_1, b \rangle$,

(14) $Seq_3(a) \wedge (a)_0 = \hat{\mathsf{d}}_{\mathsf{N}} \wedge c = \langle \hat{\mathsf{d}}_{\mathsf{N}}, (a)_1, (a)_2, b \rangle$,

(15) $Seq_4(a) \wedge (a)_0 = \hat{\mathsf{d}}_{\mathsf{N}} \wedge (a)_1 = (a)_2 \wedge c = (a)_3$,

(16) $Seq_4(a) \wedge (a)_0 = \hat{\mathsf{d}}_{\mathsf{N}} \wedge (a)_1 \neq (a)_2 \wedge c = b$,

(17) $a = \hat{\mathsf{r}}_{\mathsf{N}} \wedge c = \langle \hat{\mathsf{r}}_{\mathsf{N}}, b \rangle$,

(18) $Seq_2(a) \wedge (a)_0 = \hat{\mathsf{r}}_{\mathsf{N}} \wedge c = \langle \hat{\mathsf{r}}_{\mathsf{N}}, (a)_1, b \rangle$,

(19) $Seq_3(a) \wedge (a)_0 = \hat{\mathsf{r}}_{\mathsf{N}} \wedge c = \langle \hat{\mathsf{r}}_{\mathsf{N}}, (a)_1, (a)_2, b \rangle$,

(20) $Seq_4(a) \wedge (a)_0 = \hat{\mathsf{r}}_{\mathsf{N}} \wedge \mathsf{Rc}_R((a)_1, (a)_2, (a)_3, b, c)$,

(21) $a = \hat{\mu} \wedge c = 0 \wedge (\forall x \in \mathsf{N})(\exists y \in \mathsf{N})(y \neq 0 \wedge R(b, x, y))$,

(22) $a = \hat{\mu} \wedge R(b, c, 0) \wedge$
$(\forall x \in \mathsf{N})[x < c \rightarrow (\exists y \in \mathsf{N})(y \neq 0 \wedge R(b, x, y))]$,

(23) $a = \hat{\mathsf{E}}_1 \wedge c = 0 \wedge \mathsf{Tot}^2_R(b) \wedge$
$(\exists f : \mathsf{N} \rightarrow \mathsf{N})(\forall x \in \mathsf{N})\mathsf{Ap}^2_R(b, f(x+1), f(x), 0)$,

(24) $a = \hat{\mathsf{E}}_1 \wedge c = 1 \wedge \mathsf{Tot}^2_R(b) \wedge$
$(\forall f : \mathsf{N} \rightarrow \mathsf{N})(\exists x \in \mathsf{N})(\exists y \in \mathsf{N})(0 < y \wedge \mathsf{Ap}^2_R(b, f(x+1), f(x), y)$.

We want to remark that the clauses (1)–(22) are identical to the clauses of the inductive definition in Feferman and Jäger [11] and suffice for the treatment of $\mathsf{BON}(\mu)$. The additional clauses (23) and (24) are needed for coping with E_1. Clearly, the $\mathcal{L}^*(R)$ formula $\mathfrak{A}(R, a, b, c)$ is positive in R.

Observe that we quantify in these two clauses over proper *set-theoretic* functions from N to N, not over codes for total functions on N in the sense of

R. Although formulated in the language of set theory, $\mathfrak{A}(R,a,b,c)$ is easily reformulated in the language of second order arithmetic. Then it is a very special R positive Δ^1_2 formula, more precisely arithmetical in Π^1_1.

We now employ $\mathfrak{A}(R,a,b,c)$ as a definition clause of a positive Δ^1_2 inductive definition and show several other properties needed for interpreting SUS. Since $\mathsf{KPi^r}$ is much too weak in order to deal with positive Δ^1_2 inductive definitions in general, some extra efforts have to be made.

In the following we freely replace the relation variable R in the formula $\mathfrak{A}(R,a,b,c)$ by sets or $\mathcal{L}^*$ formulas. So, for example, $\mathfrak{A}(u,a,b,c)$ is obtained from $\mathfrak{A}(R,a,b,c)$ by replacing all subformulas $R(\mathfrak{r},\mathfrak{s},\mathfrak{t})$ by $((\mathfrak{r},\mathfrak{s},\mathfrak{t})\in u)$.

Lemma 12 *There are a Σ formula $B(u,a,b,c)$ and a Π formula $C(u,a,b,c)$ of the language $\mathcal{L}^*$ with exactly the free variables u,a,b,c so that* $\mathsf{KPi^r}$ *proves:*

1. $u \subset \mathsf{N}^3 \land a,b,c \in \mathsf{N} \;\rightarrow\; (\mathfrak{A}(u,a,b,c) \leftrightarrow B(u,a,b,c))$,
2. $u \subset \mathsf{N}^3 \land a,b,c \in \mathsf{N} \;\rightarrow\; (\mathfrak{A}(u,a,b,c) \leftrightarrow C(u,a,b,c))$.

This lemma is an immediate consequence of the quantifier theorem in Jäger [14] (see also, e.g., [15, 17]); it says that Σ^1_2 formulas of $\mathcal{L}_2$ are provably equivalent in $\mathsf{KPi^r}$ to Σ_1 formulas of $\mathcal{L}^*$.

Since Σ recursion is not available in $\mathsf{KPi^r}$ and because of the complexity of our definition clause $\mathfrak{A}(R,a,b,c)$ we introduce hierarchies to describe the stages of the corresponding inductive definition. Accordingly, we set

$$\mathfrak{H}(f,\alpha) \;:=\; \begin{cases} \mathsf{Fun}(f) \land \mathsf{dom}(f) = \alpha \land \\ (\forall \beta < \alpha)[f(\beta) = \{(x,y,z) \in \mathsf{N}^3 : \mathfrak{A}(\bigcup_{\gamma<\beta} f(\gamma), x,y,z)\}]. \end{cases}$$

Straightforward $\in$ induction for Δ_0 formulas yields that these hierarchies are unique and increasing. Following the argument in Feferman and Jäger [11] concerning the clauses (1)–(22) of the formula $\mathfrak{A}(R,a,b,c)$ and by a simple inspection of clauses (23) and (24) it is also clear that we have functionality in the third argument. I.e. we can prove in $\mathsf{KPi^r}$ that

(i) $\mathfrak{H}(f,\alpha) \land \mathfrak{H}(g,\beta) \land \gamma < \min(\alpha,\beta) \;\rightarrow\; f(\gamma) = g(\gamma)$,

(ii) $\mathfrak{H}(f,\alpha) \land \gamma < \beta < \alpha \;\rightarrow\; f(\gamma) \subset f(\beta)$,

(iii) $\mathfrak{H}(f,\alpha) \land \beta < \alpha \land (a,b,c_1) \in f(\beta) \land (a,b,c_2) \in f(\beta) \;\rightarrow\; c_1 = c_2$.

In view of the preceding lemma it is immediate that $\mathfrak{H}(f,\alpha)$ is provably equivalent in $\mathsf{KPi^r}$ to a Σ formula, with which we often identify it in the sequel. Thus the following predicate $\mathbb{S}(a,b,c)$ can be regarded as a Σ formula,

$$\mathbb{S}(a,b,c) \;:=\; (\exists f)(\exists \alpha)[\mathfrak{H}(f,\alpha{+}1) \wedge (a,b,c) \in f(\alpha)].$$

Since only weak $\in$ induction is available in $\mathsf{KPi^r}$, we cannot prove there that for all ordinals α there is a function f so that $\mathfrak{H}(f,\alpha)$. Nevertheless, we can show that $\mathbb{S}$ is a fixed point of $\mathfrak{A}(R,a,b,c)$ which is functional in its third argument.

Theorem 13 *We can prove in* $\mathsf{KPi^r}$ *that*

1. $(\forall x,y,z \in \mathsf{N})[\mathbb{S}(x,y,z) \leftrightarrow \mathfrak{A}(\mathbb{S},x,y,z)]$,
2. $(\forall x,y,z_1,z_2 \in \mathsf{N})[\mathbb{S}(x,y,z_1) \wedge \mathbb{S}(x,y,z_2) \rightarrow z_1 = z_2]$.

Proof Because of the properties (i)–(iii) of our hierarchies the second assertion is completely obvious. Moreover, the R positivity of $\mathfrak{A}$ immediately yields the direction from left to right of the first assertion. The crucial step thus is to show that $\mathbb{S}(a,b,c)$ follows from $\mathfrak{A}(\mathbb{S},a,b,c)$ for all natural numbers a,b and c.

Assume $\mathfrak{A}(\mathbb{S},a,b,c)$. Now distinction by cases according to clauses (1)–(24) of $\mathfrak{A}$ has to be carried through. Mostly the arguments are straightforward, and we confine ourselves to discussing (23) in detail. Hence, we know that

(1) $\quad a = \hat{\mathsf{E}}_1 \wedge c = 0,$

(2) $\quad (\forall x,y \in \mathsf{N})(\exists u,v \in \mathsf{N})[\mathbb{S}(b,x,u) \wedge \mathbb{S}(u,y,v)],$

(3) $\quad (\exists f : \mathsf{N} \rightarrow \mathsf{N})(\forall x \in \mathsf{N})\mathsf{Ap}^2_{\mathbb{S}}(b,f(x{+}1),f(x),0).$

Now we apply Σ reflection to (2). Because of the definition of $\mathbb{S}$ and the properties (i)–(iii) of our hierarchies we may concluded that there are a function g and an ordinal α so that $\mathfrak{H}(g,\alpha{+}1)$ and

(4) $\quad (\forall x,y \in \mathsf{N})(\exists u,v \in \mathsf{N})((b,x,u) \in g(\alpha) \wedge (u,y,v) \in g(\alpha)).$

Hence, $g(\alpha)$ can also play the role of $\mathbb{S}$ in (3) and we obtain

(5) $\quad (\exists f : \mathsf{N} \rightarrow \mathsf{N})(\forall x \in \mathsf{N})\mathsf{Ap}^2_{g(\alpha)}(b,f(x{+}1),f(x),0).$

In view of (1),(4) and (5) we conclude $\mathfrak{A}(g(\alpha),a,b,c)$. Using Δ separation gives a function h so that $\mathfrak{H}(h,\alpha{+}2)$ and $(a,b,c) \in h(\alpha{+}1)$. This implies $\mathbb{S}(a,b,c)$ and concludes our proof. $\square$

For the argument in the previous proof it was crucial that we had $\mathsf{Tot}^2_{\mathbb{S}}(b)$ (i.e. line (2) above) at our disposal. Then we could apply Σ reflection and thus localize the Σ relation $\mathbb{S}$ to a specific set $g(\alpha)$. Thus it is essential that we are interpreting the functional E_1 which is applied to total operations on N; a similar argument cannot work for the partial functional $\mathsf{E}_1^{\#}$.

A last central step in showing that $\mathbb{S}$ provides a adequate application relation for interpreting SUS in $\mathsf{KPi}^{\mathsf{r}}$ is to proof the following *inside-outside* property: given an a coding (the characteristic function of) a binary relation on N with respect to $\mathbb{S}$, we have to show that there exists a set-theoretic function which is a descending chain for a if and only if there exists a natural number coding a total function on N in the sense of $\mathbb{S}$ which is also a descending chain for a.

For the formulation (and the proof) of the following theorem some auxiliary notations are useful. First we set

$$\{a\}^{\mathbb{S}}(b_1, \dots, b_n) \simeq c \;:=\; \mathsf{Ap}^n_{\mathbb{S}}(a, b_1, \dots, b_n, c)$$

and follow the usual conventions of recursion theory when using expressions like $\{a\}^{\mathbb{S}}(\vec{b})$. Later sometimes $\mathbb{S}$ will be replaced by a subset u of N^3 with $\{a\}^u(\vec{b})$ having its obvious meaning.

Then we write $\mathsf{CDC}(a, b)$ in order to express that b is a *code of a total function* on N providing a descending chain for a, everything in the sense of $\mathbb{S}$,

$$\mathsf{CDC}_{\mathbb{S}}(a, b) \;:=\; \mathsf{Tot}^1_{\mathbb{S}}(b) \,\wedge\, (\forall x \in \mathsf{N})\mathsf{Ap}^2_{\mathbb{S}}(a, \{b\}^{\mathbb{S}}(x{+}1), \{b\}^{\mathbb{S}}(x), 0).$$

This is in contrast to $\mathsf{FDC}_{\mathbb{S}}(a, f)$,

$$\mathsf{FDC}_{\mathbb{S}}(a, f) \;:=\; (f : \mathsf{N} \to \mathsf{N}) \,\wedge\, (\forall x \in \mathsf{N})\mathsf{Ap}^2_{\mathbb{S}}(a, f(x{+}1), f(x), 0),$$

which states that f is a *set-theoretic function* providing a descending chain for a with respect to $\mathbb{S}$.

Given a hierarchy f which describes the stages of our inductive definition up to the ordinal α, i.e. $\mathfrak{H}(f, \alpha)$, it is easy to see that it can be extended up to $\alpha{+}k$ for each fixed natural number k. Suppose now that we have natural numbers which are known to code certain relations on the natural numbers at level α. Then the following two lemmas state that certain operations on these relations can be carried through and coded by making use of a fixed *finite* number of additional levels only.

For their formulation it is convenient to write $\mathsf{ext}^n_b(a)$ for the extension of a in the sense of b,

$$\mathsf{ext}^n_b(a) \;:=\; \{(x_1, \dots, x_n) \in \mathsf{N}^n : \mathsf{Ap}^n_b(a, x_1, \dots, x_n, 0)\}.$$

Since Ap_b^n is a Δ_0 formula, $\mathsf{ext}_b^n(a)$ defines in KPi^r a subset of N^n by means of Δ_0 separation.

Lemma 14 (Finite extension property I) *Let R be a fresh m-ary relation symbol and assume that $A(R, u, \vec{v})$ is a formula of $\mathcal{L}_1(R)$ with at most the variables u and $\vec{v} = v_1, \dots, v_n$ free. Then there exist a natural number k and primitive recursive functions $\mathcal{F}$ and $\mathcal{G}$ so that KPi^r proves:*

1. $\mathfrak{H}(f, \alpha{+}k{+}1) \wedge a, u \in \mathsf{N} \wedge \mathsf{Tot}^m_{f(\alpha)}(a) \;\rightarrow$

   $\mathsf{Tot}^n_{f(\alpha+k)}(\mathcal{F}(a,u)) \wedge \mathsf{Tot}^1_{f(\alpha+k)}(\mathcal{G}(a)),$

2. $\mathfrak{H}(f, \alpha{+}k{+}1) \wedge a, u, \vec{v} \in \mathsf{N} \wedge \mathsf{Tot}^m_{f(\alpha)}(a) \;\rightarrow$

   $A^\mathsf{N}(\mathsf{ext}^m_{f(\alpha)}(a), u, \vec{v}) \leftrightarrow \{\mathcal{F}(a,u)\}^{f(\alpha+k)}(\vec{v}) \simeq 0,$

3. $\mathfrak{H}(f, \alpha{+}k{+}1) \wedge a, u \in \mathsf{N} \wedge \mathsf{Tot}^m_{f(\alpha)}(a) \;\rightarrow$

   $(\hat{\mathsf{E}}_1, \mathcal{F}(a,u), 0) \in f(\alpha{+}k) \leftrightarrow \{\mathcal{G}(a)\}^{f(\alpha+k)}(u) \simeq 0.$

Lemma 15 (Finite extension property II) *Let R be a fresh m-ary and S a fresh n-ary relation symbol and assume that $B(R, S, u, v)$ is a formula of $\mathcal{L}_1(R, S)$ with at most the variables u and v free. Further assume that*

$$\mathsf{KPi}^\mathsf{r} \vdash (\forall x_1, x_2 \subset \mathsf{N})(\forall y \in \mathsf{N})(\exists! z \in \mathsf{N}) B^\mathsf{N}(x_1, x_2, y, z)$$

and let $\mathcal{F}_B$ denote the (class) function defined by B^N. Then there exists a natural number k and a primitive recursive function $\mathcal{H}$ so that KPi^r proves:

1. $\mathfrak{H}(f, \alpha{+}k{+}1) \wedge a, b, u \in \mathsf{N} \wedge \mathsf{Tot}^m_{f(\alpha)}(a) \wedge \mathsf{Tot}^n_{f(\alpha)}(b) \;\rightarrow$

   $\mathsf{Tot}^1_{f(\alpha+k)}(\mathcal{H}(a,b,u)),$

2. $\mathfrak{H}(f, \alpha{+}k{+}1) \wedge a, b, u, v \in \mathsf{N} \wedge \mathsf{Tot}^m_{f(\alpha)}(a) \wedge \mathsf{Tot}^n_{f(\alpha)}(b) \;\rightarrow$

   $\{\mathcal{H}(a,b,u)\}^{f(\alpha+k)}(0) \;\simeq\; u \;\wedge$

   $\{\mathcal{H}(a,b,u)\}^{f(\alpha+k)}(v{+}1) \;\simeq$

   $\qquad \mathcal{F}_B(\mathsf{ext}^m_{f(\alpha)}(a), \mathsf{ext}^n_{f(\alpha)}(b), \{\mathcal{H}(a,b,u)\}^{f(\alpha+k)}(v)).$

The proofs of these two finite extension lemmas are straightforward, although a bit tedious. It is essential that primitive recursion is directly built into our inductive definition. They also exploit the fact that combinatory completeness is available due to our coding of k and s. After these preparatory steps we are now ready to turn to the theorem about the *inside-outside property* mentioned above.

Theorem 16 *We can prove in* $\mathsf{KPi}^{\mathsf{r}}$ *that*

1. $(\exists b \in \mathsf{N})\mathsf{CDC}_{\mathbb{S}}(a,b) \;\rightarrow\; (\exists f)\mathsf{FDC}_{\mathbb{S}}(a,f)$,

2. $\mathsf{Tot}^2_{\mathbb{S}}(a) \wedge (\exists f)\mathsf{FDC}_{\mathbb{S}}(a,f) \;\rightarrow\; (\exists b \in \mathsf{N})\mathsf{CDC}_{\mathbb{S}}(a,b)$.

In particular, we have in $\mathsf{KPi}^{\mathsf{r}}$ *the following equivalence*

$$\mathsf{Tot}^2_{\mathbb{S}}(a) \;\rightarrow\; [(\exists f)\mathsf{FDC}_{\mathbb{S}}(a,f) \leftrightarrow (\exists b \in \mathsf{N})\mathsf{CDC}_{\mathbb{S}}(a,b)].$$

Proof For the first assertion of our theorem choose a natural number b such that $\mathsf{CDC}_{\mathbb{S}}(a,b)$. Hence, we have $\mathsf{Tot}^1_{\mathbb{S}}(b)$, i.e.

$$(1) \qquad (\forall x \in \mathsf{N})(\exists y \in \mathsf{N})\mathbb{S}(b,x,y).$$

Because of Σ reflection and the functionality of $\mathbb{S}$ in its third argument we know that

$$(2) \qquad f \;:=\; \{(x,y) \in \mathsf{N}^2 : \mathbb{S}(b,x,y)\}$$

is a set-theoretic function from N to N. It follows immediately that this f satisfies $\mathsf{FDC}(a,f)$.

Throughout the proof of the second assertion, we let a be an arbitrary natural number so that $\mathsf{Tot}^2_{\mathbb{S}}(a)$, i.e.

$$(3) \qquad (\forall x,y \in \mathsf{N})(\exists u,v \in \mathsf{N})[\mathbb{S}(a,x,u) \wedge \mathbb{S}(u,y,v)].$$

Hence, we can employ Σ reflection and find an ordinal α and a function g so that $\mathfrak{H}(g,\alpha{+}1)$ and $\mathsf{Tot}^2_{g(\alpha)}(a)$, i.e.

$$(4) \qquad (\forall x,y \in \mathsf{N})(\exists u,v \in \mathsf{N})((a,x,u) \in g(\alpha) \wedge (u,y,v) \in g(\alpha)).$$

Let us now assume that we are given a function f so that $\mathsf{FDC}_{\mathbb{S}}(a,f)$. Indeed, by (4) we have $\mathsf{FDC}_{g(\alpha)}(a,f)$, i.e. $(f : \mathsf{N} \rightarrow \mathsf{N})$ and

$$(5) \qquad (\forall x \in \mathsf{N})\mathsf{Ap}^2_{g(\alpha)}(a, f(x{+}1), f(x), 0).$$

We want to find a natural number b so that $\mathsf{CDC}_{\mathbb{S}}(a,b)$. Towards this aim, we have to formalize a standard *leftmost branch argument* in $\mathsf{KPi}^{\mathsf{r}}$ and make substantial use of the finite extension properties I and II (Lemma 14 and 15). For definiteness, assume that k is a natural number which is big enough to carry through all the finite extensions of g needed below; accordingly, let h be a function such that $\mathfrak{H}(h,\alpha{+}2k{+}1)$ holds.

First, we set $A(R, u, v_1, v_2)$ to be the following $\mathcal{L}_1(R)$ formula

$$A(R, u, v_1, v_2) \ := \left\{ \begin{array}{l} (\exists x, y)[Seq(x) \wedge lh(x) = y \wedge (x)_0 = u \wedge (x)_{y \dot{-} 1} = v_2 \\ \quad \wedge (\forall z < y \dot{-} 1) R((x)_{z+1}, (x)_z)] \wedge R(v_1, v_2). \end{array} \right.$$

$A(R, u, v_1, v_2)$ expresses that $R(v_1, v_2)$ and v_2 is accessible from u by means of R. We can now apply Lemma 14 and find primitive recursive functions $\mathcal{F}$ and $\mathcal{G}$ so that

$$(6) \qquad \mathsf{Tot}^2_{h(\alpha+k)}(\mathcal{F}(a, u)) \quad \text{and} \quad \mathsf{Tot}^1_{h(\alpha+k)}(\mathcal{G}(a))$$

for all natural numbers u and, moreover,

$$(7) \qquad A^{\mathsf{N}}(\mathsf{ext}^2_{h(\alpha)}(a), u, v_1, v_2) \ \leftrightarrow \ \{\mathcal{F}(a, u)\}^{h(\alpha+k)}(v_1, v_2) \simeq 0,$$

$$(8) \qquad (\hat{\mathsf{E}}_1, \mathcal{F}(a, u), 0) \in h(\alpha{+}k) \ \leftrightarrow \ \{\mathcal{G}(a)\}^{h(\alpha+k)}(u) \simeq 0$$

for all natural numbers u and v_1, v_2. Further, we let $B(R, S, u, v)$ denote the $\mathcal{L}_1(R, S)$ formula given by

$$B(R, S, u, v) \ := \left\{ \begin{array}{l} R(v, u) \wedge S(v) \wedge (\forall w < v)(\neg R(w, u) \vee \neg S(w)) \\ \quad \vee \neg(\exists w)(R(w, u) \wedge S(w)) \wedge v = 0. \end{array} \right.$$

Clearly, we have that $(\forall x_1, x_2 \subset \mathsf{N})(\forall y \in \mathsf{N})(\exists! z \in \mathsf{N}) B^{\mathsf{N}}(x_1, x_2, y, z)$. Hence, we are in a position to apply Lemma 15 and obtain a primitive recursive function $\mathcal{H}$ so that we have for $b = \mathcal{H}(a, \mathcal{G}(a), f(0))$,

$$(9) \qquad \mathsf{Tot}^1_{h(\alpha+2k)}(b),$$

and, further, for all natural numbers v,

$$(10) \qquad \{b\}^{h(\alpha+2k)}(0) \ \simeq \ f(0),$$

$$(11) \quad \{b\}^{h(\alpha+2k)}(v{+}1) \ \simeq \ \mathcal{F}_B(\mathsf{ext}^2_{h(\alpha)}(a), \mathsf{ext}^1_{h(\alpha+k)}(\mathcal{G}(a)), \{b\}^{h(\alpha+2k)}(v)).$$

The proof so far has been set up in such a manner that we can now apply a simple form of Δ_0 induction on N to show that indeed

$$(12) \qquad (\forall x \in \mathsf{N}) \mathsf{Ap}^2_{h(\alpha)}(a, \{b\}^{h(\alpha+2k)}(x{+}1), \{b\}^{h(\alpha+2k)}(x), 0)$$

follows from (5) and the construction of b. Together with (9) we have thus established $\mathsf{CDC}_{h(\alpha+2k)}(a, b)$ and, hence, also $\mathsf{CDC}_{\mathbb{S}}(a, b)$. This is as desired an finishes the proof of our theorem. □

Now the stage is set in order to describe a translation $*$ from L into $\mathcal{L}^*$. The central idea is to interpret the L formula $(ab \simeq c)$ by the $\mathcal{L}^*$ formula $\mathbb{S}(a, b, c)$. More precisely, let us first define an $\mathcal{L}^*$ formula $\mathbb{V}_t(u)$ for each individual term t of L so that the variable u does not occur in t. The formula $\mathbb{V}_t(u)$ expresses that u is the value of t in the sense of $\mathbb{S}$. The exact definition is by induction on the complexity of t:

1. If t is an individual variable, then $\mathbb{V}_t(u)$ is $(t = u)$.

2. If t is an individual constant, then $\mathbb{V}_t(u)$ is $(\hat{t} = u)$.

3. If t is the individual term (rs), then

$$\mathbb{V}_t(u) \ := \ (\exists x, y \in \mathsf{N})(\mathbb{V}_r(x) \land \mathbb{V}_s(y) \land \mathbb{S}(x, y, u)).$$

In a second step we define the $*$ translation of an L formula A inductively as follows:

4. If A is the formula $\mathsf{N}(t)$ or $t\downarrow$, then A^* is $(\exists x \in \mathsf{N})\mathbb{V}_t(x)$.

5. If A is the formula $(s = t)$, then A^* is $(\exists x \in \mathsf{N})(\mathbb{V}_s(x) \land \mathbb{V}_t(x))$.

6. If A is the formula $\neg B$, then A^* is $\neg(B^*)$.

7. If A is the formula $(B\ j\ C)$ for $j \in \{\lor, \land, \rightarrow\}$, then A^* is $(B^*\ j\ C^*)$.

8. If A is the formula $(\mathcal{Q}x)B$ for $\mathcal{Q} \in \{\exists, \forall\}$, then A^* is $(\mathcal{Q}x \in \mathsf{N})B^*$.

Given this interpretation of L into $\mathcal{L}^*$ we can now turn to the desired embedding of SUS with set induction, N induction and formula induction on N into our set theories.

Theorem 17 *If $A(\vec{a})$ is an* L *formula with all its free variables indicated, then we have:*

1. $\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I}_\mathsf{N}) \vdash A(\vec{a}) \quad \Longrightarrow \quad \mathsf{KPi}^\mathsf{r} \vdash \vec{a} \in \mathsf{N} \rightarrow A^*(\vec{a})$,

2. $\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I}_\mathsf{N}) \vdash A(\vec{a}) \quad \Longrightarrow \quad \mathsf{KPi}^\mathsf{r} + (\Sigma_1\text{-}\mathsf{I}_\mathsf{N}) \vdash \vec{a} \in \mathsf{N} \rightarrow A^*(\vec{a})$,

3. $\mathsf{SUS} + (\mathsf{L}\text{-}\mathsf{I}_\mathsf{N}) \vdash A(\vec{a}) \quad \Longrightarrow \quad \mathsf{KPi}^\mathsf{w} \vdash \vec{a} \in \mathsf{N} \rightarrow A^*(\vec{a})$.

Proof For showing the first assertion, we want to point out that all axioms of SUS plus set induction except for the axioms about E_1 can be treated exactly as in Feferman and Jäger [11], in which the theory $\mathsf{BON}(\mu)$ plus set induction is treated.

The axiom $(\mathsf{E}_1.1)$ is an immediate consequence of the fixed point property of $\mathbb{S}$ formulated in Theorem 13. In order to obtain axiom $(\mathsf{E}_1.2)$ in KPi^r choose a natural number f so that $(f \in (\mathsf{N} \to \mathsf{N}))^*$. Then the fixed point property of $\mathbb{S}$ nearly yields the $*$ translation of the conclusion of $(\mathsf{E}_1.2)$. There is still one serious defect: the existential quantifier on the left hand still ranges over set-theoretic functions from N to N rather than functions in the sense of $\mathbb{S}$. By the *inside-outside property* proved in Theorem 16, however, we can now internalize this set-theoretic function into a function in the sense of $\mathbb{S}$ and vice versa.

The second assertion follows form the first assertion simply by observing that formulas of the form $(t \in \mathsf{N})$ translate into Σ formulas under our translation $*$. Hence, $(\mathsf{N}\text{-}\mathsf{I_N})$ in the language L directly carries over to Σ_1 induction in $\mathcal{L}^*$. Clearly, the third assertion also immediately follows from the first. □

Everything is available now in order to state the final proof-theoretic characterization of the theories SUS plus set induction, N induction and formula induction. Theorem 7, Theorem 9 and Theorem 10 provide their lower bounds in terms of subsystems of second order arithmetic, Theorem 17 gives their upper bounds in terms of subsystems of set theory and Theorem 11 yields the proof-theoretic equivalence of the appropriate subsystems of second order arithmetic and set theory.

Corollary 18 *We have the following proof-theoretic equivalences:*

1. $\mathsf{SUS} + (\mathsf{S}\text{-}\mathsf{I_N}) \equiv (\Pi^1_1\text{-}\mathsf{CA})\!\restriction \equiv (\Delta^1_2\text{-}\mathsf{CA})\!\restriction \equiv \mathsf{KPi}^\mathsf{r}$,
2. $\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N}) \equiv (\Pi^1_1\text{-}\mathsf{CA})_{<\omega^\omega} \equiv (\Delta^1_2\text{-}\mathsf{CR}) \equiv \mathsf{KPi}^\mathsf{r} + (\Sigma_1\text{-}\mathsf{I_N})$,
3. $\mathsf{SUS} + (\mathrm{L}\text{-}\mathsf{I_N}) \equiv (\Pi^1_1\text{-}\mathsf{CA})_{<\varepsilon_0} \equiv (\Delta^1_2\text{-}\mathsf{CA}) \equiv \mathsf{KPi}^\mathsf{w}$.

Besides $(\mathsf{N}\text{-}\mathsf{I_N})$ there exists a further form of complete induction on the natural numbers which comprises $(\mathsf{S}\text{-}\mathsf{I_N})$ but is strictly weaker than $(\mathrm{L}\text{-}\mathsf{I_N})$, the so-called operation induction $(\mathsf{O}\text{-}\mathsf{I_N})$, cf. [18, 21, 20]. However, we do not have to consider $(\mathsf{O}\text{-}\mathsf{I_N})$ in our present context since Kahle [21] shows that over $\mathsf{BON}(\mu)$ operation induction $(\mathsf{O}\text{-}\mathsf{I_N})$ and N induction $(\mathsf{N}\text{-}\mathsf{I_N})$ are equivalent. Hence, in the second assertion of this corollary we can replace $\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N})$ by $\mathsf{SUS} + (\mathsf{O}\text{-}\mathsf{I_N})$.

We conclude this article with mentioning the proof-theoretic ordinals of these systems, following the notation system of Buchholz and Schütte [3], which are know since a long time for the corresponding subsystems of second order arithmetic and set theory: the proof-theoretic ordinal of $\mathsf{SUS}+(\mathsf{S}\text{-}\mathsf{I_N})$ is $\Psi 0 \Omega_\omega$, that of $\mathsf{SUS} + (\mathsf{N}\text{-}\mathsf{I_N})$ is $\Psi 0 \Omega_{\omega^\omega}$, and the one of $\mathsf{SUS} + (\mathrm{L}\text{-}\mathsf{I_N})$ is $\Psi 0 \Omega_{\varepsilon_0}$.

References

[1] Beeson, M. J. *Foundations of Constructive Mathematics: Metamathematical Studies*. Springer, Berlin, 1985.

[2] Buchholz, W., Feferman, S., Pohlers, W., and Sieg, W. *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*, vol. 897 of *Lecture Notes in Mathematics*. Springer, Berlin, 1981.

[3] Buchholz, W., and Schütte, K. *Proof Theory of Impredicative Subsystems of Analysis*. Bibliopolis, Naples, 1988.

[4] Feferman, S. A language and axioms for explicit mathematics. In *Algebra and Logic*, J. Crossley, Ed., vol. 450 of *Lecture Notes in Mathematics*. Springer, Berlin, 1975, pp. 87–139.

[5] Feferman, S. A theory of variable types. *Revista Colombiana de Matemáticas 19* (1975), 95–105.

[6] Feferman, S. Theories of finite type related to mathematical practice. In *Handbook of Mathematical Logic*, J. Barwise, Ed. North Holland, Amsterdam, 1977, pp. 913–971.

[7] Feferman, S. Recursion theory and set theory: a marriage of convenience. In *Generalized recursion theory II, Oslo 1977*, J. E. Fenstad, R. O. Gandy, and G. E. Sacks, Eds., vol. 94 of *Stud. Logic Found. Math.* North Holland, Amsterdam, 1978, pp. 55–98.

[8] Feferman, S. Constructive theories of functions and classes. In *Logic Colloquium '78*, M. Boffa, D. van Dalen, and K. McAloon, Eds. North Holland, Amsterdam, 1979, pp. 159–224.

[9] Feferman, S. Weyl vindicated: "Das Kontinuum" 70 years later. In *Temi e prospettive della logica e della filosofia della scienza contemporanee*. CLUEB, Bologna, 1988, pp. 59–93.

[10] Feferman, S. Definedness. *Erkenntnis 43* (1995), 295–320.

[11] Feferman, S., and Jäger, G. Systems of explicit mathematics with non-constructive μ-operator. Part I. *Annals of Pure and Applied Logic 65*, 3 (1993), 243–263.

[12] Feferman, S., and Jäger, G. Systems of explicit mathematics with non-constructive μ-operator. Part II. *Annals of Pure and Applied Logic 79*, 1 (1996), 37–52.

[13] Glass, T., and Strahm, T. Systems of explicit mathematics with non-constructive μ-operator and join. *Annals of Pure and Applied Logic 82* (1996), 193–219.

[14] Jäger, G. *Die konstruktible Hierarchie als Hilfsmittel zur beweistheoretischen Untersuchung von Teilsystemen der Mengenlehre und Analysis*. PhD thesis, Universität München, 1979.

[15] Jäger, G. Iterating admissibility in proof theory. In *Logic Colloquium '81. Proceedings of the Herbrand Symposion*. North Holland, Amsterdam, 1982.

[16] Jäger, G. A well-ordering proof for Feferman's theory T_0. *Archiv für mathematische Logik und Grundlagenforschung 23* (1983), 65–77.

[17] Jäger, G. *Theories for Admissible Sets: A Unifying Approach to Proof Theory*. Bibliopolis, Napoli, 1986.

[18] Jäger, G., Kahle, R., and Strahm, T. On applicative theories. In *Logic and Foundations of Mathematics*, A. Cantini, E. Casari, and P. Minari, Eds. Kluwer, 1999. To appear.

[19] Jäger, G., and Pohlers, W. Eine beweistheoretische Untersuchung von $(\Delta^1_2\text{-}\mathsf{CA}) + (\mathsf{BI})$ und verwandter Systeme. In *Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse*. 1982, pp. 1–28.

[20] Jäger, G., and Strahm, T. Some theories with positive induction of ordinal strength $\varphi\omega 0$. *Journal of Symbolic Logic 61*, 3 (1996), 818–842.

[21] Kahle, R. N-strictness in applicative theories. *Archive for Mathematical Logic*. to appear.

[22] Simpson, S. G. *Subsystems of Second Order Arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag, 1998.

[23] Troelstra, A., and van Dalen, D. *Constructivism in Mathematics*, vol. I. North-Holland, Amsterdam, 1988.

[24] Troelstra, A., and van Dalen, D. *Constructivism in Mathematics*, vol. II. North Holland, Amsterdam, 1988.

Address
Gerhard Jäger, Thomas Strahm
Institut für Informatik und angewandte Mathematik, Universität Bern
Neubrückstrasse 10, CH-3012 Bern, Switzerland
{jaeger,strahm}@iam.unibe.ch