ELSEVIER

Annals of Pure and Applied Logic 85 (1997) 1–46

ANNALS OF
PURE AND
APPLIED LOGIC

On the proof-theoretic strength of monotone induction in explicit mathematics

Thomas Glaß [a], Michael Rathjen [b,1], Andreas Schlüter [b,*,2]

[a] *Siemens AG, München, Germany*

[b] *Stanford University, Stanford, CA 94305, USA*

Abstract

We characterize the proof-theoretic strength of systems of explicit mathematics with a general principle (MID) asserting the existence of least fixed points for monotone inductive definitions, in terms of certain systems of analysis and set theory. In the case of analysis, these are systems which contain the Σ^1_2-axiom of choice and Π^1_2-comprehension for formulas without set parameters. In the case of set theory, these are systems containing the Kripke–Platek axioms for a recursively inaccessible universe together with the existence of a stable ordinal. In all cases, the exact strength depends on what forms of induction are admitted in the respective systems.

Keywords: Proof theory; Explicit mathematics; Monotone induction; Stability; Admissible sets; Asymmetric interpretation

AMS classification: 03F50; 03F35; 03F25; 03D60; 03D70

1. Introduction

Explicit mathematics has been devised by Feferman in [6, 7] as a framework in which to develop mathematics based on constructive grounds. Furthermore, systems of explicit mathematics have been used in proof-theoretic reductions of subsystems of second-order arithmetic and set theory.

Systems of explicit mathematics are theories of operations and classifications in which the latter are members of the universe of discourse and hence may be taken as arguments and/or values of operations. They are *explicit* in the sense that functions and classes are regarded as explicitly represented entities in the universe of discourse, hence the theories are intensional in that respect. They are *constructive* in the sense that all

* Correspondence address: Universität Münster, Institut für Mathematische Logik, Einsteinstr. 62, D-48149 Münster, Germany.

[1] A Heisenberg Fellow of the German Science Foundation, Deutsche Forschungsgemeinschaft.

[2] Supported by a grant of the Deutscher Akademischer Austauschdienst (DAAD).

PII S0168-0072(96)00040-1

operations may be interpreted as partial recursive functions, even if the underlying logic is taken to be classical. More importantly, the degree to which they are constructively justified hinges on the construction principles for classifications they contain. In that respect, the systems we consider are somewhat on the verge of constructivity, since they postulate least fixed points of very general operations. This requires at least a somewhat broadened understanding of what a constructive process should be.

Systems of explicit mathematics may be considered in either classical or intuitionistic logic; in this paper they are treated within classical logic only. It is an open question whether the systems we deal with retain the same proof-theoretical strength if used with intuitionistic logic instead. If not, that would be a radically different situation from what was met before with the system T_0 of explicit mathematics and its subsystems.

Proof-theoretic reductions using these systems are of two-fold use: on the one hand they illuminate the principles of the "constructive" theories of explicit mathematics and the extent to which they allow the development of constructive mathematics. On the other hand, they provide a constructive justification for prima facie non-constructive subsystems of second-order arithmetic and set theory in the sense of reductive proof theory.

The object of investigation from the first of these viewpoints is the constructive theory in its own right, its principles are explained in terms of more familiar, but in general non-constructive systems. This approach has been predominant in the investigations of proper subsystems of T_0, cf. e.g. [6, 7, 9, 2, 12, 13]. Here one usually obtains a conservation result of a theory of explicit mathematics over a system of second-order arithmetic or inductive definitions which is more or less obviously contained in it.

In contrast to this, the object of study in reductive proof theory is the non-constructive system. The goals of this approach are of a foundational rather than a technical nature. Using explicit mathematics as the constructive framework, the most prominent example of this is the proof theoretic reduction of Σ^1_2-AC+(BI) (or KPi for that matter) to T_0 by Jäger [17] and Jäger and Pohlers [19]. Similar results in a different constructive framework, namely that of Martin–Löf type theory, have recently been obtained by Griffor and Rathjen in [14] and independently by Setzer in [24].

The results of the present paper contribute to both aspects of these proof-theoretic reductions, but in contrast to the previous cases it is hard to say which aspect is more interesting: Do we learn more about the working of the constructive system or do we gain more confidence into the non-constructive systems used in the characterization? Both of these aspects will be considered and both are equally important.

The subject of our investigations are extensions of explicit mathematics by the principle (MID) which asserts the existence of least fixed points of arbitrary monotone operations. Since inductive definitions form a very powerful yet still constructively acceptable principle, the interest to understand this principle in the context of explicit mathematics is obvious. We quote from Feferman's article [8, p. 88]:

> What is the strength of T_0 + (MID)? [...] I have tried, but did not succeed, to extend my interpretation of T_0 in Σ^1_2-AC + (BI) to include the statement

(MID). The theory $T_0 + (MID)$ includes all constructive formulations of iteration of monotone inductive definitions of which I am aware, while T_0 (in its (IG) axiom) is based squarely on the general iteration of accessibility inductive definitions. Thus it would be of great interest for the present subject to settle the relationship between these theories.

We provide (the major part of) an answer to this question by proving the following theorem (cf. Theorems 9.1 and 10.1):

Theorem. *Let F be a Σ^1_2-sentence.*

(a) *The following are equivalent.*

1. $EM_0{\restriction} + (Join) + (IG){\restriction} + (MID) \vdash F$
2. $\Sigma^1_2\text{-}AC_0 + (\Pi^1_2\text{-}CA^-) \vdash F$
3. $KPi^r + \exists\gamma(L_\gamma \prec_1 L) \vdash F$

(b) *Analogously, the following are equivalent.*

1. $EM_0 + (Join) + (IG){\restriction} + (MID) \vdash F$
2. $\Sigma^1_2\text{-}AC + (\Pi^1_2\text{-}CA^-) \vdash F$
3. $KPi^w + \exists\gamma(L_\gamma \prec_1 L) \vdash F$.

Here $(\Pi^1_2\text{-}CA^-)$ asserts comprehension for Π^1_2-formulas without set parameters. The systems KPi^r and KPi^w of Kripke–Platek axioms for a recursively inaccessible universe, with restrictions on induction principles, which were introduced in [19], are described in Section 2.3. Actually, our results also show that the proof-theoretic strength of the theories does not increase if we add Church's Thesis since the models we use are based on the model of partial recursive indices for the operations. Also, there are some extensions of our results which are discussed in Remark 9.3.

Looking at these results from the perspective of reductive proof theory, we have obtained a reduction of axiom systems of second-order arithmetic and set theory with very strong non-constructive axioms to explicit mathematics. Actually, the strength of these theories is so big that until now there have not been any constructive justifications for systems of that strength in the literature.

Next, we want to go briefly into the history of investigations of proof-theoretic strength for systems of explicit mathematics with (MID). First investigations (after the problem had been posed by Feferman) in this direction were begun by Takahashi, cf. [27]. It turned out that even the construction of models of $T_0 + (MID)$ was surprisingly difficult. Takahashi showed that $T_0 + (MID)$ can be interpreted in the fragment of analysis with Π^1_2-comprehension and bar induction. The question whether $T_0 + (MID)$ is stronger than T_0 remained open.

Using Takahashi's models for proof-theoretic reductions using asymmetric interpretations, Glaß in [12] obtained a characterization of many subsystems of $T_0 + (MID)$ in terms of theories of second-order arithmetic. As already noted by Takahashi, in absence of (restricted) (IG) the strength of (MID) collapses to the strength of accessibility inductive definitions. When this axiom is present, Glaß' work, which is one source of a part of the present paper, uses strong monotone inductive definitions in the context of second-order arithmetic for a characterization. However, a direct estimate for the strength of these systems was not achieved there.

The actual reason for the strength of the axiom (MID) in the presence of (restricted) (IG) was illuminated by Rathjen in [21], thereby providing the key to a computation of the proof-theoretic strength of theories containing (MID) in terms of well-known principles in second-order arithmetic and, more importantly, in set theory. Namely, it was shown that, in the presence of (restricted) (IG) and (Join), (MID) allows to prove the existence of inductively generated sub-fixpoints of non-monotonic inductive definitions. To do this, sophisticated techniques used by Harrington and Kechris, cf. [15], to reduce non-monotonic induction to suitable forms of monotonic induction were applied.

In the present paper, we again use the machinery of generalized recursion theory for non-monotonic inductive definitions. Drawing on ideas from Cenzer's paper [4], we show that the non-monotonic inductive definitions can be used to construct sets with strong stability properties. These stability properties in turn imply the existence of sub-fixpoints of the respective non-monotonic inductive definitions, so the characterization is faithful.

Let us present a short overview of how the paper is organized. This will show an, in our opinion, interesting interplay of techniques from the areas of proof theory, set theory, second-order arithmetic, explicit mathematics, inductive definitions and generalized recursion theory, which is, in addition to the above intentions of this investigation, one attractive feature that triggered our interest in these questions.

After assembling the necessary preliminaries on theories of second-order arithmetic, explicit mathematics and set theory in Section 2, we consider stability notions and show that they imply the existence of sub-fixpoints of non-monotonic inductive definitions in Section 3. Section 4 then gives a careful account of Takahashi's models for (subsystems of) $T_0 +$ (MID) which is used in the following section to give a reduction of these systems into systems of set theory axiomatizing the existence of a stable ordinal.

Sections 6–8 then serve to prepare the converse reduction: In Section 6 we show in a purely proof-theoretic manner that the stability axiom in our theory can be reduced to some "local instances" of stability. Section 7 forms the recursion-theoretic heart of the proof in that it shows that closure ordinals of non-monotonic inductive definitions give rise to ordinals with strong stability properties. Section 8 then simply uses the well-known technique of representation trees to transfer the situation to second-order arithmetic, which is more easily interpreted in systems of explicit mathematics. In Section 9 we combine all the previous work to give our first main result, namely, the proof-theoretical equivalence between the systems of explicit mathematics and those of set theory as given by the theorem mentioned above. The main new ingredient in this section is an application of the main theorem of Rathjen's [21] to achieve the reduction back into explicit mathematics. Finally, in Section 10 we provide a characterization of the strength of these theories in terms of axiom systems for second-order arithmetic with comprehension for Π^1_2-formulas without set parameters, before we close the paper with some outlook on future work and open questions.

We would like to thank Prof. S. Feferman for his interest in the present work, and for his hospitality during a visit at Stanford by two of the authors.

2. Preliminaries

2.1. Subsystems of second-order arithmetic

The language $\mathscr{L}_2$ of second-order arithmetic contains (free and bound) number variables $a, b, a_0, a_1, \ldots, x, y, z, \ldots$, (free and bound) set variables $A, B, A_0, A_1, \ldots, X, Y, Z, \ldots$, constants $0, 1$, function symbols $+, \cdot$, and relation symbols $=, <, \in$.

Terms are built as usual, formulas are built from the prime formulas $s = t$, $s < t$, and $s \in A$ using $\wedge, \vee, \neg, \forall x, \exists x, \forall X$ and $\exists X$ where s, t are terms.

As usual, number quantifiers are called bounded if they occur in the context $\forall x(x < s \rightarrow \ldots)$ or $\exists x(x < s \wedge \ldots)$ for a term s which does not mention x. The Δ^0_0-formulas are those formulas in which all quantifiers are bounded number quantifiers, Σ^0_k-formulas are formulas of the form $\exists x_1 \forall x_2 \ldots Qx_k F$, where F is Δ^0_0, Π^0_k-formulas are those of the form $\forall x_1 \exists x_2 \ldots Qx_k F$. The union of all Π^0_k- and Σ^0_k-formulas for all $k \in \mathbb{N}$ is the class of arithmetical or Π^0_∞-formulas. The $\Sigma^1_k(\Pi^1_k)$-formulas are the formulas $\exists X_1 \forall X_2 \ldots QX_k F$ (resp. $\forall X_1 \exists X_2 \ldots Qx_k F$) for arithmetical F.

When arguing in formal theories we also say that a formula belongs to one of these formula classes if it is equivalent to one formula of the class over the basic theory Π^0_∞-CA_0. But we will comment on these identifications when they are used in a non-obvious way.

The basic axioms in all theories of second-order arithmetic are the defining axioms of $0, 1, +, \cdot, <$ and the *induction axiom*

$$\forall X(0 \in X \wedge \forall x(x \in X \rightarrow x + 1 \in X) \rightarrow \forall x(x \in X)),$$

respectively, the *schema of induction*

$$F(0) \wedge \forall x(F(x) \rightarrow F(x + 1)) \rightarrow \forall x F(x),$$

where F is an arbitrary $\mathscr{L}_2$-formula.

We consider the axiom schema of *$\mathscr{C}$-comprehension* for formula classes $\mathscr{C}$ which is given by

$$(\mathscr{C}\text{-CA}) \quad \exists X \forall x(x \in X \leftrightarrow F(x))$$

for all formulas $F \in \mathscr{C}$.

We only will consider theories containing at least (Π^0_∞-CA). For each axiom schema (Ax) we denote by Ax the theory consisting of the basic arithmetical axioms, the schema of (Π^0_∞-CA), the schema of induction and the schema (Ax). If we replace the schema of induction by the induction axiom, we call the resulting theory Ax_0.

An example for these notations is the theory Π^1_1-CA which contains the induction schema, whereas Π^1_1-CA$_0$ only contains the induction axiom in addition to the comprehension schema for Π^1_1-formulas.

In the framework of these theories we can introduce defined symbols for all primitive recursive functions, especially we can define a pairing function $\langle .,. \rangle$ along with its inverses.

Using these pairing functions, we can consider (within our theory) a set A as a sequence $(A_n)_n$ of sets, where $A_n = \{m : \langle n,m \rangle \in A\}$. Also, for a binary relation $\prec$, we can define $A_{\prec n} = \{m : \exists n' \prec n \langle n', m \rangle \in A\}$.

Using this, we can formulate another axiom schema we will encounter, namely the axiom of choice for formulas in $\mathscr{C}$ given by

$$(\mathscr{C}\text{-AC}) \quad \forall x \exists Y F(x,Y) \to \exists Y \forall x F(x, Y_x).$$

Furthermore, we can introduce a primitive recursive standard wellordering $\lhd$ of order type ε_0, cf. e.g. [23, 22, 20]. W.l.o.g. 0 is the least element of this wellordering whose elements we denote by lower case Greek letters. Also, we can define ordinal addition, multiplication and exponentiation on this order relation as primitive recursive functions. Since all our theories contain Peano Arithmetic, we have $\mathrm{TI}(\lhd_\alpha, F)$ for all Π^0_∞-formulas F and all elements α of $\lhd$ denoting ordinals below ε_0. Here

$$\mathrm{TI}(\lhd_\alpha, F) \equiv \forall x (\forall y \lhd x F(y) \to F(x)) \to \forall x \lhd \alpha F(x).$$

If the theory additionally contains the induction schema, $\mathrm{TI}(\lhd_\alpha, F)$ can be proved for all $\alpha \lhd \varepsilon_0$ and all formulas F.

Using these notations for ordinals we can define another principle which will be used in this paper.

Definition 2.1. (a) Let $U_{\Pi^1_1}$ be a universal Π^1_1-formula, cf. e.g. [25]. Then the *hyperjump* of a set A is $\mathrm{HJ}(A) = \{\langle x,y \rangle : U_{\Pi^1_1}[x,y,A]\}$.

(b) For a formula F, let the formula $\mathscr{H}_F(\alpha, A)$ be given by

$$\mathscr{H}_F(\alpha, A) \Leftrightarrow \alpha \in \mathrm{field}(\lhd) \wedge \forall x \forall y (y \in A_x \leftrightarrow x \lhd \alpha \wedge F(x,y,A_{\lhd x})).$$

$\mathscr{H}_F(\alpha, A)$ expresses that A is the set obtained by iterating the formula F along $\lhd$ up to α.

(c) The axiom schema of α-times iterated Π^1_1-comprehensions is given as the universal closure of

$$(\Pi^1_1\text{-CA}_\alpha) \qquad \exists X \mathscr{H}_F(\alpha, X)$$

for all Π^1_1-formulas F.

(d) The axiom schema $(\Pi^1_1\text{-CA}_{<\beta})$ consists of all axioms $(\Pi^1_1\text{-CA}_\alpha)$ for $\alpha \lhd \beta$.

Remark 2.2. The most important application of the above definition is that it allows to prove the existence of iterated hyperjumps. This is the special case for the formula

F (with parameter A) given by

$$F(x,y,X) \equiv (x = 0 \wedge y \in A) \vee (\mathrm{Suc}(x) \wedge y \in \mathrm{HJ}(X)) \vee (\mathrm{Lim}(x) \wedge y \in X),$$

where $\mathrm{Suc}(x)$ indicates that x denotes a successor ordinal in $\lhd$, and $\mathrm{Lim}(x)$ that it denotes a limit ordinal. The (unique) set B such that $\mathscr{H}_F(\alpha',B)$ holds is denoted by $\mathrm{HJ}(\alpha,A)$. Here α' is the successor of α w.r.t. $\lhd$.

The following connections between subsystems of second-order arithmetic will be used.

Proposition 2.3. (a) *On the basis of* (*say*) Π^0_∞-CA_0, (Σ^1_{n+1}-AC) *implies* (Δ^1_{n+1}-CA), *especially it implies* (Π^1_n-CA).

(b) *The theory* Σ^1_2-AC *proves each instance of* Π^1_1-CA_α *for* α *denoting an ordinal less than* ε_0, *therefore it contains the theory* Π^1_1-$\mathrm{CA}_{<\varepsilon_0}$.

(c) Σ^1_2-AC_0 *is conservative over* Π^1_1-CA_0 *for* Π^1_3-*sentences.*

(d) Σ^1_2-AC *is conservative over* Π^1_1-$\mathrm{CA}_{<\varepsilon_0}$ *for* Π^1_3-*sentences.*

Proof. This was originally proved by Friedman in [11], but it can also be found in [9, Theorem 2.2.1]. □

2.2. Systems of explicit mathematics

The language of EM_0, $\mathscr{L}_{\mathrm{EM}_0}$, has two sorts of variables. The free and bound variables $a,b,c,\dots$ and $x,y,z\dots$ are conceived to range over the whole constructive universe which comprises *operations* and *classifications* among other kinds of entities; while upper-case versions of these $A,B,C,\dots$ and $X,Y,Z,\dots$ are used to represent free and bound classification variables.

$0, \mathrm{s_N}$ and $\mathrm{p_N}$ are operation constants whose intended interpretations are the natural number 0 and the successor and predecessor operations. Additional operation constants are $\mathrm{k}, \mathrm{s}, \mathrm{d}, \mathrm{p}, \mathrm{p_0}$, and $\mathrm{p_1}$ for the two basic combinators, definition by cases on the natural numbers, pairing and the corresponding two projections. Additional for the uniform formulation of classification existence axioms the constants j, i and c_n with $n \in \mathbb{N}$ are used for *join, inductive generation*, and *comprehension.*

The *terms* of EM_0 are just the variables and constants of the two sorts. The atomic formulas of EM_0 are built up using the terms and four primitive relation symbols $=$, N, App, and ε as follows: if q,r,r_1,r_2 are terms, then $q = r$, $\mathrm{N}(q)$, $\mathrm{App}(q,r_1,r_2)$, and $q \mathrel{\varepsilon} P$ (where P has to be a classification variable) are atomic formulas. $\mathrm{App}(q,r_1,r_2)$ expresses that the operation q applied to r_1 yields the value r_2; $q \mathrel{\varepsilon} P$ asserts[3] that q is in P or that q is classified under P.

We write $t_1t_2 \simeq t_3$ for $\mathrm{App}(t_1,t_2,t_3)$.

[3] We use the symbol "ε" instead of "$\in$", the latter being reserved for the set-theoretic elementhood relation.

The set of formulas is then obtained from these using the propositional connectives and the two quantifiers of each sort.

In order to facilitate the formulation of the axioms, the language of EM_0 is expanded definitionally with the symbol $\simeq$, and the auxiliary notion of an *application term* is introduced. The set of application terms is given by two clauses: all terms of EM_0 are application terms, and if s and t are application terms, then (st) is an application term.

For s and t application terms, we have auxiliary, defined formulas of the form:

$$s \simeq t := \forall y(s \simeq y \leftrightarrow t \simeq y),$$

if t is not a variable. Here $s \simeq a$ (for a a free variable) is inductively defined by $s \simeq a$ is $s = a$, if s is a variable or a constant, and $s \simeq a$ is $\exists x, y[s_1 \simeq x \wedge s_2 \simeq y \wedge \mathrm{App}(x, y, a)]$, if s is an application term $(s_1 s_2)$.

Some abbreviations are $t_1 t_2 \ldots t_n$ for $((\ldots(t_1 t_2)\ldots)t_n)$; $t\downarrow$ for $\exists y(t \simeq y)$ and $\phi(t)$ for $\exists y(t \simeq y \wedge \phi(y))$. If s, t are application terms, where t is not a classification variable, then $s \,\varepsilon\, t$ is short for $\exists X[t \simeq X \wedge s \,\varepsilon\, X]$.

Some further conventions are useful. Systematic notation for n-tuples is introduced as follows: (t) is t, (s, t) is $\mathrm{p}st$, and $(t_1, \ldots, t_n)$ is defined by $((t_1, \ldots, t_{n-1}), t_n)$. t' is written for the term $\mathrm{s_N}t$. $s \not\varepsilon X$ and $s \neq t$ are short for $\neg(s \,\varepsilon\, X)$ and $\neg(s = t)$, respectively. $\forall x \,\varepsilon\, Y(\ldots)$ stands for $\forall x(x \,\varepsilon\, Y \to \ldots)$. Similar conventions apply to $\exists$. Variables k, n, m are supposed to range over N; $\forall x \in \mathrm{N}$ and $\exists x \in \mathrm{N}$ are short for $\forall x(\mathrm{N}(x) \to \ldots)$ and $\exists x(\mathrm{N}(x) \wedge \ldots)$, respectively.

Gödel numbers for formulae play a key role in the axioms introducing the constants c_n. A formula is said to be *elementary* if it contains only free occurrences of classification variables A (i.e. only as *parameters*), and even those free occurrences of A are restricted: A must occur only to the right of ε in atomic formulas. We assume that a standard Gödel numbering has been chosen for $\mathscr{L}_{\mathrm{EM}_0}$; if ϕ is an elementary formula and $a, b_1, \ldots, b_n, A_1, \ldots, A_l$ is a list of variables which includes all parameters of ϕ, then $\{x : \phi(x, b_1, \ldots, b_n, A_1, \ldots, A_l)\}$ stands for $\mathrm{c}_m(b_1, \ldots, b_n, A_1, \ldots, A_l)$, where m is the numeral that codes the pair of Gödel numbers $\langle \ulcorner\phi\urcorner, \ulcorner(a, b_1, \ldots, b_n, A_1, \ldots, A_l)\urcorner \rangle$; m is called the *index* of ϕ and the list of variables.

In this paper, the logic of all subsystems of T_0 is assumed to be that of classical two-sorted predicate logic with identity. The non-logical axioms of EM_0 are the following:

I. *Basic Axioms*

(a) $\forall X \exists x(X = x)$,

(b) $\mathrm{App}(a, b, c_1) \wedge \mathrm{App}(a, b, c_2) \to c_1 = c_2$.

II. *Applicative Axioms*

(a) $\mathrm{k}ab \simeq a$,

(b) $(\mathrm{s}ab)\downarrow \wedge\, \mathrm{s}abc \simeq ac(bc)$,

(c) $\mathrm{p_0}(\mathrm{p}a_0a_1) \simeq a_0 \wedge \mathrm{p_1}(\mathrm{p}a_0a_1) \simeq a_1$,

(d) $(c = d \to \mathrm{d}abcd \simeq a) \wedge (c \neq d \to \mathrm{d}abcd \simeq b)$,

(e) $\mathrm{N}(a) \wedge \mathrm{N}(b) \to [\mathrm{p_N}(a') \simeq a \wedge \neg(a' = 0) \wedge (a' \simeq b' \to a \simeq b)]$.

III. *Elementary Comprehension Axiom*

(ECA) $\exists X[X \simeq \{x : \psi(x)\} \wedge \forall x(x \,\varepsilon\, X \leftrightarrow \psi(x))]$

for each elementary formula $\psi(a)$, which may contain additional parameters.

IV. *Natural Numbers*

(N1) $\mathrm{N}(0) \wedge \forall x(\mathrm{N}(x) \rightarrow \mathrm{N}(x'))$

(N2) $\phi(0) \wedge \forall x(\phi(x) \rightarrow \phi(x')) \rightarrow \forall x \in \mathrm{N}\phi(x)$

for each formula ϕ of $\mathscr{L}_{\mathrm{EM}_0}$.

Definition 2.4. The following axioms will be considered in extensions to EM_0:

- The join axiom (Join):

$$\forall x \,\varepsilon\, A \exists Y(fx \simeq Y) \rightarrow \exists X[X \simeq \mathrm{j}(A,f) \wedge \forall z(z \,\varepsilon\, X \leftrightarrow \exists x \,\varepsilon\, A \exists y(z \simeq (x,y) \wedge y \,\varepsilon\, fx))]$$

- Inductive Generation (IG):

$$\exists X[X \simeq \mathrm{i}(A,B) \wedge \mathrm{Prog}_A(B,X) \wedge (\mathrm{Prog}_A(B,\{x : F(x)\}) \rightarrow \forall x \,\varepsilon\, X F(x))],$$

where F is an arbitrary formula of EM_0 and

$$\mathrm{Prog}_A(B,X) :\equiv \forall x \,\varepsilon\, A(\forall y[(y,x) \,\varepsilon\, B \rightarrow y \,\varepsilon\, X] \rightarrow x \,\varepsilon\, X).$$

- Restricted Inductive Generation $(\mathrm{IG})\!\upharpoonright$:

$$\exists X[X \simeq \mathrm{i}(A,B) \wedge \mathrm{Prog}_A(B,X) \wedge \forall Z(\mathrm{Prog}_A(B,Z) \rightarrow \forall a \,\varepsilon\, X(x \,\varepsilon\, Z)].$$

Definition 2.5. (a) $\mathrm{EM}_0\!\upharpoonright$ is EM_0 where N-induction, i.e. (N2), is replaced by the N-induction axiom

$$\forall Z[0 \,\varepsilon\, Z \wedge \forall x(x \,\varepsilon\, Z \rightarrow x' \,\varepsilon\, Z) \rightarrow \forall x \in \mathrm{N}(x \,\varepsilon\, Z)].$$

(b) T_0 is $\mathrm{EM}_0 + (\mathrm{Join}) + (\mathrm{IG})$, $\mathrm{T}_0\!\upharpoonright$ is $\mathrm{EM}_0\!\upharpoonright + (\mathrm{Join}) + (\mathrm{IG})\!\upharpoonright$.

Next, we state two important tools for obtaining operations in $\mathrm{EM}_0\!\upharpoonright$. Both results can already be proved in the fragment of $\mathrm{EM}_0\!\upharpoonright$ without classification axioms.

Employing the axioms for the combinators k and s one can deduce an abstraction lemma yielding λ-terms of one argument. This can be generalized using n-tuples and projections.

Lemma 2.6 (Abstraction Lemma, cf. Federman [6]). *For each application term t there is a new application term t^* such that the parameters of t^* are among the parameters of t minus $x_1,\ldots,x_n$ and such that*

$$\mathrm{EM}_0\!\upharpoonright\, \vdash t^*\!\downarrow \wedge\, t^*(x_1,\ldots,x_n) \simeq t.$$

$\lambda(x_1,\ldots,x_n).t$ is written for t^*.

The most important consequence of the abstraction lemma is the recursion theorem. It can be derived in the same way as for the λ-calculus, cf. [6, 7, 2, VI.2.7]. Actually, one can prove a uniform version of the recursion theorem (with a recursion operator) in the applicative fragment of $\mathrm{EM}_0\!\restriction$.

Corollary 2.7 (Recursion theorem).

$$\forall f\,\exists g\,\forall x_1\ldots\forall x_n\, g(x_1,\ldots,x_n)\simeq f(g,x_1,\ldots,x_n).$$

Now we describe the monotone inductive definition principle and its uniform version in $\mathscr{L}_{\mathrm{EM}_0}$. Several other principles considered in this paper will also be described.

Definition 2.8. For classifications, "$\mathrel{\mathring{=}}$" denotes *extensional equality*, i.e.

$$X \mathrel{\mathring{=}} Y := \forall v(v\,\varepsilon\,X \leftrightarrow v\,\varepsilon\,Y).$$

Further, let $X\subseteq Y := \forall v(v\,\varepsilon\,X \rightarrow v\,\varepsilon\,Y)$ and

$$\begin{array}{ll}
\mathrm{Clop}(f) & \equiv \forall X\,\exists Y\, fX\simeq Y\\
\mathrm{Ext}(f) & \equiv \forall X\,\forall Y\,[X\mathrel{\mathring{=}} Y\rightarrow fX\mathrel{\mathring{=}} fY]\\
\mathrm{Mon}(f) & \equiv \forall X\,\forall Y[X\subseteq Y\rightarrow fX\subseteq fY]\\
\mathrm{Lfp}(Y,f) & \equiv fY\subseteq Y\wedge\forall X[fX\subseteq X\rightarrow Y\subseteq X]\\
\mathrm{Elfp}(f) & \equiv \exists Y\,\mathrm{Lfp}(Y,f).
\end{array}$$

When f satisfies $\mathrm{Clop}(f)$, we call f a *classification operation*. When f satisfies $\mathrm{Clop}(f)$ and $\mathrm{Ext}(f)$, we call f *extensional* or an *extensional operation*. When f satisfies $\mathrm{Clop}(f)$ and $\mathrm{Mon}(f)$, we call f *monotone* or a *monotone operation*. Since monotonicity entails extensionality, a monotone operation is always extensional.

Now we state (MID) and (UMID).

Definition 2.9. MID (*Monotone Inductive Definition*) is the axiom

$$\forall f[\mathrm{Clop}(f)\wedge\mathrm{Mon}(f)\rightarrow\mathrm{Elfp}(f)].$$

An axiom which seems to be more in keeping with the spirit of explicit mathematics can be formulated by adding a constant lfp to $\mathscr{L}_{\mathrm{EM}_0}$ which names a fixed point when applied to a monotone operation. This leads to the principle (UMID) (Uniform Monotone Inductive Definition):

$$\forall f[\mathrm{Clop}(f)\wedge\mathrm{Mon}(f)\rightarrow\mathrm{Lfp}(\mathrm{lfp}(f),f)].$$

(MID) states that if f is monotone, there is a least fixed point. (UMID) states that if f is monotone, $\mathrm{lfp}(f)$ is a least fixed point of f.

2.3. Subsystems of set theory

The axiom systems for set theory considered in this paper are formulated in (definitorial extensions of) the usual language $\mathscr{L}_\in$ containing $\in$ as the only non-logical symbol besides $=$. Formulas are built from prime formulas $a \in b$ and $a = b$ by use of $\wedge$, $\vee$, $\neg$, bounded quantifiers $\forall x \in a$, $\exists x \in a$ and unbounded quantifiers $\forall x, \exists x$. As usual, Δ_0-formulas are the formulas without unbounded quantifiers, Σ_1-formulas are those of the form $\exists x \varphi(x)$ where $\varphi(a)$ is a Δ_0-formula. Π_n-formulas are the formulas with a leading sequence of n alternating unbounded quantifiers starting with a universal one followed by a Δ_0-formula. The class of Σ-formulas is the smallest class of formulas containing the Δ_0-formulas and closed under $\wedge$, $\vee$, bounded quantification, and unbounded existential quantification.

The exact details of the formulation do not really matter for the purpose of this paper, any standard formulation will work. Also, we use the standard Δ_0-definitions of predicates like $x = \emptyset$, $\mathrm{Tran}(x)$, $\mathrm{On}(x)$ and the like.

Definition 2.10. We use Kripke–Platek set theory KP, cf. [1], as our basic theory. It consists of the axioms of extensionality, pairing, union, of the axiom schemas of separation and collection for Δ_0-formulas and of the foundation schema for arbitrary formulas. KPω arises from KP when adding the infinity axiom
$\exists x(x \neq \emptyset \wedge \forall y \in x \exists z \in x(y \in z))$.

Definition 2.11. (a) The language $\mathscr{L}_{\mathrm{Ad}}$ contains, in addition to $\mathscr{L}_\in$, a unary predicate symbol Ad.

(b) The Ad-axioms are the following:

- $\mathrm{Ad}(x) \rightarrow x \neq \emptyset \wedge \mathrm{Tran}(x)$.
- $\mathrm{Ad}(x) \rightarrow F^x$ for all axioms F of KPω.

(c) KPl is the theory containing extensionality, the axiom schema of foundation, the axioms for Ad and the axiom $\forall x \exists y(x \in y \wedge \mathrm{Ad}(y))$. Since the axioms of KP$\omega$ apart from Δ_0-collection are provable in KPl, we may consider them as axioms if useful.

(d) KPi is KP + KPl.

(e) KPl^r (KPi^r) arises from KPl (KPi) when replacing the axiom schema of foundation by the foundation axiom

$$\forall x(\exists y(y \in x) \rightarrow \exists y(y \in x \wedge \forall z \in x(z \notin y))).$$

(f) KPl^w (KPi^w) is obtained when adding the schema

$$\forall x \in \omega(\forall y \in x F(y) \rightarrow F(x)) \rightarrow \forall x \in \omega F(x)$$

of complete induction to KPl^r (KPi^r).

Remark 2.12. We will use Gödel's constructible hierarchy $L = (L_\alpha)_{\alpha \in \mathrm{On}}$ in one of its usual formulations. For definiteness let

$$L_0 = \emptyset, \qquad L_{\alpha+1} = \mathrm{Def}(L_\alpha), \qquad L_\lambda = \bigcup_{\alpha<\lambda} L_\alpha \quad \text{for } \lambda \in \mathrm{Lim}.$$

Here $\mathrm{Def}(x)$ is the set of all definable subsets of x.

For subsets $X \subseteq \omega$ we will also consider the relativized constructible hierarchy $L(X) = (L_\alpha(X))_{\alpha \in \mathrm{On}}$ which is defined as

$$L_0(X) = \emptyset, \qquad L_{\alpha+1}(X) = \mathrm{Def}_X(L_\alpha(X)), \qquad L_\lambda(X) = \bigcup_{\alpha<\lambda} L_\alpha(X) \quad \text{for } \lambda \in \mathrm{Lim}.$$

Here $\mathrm{Def}_X(x)$ is the set of all subsets definable over the structure $(x, \in\restriction x^2, X \cap x)$ in the language $\mathscr{L}_\in(\mathsf{R})$ which contains an additional relation symbol R.

Definition 2.13. (a) A non-empty, transitive set a is called an admissible set, if $\langle a, \in \rangle \models \mathrm{KP}$.

(b) An ordinal α is called admissible, if L_α is an admissible set.

(c) We use $\lambda\xi.\Omega_\xi$ to denote the enumerating function of the class of admissible ordinals and their limits. ω_1^{CK} is another name for the first admissible ordinal beyond ω.

(d) For $\gamma, \delta \in \mathrm{On}$, the ordinal $\gamma^{+(\delta)}$ is the δ-th ordinal greater than γ which is admissible or a limit of admissibles.

Remark 2.14. The above notions can be formalized in KPl^r. Especially we get a Δ_0-predicate Ad' defining the admissible sets. For this obviously

$$\mathrm{KPl}^r \vdash \forall x(\mathrm{Ad}(x) \to \mathrm{Ad}'(x))$$

is provable, where Ad is the basic predicate symbol of $\mathscr{L}_{\mathrm{Ad}}$ and Ad' is the defined symbol. The converse direction is not provable, since the Ad-axioms do not enforce that every admissible set really falls under Ad. Nevertheless, we will identify Ad and Ad' in the following since each model of KPl^r or of any of the other theories we will encounter can be transformed to a model of the same theory in addition satisfying $\mathrm{Ad}(x) \leftrightarrow \mathrm{Ad}'(x)$ by reinterpreting Ad by the set of all x satisfying Ad'.

We also will use a theory which is similar to the subsystem $\Pi^1_1\text{-}\mathrm{CA}_{<\varepsilon_0}$ of second-order arithmetic. For its formulation, we notice that in KPl^r ω-exponentiation of ordinals can easily be shown to be total. Therefore we can assume that for all ordinals $\delta < \varepsilon_0$ there is a constant in KPl^r.

Definition 2.15. The theory $\mathrm{KPl}^r_{<\varepsilon_0}$ is the extension of KPl^r by the axiom schema which says that for each (meta) $\delta < \varepsilon_0$ and each γ the ordinal $\gamma^{+(\delta)}$ exists. Moreover, it can be shown that there is an order isomorphism from the primitive recursive ordering $\lhd$ onto the set of ordinals less than ε_0, so we may switch freely between these two notions.

Remark 2.16. As in the case of Σ^1_2-AC and Π^1_1-CA$_{<\varepsilon_0}$, KPi^w proves the axioms of KPl$^r_{<\varepsilon_0}$ and so KPl$^r_{<\varepsilon_0}$ can be regarded as a subtheory of KPi^w. This inclusion is conservative for Σ_1-sentences as we will see in Section 6.

We will regard the language of second-order arithmetic as a sublanguage of set theory via the translation mapping numerical quantifiers $\exists x$ to $\exists x(x \in \omega \wedge \ldots)$ and set quantifiers $\exists X$ to $\exists X(X \subseteq \omega \wedge \ldots)$. Here we already used the convention to use upper case letters also for variables in set theory, if they are intended to denote a subset of ω. Also, it will be convenient to be able to perform generalized recursion theory on ω directly within our systems of set theory. For this, we provide the following notions.

Definition 2.17. Let (for convenience) $\delta < \varepsilon_0$. For a set $X \subseteq \omega$, $\omega^X_{\delta+1}$ is the least ordinal which is not the order type of a well-ordering recursive in HJ(δ, X).

Using this, we define $\omega^X_\lambda = \sup\{\omega^X_{\delta+1} : \delta < \lambda\}$ for limit ordinals λ.

Proposition 2.18. *The following are provable in* KPl^r.

(a) *If* $\alpha = \omega^X_{\delta+1}$ *for some* X, *then* $\alpha \in \mathrm{Ad}$.

(b) *Conversely, if* $X \in L_\alpha$ *where* $\alpha \in \mathrm{Ad}$, *then* $\omega^X_1 \leqslant \alpha$.

Proof. The usual proof, cf. e.g. the relativization of [1, Theorem V.5.10] to HJ(δ, X) for (a) and [1, Theorem V.3.3] for (b), can be carried out in KPl^r. □

We finish the section by recalling two theorems which are provable in KPl^r.

Proposition 2.19 (Quantifier Theorem). KPl^r *proves that each (translation of a)* Σ^1_2-*formula is equivalent to a* Σ-*formula.*

Proof. Cf. e.g. [18, Theorem 7.1]. □

Proposition 2.20 (Shoenfield absoluteness). KPl^r *proves that for each* Σ_1-*formula* F *without parameters,* F *is equivalent to* F^L.

Proof. The usual proof, cf. e.g. [1, Theorem V.8.11], can again be formalized in KPl^r. □

3. Stable ordinals and inductive definitions

In this section we introduce the notions of stable ordinals and a special class of inductive operators on the power set of ω. Then we show that the stability properties allow to construct sub-fixed points of these operators. A sort of converse of this will be shown in Section 7.

Definition 3.1. (a) An ordinal γ is *stable* if $L_\gamma \prec_1 L$, i.e. for all $a_1, \ldots, a_n \in L_\gamma$ and all Σ_1-formulas F

$$L \models F[a_1, \ldots, a_n] \Rightarrow L_\gamma \models F[a_1, \ldots, a_n].$$

(b) An ordinal γ is δ-stable if $\gamma \leqslant \delta$ and $L_\gamma \prec_1 L_\delta$.

(c) An ordinal γ is *weakly (δ)-stable* if $L_\gamma \prec_1^- L (L_\gamma \prec_1^- L_\delta)$, i.e. if the above implication holds for parameter-free Σ_1-formulas F.

Remark 3.2. (a) The above definition can be formalized in KPl^r using a universal Σ_1-formula Sat_Σ:

$$L_\gamma \prec_1 L \Leftrightarrow \forall e \in \omega \forall x \in L_\gamma (\mathrm{Sat}_\Sigma(e,x)^L \to \mathrm{Sat}_\Sigma(e,x)^{L_\gamma}).$$

(b) If γ is $\gamma+1$-stable, then it is a first-order reflecting ordinal, from which it easily follows that L_γ is a model of KPi.

The strength of the assumption of stability in the context of a given theory greatly depends on the strength of that theory. Obviously, if the theory can prove strong closure properties of L, it can prove strong closure properties of L_γ for stable ordinals γ.

Proposition 3.3. (a) *For each $n \in \mathbb{N}$, $\mathrm{KPi}^r + L_\gamma \prec_1 L$ proves that for all $\alpha < \gamma$ there is a $\beta < \gamma$ such that $\alpha \leqslant \beta$ and $L_\beta \prec_1 L_{\beta^{+(n)}}$.*

(b) *For each (meta) $\delta < \varepsilon_0$, $\mathrm{KPi}^w + L_\gamma \prec_1 L$ proves that for all $\alpha < \gamma$ there is a $\beta < \gamma$ such that $\alpha \leqslant \beta$ and $L_\beta \prec_1 L_{\beta^{+(\delta)}}$.*

Proof. (a) Choose $\alpha < \gamma$. Using the limit axiom n times, KPi^r proves the existence of $\gamma^{+(n)}$. Since $L_\gamma \prec_1 L$ we especially have $L_\gamma \prec_1 L_{\gamma^{+(n)}}$. So we have

$$L \models \exists \beta \exists \delta (\alpha < \beta \wedge \delta = \beta^{+(n)} \wedge L_\beta \prec_1 L_\delta),$$

where β, δ can be chosen as $\gamma, \gamma^{+(n)}$, respectively. This formula can easily be seen to be equivalent to a Σ_1-formula and therefore stability of γ gives

$$L_\gamma \models \exists \beta \exists \delta (\alpha < \beta \wedge \delta = \beta^{+(n)} \wedge L_\beta \prec_1 L_\delta).$$

A $\beta < \gamma$ satisfying this formula is as required for part (a) of the proposition.

(b) is proved completely analogously using the fact that using the induction scheme up to $\delta < \varepsilon_0$ for arbitrary set theoretic formulas (which can be proved from the schema of complete induction) we can show $\forall \gamma \exists \eta (\eta = \gamma^{+(\delta)})$, especially this holds for γ as in the assertion. Then proceed as in (a). $\square$

Definition 3.4. (a) An operator $\Gamma : \mathrm{Pow}(\omega) \to \mathrm{Pow}(\omega)$ is called a $\Pi^1_{1,\delta}$-operator iff

$$\forall n \in \omega (n \in \Gamma(X) \Leftrightarrow L_{\omega_{\delta+1}^{(X, X_1, \ldots, X_k)}}(\langle X, X_1, \ldots, X_k \rangle) \models F[n, \mathsf{R}])$$

for some Σ_1-formula F. The sets $X_1, \ldots, X_k \subseteq \omega$ are called the parameters of Γ.

(b) The iteration stages of an operator Γ are defined as

$$I_\Gamma^\alpha = \Gamma(I_\Gamma^{<\alpha}) \cup I_\Gamma^{<\alpha} \quad \text{where } I_\Gamma^{<\alpha} = \bigcup_{\beta<\alpha} I_\Gamma^\beta.$$

(c) A set $X \subseteq \omega$ is called a *sub-fixpoint* of an operator Γ if $\Gamma(X) \subseteq X$.

Remark 3.5. The definition of $Y = \Gamma(X)$ is a Σ-statement as it stands, namely it expresses that there is a set which is equal to $L_{\omega_{\delta+1}^{\langle X,X_1,\dots,X_n\rangle}}(\langle X, X_1, \dots, X_n\rangle)$ and in which $F[n]$ is evaluated. But we will in each case only use $\Pi^1_{1,\delta}$-operators such that our meta-theory allows to prove $\forall X \subseteq \omega \exists y$("$y = L_{\omega^X_{\delta+1}}(X)$"), thus turning the operators into Δ-form. For the meta-theory KPi^r these are the $\Pi^1_{1,n}$-operators for $n \in \mathbb{N}$ and for KPi^w the $\Pi^1_{1,\delta}$-operators for $\delta < \varepsilon_0$. Moreover, this also leads to the definition of a Δ-predicate of x and γ which says that $x = (I_\Gamma^\alpha)_{\alpha<\gamma}$.

When working in axiom systems of set theory without the full foundation scheme, it is not always possible to prove the existence of the sequences $(I_\Gamma^\alpha)_{\alpha<\gamma}$ for arbitrary γ. But it will be enough to prove existence in the "local" form of the following lemma.

Lemma 3.6 (KPl^r). *If $L_\gamma \models$ KPi and Γ is a $\Pi^1_{1,\delta}$-operator in parameters from L_γ for some (meta) $\delta < \varepsilon_0$, then $(I_\Gamma^\alpha)_{\alpha<\gamma}$ can be defined by Σ-recursion in L_γ and therefore is an element of $L_{\gamma+1}$.*

Moreover, the definition of I_Γ^α is absolute w.r.t. all models of KPi *containing the parameters of Γ.*

Proof. Standard. □

Lemma 3.7 (KPl^r). *Assume Γ is monotone on L_γ where $L_\gamma \models$ KPi and Γ is a $\Pi^1_{1,\delta}$-operator in parameters from L_γ for some (meta) $\delta < \varepsilon_0$. Then for all $X \in L_\gamma$ such that $\Gamma(X) \subseteq X$ and all $\alpha < \gamma$ we have $I_\Gamma^\alpha \subseteq X$.*

Proof. Obvious induction on α. □

Proposition 3.8 (KPl^r). *Assume $L_\gamma \prec_1 L_{\gamma+(\delta+1)}$ where (meta) $\delta < \varepsilon_0$. If Γ is a $\Pi^1_{1,\delta}$-operator with parameters from L_γ, then $\Gamma(I_\Gamma^{<\gamma}) \subseteq I_\Gamma^{<\gamma}$.*

Proof. Assume $n \in \Gamma(I_\Gamma^{<\gamma})$. Let $Y := \langle I_\Gamma^{<\gamma}, X_1, \dots, X_n\rangle$ where $X_1, \dots, X_n$ are the parameters of Γ. This means

$$L_{\omega^Y_{\delta+1}}(Y) \models F[n, \mathsf{R}]$$

for the corresponding Σ_1-formula F from Definition 3.4. Since $I_\Gamma^{<\gamma} \in L_{\gamma+1}$ holds by Lemma 3.6, we have $\omega^Y_{\delta+1} \leqslant \gamma^{+(\delta+1)}$ and so it holds

$$L_{\gamma^{+(\delta+1)}} \models \exists z \exists \alpha \exists \beta \exists X \exists Y (X = I_\Gamma^{<\alpha} \wedge Y = \langle X, X_1, \dots, X_n\rangle \wedge$$
$$\beta < \omega^Y_{\delta+1} \wedge z = L_\beta(Y) \wedge z \models F[n]).$$

This can be formalized by a Σ_1-formula with parameters from L_γ. So the stability property of γ gives

$$L_\gamma \models \exists z \exists \alpha \exists \beta \exists X \exists Y (X = I_\Gamma^{<\alpha} \wedge Y = \langle X, X_1, \ldots, X_n \rangle \wedge \beta < \omega_{\delta+1}^{Y} \wedge z = L_\beta(Y) \wedge z \models F[n]).$$

Picking such an ordinal α, we can conclude from this formula that

$$L_{\langle I_\Gamma^{<\alpha}, X_1, \ldots, X_n \rangle}(\langle I_\Gamma^{<\alpha}, X_1, \ldots, X_n \rangle) \models F[n]$$

and therefore $n \in \Gamma(I_\Gamma^{<\alpha}) \subseteq I_\Gamma^{<\gamma}$. □

Corollary 3.9. (a) *For each* $n \in \mathbb{N}$ KPi^r *proves that each* $\Pi^1_{1,n}$*-operator* Γ *with parameters from a set* L_γ *such that* $L_\gamma \prec L$ *has a sub-fixpoint in* L_γ.

Moreover, if Γ *is monotone on* L_γ, *it has a* minimal *sub-fixpoint in* L_γ.

(b) KPi^w *proves the above for all* $\Pi^1_{1,\delta}$*-operators where* $\delta < \varepsilon_0$.

Proof. Work in KPi^r under the assumption that γ is stable. Then for each $\alpha < \gamma$ there is a $\alpha \leqslant \beta < \gamma$ such that $L_\beta \prec_1 L_{\beta^{+(n+1)}}$. Using the previous proposition for β, the assertion follows. The minimality condition follows applying Lemma 3.7 to L_γ itself.

Similarly, KPi^w proves that for each $\alpha < \gamma$ there is a $\alpha \leqslant \beta < \gamma$ such that $L_\alpha \prec_1 L_{\alpha^{+(\delta+1)}}$. Then again use the previous proposition. □

4. Modeling T_0 in set theory

4.1. Applicative structures

Modeling the applicative part of T_0 can be done in very weak systems of set theory, since only recursively enumerable sets are necessary. Nevertheless, we again use KPI^r as our base system in this subsection since we do not need more exact information in the following part. Since the models we use are already well described in the literature, cf. [7, 27], we do not present the full details.

We start off with the pair structure $\mathfrak{S}^{\text{pair}} = (S, \pi, \pi_0, \pi_1, 0)$ where $S = \omega$ and $\pi : S^2 \to S \setminus \{0\}$ is an injective (recursive) pairing function with (recursive) inverses π_0, π_1 such that $\pi_0(0) = \pi_1(0) = 0$. For technical reasons we moreover fix a special such function π, namely $\pi(x, y) = 2^x \cdot 3^y$. As its inverses, we fix π_0, π_1 where $\pi_0(z) = x$ and $\pi_1(z) = y$ if $z = 2^x \cdot 3^y$ and $\pi_0(z) = \pi_1(z) = z$ if z cannot be written in this form.

We call the base set S (and not ω) since we will have "natural numbers" in this model and we want to avoid confusion between those two sets. Moreover, the intuition about S is that S consists of general objects and not only of the natural numbers.

For each $n \in \omega$, the representation $n^\circ \in S$ of n in the structure $\mathfrak{S}^{\text{pair}}$ is defined inductively by $0^\circ = 0$, $(n+1)^\circ = \pi(0, n^\circ)$. Then let $N_S \subseteq S$ be the set of all n° for

$n \in \omega$. More generally, for $X \subseteq \omega$ let $X^\circ = \{n^\circ : n \in X\}$. In the following, we use the codes

$$\mathsf{k} = 1^\circ, \quad \mathsf{s} = 2^\circ, \quad \mathsf{p} = 3^\circ, \quad \mathsf{p}_0 = 4^\circ, \quad \mathsf{p}_1 = 5^\circ, \quad \mathsf{d} = 6^\circ,$$

$$\mathsf{s}_N = 7^\circ, \quad \mathsf{p}_N = 8^\circ, \quad \mathsf{i} = 9^\circ, \quad \mathsf{j} = 10^\circ, \quad \text{and } \mathsf{c}_m = (11+m)^\circ.$$

The relation $App \subseteq S^3$ then is inductively defined by the following clauses, where we use the abbreviations $xy \simeq z :\equiv App(x,y,z)$ and (x,y) for $\pi(x,y)$.

- $\mathsf{k}x \simeq (\mathsf{k},x), (\mathsf{k},x)y \simeq x, \mathsf{s}x \simeq (\mathsf{s},x), (\mathsf{s},x)y \simeq ((\mathsf{s},x),y)$
- $\mathsf{p}x \simeq (\mathsf{p},x), (\mathsf{p},x)y \simeq \pi(x,y), \mathsf{p}_0 x \simeq \pi_0(x), \mathsf{p}_1 x = \pi_1(x)$
- $\mathsf{d}x \simeq (\mathsf{d},x), (\mathsf{d},x)y \simeq ((\mathsf{d},x),y), ((\mathsf{d},x),y)z_1 \simeq (((\mathsf{d},x),y),z_1)$
  $(((\mathsf{d},x),y),z_1)z_2 = \begin{cases} x & \text{if } z_1 = z_2 \\ y & \text{if } z_1 \neq z_2 \end{cases}$
- $\mathsf{s}_N x \simeq (0,x), \mathsf{p}_N(0,x) \simeq x$
- If $xz \simeq u, yz \simeq v$ and $uv \simeq w$, then $((\mathsf{s},x),y)z \simeq w$.

This defines an applicative structure $\mathfrak{S}^{\text{app}} = (S, App, N_S, \mathsf{k}, \mathsf{s}, \mathsf{p}, \mathsf{p}_0, \mathsf{p}_1, \mathsf{d}, \mathsf{s}_N, \mathsf{p}_N, 0)$ such that $\mathfrak{S}^{\text{app}}$ models the applicative part of T_0. $\mathfrak{S}^{\text{app}}$ can be shown to be an element of $L_{\omega_1^{CK}}$, actually, it can be shown to be coded by recursively enumerable sets.

Definition 4.1. (a) Let $B \subseteq S$ be defined as $B := S \setminus \pi[S^2]$. Denoting the closure of B under π by Gen(B), we see that $\pi : S^2 \to S \setminus B$, Gen$(B) = S$ and $\pi_0(x) = \pi_1(x) = x$ for all $x \in B$. We therefore say that B is an *atomic base* for S.

(b) For $x \in S$ then $\text{supp}_B(x) \subseteq B$ is defined by recursion on the definition of Gen(B) by $\text{supp}_B(x) = \{x\}$ for $x \in B$ and $\text{supp}_B(\pi(x,y)) = \text{supp}_B(x) \cup \text{supp}_B(y)$.

(c) For finite sets $F \subseteq B$ let

$$\text{Aut}(B/F) := \{\sigma : B \to B : \sigma \text{ is bijective and } \forall x \in F \cup \{0\}(\sigma(x) = x)\}$$

We identify each $\sigma \in \text{Aut}(B/F)$ with the mapping $\sigma : S \to S$ it induces via $\sigma(\pi(x,y)) = \pi(\sigma(x), \sigma(y))$.

(d) Let $\text{Pow}(S/F) = \{X \in \text{Pow}(S) : \sigma[X] = X \text{ for all } \sigma \in \text{Aut}(B/F)\}$.

Lemma 4.2. (a) *If* $xy \simeq z$, *then* $\text{supp}_B(z) \subseteq \text{supp}_B(x) \cup \text{supp}_B(y)$.

(b) *If* $\sigma \in \text{Aut}(B/\emptyset)$, *then* $xy \simeq z \Leftrightarrow \sigma(x)\sigma(y) \simeq \sigma(z)$.

(c) *If* $f \in S$ *and* $b, c \in B \setminus (\text{supp}_B(f) \cup \{0\})$, *then*

$$\forall x \in S(fb \simeq x \to fc \simeq x[b := c]),$$

where $x[b := c]$ *is the obvious substitution of one base element* b *by an element* $c \in S$ *in* $x \in S = \text{Gen}(B)$.

(d) *If* $X \subseteq \omega$, *then* $X^\circ \in \text{Pow}(S/F)$ *for all finite* $F \subseteq B$.

Proof. (a), (b), (c) can be proved by induction over the definition of App, (d) is obvious from the definitions. For details cf. [27]. □

Definition 4.3. The *trace* $\mathrm{tr}_{B/F}(x)$ of an element $x \in S$ over the finite set $F \subseteq B$ is defined by $\mathrm{tr}_{B/F}(x) = \{\sigma(x) : \sigma \in \mathrm{Aut}(B/F)\}$.

Lemma 4.4. (a) *The predicate* $\mathrm{tr}_{B/F}(x) \subseteq X$ *is arithmetical in* x, X.
(b) *For any set* $X \subseteq S$ *it holds* $\bigcup\{\mathrm{tr}_{B/F}(x) : x \in X\} \in \mathrm{Pow}(S/F)$.
(c) *For any set* $X \subseteq S$ *it holds* $\bigcup\{x : \mathrm{tr}_{B/F}(x) \subseteq X\} \in \mathrm{Pow}(S/F)$.

Proof. (a) follows because in the definition of $\mathrm{tr}_{B/F}(x) \subseteq X$ by

$$\mathrm{tr}_{B/F}(x) \subseteq X \Leftrightarrow \forall \sigma \in \mathrm{Aut}(B/F)(\sigma(x) \in X)$$

we can replace the quantification over $\mathrm{Aut}(B/F)$ by quantification over finite sequences in $\mathrm{supp}_B(x)$.

(b), (c) are verified straightforwardly. □

We will need a set $M \subseteq B$ which will provide names for parameters we want to code into our models for T_0. For this set M some special technical conditions are needed. We now describe this set M and an atomic base B. The idea again stems from [27].

Since B is the set of non-pairs with respect to π, we can construct

- a partition $B = \sum_{n \in \omega} B_n$.
- for each finite $F \subseteq B$ infinite sets $M_F^{(n)}$ such that the $M_F^{(n)}$ are pairwise disjoint and the following property holds:
  If $F \not\subseteq \bigcup_{m<n} B_m$ with $n > 0$, then $M_F^{(n)} = \emptyset$ and if $F \subseteq \bigcup_{m<n} B_m$ for $n > 0$, then $M_F^{(n)} \subseteq B_n$.

Namely, for $n \in \omega$ we define $B_{n+1} = \{(p_{n+2})^{x+1} : x \in \omega\} \subseteq B$, $B_0 = B \backslash \bigcup_{n>0} B_n$. If we identify $F = \{b_0, \ldots, b_k\}$ where $b_0 < \cdots < b_k$ with the sequence $\langle b_0, \ldots, b_k \rangle$ we can define for $F \subseteq B$

$$M_F^{(n)} = \begin{cases} \{p_{n+1}^{\langle F,x \rangle} : x \in \omega\} & \text{if } F \subseteq \bigcup\limits_{m<n} B_m \text{ and } n > 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

These sets are as required.

We get the set $M \subseteq B$ of names for the parameters by defining $M_F = \bigcup_{n>0} M_F^{(n)}$ and $M = \bigcup_{F \subseteq B \text{ finite}} M_F$. Since $B_0 \subseteq B \backslash M$, we see that $B \backslash M$ is infinite.

4.2. *Models with finitely many parameters*

We work in KPi in this subsection. We want to formalize the standard models of T_0, originally defined in [6], using a finite set $M_0 \subseteq M$ (where M from now on is the set defined above) to denote parameters.

So fix $M_0 \subseteq M$ and sets $\widehat{b} \subseteq S$ for $b \in M_0$, the parameters of the construction.

Definition 4.5. By induction on α we define structures

$$\mathfrak{S}_{M_0,\alpha} = (S, \mathrm{Cl}_{M_0,\alpha}, \varepsilon_{M_0,\alpha}, App, N_S, \mathsf{k}, \mathsf{s}, \mathsf{p}, \mathsf{p}_0, \mathsf{p}_1, \mathsf{d}, \mathsf{s}_{\mathrm{N}}, \mathsf{p}_{\mathrm{N}}, 0, (\mathsf{c}_m)_{m\in\omega}, \mathsf{j}, \mathsf{i})$$

extending $\mathfrak{S}^{\mathrm{app}}$. So we only have to define $\mathrm{Cl}_{M_0,\alpha} \subseteq S$ and $\varepsilon_{M_0,\alpha} \subseteq S \times \mathrm{Cl}_{M_0,\alpha}$.

(a) $\mathrm{Cl}_{M_0,0} = M_0$ and for $b \in M_0$ and $z \in S$ let $z\ \varepsilon_{M_0,0}\ b \Leftrightarrow z \in \widehat{b}$.

(b) If $\alpha = \beta + 1$ is a successor, then let $\mathrm{Cl}_{M_0,\beta} \subseteq \mathrm{Cl}_{M_0,\alpha}$ and $\varepsilon_{M_0,\beta} \subseteq \varepsilon_{M_0,\alpha}$ and additionally:

- If F is an elementary formula with Gödelnumber m, then let $\mathsf{c}_m(\vec{x},\vec{a}) \in \mathrm{Cl}_{M_0,\alpha}$ for all $\vec{x} \in S$ and $\vec{a} \in \mathrm{Cl}_{M_0,\beta}$. Further define

$$z\ \varepsilon_{M_0,\alpha}\ \mathsf{c}_m(\vec{x},\vec{a}) \Leftrightarrow \mathfrak{S}_{M_0,\beta} \models F[z,\vec{x},\vec{a}].$$

- If $a \in \mathrm{Cl}_{M_0,\beta}$, $f \in S$ and $\mathfrak{S}_{M_0,\beta} \models \forall x\ \varepsilon\ a \exists Y (fx \simeq Y)$, then let $\mathsf{j}(f,a) \in \mathrm{Cl}_{M_0,\alpha}$ and for $z \in S$

$$z\ \varepsilon_{M_0,\alpha}\ \mathsf{j}(f,a) \Leftrightarrow \mathfrak{S}_{M_0,\beta} \models \exists x\ \varepsilon\ a \exists y\ \varepsilon\ fx (z = (x,y)).$$

- For $a, b \in \mathrm{Cl}_{M_0,\beta}$ let $\mathsf{i}(a,b) \in \mathrm{Cl}_{M_0,\alpha}$ and for $z \in S$ let

$$z\ \varepsilon_{M_0,\alpha}\ \mathsf{i}(a,b) \Leftrightarrow \forall X \subseteq S (\mathrm{Prog}(a,b,X) \to z \in X)$$

  where $\mathrm{Prog}(a,b,X) :\equiv$

$$\forall x\ \varepsilon_{M_0,\beta}\ a (\forall y \in S ((y,x)\ \varepsilon_{M_0,\beta}\ b \to y \in X) \to x \in X).$$

(c) If α is a limit ordinal, then let $\mathrm{Cl}_{M_0,\alpha} = \bigcup_{\beta<\alpha} \mathrm{Cl}_{M_0,\beta}$ and $\varepsilon_{M_0,\alpha} = \bigcup_{\beta<\alpha} \varepsilon_{M_0,\beta}$.

Remark 4.6. The formalization of the preceding definition in KPi needs some care. On the one hand, a truth definition for the structures $\mathfrak{S}_{M_0,\alpha}$ is necessary. The definition of $\mathfrak{S}_{M_0,\alpha} \models F$ is given by recursion on the set of formulas. This recursion can be performed in any admissible set containing $\mathfrak{S}_{M_0,\alpha}$ as an element therefore leading to a Δ-notion on this admissible set.

On the other hand, in the clause for (IG), the definition of $z\ \varepsilon_{M_0,\alpha}\ a$ uses a condition which is Π^1_1 on $S = \omega$ in the parameter $\varepsilon_{M_0,\beta}$ where $\alpha = \beta + 1$. As is known from generalized recursion theory, cf. e.g. [1, Theorem IV.3.1], this is equivalent to a condition which is Σ_1 on the next admissible set, namely

$$\begin{aligned} z\ \varepsilon_{M_0,\alpha}\ \mathsf{i}(a,b) &\Leftrightarrow \forall X \subseteq S (\mathrm{Prog}(a,b,X) \to z \in X) \\ &\Leftrightarrow \exists \beta \exists f (\mathrm{fun}(f) \wedge \mathrm{dom}(f) \subseteq a \wedge \mathrm{rng}(f) \subseteq \mathrm{On} \wedge z \in \mathrm{dom}(f) \wedge \\ &\qquad \forall x, y \in S ((x,y)\ \varepsilon_{M_0,\beta}\ b \wedge y \in \mathrm{dom}(f)) \\ &\qquad \to x \in \mathrm{dom}(f) \wedge f(x) < f(y)) \\ &\Leftrightarrow \exists \beta \in u \exists f \in u (\ldots) \end{aligned}$$

for all u such that $\mathrm{Ad}(u)$ and $\varepsilon_{M_0,\beta} \in u$.

So the above definition can be given by Σ-recursion in the theory KPi. Moreover, we see that in each inductive step, the definition uses a Σ_1-predicate over the next admissible set. More exactly, we get the following result:

Proposition 4.7. *There are Σ-predicates $P_0(b, M_0, \alpha, \mathsf{R})$, $P_1(x, b, M_0, \alpha, \mathsf{R})$ with the following property: for every finite $M_0 \subseteq M$, every set $X \subseteq S \times M_0$ let $\widehat{\ } : M_0 \to \mathrm{Pow}(S)$ be defined by $\widehat{a} = \{x \in S : \langle x, a\rangle \in X\}$. Then we have*

$$b \in \mathrm{Cl}_{M_0,\alpha} \Leftrightarrow P_0(b, M_0, \alpha, X)$$

and

$$x \, \varepsilon_{M_0,\alpha} \, b \Leftrightarrow P_1(x, b, M_0, \alpha, X).$$

This means that the inductive definition clauses for $\mathrm{Cl}_{M_0,\alpha}$ and $\varepsilon_{M_0,\alpha}$ (built on the basis of $\widehat{\ } \restriction M_0$) of the above definition can be proved. Moreover, it holds

$$P_0(b, M_0, \alpha, X) \Leftrightarrow L_{\omega^X_{\alpha+1}}(X) \models P_0(b, M_0, \alpha, \mathsf{R})$$

and

$$P_1(x, b, M_0, \alpha, X) \Leftrightarrow L_{\omega^X_{\alpha+1}}(X) \models P_1(x, b, M_0, \alpha, \mathsf{R}).$$

For finite subsets $M_0 \subseteq M$ we identify mappings $\widehat{\ } : M_0 \to \mathrm{Pow}(S)$ with the corresponding set X as in the definition and use the more suggestive notions $b \in \mathrm{Cl}_{M_0,\alpha}$ and $x \, \varepsilon_{M_0,\alpha} \, b$ where the mapping $\widehat{\ }$ is to be understood by the context.

Lemma 4.8. *Let $M_1 \subseteq M_0$ be finite subsets of M and $\widehat{\ } : M_0 \to \mathrm{Pow}(S)$ be given.*

(a) *For all α we have $\mathrm{Cl}_{M_1,\alpha} \subseteq \mathrm{Cl}_{M_0,\alpha}$.*

(b) *For all $x \in S$ and $a \in \mathrm{Cl}_{M_1,\alpha}$ it is $x \, \varepsilon_{M_1,\alpha} \, a \Leftrightarrow x \, \varepsilon_{M_0,\alpha} \, a$.*

(c) *For all $\alpha \leqslant \beta$, all $a \in \mathrm{Cl}_{M_0,\alpha}$ and all $x \in S$ it is $x \, \varepsilon_{M_0,\alpha} \, a \Leftrightarrow x \, \varepsilon_{M_0,\beta} \, a$.*

Proof. Induction on α. $\square$

In the following we may leave out the indices M_0, α from the relation $x \, \varepsilon \, a$ since the preceding lemma shows that the relation is independent of these parameters as long as $a \in \mathrm{Cl}_{M_0,\alpha}$.

Lemma 4.9. *Let $M_1 \subseteq M_0 \subseteq M$ and $\widehat{\ } : M_0 \to \mathrm{Pow}(S/F)$ for a finite set $F \subseteq S$. If $\sigma \in \mathrm{Aut}(B/F)$ satisfies both $\sigma \restriction M_1 : M_1 \to M_0$ and $\widehat{b} = \widehat{\sigma(b)}$ for all $b \in M_1$, then for all $a \in \mathrm{Cl}_{M_1,\alpha}$ and $x \in S$:*

(a) *$a \in \mathrm{Cl}_{M_1,\alpha} \Leftrightarrow \sigma(a) \in \mathrm{Cl}_{\sigma[M_1],\alpha}$*

(b) *$x \, \varepsilon \, a \Leftrightarrow \sigma(x) \, \varepsilon \, \sigma(a)$.*

Proof. Induction on α, cf. [27, Lemma 3.1]. $\square$

Corollary 4.10. *Let $F \subseteq S$ and $M_0 \subseteq M$ be finite,* $\hat{} : M_0 \to \mathrm{Pow}(S/F)$ *and* $\sigma \in \mathrm{Aut}(B/M_0 \cup F)$. *Then*

$$a \in \mathrm{Cl}_{M_0,\alpha} \Leftrightarrow \sigma(a) \in \mathrm{Cl}_{M_0,\alpha} \quad \textit{and} \quad x \,\varepsilon\, a \Leftrightarrow \sigma(x) \,\varepsilon\, \sigma(a).$$

Proof. Take $M_0 = M_1$ in Lemma 4.9. □

Corollary 4.11. *Let* $\hat{} : M_0 \to \mathrm{Pow}(S/F)$ *for finite* $M_0 \subseteq M$ *and* $F \subseteq S$. *Further assume* $G \subseteq S$ *is finite and* $(F \cup G) \cap M \subseteq M_0$.

For all $a \in \mathrm{Cl}_{M_0,\alpha}$ *satisfying* $\mathrm{supp}_B(a) \cap M \subseteq M_0$ *and all* $x \,\varepsilon\, a$ *there is an automorphism* $\sigma \in \mathrm{Aut}(B/M_0 \cup F \cup G)$ *such that* $\sigma(x) \in a$ *and* $\mathrm{supp}_B(\sigma(x)) \cap M \subseteq M_0$.

Proof. In this situation let $H = (\mathrm{supp}_B(x) \cap M) \setminus M_0$. By hypothesis

$$(M_0 \cup F \cup G \cup \mathrm{supp}_B(a)) \cap M \subseteq M_0$$

and so $H \cap (M_0 \cup F \cup G \cup \mathrm{supp}_B(a)) \subseteq H \cap M_0 = \emptyset$. Since $B \setminus M$ is infinite, cf. subsection 4.1, there is a $\sigma \in \mathrm{Aut}(B/M_0 \cup F \cup G \cup \mathrm{supp}_B(a))$ such that $\sigma : H \to B \setminus M$. For this σ it is $\sigma(a) = a$ since σ is the identity on $\mathrm{supp}_B(a)$. By Corollary 4.10 we therefore have $\sigma(x) \,\varepsilon\, a$. Moreover, we have

$$\mathrm{supp}_B(\sigma(x)) \cap M = \sigma[\mathrm{supp}_B(x)] \cap M \subseteq \sigma[M_0] \subseteq M_0$$

as desired. □

Proposition 4.12. *Let* $M_0 \subseteq M_1$ *be finite subsets of* M *and* $\hat{} : M_1 \to \mathrm{Pow}(S)$. *If* $\hat{} \restriction M_0 : M_0 \to \mathrm{Pow}(S/F)$ *for a finite set* F *such that* $F \cap M \subseteq M_0$, *then we have for all* $a \in \mathrm{Cl}_{M_1,\alpha}$

$$\mathrm{supp}_B(a) \cap M \subseteq M_0 \Rightarrow a \in \mathrm{Cl}_{M_0,\alpha}.$$

Proof. Again, we use induction on α. In the most important case that $a = \mathrm{j}(f, a_0)$, we use the following lemma. The induction hypothesis guarantees the assumption of that lemma. □

Lemma 4.13. *Let* $M_0, M_1, F, \hat{}$ *be as in Proposition* 4.12. *If* $a_0 \in \mathrm{Cl}_{M_0,\alpha}$,

$$(\mathrm{supp}_B(a_0) \cup \mathrm{supp}_B(f)) \cap M \subseteq M_0 \quad \textit{and} \quad \forall x \,\varepsilon\, a_0 \exists y \in \mathrm{Cl}_{M_1,\alpha}[fx \simeq y]$$

and, furthermore,

$$\forall x \,\varepsilon\, a_0 \forall y \in \mathrm{Cl}_{M_1,\alpha}[fx \simeq y \wedge \mathrm{supp}_B(y) \cap M \subseteq M_0 \Rightarrow y \in \mathrm{Cl}_{M_0,\alpha}],$$

then for $x \,\varepsilon\, a_0$ *and* $fx \simeq y$ *we have* $y \in \mathrm{Cl}_{M_0,\alpha}$.

Proof. Let $G := \mathrm{supp}_B(a_0) \cup \mathrm{supp}_B(f)$ and $a := a_0$ in Corollary 4.11. Then the hypotheses of this corollary are satisfied and so there is an automorphism $\sigma \in \mathrm{Aut}(B/M_0 \cup F \cup \mathrm{supp}_B(a_0) \cup \mathrm{supp}_B(f))$ such that $\sigma(x) \in a_0$ and $\mathrm{supp}_B(\sigma(x)) \cap M \subseteq M_0$.

From $fx \simeq y$ we therefore conclude

$$\text{supp}_B(\sigma(y)) \subseteq \text{supp}_B(\sigma(f)) \cup \text{supp}_B(\sigma(x)) \subseteq \text{supp}_B(f) \cup \text{supp}_B(\sigma(x))$$

since $\sigma(f) = f$. This gives $\text{supp}_B(\sigma(y)) \cap M \subseteq M_0$. Using $\sigma(y)$ instead of y in the additional hypothesis of the lemma we get $\sigma(y) \in \text{Cl}_{M_0,\alpha}$ and from this $y \in \text{Cl}_{M_0,\alpha}$ applying Corollary 4.10 to σ^{-1}. □

4.3. Models with infinitely many parameters

Up to now we have only used finitely many parameters in our models $\mathfrak{S}_{M_0}$. But we will need to use infinitely many parameters in our models, actually we will use all suitable sets of integers in a certain set theoretic universe, namely all that are in L_γ for a countable stable ordinal γ, as parameters. The idea is to choose a surjection

$$\widehat{\ }: M \to \bigcup\{X \in L_\gamma : X \in \text{Pow}(S/F) \text{ for some finite } F \subseteq B\}$$

and to define $\text{Cl}_{M,\alpha} = \bigcup\{\text{Cl}_{M_0,\alpha} : M_0 \subseteq M \text{ is finite}\}$.

The problem with this approach obviously is that the function $\widehat{\ }$ will not be an element of L_γ and that we are already lucky if we manage to show that it is definable at all over L_γ. It will definitely not be Σ_1-definable and so we cannot hope to have it available for arguments in L_γ. But the levels of its finite approximations $\text{Cl}_{M_0,\alpha}$ are elements of L_γ and so we will have to get by with those.

In any case, we will have to step out of KPi. But for the moment it will be enough to assume that we are given some countable γ such that $L_\gamma \models \text{KPi}$.[4] Apart from the assumption of the existence of such an ordinal, we work in KPl^r in this section. Within L_γ we can use the results from the preceding section.

First we construct the mapping $\widehat{\ }: M \to \text{Pow}(S)$. For this we also use the sets $M_F^{(n)}$ which were used in subsection 4.1 to define $M_F = \bigcup\{M_F^{(n)} : n \in \omega\}$ and $M = \bigcup\{M_F : F \subseteq B \text{ is finite}\}$.

Let for $F \subseteq B$ the set $\mathscr{M}_F$ be defined as

$$\mathscr{M}_F = \{X \in L_\gamma : X \in \text{Pow}(S/F)\}$$

and $\mathscr{M} = \bigcup\{\mathscr{M}_F : F \subseteq B \text{ is finite}\}$. Choose then a surjection $\widehat{\ }: M \to \mathscr{M}$ such that for all n

$$\widehat{\ } \restriction M_F^{(n)} : M_F^{(n)} \to \mathscr{M}_F$$

is a surjective mapping onto $\mathscr{M}_F$ if $F \subseteq \bigcup_{m<n} B_m$. Note that for each set $X \subseteq \omega$ in L_γ the set $X^\circ = \{n^\circ : n \in X\}$ is also in L_γ and so it is in $\mathscr{M}$.

[4] In the proof of Theorem 9.1 we will see how to construct an appropriate γ.

Lemma 4.14. (a) *For each $b \in M$ the set $\{b' \in M : \widehat{b} = \widehat{b'}\}$ is infinite.*

(b) *For each finite set $F_0 \subseteq B$ there is a finite set F_1 such that $F_0 \subseteq F_1 \subseteq B$ and $\widehat{\ } \upharpoonright F_1 \cap M : F_1 \cap M \to \mathscr{M}_F$.*

Proof. (a) Let $b \in M$, i.e. $\widehat{b} \in \mathscr{M}_F$ for some finite $F \subseteq B$. For each of the infinitely many n such that $F \subseteq \bigcup_{m<n} B_m$, $\widehat{b}$ is in the range of $\widehat{\ } \upharpoonright M_F^{(n)}$. This gives infinitely many b' with $\widehat{b} = \widehat{b'}$.

(b) can be verified by elementary computation, cf. [27, Lemma 3.11]. □

Now we use the mapping $\widehat{\ }$ given above to define a model $\mathfrak{S}_M$. All we have to do is to define the stages $\mathrm{Cl}_{M,\alpha}$ and $\varepsilon_{M,\alpha}$ of the classifications and the ε-relation of the model. Let for $\alpha < \gamma$ and finite sets $M_0 \subseteq M$ the sets $\mathrm{Cl}_{M_0,\alpha}$ and $\varepsilon_{M_0,\alpha}$ be defined as in subsection 4.2 based on the restrictions of the mapping $\widehat{\ }$. Let then $\mathrm{Cl}_{M,\alpha} = \bigcup\{\mathrm{Cl}_{M_0,\alpha} : M_0 \subseteq M \text{ is finite}\}$ and $\varepsilon_{M,\alpha} = \bigcup\{\varepsilon_{M_0,\alpha} : M_0 \subseteq M \text{ is finite}\}$. Although we will not use it, it is good for motivational purposes to define $\mathrm{Cl}_M = \bigcup\{\mathrm{Cl}_{M,\alpha} : \alpha < \gamma\}$ and $\varepsilon_M = \bigcup\{\varepsilon_{M,\alpha} : \alpha < \gamma\}$. $\mathfrak{S}_M$ is then as usual $\mathfrak{S}_M = (S, \mathrm{Cl}_M, \varepsilon_M, \ldots)$.

First we note one important point in the construction of our models. Namely, each classification in the model is extensionally equal to one in the "basis" $\mathrm{Cl}_{M,0}$. Although we will only need this for elements in levels $\mathrm{Cl}_{M,\delta}$ where $\delta < \varepsilon_0$, we formulate it in full generality here.

Lemma 4.15. *Let $a \in \mathrm{Cl}_{M,\delta}$ for some $\delta < \gamma$. Then there exists a set $Y \in \mathscr{M} = \{X \in L_\gamma : X \in \mathrm{Pow}(S/F)$ for some finite $F \subseteq M\}$ such that*

$$\mathfrak{S}_{M,\delta} \models x \,\varepsilon\, a \Leftrightarrow x \in Y.$$

Proof. Let $a \in \mathrm{Cl}_{M,\delta}$ and the finite set F_0 defined by $F_0 := \mathrm{supp}_B(a)$. By Lemma 4.14 (b), there is a finite $F \subseteq B$ containing F_0 such that $\widehat{\ } : F \cap M \to \mathscr{M}_F$. Defining $M_0 := F \cap M = \{b_0, \ldots, b_n\}$ we have $a \in \mathrm{Cl}_{M_0,\delta}$ by Lemma 4.12 and so Lemma 4.7 yields

$$x \,\varepsilon\, a \Leftrightarrow L_{\omega^X_{\delta+1}}(X) \models P_1(x, a, M_0, \delta, \mathsf{R})$$

for X with $X_{b_i} = \widehat{b_i}$ for $i = 0, \ldots, n$. The set

$$Y = \{x \in S : L_{\omega^X_{\delta+1}}(X) \models P_1(x, a, M_0, \delta, \mathsf{R})\}$$

is in L_γ. Moreover, $Y \in \mathscr{M}_F$ since for $\sigma \in \mathrm{Aut}(B/F)$ we can use the equivalences

$$x \in Y \Leftrightarrow x \,\varepsilon\, a \Leftrightarrow \sigma(x) \,\varepsilon\, \sigma(a) = a \Leftrightarrow \sigma(x) \in X.$$

So $Y \in \mathscr{M}$ as desired. □

The following lemma is central to our embedding. It is a refinement of Takahashi's Theorem 3.8 in [27].

Lemma 4.16. *Let $L_\gamma \models \mathrm{KPi}$, $\widehat{\ }: M \to \mathrm{Pow}(S)$ and $\mathfrak{S}_M$ be as above. Assume that for some $f \in S$ and for all $b \in M$*

$$\mathfrak{S}_{M,\delta} \models \exists Y(fb \simeq Y)$$

as well as for all $b_1, b_2 \in M$

$$\mathfrak{S}_{M,\delta} \models b_1 \mathrel{\mathring{=}} b_2 \to fb_1 \mathrel{\mathring{=}} fb_2.$$

Then there is a finite $F \subseteq B$ and a $\Pi^1_{1,\delta}$-functional $\Gamma : \mathrm{Pow}(\omega) \to \mathrm{Pow}(\omega)$ such that for all $b \in M$ with $\widehat{b} \in \mathrm{Pow}(S/F)$

$$\Gamma(\widehat{b}) \in \mathrm{Pow}(S/F) \quad \text{and} \quad \mathfrak{S}_{M,\delta} \models x \,\varepsilon\, fb \Leftrightarrow x \in \Gamma(\widehat{b}).$$

Proof. By Lemma 4.14 choose $F \supseteq \mathrm{supp}_B(f)$ such that $\widehat{\ }: F \cap M \to \mathrm{Pow}(S/F)$. Write $F \cap M = \{b_1, \ldots, b_n\}$ and choose $b_0 \in M \backslash F$. Then it is $\mathfrak{S}_{M,\delta} \models fb_0 \simeq a$ for some $a \in \mathrm{Cl}_{M,\delta}$. Let $M_0 := \{b_0, \ldots, b_n\}$.

Then $\mathrm{supp}_B(a) \cap M \subseteq \mathrm{supp}_B(f) \cup \mathrm{supp}_B(b_0) \subseteq \{b_0, \ldots, b_n\}$ and so by 4.12 (for some $M_1 \supseteq M_0$ such that $a \in \mathrm{Cl}_{M_1,\delta}$) we see that $a \in \mathrm{Cl}_{M_0,\delta}$.

Define the operator Γ by

$$\begin{aligned}\Gamma(X) = \{x : L_{\omega^U_{\delta+1}}(U) \models x \,\varepsilon_{M_0,\delta}\, a \\ &\text{where } U_{b_0} = X, U_{b_i} = \widehat{b_i} \text{ for } i = 1, \ldots, n \text{ and } (U)_x = \emptyset \text{ otherwise}\} \\ = \{x : x \,\varepsilon_{M_0,\delta}\, a \text{ w.r.t. } \widetilde{\ } : M_0 \to \mathrm{Pow}(S) \\ &\text{where } \widetilde{b_0} = X, \widetilde{b_i} = \widehat{b_i} \text{ for } i = 1, \ldots, n\}\end{aligned}$$

Obviously, Γ is a $\Pi^1_{1,\delta}$-operator in the parameters $\widehat{b_1}, \ldots, \widehat{b_n}$.

Claim 1. *If $\widehat{b} \in \mathrm{Pow}(S/F)$ and $b \notin \{b_1, \ldots, b_n\}$, and if $\sigma \in \mathrm{Aut}(B/F)$, then*

$$\mathfrak{S}_{M,\delta} \models x \,\varepsilon\, a[b_0 := b] \Leftrightarrow \sigma(x) \,\varepsilon\, a[b_0 := b].$$

Proof of Claim 1. Let $M_0' := \{b, b_1, \ldots, b_n\}$. Then $b, b_0 \notin \mathrm{supp}_B(f)$ and therefore $fb \simeq a[b_0 := b]$, consequently $\mathrm{supp}_B(a[b_0 := b]) \cap M \subseteq M_0'$, which by Lemma 4.12 gives $a[b_0 := b] \in \mathrm{Cl}_{M_0',\delta}$.

For $\sigma \in \mathrm{Aut}(B/F)$ and $x \in S$ choose some $b' \in M$ such that $\widehat{b} = \widehat{b'}$ but $b' \notin \mathrm{supp}_B(x) \cup \mathrm{supp}_B(\sigma(x)) \cup \{b_1, \ldots, b_n\}$. Use this to define $\sigma' \in \mathrm{Aut}(B/\{b'\} \cup F)$ which agrees with σ on $\mathrm{supp}_B(x)$. Then we easily compute:

$$\begin{aligned}\mathfrak{S}_{M,\delta} \models x \,\varepsilon\, a[b_0 := b] &\Leftrightarrow \mathfrak{S}_{M,\delta} \models x \,\varepsilon\, a[b_0 := b'] && \text{because } fb' \simeq a[b_0 := b'] \\ & && \text{and } f \text{ is ext. on } M. \\ &\Leftrightarrow \mathfrak{S}_{M,\delta} \models \sigma'(x) \,\varepsilon\, a[b_0 := b'] && \text{because of Lemma 4.9} \\ &\Leftrightarrow \mathfrak{S}_{M,\delta} \models \sigma(x) \,\varepsilon\, a[b_0 := b] && \text{because of ext. of } f \text{ on } M\end{aligned}$$

Claim 2. *For* $\widehat{b} \in \mathrm{Pow}(S/F)$ *it is*

$$\mathfrak{S}_{M,\delta} \models x \mathrel{\varepsilon} f b \text{ if and only if } x \in \Gamma(\widehat{b}).$$

Proof of Claim 2. First we consider the case that $b \notin \{b_1,\dots,b_n\}$. We have to show

$$\mathfrak{S}_{M,\delta} \models x \mathrel{\varepsilon} a[b_0 := b](= f b) \Leftrightarrow \widetilde{\mathfrak{S}}_{M_0,\delta} \models x \mathrel{\varepsilon} a$$

where the latter model is based on $\widetilde{\ } : M_0 \to \mathrm{Pow}(S/F)$ with $\widetilde{b_0} = \widehat{b}$ and $\widetilde{b_i} = \widehat{b_i}$ for $i = 1,\dots,n$. To this end, choose $\sigma \in \mathrm{Aut}(B/F)$ with $\sigma(b) = b_0$ and let $M_0' = \{b, b_1, \dots, b_n\}$ and $M_1 = \{b, b_0, b_1, \dots, b_n\}$. We can extend $\widehat{\ }$ from M_0' to a mapping $\widetilde{\ } : M_1 \to \mathrm{Pow}(S/F)$ by additionally defining $\widetilde{b_0} := \widehat{b}$. Since $a[b_0 := b]$ in $\mathrm{Cl}_{M_0',\delta}$ by Lemma 4.9 we have

$$\mathfrak{S}_{M,\delta} \models x \mathrel{\varepsilon} a[b_0 := b] \Leftrightarrow \mathfrak{S}_{M_0',\delta} \models x \mathrel{\varepsilon} a[b_0 := b]$$

$$\Leftrightarrow \mathfrak{S}_{M_0',\delta} \models \sigma^{-1}(x) \mathrel{\varepsilon} a[b_0 := b] \quad \text{by Claim 1.}$$

$$\Leftrightarrow \widetilde{\mathfrak{S}}_{M_0,\delta} \models x \mathrel{\varepsilon} \sigma(a[b_0 := b]) = a,$$

where in the final equivalence we used Lemma 4.9 for $M_0' \subseteq M_1$ and the mapping $\widetilde{\ } : M_1 \to \mathrm{Pow}(S/F)$.

This finishes the case that $b \notin \{b_1,\dots,b_n\}$. If on the other hand $b \in \{b_1,\dots,b_n\}$ holds, then choose $b' \notin \{b_1,\dots,b_n\}$ such that $\widehat{b'} = \widehat{b}$. By extensionality of f on M, we conclude

$$\mathfrak{S}_{M,\delta} \models x \mathrel{\varepsilon} f b \Leftrightarrow x \mathrel{\varepsilon} f b'.$$

The claim now follows from the first case.

Claim 3. *If* $\widehat{b} \in \mathrm{Pow}(S/F)$, *then* $\Gamma(\widehat{b}) \in \mathrm{Pow}(S/F)$.

Proof of Claim 3. Assume $\widehat{b} \in \mathrm{Pow}(S/F)$, $\sigma \in \mathrm{Aut}(B/F)$, and choose $b' \notin \{b_1,\dots,b_n\}$ such that $\widehat{b} = \widehat{b'}$. Then we have

$$x \in \Gamma(\widehat{b}) = \Gamma(\widehat{b'}) \Leftrightarrow \mathfrak{S}_{M,\delta} \models x \mathrel{\varepsilon} f b'$$

$$\Leftrightarrow \mathfrak{S}_{M,\delta} \models x \mathrel{\varepsilon} a[b_0 := b']$$

$$\Leftrightarrow \mathfrak{S}_{M,\delta} \models \sigma(x) \mathrel{\varepsilon} a[b_0 := b']$$

$$\Leftrightarrow \mathfrak{S}_{M,\delta} \models \sigma(x) \mathrel{\varepsilon} f b'$$

$$\Leftrightarrow \sigma(x) \in \Gamma(\widehat{b'}) = \Gamma(\widehat{b}). \qquad \square$$

Lemma 4.17. *Let $F \subseteq B$ be finite and Γ a $\Pi^1_{1,\delta}$-operator as in the preceding lemma such that for all $b \in M$ with $\widehat{b} \in \mathrm{Pow}(S/F)$*

$$x \in \Gamma(\widehat{b}) \quad \textit{and} \quad \mathfrak{S}_{M,\delta} \models x \,\varepsilon\, fb \Leftrightarrow x \in \Gamma(\widehat{b}).$$

Define the operator $\Gamma' : \mathrm{Pow}(\omega) \to \mathrm{Pow}(\omega)$ by

$$\Gamma'(X) = \Gamma\left(\bigcup\{\mathrm{tr}_{B/F}(x) : x \in X\}\right).$$

(a) *Γ' is a $\Pi^1_{1,\delta}$-operator.*

(b) *If f is monotone, then Γ is monotone on $\mathscr{M}_F = L_\gamma \cap \mathrm{Pow}(S/F)$ and Γ' is monotone on L_γ.*

(c) *Let f be monotone on M. Let $X' \subseteq \omega$ be minimal in L_γ such that $\Gamma'(X') \subseteq X'$, which exists by Corollary 3.9. Then $X = \bigcup\{\mathrm{tr}_{B/F}(x) : x \in X'\} \in \mathrm{Aut}(B/F)$, therefore there is some $b \in M_F$ such that $x = \widehat{b}$. For this b*

$$\mathfrak{S}_M \models fb \subseteq b.$$

Moreover, for all $a \in \mathfrak{S}_M$ we can conclude

$$\mathfrak{S}_M \models fa \subseteq a \to b \subseteq a.$$

Proof. (a) Follows from Lemma 4.4(a).

(b) The monotonicity of Γ follows from that of f using the equivalence characterizing Γ. From this, the monotonicity of Γ' is obvious since $\bigcup\{\mathrm{tr}_{B/F}(x) : x \in X\} \in \mathrm{Pow}(S/F)$ by Lemma 4.4(b).

(c) Since $\mathrm{tr}_{B/F}(x) \subseteq X'$ is arithmetical in x, X' by Lemma 4.4(a), the set X is in L_γ if X' is and moreover it is in $\mathrm{Pow}(S/F)$.

Note that

$$\Gamma(X) = \Gamma'(X') \subseteq X' \subseteq X$$

from which $\mathfrak{S}_M \models fb \subseteq b$ follows since $b \in M_F$.

Now assume $\mathfrak{S}_M \models fa \subseteq a$. By Lemma 4.15 we have

$$\mathfrak{S}_M \models x \,\varepsilon\, a \Leftrightarrow x \in Y$$

for some $Y \in \mathscr{M} = \{X \in L_\gamma : X \in \mathrm{Pow}(S/F) \text{ for some finite } F \subseteq B\}$. Then the set $Y' = \{x \in S : \mathrm{tr}_{B/F}(x) \subseteq Y\}$ is in $L_\gamma \cap \mathrm{Pow}(S/F)$ by Lemma 4.4. So there is some $b' \in M_F$ such that $\widehat{b'} = Y'$. Obviously also $\mathfrak{S}_M \models b' \subseteq a$ and so by monotonicity of f we have $\mathfrak{S}_M \models fb' \subseteq fa \subseteq a$. Since $b' \in M_F$ this means $\Gamma(Y') \subseteq Y$ and $\Gamma(Y') \in \mathscr{M}_F$.

Therefore for all $x \in \Gamma(Y')$ and $\sigma \in \mathrm{Aut}(B/F)$ we have $\sigma(x) \in \Gamma(Y') \subseteq Y$, which means $\mathrm{tr}_B(x) \subseteq Y$, leading to $\Gamma(Y') \subseteq Y'$. Since $Y' \in \mathscr{M}_F$, we moreover have $\Gamma'(Y') = \Gamma(Y') \subseteq Y'$. The minimality of X' yields $X' \subseteq Y'$ and thus $X \subseteq X' \subseteq Y' \subseteq Y$. But this means $\mathfrak{S}_M \models b \subseteq a$. $\square$

5. Proof-theoretic reduction to systems of set theory

5.1. A Tait-style calculus for explicit mathematics

The Tait-style calculus to be developed in this subsection relies on a slightly different account of the language of explicit mathematics. Namely, the Tait language $\mathscr{L}_T$ only contains the logical symbols $\wedge$, $\vee$, $\forall$, $\exists$, but has the relation symbols N, $\sim N$, $=$, $\neq$, App, $\sim$App, ε, $\not\varepsilon$. Negation in this language is defined in the obvious way using the de Morgan laws to push it down to the prime formulas.

Definition 5.1. The Σ^{EM}-formulas form the least class of formulas containing the quantifier-free formulas which is closed under $\wedge$, $\vee$, object quantification, and $\exists$-quantification over classifications.

The Π^{EM}-formulas form the least class of formulas containing the quantifier-free formulas which is closed under $\wedge$, $\vee$, object quantification, and $\forall$-quantification over classifications.

Δ^{EM}-formulas of $\mathscr{L}_T$ are formulas which are both Σ^{EM}- and Π^{EM}-formulas, i.e. which do not contain any unbounded classification quantifiers.

Σ_1^{EM}-formulas are formulas of the form $\exists X_1 \ldots \exists X_k F(X_1,\ldots,X_k)$ where F is a Δ^{EM}-formula. Similarly for Π_1^{EM}-formulas.

The idea now is to embed theories from explicit mathematics into the Tait-calculus and then to perform a partial cut-elimination which only leaves us with cuts on Σ_1^{EM}- (and Π_1^{EM}-) formulas. For this to work we have to use some minor adjustments. First, we need an adequate definition of the rank of a formula.

Definition 5.2. The rank of an $\mathscr{L}_T$-formula is the rank over its Σ_1^{EM}- and Π_1^{EM}-subformulas. Formally:

(a) If F is a Σ_1^{EM}- or Π_1^{EM}-formula, then $\mathrm{rk}(F) = 0$.
(b) Otherwise, if F is $F_0 \wedge F_1$ or $F_0 \vee F_1$, then $\mathrm{rk}(F) = \max\{\mathrm{rk}(F_0), \mathrm{rk}(F_1)\} + 1$.
(c) Otherwise, if F is $\exists x G(x), \forall x G(x), \exists X G(X), \forall X G(X)$, then $\mathrm{rk}(F) = \mathrm{rk}(G) + 1$.

The second adjustment is to make sure that all formulas introduced by non-logical axioms and rules are Σ_1^{EM}. For this it is necessary to switch to a slightly different formulation of the join axiom which has a syntactically simpler form.

Lemma 5.3. *The applicative fragment of* $\mathrm{EM}_0 \restriction$ *proves that under the hypothesis* $\forall x \,\varepsilon\, A \exists X (fx \simeq X)$ *the following assertions are equivalent*:

(a) $\exists Z\, \mathrm{Join}(f,A,Z)$, i.e. $\exists Z(Z \simeq \mathrm{j}(f,A) \wedge \forall z(z \,\varepsilon\, Z \leftrightarrow \exists x \,\varepsilon\, A \exists y (z \simeq (x,y) \wedge y \,\varepsilon\, fx))$.
(b) $\forall z \exists Z\, \mathrm{Join}'(f,z,A,Z)$ *where*

$$\begin{aligned}\mathrm{Join}'(f,z,A,Z) \equiv \exists Y \exists X (Z \simeq \mathrm{j}(f,a) \wedge \\ (z \,\varepsilon\, Z \rightarrow \mathrm{p}_0 z \,\varepsilon\, A \wedge Y \simeq f(\mathrm{p}_0 z) \wedge \mathrm{p}_1 z \,\varepsilon\, Y) \wedge \\ (\mathrm{p}_0 z \,\varepsilon\, A \wedge (X \simeq f(\mathrm{p}_0 z) \rightarrow \mathrm{p}_1 z \,\varepsilon\, X) \rightarrow z \,\varepsilon\, Z)).\end{aligned}$$

Proof. Argue in the applicative fragment of $\mathrm{EM}_0 \restriction$. If $\forall x \,\varepsilon\, A(\exists X(fx \simeq X))$, then these X are unique. Therefore

$$\begin{aligned}\exists Z\,\mathrm{Join}(f,A,Z) &\Leftrightarrow \forall z \exists Z(Z \simeq \mathrm{j}(f,A) \wedge (z \,\varepsilon\, Z \leftrightarrow \exists x \,\varepsilon\, A \exists y(z \simeq (x,y) \wedge y \,\varepsilon\, fx))\\ &\Leftrightarrow \forall z \exists Z(Z \simeq \mathrm{j}(f,A) \wedge\\ &\quad (z \,\varepsilon\, Z \rightarrow \mathrm{p}_0 z \,\varepsilon\, A \wedge \exists Y(Y \simeq f(\mathrm{p}_0 z) \wedge \mathrm{p}_1 z \,\varepsilon\, Y)) \wedge\\ &\quad (\mathrm{p}_0 z \,\varepsilon\, A \wedge \forall X(X \simeq f(\mathrm{p}_0 z) \rightarrow \mathrm{p}_1 z \,\varepsilon\, X) \rightarrow z \,\varepsilon\, Z))\\ &\Leftrightarrow \forall z \exists Z\,\mathrm{Join}'(f,z,A,Z). \qquad \square\end{aligned}$$

Definition 5.4. The calculus $\mathscr{T}$ is defined as follows:

(a) Logical axioms

(Ax) $\Gamma, \neg F, F$ where $\mathrm{rk}(F) = 0$.

(b) Equality axioms

(Eq1) $\Gamma, t = t$ for object terms t.

(Eq2) $\Gamma, s \neq t, \neg F(s), F(t)$ where $\mathrm{rk}(F) = 0$.

(c) Logical rules

$$(\wedge)\ \frac{\Gamma, F_0 \quad \Gamma, F_1}{\Gamma, F_0 \wedge F_1} \qquad (\vee)\ \frac{\Gamma, F_i}{\Gamma, F_0 \vee F_1} \quad i = 0,1$$

$$(\forall^0)\ \frac{\Gamma, F(a)}{\Gamma, \forall x F(x)}* \qquad (\exists^0)\ \frac{\Gamma, F(t)}{\Gamma, \exists x F(x)}$$

$$(\forall^1)\ \frac{\Gamma, F(A)}{\Gamma, \forall X F(X)}* \qquad (\exists^1)\ \frac{\Gamma, F(A)}{\Gamma, \exists X F(X)}$$

The variables a and A in the $\forall$-inferences may not occur in the conclusion of the inferences.

(d) Non-logical axioms

Γ, F

where F is one of the following:

- an instance of an applicative axiom.
- an instance of (ECA), i.e. $\exists X(X \simeq c_m(\vec{t},\vec{A}) \wedge \forall x(x \,\varepsilon\, X \leftrightarrow F(x,\vec{t},\vec{A})))$ for certain terms $\vec{t}$ and classification variables $\vec{A}$.
- the *induction axiom*

$$0 \,\varepsilon\, A \wedge \forall x \in \mathrm{N}(x \,\varepsilon\, A \rightarrow \mathrm{s_N} x \,\varepsilon\, A) \rightarrow \forall x \in \mathrm{N}(x \,\varepsilon\, A).$$

- the open form of (IG)$\restriction$, which is separated into two axioms,

(IG1) $\Gamma, \exists X(X \simeq \mathrm{i}(A,B) \wedge \mathrm{Prog}_A(B,X))$.

and

$$\text{(IG2)} \restriction \quad \Gamma, \mathrm{i}(A,B) \simeq D \wedge \mathrm{Prog}_A(B,C) \to \forall x \,\varepsilon\, D(x \,\varepsilon\, C).$$

(e) the rule for join

$$\text{(Join)} \quad \frac{\Gamma, \forall x \,\varepsilon\, A \exists X(fx \simeq X)}{\Gamma, \exists Z \mathrm{Join}'(f,t,A,Z)}$$

for terms f and t.

(f) the ω-rule

$$(\omega) \quad \frac{\ldots \Gamma, n \neq t \ldots}{\Gamma, \neg \mathrm{N}(t)}$$

In the following we write $\mathscr{T} \,|\!\frac{\alpha}{k}\, \Gamma$ for the existence of a derivation in $\mathscr{T}$ in which all cut-formulas have rank less than k and which is of length $\leqslant \alpha$. We further assume that for a derivation that uses the ω-rule we always have $\alpha \geqslant \omega$.

The definition of the calculus $\mathscr{T}$ is tailored so that the following proposition holds:

Proposition 5.5. (a) *If* $\mathrm{EM}_0 \restriction +(\mathrm{IG}) \restriction +(\mathrm{Join}) \vdash F$, *then there are* $n,k < \omega$ *such that* $\mathscr{T} \,|\!\frac{n}{k}\, F$.

(b) *If* $\mathrm{EM}_0 + (\mathrm{IG}) \restriction +(\mathrm{Join}) \vdash F$, *then* $\mathscr{T} \,|\!\frac{\alpha}{k}\, F$ *for some* $\alpha < \omega \cdot 2$ *and* $k \in \mathbb{N}$.

Proof. The only noteworthy point is that in part (b) the usual ω-rule

$$\frac{\ldots \Gamma, F(n) \ldots}{\Gamma, \forall x \in NF(x)}$$

is derivable. Indeed, using cuts with $\Gamma, n \neq a, \neg F(n), F(a)$ (derivable from the equality axioms), the premises of the rule give $\Gamma, n \neq a, F(a)$ for a new a, from which we get $\Gamma, \neg \mathrm{N}(a), F(a)$ by the ω-rule which in turn leads to the conclusion using $(\vee)$ and $\forall^0$-inferences. □

Since all non-logical axioms and rules only introduce formulas of rank 0, we can eliminate all cuts of higher complexity from our derivations. In other words:

Proposition 5.6. *If* $\mathscr{T} \,|\!\frac{\alpha}{k}\, \Gamma$, *then there is some* β *such that* $\mathscr{T} \,|\!\frac{\beta}{1}\, \Gamma$.

More exactly, it is $\beta \leqslant 2_{k-1}(\alpha)$ *where* $2_0(\alpha) = \alpha$ *and* $2_{n+1}(\alpha) = 2^{2_n(\alpha)}$.

Proof. Standard cut-elimination. □

Putting the previous propositions together, we obtain:

Proposition 5.7. (a) *If* $\mathrm{EM}_0 \restriction +(\mathrm{IG}) \restriction +(\mathrm{Join}) \vdash F$, *then there is some* $n < \omega$ *such that* $\mathscr{T} \,|\!\frac{n}{1}\, F$.

(b) *If* $\mathrm{EM}_0 + (\mathrm{IG}) \restriction +(\mathrm{Join}) \vdash F$, *then* $\mathscr{T} \,|\!\frac{\alpha}{1}\, F$ *for some* $\alpha < \varepsilon_0$.

To treat (MID) in this context we again (as in the case of (Join)) have to use a slight variant of the axiom which is in a syntactic form that can be dealt with in an easier way in the following.

Lemma 5.8. *The applicative fragment of* $\mathrm{EM}_0 \restriction$ *proves: If* $\mathrm{Clop}(f)$, *then the following formulations of the least fixed-point axiom are equivalent.*

(a) $\mathrm{Lfp}(f,A)$.

(b) $\mathrm{Lfp}'(f,A) \equiv \forall X \forall Y \forall Z (Y \simeq fA \wedge Z \simeq fX \to Y \subseteq A \wedge (Z \subseteq X \to A \subseteq X))$.

Therefore, the axiom (MID) *is equivalent to*

$$(\mathrm{Mid}) \qquad \forall f(\mathrm{Clop}(f) \wedge \mathrm{Mon}(f) \to \exists X \mathrm{Lfp}'(f,X)).$$

Proof. Similar to Lemma 5.3. □

Remark 5.9. The above propositions can be proved in KPl^r (actually in much weaker theories). We will use this fact later on, which is especially important in the case of $\mathrm{EM}_0\restriction + (\mathrm{IG})\restriction + (\mathrm{Join}) + (\mathrm{MID})$.

5.2. Asymmetric interpretations

In this subsection we actually reduce theories for explicit mathematics containing (MID) to systems of set theory which axiomatize the existence of a stable ordinal. To this end we will use asymmetric interpretations of the quasi cut-free derivations of the previous subsection into the model of T_0 as defined in subsection 4.3.

In the following, we work in theories which assume

$$L_\gamma \prec_1 L \wedge \gamma \text{ is countable}$$

in addition to KPi^r (resp. KPi^w) when treating $\mathrm{EM}_0\restriction + (\mathrm{Join}) + (\mathrm{IG})\restriction + (\mathrm{MID})$ (resp. $\mathrm{EM}_0 + (\mathrm{Join}) + (\mathrm{IG})\restriction + (\mathrm{MID})$).

Let the set M be defined as in subsection 4.1. Using this, we define the models $\mathfrak{S}_M = \bigcup_{\alpha<\gamma} \mathfrak{S}_{M,\alpha}$ as in subsection 4.3 based on the mapping $\widehat{\ }: M \to \mathrm{Pow}(S)$ given there.

The importance of Σ^{EM}- and Π^{EM}-formulas in our context rests on the fact that they satisfy persistency properties in these models for T_0 in the following sense.

Definition 5.10. (a) A formula $F[\vec{a},\vec{A}]$ is called *upwards persistent* (w.r.t. the model $\mathfrak{S}_M = \bigcup_\alpha \mathfrak{S}_{M,\alpha}$) if for all $\alpha \leqslant \beta$

$$\forall \vec{a} \in \mathrm{Cl}_{M,\alpha} \forall \vec{x} \in S(\mathfrak{S}_{M,\alpha} \models F[\vec{x},\vec{a}] \Rightarrow \mathfrak{S}_{M,\beta} \models F[\vec{x},\vec{a}]).$$

(b) A formula $F[\vec{a},\vec{A}]$ is called *downwards persistent* if in the above situation the converse implication holds.

(c) A formula is called *absolute* if it is both upwards and downwards persistent.

Proposition 5.11 (Persistency). *Σ^{EM}-formulas are upwards persistent, Π^{EM}-formulas are downwards persistent, and Δ^{EM}-formulas are absolute.*

Proof. Straightforward induction on the definition of Σ^{EM}- and Π^{EM}-formulas. □

In the following proposition, we use the convention to use X, Y, Z as notations for elements of $\mathrm{Cl}_{M,\alpha}$ (instead of a, b as before) in order to avoid confusion with free object variables.

Proposition 5.12 (Asymmetric interpretation). (a) *For each (meta) n and m the theory* $\mathrm{KPi}^r + \exists\gamma(L_\gamma \prec_1 L \wedge$ "γ *is countable*") *proves*:
If $\mathcal{T} \,|\!\frac{n}{1}\, \neg(\mathrm{Mid}), \Gamma[\vec{a},\vec{A}]$, *$\Gamma$ a set of Σ^{EM}-formulas, then for all (meta) m*

$$\forall \vec{X} \in \mathrm{Cl}_{M,m} \forall \vec{x} \in S(\mathfrak{S}_{M,m+2^n} \models \Gamma[\vec{x},\vec{X}]).$$

(b) *For each (meta) $\delta < \varepsilon_0$ the theory* $\mathrm{KPi}^w + \exists\gamma(L_\gamma \prec_1 L \wedge$ "γ *is countable*") *proves*: *If* $\mathcal{T} \,|\!\frac{\alpha}{1}\, \neg(\mathrm{Mid}), \Gamma[\vec{a},\vec{A}]$ *for some $\alpha < \omega^\delta$ and a set Γ of Σ^{EM}-formulas, then*

$$\forall \beta < \omega^\delta \forall \vec{X} \in \mathrm{Cl}_{M,\beta} \forall \vec{x} \in S(\mathfrak{S}_{M,\beta+2^\alpha} \models \Gamma[\vec{x},\vec{X}]).$$

Proof. We prove part (a) by induction on n. We restrict our attention to the most important cases, as the remaining ones easily follow using the i.h.

If Γ is an axiom, then there are two subcases.

In the first one, Γ is a Δ^{EM}-formula (in the cases of (Ax), (Eq), applicative axioms, induction axiom and (IG2)$\restriction$). Then the assertion holds by construction of $\mathfrak{S}_{M,\gamma}$. In the case of the induction axiom we have to note that for each $X \in \mathrm{Cl}_{M,m}$ the set $\{x \in S : \mathfrak{S}_{M,m+2^n} \models x \,\varepsilon\, X\}$ is in L_γ and therefore we can use induction in L_γ (on the set $\{n^\circ : n \in \omega\}$) to prove the instance of the induction axiom.

In the second axiom case we have an instance of (ECA) or one of (IG1) in its open formulation. For example, let us treat (IG1). For arbitrary m and $X_0, X_1 \in \mathrm{Cl}_{M,m}$ we have $\mathsf{i}(X_0,X_1) \in \mathrm{Cl}_{M,m+1} \subseteq \mathrm{Cl}_{M,m+2^n}$ and so the assertion is established.

We leave out the propositional, quantifier and equality rules, since they can be treated using the i.h. But note that it is important that there are no $(\forall^1)$-rules because of the fact that Γ consists of Σ^{EM}-formulas.

Now assume the last inference was a cut with formulas of rank 0. Then we have the premises

$$\mathcal{T} \,|\!\frac{n_0}{1}\, \Gamma[\vec{a},\vec{A}], \exists \vec{Y} F[\vec{a},\vec{b},\vec{Y},\vec{A},\vec{B}]$$

and

$$\mathcal{T} \,|\!\frac{n_1}{1}\, \Gamma[\vec{a},\vec{A}], \forall \vec{Y} \neg F[\vec{a},\vec{b},\vec{Y},\vec{A},\vec{B}]$$

where F is a Δ^{EM}-formula and $n_0, n_1 < n$. Application of the induction hypothesis to the first premise yields

$$\forall \vec{X} \in \mathrm{Cl}_{M,m}(\mathfrak{S}_{M,m+2^{n_0}} \models \Gamma[\vec{x},\vec{X}], \exists \vec{Y} F[\vec{x},\vec{0},\vec{Y},\vec{X},\vec{Z}])$$

for all m, $\vec{x} \in S$ and $\vec{X}, \vec{Z} \in \mathrm{Cl}_{M,m}$. Using inversion on the second premise we get

$$\mathcal{T} \mathrel{\vert\!\frac{n_1}{1}} \Gamma[\vec{a},\vec{A}], \neg F[\vec{a},\vec{b},\vec{C},\vec{A},\vec{B}]$$

for new classification variables $\vec{C}$. Applying the i.h. to this derivation we get

$$\forall \vec{X} \in \mathrm{Cl}_{M,m'}(\mathfrak{S}_{M,m'+2^{n_1}} \models \Gamma[\vec{x},\vec{X}], \forall \vec{Y} \neg F[\vec{x},\vec{0},\vec{Y},\vec{X},\vec{Z}])$$

for all m' and appropriate $\vec{x}, \vec{Y}, \vec{X}, \vec{Z}$.

Now assume that there are $\vec{x} \in S$, $\vec{X} \in \mathrm{Cl}_{M,m}$ such that $\mathfrak{S}_{M,m+2^n} \not\models \Gamma[\vec{x},\vec{X}]$. Using persistency, the above conclusion from the i.h. for the first premise supplies us with $\vec{Y} \in \mathrm{Cl}_{M,m+2^{n_0}}$ such that

$$\mathfrak{S}_{M,m+2^{n_0}} \models F[\vec{x},\vec{0},\vec{Y},\vec{X},\vec{Z}].$$

Using the conclusion from the i.h. for $m' = m + 2^{n_0}$ we get

$$\mathfrak{S}_{M,m'+2^{n_1}} \models \Gamma[\vec{x},\vec{X}], \neg F[\vec{x},\vec{0},\vec{Y},\vec{X},\vec{Z}],$$

so using the choice of $\vec{Y}$ this means

$$\mathfrak{S}_{M,m'+2^{n_1}} \models \Gamma[\vec{x},\vec{X}]$$

which by persistency contradicts the assumption $\mathfrak{S}_{M,m+2^n} \not\models \Gamma[\vec{x},\vec{X}]$, so that this must be false and the assertion is shown in this case.

If the last inference is (Join), the formula $\exists Z \mathrm{Join}'(f, t[\vec{a},\vec{A}], A, Z)$ is in Γ and the premise of the inference is

$$\mathcal{T} \mathrel{\vert\!\frac{n_0}{1}} \Gamma[\vec{a},\vec{A}], \forall x \,\varepsilon\, A \exists X (fx \simeq X).$$

Fix $\vec{X} \in \mathrm{Cl}_{M,m}$ and $\vec{x} \in X$ and identify $f = f(\vec{x},\vec{X})$. Assume $\mathfrak{S}_{M,m+2^n} \not\models \Gamma[\vec{x},\vec{X}]$. The i.h. gives, using persistency again, $\mathfrak{S}_{M,m+2^{n_0}} \models \forall x \,\varepsilon\, X \exists Y (fx \simeq Y)$ and therefore $\mathrm{j}(f,X) \in \mathrm{Cl}_{M,m+2^{n_0}+1} \subseteq \mathrm{Cl}_{M,m+2^n}$. Consequently, $\mathfrak{S}_{M,m+2^n} \models \exists Z \mathrm{Join}(f,X,Z)$ and therefore $\mathfrak{S}_{M,m+2^n} \models \forall z \exists Z \mathrm{Join}'(f,z,X,Z)$, a contradiction establishing the assertion also in this case.

Assume, and this is the central case, that the last inference was an $(\exists^0)$-inference with main formula $\neg(\mathrm{Mid})$. Then we have the premise

$$\mathcal{T} \mathrel{\vert\!\frac{n_0}{1}} \neg(\mathrm{Mid}), \Gamma[\vec{a},\vec{A}], \mathrm{Clop}(t) \wedge \mathrm{Mon}(t) \wedge \forall X \neg \mathrm{Lfp}'(t,X)$$

for $n_0 < n$ and an object term t which w.l.o.g. has no free variables not in $\vec{a},\vec{A}$. Using inversions, we get the following derivations

(1) $\mathcal{T} \mathrel{\vert\!\frac{n_0}{1}} \neg(\mathrm{Mid}), \Gamma[\vec{a},\vec{A}], A \subseteq B \to tA \subseteq tB$

(2) $\mathcal{T} \mathrel{\vert\!\frac{n_0}{1}} \neg(\mathrm{Mid}), \Gamma[\vec{a},\vec{A}], \exists Y (tA \simeq Y)$

(3) $\mathcal{T} \mathrel{\vert\!\frac{n_0}{1}} \neg(\mathrm{Mid}), \Gamma[\vec{a},\vec{A}], \neg \mathrm{Lfp}'(t,A)$

where A, B are new free variables.

Now assume $m \in \mathbb{N}$, $\vec{X} \in \mathrm{Cl}_{M,m}$ and $\vec{x} \in S$. Define $f := t(\vec{x},\vec{X}) \in S$ and $k := m+2^{n_0}$. Then $k + 2^{n_0} \leqslant m + 2^n$ and so using persistency we see that if $\mathfrak{S}_{M,k+2^{n_0}} \models \Gamma[\vec{x},\vec{X}]$, then also $\mathfrak{S}_{M,m+2^n} \models \Gamma[\vec{x},\vec{X}]$ and we are done. Otherwise the i.h. for (1) and (2) leads to

(4) $\quad \forall X, Y \in \mathrm{Cl}_{M,k} \mathfrak{S}_{M,k+2^{n_0}} \models X \subseteq Y \to fX \subseteq fY,$

(5) $\quad \forall X \in \mathrm{Cl}_{M,k} \mathfrak{S}_{M,k+2^{n_0}} \models \exists Y(fX \simeq Y).$

Since $M = \mathrm{Cl}_{M,0} \subseteq \mathrm{Cl}_{M,k}$, (4) and (5) imply especially that f satisfies the hypotheses of Lemma 4.16 for $\delta = k + 2^{n_0}$. Therefore by this lemma there is a $\Pi^1_{1,k+2^{n_0}}$-operator $\Gamma : \mathrm{Pow}(\omega) \to \mathrm{Pow}(\omega)$ with parameters in L_γ and some finite F such that for all $b \in M$ with $\widehat{b} \in \mathrm{Pow}(S/F)$

(6) $\quad \Gamma(\widehat{b}) \in \mathrm{Pow}(S/F)$ and $\mathfrak{S}_{M,k+2^{n_0}} \models x \,\varepsilon\, fb \Leftrightarrow x \in \Gamma(\widehat{b}).$

From this operator Γ we define again the variant Γ' by

$$\Gamma'(X) = \Gamma\left(\bigcup\{\mathrm{tr}_{B/F}(x) : x \in X\}\right).$$

By Corollary 3.9 and Lemma 4.17 Γ' has a sub-fixpoint Y' in L_γ. For

$$Y = \{\mathrm{tr}_{B/F}(x) : x \in Y'\},$$

this lemma moreover yields

$$\mathfrak{S}_{M,k+2^{n_0}} \models \forall z(z \,\varepsilon\, fb \to z \,\varepsilon\, b)$$

and

(7) $\quad \forall X \in \mathrm{Cl}_{M,k} \mathfrak{S}_{M,k+2^{n_0}} \models fX \subseteq X \to b \subseteq X$

for some $b \in M_F$ such that $\widehat{b} = Y$. On the other hand it follows that

(8) $\quad \mathfrak{S}_{M,m+2^{n_0}} \models fb \subseteq b \to \exists X(fX \subseteq X \wedge b \not\subseteq X)$

when choosing $b \in M = \mathrm{Cl}_{M,0}$ for A in the i.h. for (3). Fixing X as in (8) contradicts (7) since because of $k = m + 2^{n_0}$ we also have $X \in \mathrm{Cl}_{M,k}$.

(b) can be proved analogously using transfinite induction up to δ. □

Corollary 5.13. (a) *If* $\mathrm{EM}_0{\restriction} + (\mathrm{Join}) + (\mathrm{IG}){\restriction} + (\mathrm{MID}) \vdash F$ *for a* Σ^{EM}*-formula* F*, then* $\mathfrak{S}_{M,\omega} \models F$.

(b) *If* $\mathrm{EM}_0 + (\mathrm{Join}) + (\mathrm{IG}){\restriction} + (\mathrm{MID}) \vdash F$ *for a* Σ^{EM}*-formula* F*, then* $\mathfrak{S}_{M,\varepsilon_0} \models F$.

6. Reductions of subsystems of KPi in the presence of stability axioms

Now we are going to prepare the second part of the reductions, namely reducing subsystems of set theory to systems of explicit mathematics. This will take some intermediate steps.

First, namely in the present section, we prove several well-known results on subsystems of KPi which will be used in the second part of the section for reductions between different theories involving stability. First we want to reduce KPi^r to KPl^r in an analogous way as Σ^1_2-AC_0 can be reduced to Π^1_1-CA_0.

Definition 6.1. The calculus $\mathrm{T}(\mathrm{KPi}^r)$ is defined as a Tait-style calculus for set theory (with equality rules) together with the rules

$$(Ax_{\mathrm{KPl}^r}) \qquad \frac{\Gamma, \neg F}{\Gamma} \qquad \text{for axioms } F \text{ of } \mathrm{KPl}^r$$

$$(\Delta_0\text{-Coll}) \qquad \frac{\Gamma, \forall x \in a \exists y F(x,y)}{\Gamma, \exists z \forall x \in a \exists y \in z F(x,y)} \qquad \text{for } F \in \Delta_0.$$

By the usual proof-theoretic arguments we obtain:

Proposition 6.2. (a) *If* $\mathrm{KPi}^r \vdash F$, *then there are* $l, r \in \mathbb{N}$ *such that* $\mathrm{T}(\mathrm{KPi}^r) \overset{l}{\underset{r}{\vdash}} F$, *where* l *is an upper bound for the length of the derivation and* r *is the cut-rank of the derivation. Here the rank of* Σ_1- *and* Π_1-*formulas is defined to be* 0 *and for other formulas it is defined from this using the usual clauses.*

(b) *If* $\mathrm{T}(\mathrm{KPi}^r) \overset{l}{\underset{r}{\vdash}} F$ *for some* $l, r \in \mathbb{N}$, *then there is a* k, $k = 2_{r-1}(l)$, *such that* $\mathrm{T}(\mathrm{KPi}^r) \overset{k}{\underset{1}{\vdash}} F$.

Definition 6.3. For a formula F let $F^{x,y}$ the formula arising from F by relativizing all unbounded universal quantifiers to x and all unbounded existential quantifiers to y (after appropriate renaming). For a finite set Γ of formulas, $\Gamma^{x,y}$ is the set of all $F^{x,y}$ where $F \in \Gamma$.

When arguing in theories which allow the definition of the constructible hierarchy, we will write $F^{\alpha,\beta}$ instead of F^{L_α, L_β}.

Proposition 6.4. *If* $\mathrm{T}(\mathrm{KPi}^r) \overset{k}{\underset{1}{\vdash}} \Gamma[\boldsymbol{a}]$ *and* Γ *only contains* $\forall\Sigma$-*formulas, then for all* $l \in \mathbb{N}$

$$\mathrm{KPl}^r \vdash \forall \alpha \forall \boldsymbol{x} \in L_{\alpha^{+(l)}} \bigvee \Gamma^{\alpha^{+(l)}, \alpha^{+(l+2^k)}}[\boldsymbol{x}].$$

Proof. Induction on k. □

Corollary 6.5. *If* $\mathrm{KPi}^r \vdash F$ *for a* Σ-*sentence* F, *then also* $\mathrm{KPl}^r \vdash F$.

Now we are going to extend these arguments to KPi^w, so that we have to treat the scheme of induction on $\mathbb{N}$. The aim is to obtain a similar reduction as that of Σ^1_2-AC to $(\Pi^1_1\text{-CA})_{<\varepsilon_0}$.

Definition 6.6. The calculus $\mathrm{T}_\omega(\mathrm{KPi}^r)$ is defined analogously to $\mathrm{T}(\mathrm{KPi}^r)$, but contains additionally the ω-rule

$$\frac{\Gamma, t \neq \underline{n} \text{ for all } n \in \mathbb{N}}{\Gamma, t \notin \omega.}$$

To formulate this precisely, the calculus derives formulas in a language extended by constants $\underline{n}$ for $n \in \mathbb{N}$ and a constant ω. It also contains a rule

$$\mathrm{T}_\omega(\mathrm{KPi}^r) \vdash \Gamma, \neg F \Rightarrow \mathrm{T}_\omega(\mathrm{KPi}^r) \vdash \Gamma$$

for $F \equiv \forall x(x \notin \underline{0})$, $F \equiv \forall x(x \in \underline{n+1} \leftrightarrow x \in \underline{n} \vee x = \underline{n})$ and the defining axiom for ω, namely $\omega \in \mathrm{On} \wedge \omega \in \mathrm{Lim} \wedge \forall x \in \omega(x \notin \mathrm{Lim})$.

The above definition can be formalized in KPl^r and therefore the following proposition can be proved (actually a much weaker theory than KPl^r would suffice).

Proposition 6.7. (a) *If* $\mathrm{KPi}^w \vdash F$, *then* KPl^r *proves that* $\mathrm{T}_\omega(\mathrm{KPi}^r) \mathrel{|\!\!\frac{\omega+\omega}{n}} F$.

(b) *For all (meta)* $\delta < \varepsilon_0$ *and* $n \in \mathbb{N}$, KPl^r *proves that if* $\mathrm{T}_\omega(\mathrm{KPi}^r) \mathrel{|\!\!\frac{\alpha}{n}} \Gamma$ *for some* $\alpha \leqslant \delta$, *then* $\mathrm{T}_\omega(\mathrm{KPi}^r) \mathrel{|\!\!\frac{2_n(\alpha)}{1}} \Gamma$ *where* $2_1(\alpha) = \alpha, 2_{n+1}(\alpha) = 2^{2_n(\alpha)}$.

Finally, we have (now really exploiting the full strength of $\mathrm{KPl}^r_{<\varepsilon_0}$):

Proposition 6.8. *For all (meta)* $\delta < \varepsilon_0$ *the theory* $\mathrm{KPl}^r_{<\varepsilon_0}$ *proves the following: If* $\mathrm{T}_\omega(\mathrm{KPi}^r) \mathrel{|\!\!\frac{\gamma}{1}} \Gamma$ *where* $\gamma < \delta$ *and* Γ *consists of* $\forall\Sigma$*-formulas, then*

$$\forall\alpha\forall\beta < \omega^\delta \forall \vec{x} \in L_{\alpha^{+(\beta)}} \bigvee \Gamma^{\alpha^{+(\beta)}, \alpha^{+(\beta+2^\gamma)}}.$$

Proof. Induction on γ. The proof is straightforward once it is established that this induction can be carried out in our meta-theory. To that end we fix, arguing in $\mathrm{KPl}^r_{<\varepsilon_0}$, an arbitrary ordinal α. By the main axiom of this theory, $\alpha^{+(\omega^\delta)}$ exists. Since the assertion of the theorem concerns validity in $L_{\alpha^{+(\omega^\delta)}}$, it can be described by a Δ_0-formula. Hence the necessary induction principle is available in $\mathrm{KPl}^r_{<\varepsilon_0}$. □

Corollary 6.9. *If* $\mathrm{KPi}^w \vdash F$ *for a sentence* $F \in \Sigma$, *then* $\mathrm{KPl}^r_{<\varepsilon_0} \vdash F$.

Now we apply the results of the asymmetric interpretations obtained in Propositions 6.4 and 6.8 in a context in which stable ordinals are present. The point of the proof is that the additional parameter α, which was not necessary to obtain the proof-theoretic reductions of the systems KPi^r and KPi^w given by Corollaries 6.5 and 6.9, is now instantiated to these stable ordinals.

Proposition 6.10. *If* $\mathrm{KPi}^r + \exists\alpha(L_\alpha \prec_1 L) \vdash F$ *for some* Σ*-sentence* F, *then there is an* $n \in \mathbb{N}$ *such that* $\mathrm{KPl}^r + \exists\alpha(L_\alpha \prec_1 L_{\alpha^{+(n)}}) \vdash F$.

Proof. Let $\mathrm{KPi}^r + \exists\alpha(L_\alpha \prec_1 L) \vdash F$. This means

$$\mathrm{KPi}^r \vdash \exists\alpha\forall x \in L_\alpha \forall e \in \omega(\mathrm{Sat}_\Sigma(e,x)^L \rightarrow \mathrm{Sat}_\Sigma(e,x)^{L_\alpha}) \rightarrow F.$$

Using embedding and cut-elimination in $\mathrm{T}(\mathrm{KPi}^r)$ we obtain

$$\mathrm{T}(\mathrm{KPi}^r) \mathrel{|\!\!\frac{k}{1}} \forall\alpha\exists x \in L_\alpha \exists e \in \omega(\mathrm{Sat}_\Sigma(e,x)^L \wedge \neg\mathrm{Sat}_\Sigma(e,x)^{L_\alpha}), F.$$

If we apply Proposition 6.4 with $l = 1$, we see that KPl^r proves:[5]

$$\forall\alpha\forall\beta \in L_{\alpha^+}\exists x \in L_\beta\exists e \in \omega(\mathrm{Sat}_\Sigma(e,x)^{\alpha^{+(1+2^k)}} \wedge \neg\mathrm{Sat}_\Sigma(e,x)^{L_\beta}) \vee F^{\alpha^{+(1+2^k)}}.$$

For each ordinal α, we can instantiate β to α in the above formula and apply persistency to F, by which we get

$$\mathrm{KPl}^r \vdash \forall\alpha(\forall x \in L_\alpha\forall e \in \omega(\mathrm{Sat}_\Sigma(e,x)^{\alpha^{+(1+2^k)}} \rightarrow \mathrm{Sat}_\Sigma(e,x)^\alpha) \rightarrow F).$$

As α does not occur in F, this amounts to

$$\mathrm{KPl}^r + \exists\alpha(L_\alpha \prec_1 L_{\alpha^{+(1+2^k)}}) \vdash F. \qquad \square$$

Similarly, we obtain:

Proposition 6.11. *If* $\mathrm{KPi}^w + \exists\alpha(L_\alpha \prec_1 L) \vdash F$ *for some Σ-formula F, there is some $\delta < \varepsilon_0$ such that* $\mathrm{KPl}^r_{<\varepsilon_0} + \exists\alpha(L_\alpha \prec_1 L_{\alpha^{+(\delta)}}) \vdash F$.

7. Non-monotonic inductive definitions give rise to stability

The aim of this section is to prove the existence of ordinals γ which are $\gamma^{+(\delta+1)}$-stable using the existence of inductively generated sub-fixpoints of certain non-monotonic $\Pi^1_{1,\delta}$-operators. For this, we have to construct an operator of maximal closure ordinal.

Fix $\delta<\varepsilon_0$. We work in $\mathrm{KPl}^r + (V = L) + \forall\gamma\exists\eta(\eta = \gamma^{+(\delta+1)})$ in this section. This theory is a subtheory of $\mathrm{KPl}^r + (V = L)$ if $\delta < \omega$ and of $\mathrm{KPl}^r_{<\varepsilon_0} + (V = L)$ otherwise.

Definition 7.1. Define $\Lambda := \Lambda_\delta : \mathrm{Pow}(\omega) \rightarrow \mathrm{Pow}(\omega)$ by

$$n \in \Lambda(X) \Leftrightarrow L_{\omega^X_{\delta+1}} \models \mathrm{Sat}_\Sigma(n,\emptyset)$$

where Sat_Σ is the Σ_1-truth predicate.

Remark 7.2. This definition gives rise to Δ-predicates

$$P(X,Y) :\equiv Y = \Lambda(X) \text{ and } Q(\alpha,X) :\equiv X = I^\alpha_\Lambda$$

because in our meta-theory we can prove that for all $X \subseteq \omega$ there is a uniquely determined ordinal $\omega^X_{\delta+1}$.

Further note that Λ is a $\Pi^1_{1,\delta}$-operator since L_α can be defined by a Σ_1-formula in $L_\alpha(X)$ if α is admissible. Therefore the condition $L_{\omega^X_{\delta+1}} \models \mathrm{Sat}_\Sigma(n,\emptyset)$ can be easily written in the form $L_{\omega^X_{\delta+1}}(X) \models F[n]$ for some Σ_1-formula F.

[5] Formally, we would have to give a Π_2-formula in the language of set theory expressing the mentioned property and apply Proposition 6.4 to this formula, but we do not bother to make explicit the necessary computations.

By Proposition 3.8 we know that $\Lambda(I_\Lambda^{<\gamma}) \subseteq I_\Lambda^{<\gamma}$ holds if γ is stable. Since Λ is defined without parameters, the proof of Proposition 3.8 gives that this also holds if γ is only weakly stable.

The other implication, namely that the closure ordinal γ of Λ is $\gamma^{+(\delta+1)}$-stable is somewhat harder, since Λ is parameter-free. Therefore, we need a further characterization of this closure ordinal. This will be taken up next.

Definition 7.3. $b \in L_\alpha$ has a good Σ_1^--definition in L_α if there is some $F[x] \in \Sigma_1$ without further parameters (free variables) such that

$$L_\alpha \models F[b] \quad \text{and} \quad L_\beta \models \exists! x F[x] \quad \text{for all } \beta \geqslant \alpha.$$

Definition 7.4. Let $\alpha = \gamma_{\delta+1}$ iff α is minimal such that α has no good Σ_1^--definition in $L_{\alpha^{+(\delta+1)}}$.

Note that the previous definition does not say that $\gamma_{\delta+1}$ exists. In fact, our metatheory does not allow to prove the existence of $\gamma_{\delta+1}$. We will show that its existence is equivalent to the existence of some γ which is $\gamma^{+(\delta+1)}$-stable, indeed if $\gamma_{\delta+1}$ exists, it satisfies this property.

Proposition 7.5. *Assume $\gamma_{\delta+1}$ exists.*

(a) *$\gamma_{\delta+1}$ is recursively inaccessible.*

(b) *If $\tau < \gamma_{\delta+1}$, then τ has a good Σ_1^--definition in $L_{\gamma_{\delta+1}^{+(\delta+1)}}$.*

(c) *If $\gamma_{\delta+1} \leqslant \sigma < \gamma_{\delta+1}^{+(\delta+1)}$, then σ has no good Σ_1^--definition in $L_{\gamma_{\delta+1}^{+(\delta+1)}}$.*

Proof. (a) Assume $\gamma_{\delta+1}$ was not admissible. Then we had some $\alpha < \gamma_{\delta+1}$ and a Σ_1-formula F (possibly containing parameters from $L_{\gamma_{\delta+1}}$) such that

$$\forall \beta < \alpha \exists \gamma < \gamma_{\delta+1} F(\beta)^{L_\gamma} \wedge \forall \gamma < \gamma_{\delta+1} \exists \beta < \alpha \neg F(\beta)^{L_\gamma}.$$

α has a good Σ_1^--definition in $L_{\alpha^{+(\delta+1)}} \subseteq L_{\gamma_{\delta+1}^{+(\delta+1)}}$ and so have all parameters of F, which we may assume to be ordinals less than $\gamma_{\delta+1}$, therefore the following can be turned into a good Σ_1^--definition of $\gamma_{\delta+1}$ in $L_{\gamma_{\delta+1}^{+(\delta+1)}}$:

$$\text{“}\gamma_{\delta+1} = \min\{\xi : \forall \beta < \alpha \exists \gamma < \xi F(\beta)^{L_\gamma}\}\text{”}.$$

So $\gamma_{\delta+1}$ is admissible. To show that it is recursively inaccessible, assume we had $\gamma_{\delta+1} = \alpha^+$ for some α. Then $\gamma_{\delta+1}$ can be Σ_1^--defined in $L_{\gamma_{\delta+1}^{+(\delta+1)}}$ as the least admissible γ above α, where α is replaced by its good Σ_1^--definition in $L_{\alpha^{+(\delta+1)}} \subseteq L_{\gamma_{\delta+1}^{+(\delta+1)}}$.

(b) Each $\tau < \gamma_{\delta+1}$ has a good Σ_1^--definition in $L_{\tau^{+(\delta+1)}}$. By (a) from $\tau < \gamma_{\delta+1}$ it also follows that $\tau^{+(\delta+1)} \leqslant \gamma_{\delta+1}$.

(c) If F were a good Σ_1^--definition of $\sigma \in [\gamma_{\delta+1}, \gamma_{\delta+1}^{+(\delta+1)}[$, then $\gamma_{\delta+1}$ itself could be defined in $L_{\gamma_{\delta+1}^{+(\delta+1)}}$ as the maximal recursively inaccessible ordinal less than σ. $\square$

Theorem 7.6. (a) *If γ is weakly $\gamma^{+(\delta+1)}$-stable, then γ has no good Σ_1^--definition in $L_{\gamma^{+(\delta+1)}}$.*

(b) *If $\gamma_{\delta+1}$ exists, then it is $\gamma_{\delta+1}^{+(\delta+1)}$-stable.*

Proof. (a) Assume γ to be weakly $\gamma^{+(\delta+1)}$-stable and there is some formula $F[x] \in \Sigma_1$ without further parameters such that $L_{\gamma^{+(\delta+1)}} \models F[\gamma] \wedge \exists! x F[x]$. From $L_{\gamma^{+(\delta+1)}} \models \exists x F[x]$ we infer by weak $\gamma^{+(\delta+1)}$-stability that $L_\gamma \models \exists x F[x]$, i.e. $L_\gamma \models F[x_0]$ for some $x_0 \in L_\gamma$. By persistence $L_{\gamma^{+(\delta+1)}} \models F[x_0]$ contradicting the uniqueness condition.

(b) Assume $L_{\gamma_{\delta+1}^{+(\delta+1)}} \models F(a_1,\ldots,a_m)$ for $a_1,\ldots,a_m \in L_{\gamma_{\delta+1}}$ and $F \in \Sigma_1$. The parameters $a_1,\ldots,a_m$ can be Σ_1-defined using certain ordinals $\alpha_1,\ldots,\alpha_n < \gamma_{\delta+1}$, which in turn can be Σ_1^--defined by Proposition 7.5. So we may as well assume that F is parameter-free. Since $L_{\gamma_{\delta+1}^{+(\delta+1)}} \models F$, for all $\beta \geqslant \gamma_{\delta+1}^{+(\delta+1)}$ we can infer $L_\beta \models \exists! \alpha(\alpha = \min\{\alpha : L_\alpha \models F\})$. This provides a good Σ_1^--definition of the minimal α such that $L_\alpha \models F$, which by Proposition 7.5(c) therefore is less than $\gamma_{\delta+1}$. This means $L_{\gamma_{\delta+1}} \models F$. □

Lemma 7.7. *Let τ_0 be such that no ordinal less than τ_0 satisfies the properties of $\gamma_{\delta+1}$. If $(I_A^\tau)_{\tau<\tau_0}$ exists, then for all $\tau \leqslant \tau_0$*

$(a)_\tau$ $I_A^{<\tau} = \{y \in \omega : L_{\Omega_{(\delta+1)\cdot\tau}} \models \mathrm{Sat}_\Sigma(y,\emptyset)\}$ *where Ω_ξ was defined as the ξ-th admissible or limit of admissibles.*

$(b)_\tau$ $\omega_1^{I_A^{<\tau}} = \Omega_{(\delta+1)\cdot\tau+1}$ *and therefore* $\omega_{\delta+1}^{I_A^{<\tau}} = \Omega_{(\delta+1)\cdot(\tau+1)}$.

Proof. We prove $(a)_\tau$ and $(b)_\tau$ simultaneously by induction on τ. A few words seem in order to justify this induction in our meta-theory. We want to show that it can be expressed by a Δ_0-induction because we can restrict attention to one fixed set in which all these sets exist.

This can be seen as follows. First, it is easy to see that τ_0 is countable because there are only countably many Σ_1^--formulas. So let $f : \omega \to \tau_0$ be a bijection. Defining $X = \{\langle n,x\rangle \in \omega : x \in I_A^{<f(n)}\}$, we have that $\omega_{\delta+1}^{I_A^{<\tau}} \leqslant \omega_{\delta+1}^X$ for all $\tau \leqslant \tau_0$ since $I_A^{<\tau}$ is recursive in X. Therefore, our inductive assertion is a property in $L_{\omega_{\delta+1}^X}$. Hence our theory allows the intended induction.

$(a)_0$ and $(b)_0$ are obvious as $I_A^{<0} = \emptyset$. Next we prove $(a)_\tau$ for $\tau > 0$. If τ is a limit ordinal, we have

$$I_A^{<\tau} = \bigcup_{\xi<\tau} I_A^\xi = \bigcup_{\xi<\tau} \{y \in \omega : L_{\Omega_{(\delta+1)\cdot\xi}} \models \mathrm{Sat}_\Sigma(y,\emptyset)\}$$

$$= \{y \in \omega : L_{\Omega_{(\delta+1)\cdot\tau}} \models \mathrm{Sat}_\Sigma(y,\emptyset)\}$$

by induction hypothesis.

If on the other hand $\tau = \tau' + 1$ holds, we get

$$\begin{aligned} I_A^{<\tau} &= I_A^{<\tau'} \cup A(I_A^{<\tau'}) \\ &= \{y \in \omega : L_{\Omega_{(\delta+1)\cdot\tau'}} \models \mathrm{Sat}_\Sigma(y,\emptyset)\} \cup \{y \in \omega : L_{\omega_{\delta+1}^{I_A^{<\tau'}}} \models \mathrm{Sat}_\Sigma(y,\emptyset)\} \\ &= \{y \in \omega : L_{\Omega_{(\delta+1)\cdot(\tau'+1)}} \models \mathrm{Sat}_\Sigma(y,\emptyset)\} \end{aligned}$$

by i.h. for $(a)_{\tau'}$ and $(b)_{\tau'}$.

To show $(b)_\tau$ let $\xi < \Omega_{(\delta+1)\cdot\tau}$. First, we show that ξ has a good Σ_1^--definition in $L_{\Omega_{(\delta+1)\cdot\tau}}$. Namely, if τ is a limit ordinal, we can conclude that $\xi^{+(\delta+1)} \leqslant \Omega_{(\delta+1)\cdot\tau}$. Since $\gamma_{\delta+1} \nleqslant \xi$, ξ has a good Σ_1^--definition in $L_{\xi^{+(\delta+1)}} \subseteq L_{\Omega_{(\delta+1)\cdot\tau}}$.

For $\tau = \tau' + 1$ the inequality $\xi < \Omega_{(\delta+1)\cdot\tau}$ means $\xi < \Omega_{(\delta+1)\cdot\tau'+\delta+1}$. If $\xi \leqslant \Omega_{(\delta+1)\cdot\tau'}$, we can argue in the same way as in the previous case. If finally $\xi > \Omega_{(\delta+1)\cdot\tau'}$, then we show by induction on ξ that ξ has a good Σ_1-definition with parameters $\leqslant \Omega_{(\delta+1)\cdot\tau'}$, and so it also has a parameter-free one if we replace the parameters by their Σ_1^--definitions from the previous paragraph. This is obvious for the ordinals $\Omega_{(\delta+1)\cdot\tau'+\gamma}$ for $\gamma < \delta$, since these γ have good Σ_1^--descriptions in $L_{\omega_1^{\mathrm{CK}}}$. If $\xi \in]\Omega_{(\delta+1)\cdot\tau'+\gamma}, \Omega_{(\delta+1)\cdot\tau'+\gamma+1}[$, then ξ is the order type of a $\Sigma_1^{L_\nu}$ well-ordering for $\nu = \Omega_{(\delta+1)\cdot\tau'+\gamma}$. This can be expressed by a $\Sigma_1^{L_{\nu'}}$-formula with parameters from ν where $\nu' = \Omega_{(\delta+1)\cdot\tau'+\gamma+1}$. Replacing these parameters by their definitions given by i.h. we get a good Σ_1^--definition of ξ.

Since we have shown that every ordinal in $L_{\Omega_{(\delta+1)\cdot\tau}}$ has a good Σ_1^--definition in $L_{\Omega_{(\delta+1)\cdot\tau}}$ the following defines a pre-wellordering $\prec$ on ω of ordertype $\Omega_{(\delta+1)\cdot\tau}$: Let $x \prec y$ if x and y are Σ_1^--formulas defining ordinals $\alpha, \beta \in L_{\Omega_{(\delta+1)\cdot\tau}}$ and $\alpha < \beta$. By the characterization of $I_A^{<\tau}$ from $(a)_\tau$ we easily see that $\prec$ is recursive in $I_A^{<\tau}$, so $\omega_1^{I_A^{<\tau}} > \Omega_{(\delta+1)\cdot\tau}$. By $(a)_\tau$ again, $I_A^{<\tau} \in L_{\Omega_{(\delta+1)\cdot\tau+1}}$ and therefore $\omega_1^{I_A^{<\tau}} = \Omega_{(\delta+1)\cdot\tau+1}$. Consequently, $\omega_\alpha^{I_A^{<\tau}} = \Omega_{(\delta+1)\cdot\tau+\alpha}$ for all α. □

Theorem 7.8. *Assume there is some γ such that $(I_A^\alpha)_{\alpha<\gamma}$ exists and $A(I_A^{<\gamma}) \subseteq I_A^{<\gamma}$. Then $\gamma_{\delta+1}$ exists and is $\leqslant \gamma$.*

Proof. If $\gamma_{\delta+1} \nless \gamma$, we can apply the previous lemma to γ. It says that

$$I_A^{<\gamma} = \{y \in \omega : L_{\Omega_{(\delta+1)\cdot\gamma}} \models \mathrm{Sat}_\Sigma(y,\emptyset)\}$$

and

$$y \in A(I_A^{<\gamma}) \Leftrightarrow L_{\Omega_{(\delta+1)\cdot\gamma+\delta+1}} \models \mathrm{Sat}_\Sigma(y,\emptyset).$$

Hence $A(I_A^{<\gamma}) \subseteq I_A^{<\gamma}$ leads to

$$L_{\Omega_{(\delta+1)\cdot\gamma+\delta+1}} \models \mathrm{Sat}_\Sigma(y,\emptyset) \Rightarrow L_{\Omega_{(\delta+1)\cdot\gamma}} \models \mathrm{Sat}_\Sigma(y,\emptyset).$$

Since Sat_Σ is universal for Σ_1^--formulas, this implies $L_{\Omega_{(\delta+1)\cdot\gamma}} \prec_1^- L_{\Omega_{(\delta+1)\cdot\gamma+\delta+1}} = L_{(\Omega_{(\delta+1)\cdot\gamma})^{+(\delta+1)}}$. Since it is the least ordinal with this stability property, $\gamma_{\delta+1}$ exists and is $\leqslant \Omega_{(\delta+1)\cdot\gamma}$. Since we also assumed $\gamma \leqslant \gamma_{\delta+1}$ and $\gamma_{\delta+1}$ is inaccessible by Proposition 7.5, this implies $\gamma_{\delta+1} = \gamma$. □

Altogether, we have shown the following in this section:

Corollary 7.9. (a) *On the basis of* $\mathrm{KPl}^r + (V = L)$, *for each* (*meta*) n *the schema*

$$\exists x(x = (I_A^\alpha)_{\alpha<\gamma} \wedge A(\bigcup \mathrm{rng}(x)) \subseteq \bigcup \mathrm{rng}(x))$$

where A *is* $\Pi^1_{1,n}$ *w.o. parameters proves the existence of some* γ *such that* $L_\gamma \prec_1 L_{\gamma^{+(n)}}$.
(b) *On the basis of* $\mathrm{KPl}^r_{<\varepsilon_0} + (V = L)$, *for each* $\delta < \varepsilon_0$ *the schema*

$$\exists x(x = (I_A^\alpha)_{\alpha<\gamma} \wedge A(\bigcup \mathrm{rng}(x)) \subseteq \bigcup \mathrm{rng}(x))$$

where A *is* $\Pi^1_{1,\delta}$ *w.o. parameters proves the existence of some* γ *such that* $L_\gamma \prec_1 L_{\gamma^{+(\delta)}}$.

Proof. (a) Let $n > 0$ be given. Theorem 7.8 yields in this case that γ_n exists, Theorem 7.6 says that it is $\gamma_n^{+(n)}$-stable as desired.
(b) is proved in the same way. □

8. Modeling set theory using representation trees

We are left with the task to reduce axiom systems for set theory postulating the existence of inductively generated sub-fixpoints of (non-monotonic) $\Pi^1_{1,\delta}$-operators to systems of explicit mathematics with (MID). The next step in this direction is to model set theory in a way such that it can be treated in constructive systems.

For this we want to use the method of representation trees, which originally was used to reduce systems of set theory to systems of second-order arithmetic. We will use the following theorem, which can be found for example in [18, Corollary 7.2].

Proposition 8.1. (a) $\Pi^1_1\text{-CA}_0 \vdash F^{REP}$ *for each axiom* F *of* KPl^r.
(b) $\Pi^1_1\text{-CA}_0 \vdash F \leftrightarrow F^{REP}$ *for each* $F \in \mathscr{L}_2$.
(c) *Analogously to* (a) *it holds* $\Pi^1_1\text{-CA}_{<\varepsilon_0} \vdash F^{REP}$ *for each axiom* F *of* $\mathrm{KPl}^r_{<\varepsilon_0}$.

Our intention is to use part (b) of this proposition to treat axioms which state the existence of inductively generated sub-fixpoints of $\Pi^1_{1,\delta}$-operators by translating them to statements of second-order arithmetic which then can be treated in the context of explicit mathematics. For this, we have to get rid of the ordinals in the formulation of the definition of inductively generated sub-fixpoints. We will replace them by pre-wellorderings.

Definition 8.2. (a) A binary relation $\prec \subset A^2$ for some set A is called a *pre-wellordering* of A if it is transitive, linear and satisfies

$$\forall x, y \in A(x \prec y \vee y \prec x \vee x \equiv_\prec y),$$

where $x \equiv_\prec y$ means $\forall u \in A((u \prec x \leftrightarrow u \prec y) \wedge (x \prec u \leftrightarrow y \prec u))$.

(b) $X = \mathscr{H}_\Gamma(\prec) \Leftrightarrow X_n = \begin{cases} \emptyset & \text{if } n \notin \text{field}(\prec) \\ \Gamma(X_{\prec n}) \cup X_{\prec n} & \text{if } n \in \text{field}(\prec) \end{cases}$

Proposition 8.3. *Assume that Γ is a $\Pi^1_{1,\delta}$-operator. The following principles are equivalent in* $\mathrm{KPl}^r + (V = L) + \forall\gamma(\gamma^{+(\delta+1)}$ *exists*):

(a) $\exists\gamma\exists x(x = I_\Gamma^{<\gamma} \wedge \Gamma(x) \subseteq x)$

(b) $\exists \prec \subseteq \omega \exists X \subseteq \omega(\mathrm{PWO}(\prec) \wedge X = \mathscr{H}_\Gamma(\prec) \wedge \Gamma(\bigcup_n X_n) \subseteq \bigcup_n X_n)$.

We abbreviate the latter principle as $\mathscr{I}(\Gamma)$.

Proof. (a) $\Rightarrow$ (b). Take the least such γ. Then γ is countable. (An injective mapping into ω is given by mapping $\alpha < \gamma$ to the least n such that $n \in \Gamma(I_\Gamma^{<\alpha}) \backslash I_\Gamma^{<\alpha}$.) So take a bijection $f : \omega \to \gamma$, neglecting the trivial case that γ is finite. Define $m \prec n \Leftrightarrow f(m) < f(n)$. This is a well-ordering, especially a pre-wellordering. Defining the set $X \subseteq \omega$ by $X_n = I_\Gamma^{f(n)}$, we see that $X = \mathscr{H}_\Gamma(\prec)$.

(b) $\Rightarrow$ (a). Let $\prec$ be a pre-wellordering on ω and $X = \mathscr{H}_\Gamma(\prec)$ as in (b). Define $\|.\| : \text{field}(\prec) \to \mathrm{On}$ by $\|n\| = \{\|m\| : m \prec n\}$. This definition can be formalized as a $\prec$-recursion in some u with $\mathrm{Ad}(u)$ and $\prec \in u$. By $\prec$-induction in u, we can show

(1) $\forall n \in \text{field}(\prec)(\|n\| \in \mathrm{On})$ and $\gamma := \bigcup_n \|n\| \in \mathrm{On}$.

(2) $\forall m, n \in \text{field}(\prec)(\|m\| = \|n\| \to X_m = X_n)$.

Using the weak inverse $g : \gamma \to \text{field}(\prec)$ with $g(\alpha) = \min\{n \in \omega : \|n\| = \alpha\}$, we can define $x_\alpha = X_{g(\alpha)}$. We show that $(x_\alpha)_{\alpha<\gamma} = (I_\Gamma^\alpha)_{\alpha<\gamma}$.

We have to show that $x_\alpha = \Gamma(\bigcup_{\beta<\alpha} x_\beta) \cup \bigcup_{\beta<\alpha} x_\beta$. For this, the first thing to note is that $\bigcup_{m\prec g(\alpha)} X_m = \bigcup_{\beta<\alpha} x_\beta$. Namely, if $m \prec g(\alpha)$, then $\|m\| < \alpha$. Since $\|g(\|m\|)\| = \|m\|$, (2) then yields $X_m = X_{g(\|m\|)} = x_{\|m\|} \subseteq \bigcup_{\beta<\alpha} x_\beta$. Conversely, from $\beta < \alpha$ we obtain $\beta = \|m\|$ for some $m \prec g(\alpha)$. But then $x_\beta = X_{g(\beta)} = X_m$ because of (2) and $\|m\| = \beta = \|g(\beta)\|$.

Therefore we have

$$x_\alpha = X_{g(\alpha)} = \Gamma(\bigcup_{m\prec g(\alpha)} X_m) \cup \bigcup_{m\prec g(\alpha)} X_m = \Gamma(\bigcup_{\beta<\alpha} x_\beta) \cup \bigcup_{\beta<\alpha} x_\beta.$$

Finally,

$$\Gamma(\bigcup_{\alpha<\gamma} x_\alpha) = \Gamma(\bigcup_{m\in \text{field}(\prec)} X_m) \subseteq \bigcup_{m\in \text{field}(\prec)} X_m = \bigcup_{\alpha<\gamma} x_\alpha$$

follows in the same way. $\square$

Proposition 8.4. *An operator $\Gamma : \mathrm{Pow}(\omega) \to \mathrm{Pow}(\omega)$ is a $\Pi^1_{1,\delta}$-operator iff it can be written in the form*

$$n \in \Gamma(X) \Leftrightarrow F[n, HJ(\delta, \langle X_1, \ldots, X_n\rangle)]$$

for some Π^1_1-formula $F[a,A]$ and sets $X_1, \ldots, X_n \subseteq \omega$. Replacing the ordinal δ with its notation in the set $\lhd$ from Section 2.1 we can regard this definition of Γ as given by an $\mathscr{L}_2$-formula.

Proof. This also follows using suitable trees and noting that $\omega_{\delta+1}^{\langle X,X_1,\dots,X_n\rangle}$ is the least ordinal not recursive in $\mathrm{HJ}(\delta,\langle X,X_1,\dots,X_n\rangle)$. □

9. Characterizations of proof-theoretic strength

Now we are ready to combine the material assembled in the previous sections to prove the following main result:

Theorem 9.1. *Let F be a Σ^1_2-sentence. Then*
(a) $\mathrm{EM}_0{\restriction} + (\mathrm{Join}) + (\mathrm{IG}){\restriction} + (\mathrm{MID}) \vdash F \Leftrightarrow \mathrm{KPi}^r + \exists\gamma(L_\gamma \prec_1 L) \vdash F$.
(b) $\mathrm{EM}_0 + (\mathrm{Join}) + (\mathrm{IG}){\restriction} + (\mathrm{MID}) \vdash F \Leftrightarrow \mathrm{KPi}^w + \exists\gamma(L_\gamma \prec_1 L) \vdash F$.

Proof. (a) Assume $\mathrm{EM}_0{\restriction} + (\mathrm{Join}) + (\mathrm{IG}){\restriction} + (\mathrm{MID}) \vdash F$ for a Σ^1_2-sentence $F \equiv \exists X F_0(X)$ where F_0 is Π^1_1. Defining an operation f that computes from a classification $X \subseteq \omega$ the tree of unsecured sequences for $F_0(X)$,[6] we see that, in $\mathrm{EM}_0{\restriction} + (\mathrm{IG}){\restriction}$, F is equivalent to the Σ_1-formula (in the sense of Definition 5.1)

$$F^\circ :\equiv \exists X \exists Y (fX \simeq Y \wedge \mathrm{i}(N,Y) \simeq N)).$$

Working in $\mathrm{KPi}^r + \exists\gamma(L_\gamma \prec_1 L)$, by Corollary 5.13 $\mathfrak{S}_{M,\omega} \models F^\circ$ and since $\mathfrak{S}_{M,\omega} \models \mathrm{EM}_0{\restriction} + (\mathrm{IG}){\restriction}$, we also have $\mathfrak{S}_{M,\omega} \models F$ for the model constructed there. (Note that the least γ such that $L_\gamma \prec_1 L$ is countable. In fact, the proof of this fact, cf. e.g. [1, V 7.8], can easily be formalized in KPl^r. Accordingly, the assumption from subsection 5.2 is satisfied.) Since moreover for $X \subseteq \omega$

$$\begin{aligned} X \in L_\gamma &\Leftrightarrow X^\circ = \{z : z \,\varepsilon\, a\} \text{ for some } a \in \mathrm{Cl}_{M,0} \\ &\Leftrightarrow X^\circ = \{z : z \,\varepsilon\, a\} \text{ for some } a \in \mathrm{Cl}_{M,\omega}, \end{aligned}$$

we also see that the second-order part of $\mathfrak{S}_{M,\omega}$ is isomorphic to $L_\gamma \cap \mathrm{Pow}(\mathrm{N})$. Therefore $L_\gamma \models F$. By the Quantifier Theorem 2.19, F is equivalent in KPl^r to a Σ_1-formula of set theory. Therefore persistency implies that F holds in the universe of $\mathrm{KPi}^r + \exists\gamma(L_\gamma \prec_1 L)$.

For the converse direction, assume $\mathrm{KPi}^r + \exists\gamma(L_\gamma \prec_1 L) \vdash F$ for a Σ^1_2-sentence F. Since F is equivalent to a Σ_1-formula, Proposition 6.10 yields $\mathrm{KPl}^r + \exists\alpha(L_\alpha \prec_1 L_{\alpha^{+(n)}}) \vdash F$ for some $n \in \mathbb{N}$. By Corollary 7.9 we get that

(9) $\mathrm{KPl}^r + (V = L) +$
$\{\exists x(x = (I^\alpha_A)_{\alpha<\gamma}) \wedge A(I^{<\gamma}_A) \subseteq I^{<\gamma}_A : A \text{ is } \Pi^1_{1,n} \text{w.o.parameters}\} \vdash F$.

By Proposition 8.3 we see that this is equivalent to

$$\mathrm{KPl}^r + (V = L) + \{\mathcal{I}(A) : A \text{ is } \Pi^1_{1,n} \text{ w.o. parameters}\} \vdash F.$$

[6] i.e. fX is a relation $<_{F_0(X)}$ defining a tree which is well founded iff $F_0(X)$ holds, cf. e.g. [16, Theorem III.3.2].

Since by Shoenfield absoluteness $\mathscr{I}(\Lambda)$ implies $\mathscr{I}(\Lambda)^L$, this implies

$$\mathrm{KPl}^r + \{\mathscr{I}(\Lambda) : \Lambda \text{ is } \Pi^1_{1,n} \text{ w.o. parameters}\} \vdash F^L$$

and

$$\mathrm{KPl}^r + \{\mathscr{I}(\Lambda) : \Lambda \text{ is } \Pi^1_{1,n} \text{ w.o. parameters}\} \vdash F$$

by absoluteness again. Since $\mathscr{I}(\Lambda)$ can be considered to be an $\mathscr{L}_2$-formula using Proposition 8.4, by Proposition 8.1 we get

$$\Pi^1_1\text{-CA}_0 + \{\mathscr{I}(\Lambda) : \Lambda \text{ is } \Pi^1_{1,n} \text{ w.o. parameters}\} \vdash F.$$

Π^1_1-CA$_0$ may be regarded as a subtheory of $\mathrm{EM}_0\restriction + (\mathrm{Join}) + (\mathrm{IG})\restriction$ and the main result of [21] says that $\mathrm{EM}_0\restriction + (\mathrm{Join}) + (\mathrm{IG})\restriction$ proves $\mathscr{I}(\Lambda)$. To see this, note that the operation $X \mapsto \Lambda(X)$ can be defined in $\mathrm{EM}_0\restriction + (\mathrm{Join}) + (\mathrm{IG})\restriction$. Namely, let $x \in \Lambda(X) \Leftrightarrow G[x, \mathrm{HJ}(n,X)]$ where $G \in \Pi^1_1$ by Proposition 8.4. Using an operation f which maps (x,Y) to the tree of unsecured sequences of $G[x,Y]$ and a g which maps X to $\mathrm{HJ}(n,X)$, we can define

$$x \in \Lambda(X) \Leftrightarrow x \in \{y : N \mathrel{\mathring{=}} \mathrm{i}(N, f(y, gX))\}$$

for $N = \{x : \mathrm{N}(x)\}$. This gives rise to an extensional operator in $\mathrm{EM}_0\restriction + (\mathrm{Join}) + (\mathrm{IG})\restriction$. So by Theorem 4.1 of [21], which for convenience is quoted below, $\mathscr{I}(\Lambda)$ is provable in $\mathrm{EM}_0\restriction + (\mathrm{Join}) + (\mathrm{IG})\restriction + (\mathrm{MID})$. Therefore

$$\mathrm{EM}_0\restriction + (\mathrm{Join}) + (\mathrm{IG})\restriction + (\mathrm{MID}) \vdash F.$$

(b) can be proved analogously. □

Theorem 9.2 (Rathjen). *In* $\mathrm{T}_0\restriction + (\mathrm{MID})$, *to any operator* Λ *there can be associated a monotone operator* Υ *and a total operation* $x \mapsto \Lambda^x$, *giving a classification* Λ^x *for all* x, *such that with* $<_\Upsilon$ *denoting the pre-wellordering pertaining to* Υ

$$\Lambda^x \mathrel{\mathring{=}} \Lambda\left(\bigcup_{y <_\Upsilon x} \Lambda^y\right) \cup \bigcup_{y <_\Upsilon x} \Lambda^y,$$

and, for the classification defined by

$$\mathrm{I}_\Lambda := \bigcup_{x \varepsilon V} \Lambda^x$$

it is

$$\Lambda(\mathrm{I}_\Lambda) \mathrel{\mathring{\subseteq}} \mathrm{I}_\Lambda.$$

Put differently: I_Λ *is a classification that arises by iterating* Λ *along* $<_\Upsilon$ *and is closed under* Λ.

Since the pre-wellordering $<_\Upsilon$ can be taken to be a subset of $\mathbb{N}$ if $\Lambda : \mathrm{Pow}(\mathbb{N}) \to \mathrm{Pow}(\mathbb{N})$ the theorem amounts to saying that the principle $\mathscr{I}(\Lambda)$ can be proved in $\mathrm{T}_0{\restriction} + (\mathrm{MID})$ for operators Λ that are operations in that theory.

Remark 9.3. The proof-theoretic strength of the theories of explicit mathematics does not increase if we add Church's Thesis to them since the model used in the previous proof satisfies this additional axiom.

Alternatively, we could also add the non-constructive μ-operator as introduced in [6] to the theories without increasing proof-strength. To see this, note that we could have used the model built on the applicative system consisting of the Δ^1_1-indices and performed literally the same proof.

Finally, the result (a) of the preceding theorem remains correct if we omit the join axiom from the system of explicit mathematics. In fact, Glaß' work in [12] shows that addition of the join axiom in this context leads to an extension which is conservative (at least) for Σ^1_2-sentences.

10. Connections to theories of second-order arithmetic

In this section we want to indicate shortly which subsystems of second-order arithmetic correspond to the theories we encountered in this paper. Let $(\Pi^1_2\text{-CA}^-)$ be the axiom scheme of comprehension for Π^1_2-formulas without parameters.

Theorem 10.1. *For all $\mathscr{L}_2$-sentences F we have*

$$\mathrm{KPi}^r + \exists\gamma(L_\gamma \prec_1 L) \vdash F \Leftrightarrow \Sigma^1_2\text{-AC}_0 + (\Pi^1_2\text{-CA}^-) \vdash F$$

and

$$\mathrm{KPi}^w + \exists\gamma(L_\gamma \prec_1 L) \vdash F \Leftrightarrow \Sigma^1_2\text{-AC} + (\Pi^1_2\text{-CA}^-) \vdash F.$$

Proof. Consider the first assertion. First we treat "$\Leftarrow$". Here it is easy to show, cf. e.g. [18, Theorem 8.2, Lemma 8.2], that $\Sigma^1_2\text{-AC}_0 \subseteq \mathrm{KPi}^r$. Hence we have to show

$$\exists X(X = \{x \in \omega : F(x)\})$$

where F is a Σ^1_2-formula without parameters (and then to use recursive comprehension to define $Y := \mathbb{N}\backslash X$). F is equivalent to a Σ_1-formula G in KPl^r. Let $L_\gamma \prec_1 L$ and define $X := \{x \in \omega : L_\gamma \models G(x)\}$. Using the stability of γ and Shoenfield-Absoluteness, we see that X is as required.

As to "$\Rightarrow$", it is a standard result, cf. again [18, Theorem 8.3], that $\Sigma^1_2\text{-AC}_0 \vdash F^{\mathrm{REP}}$ for each axiom F of KPi^r. What is left to show is that

$$\Sigma^1_2\text{-AC}_0 + (\Pi^1_2\text{-CA}^-) \vdash \exists\gamma(L_\gamma \prec^-_1 L)^{\mathrm{REP}},$$

because as in Section 7 it can be shown that the least γ such that $L_\gamma \prec^-_1 L$ already satisfies $L_\gamma \prec_1 L$. Using $\Pi^1_2\text{-CA}^-$, define $X = \{\ulcorner F \urcorner : (L \models F \text{ and } F \in \Sigma^-_1)^{\mathrm{REP}}\}$. Here we use the fact that validity of a Σ_1-formula in L is represented by a Σ^1_2-formula. Using this X, we have

$$\forall \ulcorner F \urcorner \in X \exists Y (Y = L_\alpha \wedge L_\alpha \models F)^{\mathrm{REP}}.$$

By the Σ^1_2-axiom of choice we obtain a set Z such that

$$\forall \ulcorner F \urcorner \in X (Z_{\ulcorner F \urcorner} = L_\alpha \wedge L_\alpha \models F)^{\mathrm{REP}}.$$

These Z can easily be combined to a representation tree for some set L_γ such that $L_\gamma \prec_1^- L$.

Obviously, for the second part of the assertion we only have to add induction on the integers to both theories. □

11. Related results and future research

In this paper, we did not attack the most obvious next question, namely the question about the strength of the full system T_0 + (MID). Although the results of this paper seem to suggest that the strength of this system is that of KPi + $\exists\gamma(L_\gamma \prec_1 L)$, we conjecture that $T_0 + \mu + D$ exceeds the strength of that theory. The full machinery of ordinal analysis for impredicative systems would have to be developed for systems of explicit mathematics. To include that in this paper certainly would exceed the tolerable limits for its length.

In any case, the results of this paper already show the principles at work in theories containing (MID) in addition to (Join) and (IG) $\restriction$-axioms. This picture should not change when allowing the full schema of (IG), we expect it to involve a more or less obvious iteration of these principles. The computation of the proof-theoretic strength of this system therefore is more of technical than of foundational interest.

The question about systems of explicit mathematics containing (UMID) seems to be much more interesting and challenging. The techniques of this paper do not apply to this axiom. The additional operation lfp allows to iterate the formation of fixpoints of inductive definitions in a very general way. A computation of the proof-theoretic strength of these systems would certainly lead to deep insight into these induction principles. Especially it would be interesting to compare this theory to subsystems of second-order arithmetic based on Π^1_2-comprehension and to systems of set theory axiomatizing a non-projectible ordinal.

Finally, it should be mentioned again that we only considered theories of explicit mathematics based on classical logic. It is not clear whether the theories have the same strength when formulated on the basis of intuitionistic logic. As we said in the introduction, if that is not the case, we would get a radically different situation from former characterizations of proof-theoretical strength of T_0 and various of its subsystems, which have turned out to be independent of whether the underlying logic is classical or intuitionistic.

References

[1] J. Barwise, Admissible Sets and Structures, Perspectives in Mathematical Logic (Springer, Berlin, 1975).
[2] M.J. Beeson, Foundations of Constructive Mathematics, Ergebnisse der Mathematik und ihrer Grenzgebiete (Springer, New York, 1985).
[3] M. Boffa, D. van Dalen and K. McAloon, eds., Logic Colloquium 78, Mons 1978, Studies in Logic and the Foundations of Mathematics, Vol. 78 (North-Holland, Amsterdam, 1979).
[4] D. Cenzer, Ordinal recursion and inductive definitions, in: [10, 221–264].
[5] J.N. Crossley, ed., Algebra and Logic, Lecture Notes in Mathematics, Vol. 450 (Springer, New York, 1975).
[6] S. Feferman, A language and axioms for explicit mathematics, in: [5, 87–139].
[7] S. Feferman, Constructive theories of functions and classes, in: [3, 159–224].
[8] S. Feferman, Monotone inductive definitions, in: A.S. Troelstra and D. van Dalen, eds., The L.E.J. Brouwer Centenary Symp., Studies in Logic and the Foundations of Mathematics, Vol. 110 (North-Holland, Amsterdam, 1982) 95–128.
[9] S. Feferman and W. Sieg, Proof-theoretical equivalences between classical and constructive theories for analysis, in: Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies, Lecture Notes in Mathematics, Vol. 897, Ch. II (Springer, Heidelberg/New York, 1981) 78–142.
[10] J.E. Fenstad and P.G. Hinman, eds., Generalized Recursion Theory I, Oslo 1972, Studies in Logic and the Foundations of Mathematics, Vol. 79 (North-Holland, Amsterdam, 1974).
[11] H. Friedman, Iterated inductive definitions and Σ^1_2-AC, in: Akiko Kino, John Myhill, and Richard E. Vesley, eds., Intuitionism and Proof Theory, Studies in Logic and the Foundations of Mathematics (North-Holland, Amsterdam, 1970) 435–442.
[12] T. Glaß, Standardstrukturen für Systeme Expliziter Mathematik (Eine beweistheoretische Untersuchung), Ph.D. Thesis, Universität Münster, 1993.
[13] T. Glaß, Understanding uniformity in Feferman's explicit mathematics, Ann. Pure Appl. Logic 75 (1995) 89–106.
[14] E. Griffor and M. Rathjen, The strength of some Martin-Löf type theories, Arch. Math. Logic 33 (1994) 347–385.
[15] L. Harrington and A.S. Kechris, On monotone versus nonmonotone induction, Bull. Amer. Math. Soc. 82 (1976) 888–890.
[16] P.G. Hinman, Recursion-theoretic Hierarchies, Perspectives in Mathematical Logic (Springer, Heidelberg, 1978).
[17] G. Jäger, A well ordering proof for Feferman's theory T_0, Arch. Math. Logik und Grundlagenforschung 23 (1983) 65–77.
[18] G. Jäger, Theories for Admissible Sets: A Unifying Approach to Proof Theory, Stud. Proof Theory, Lect. Notes, Vol. 2 (Bibliopolis, Naples, 1986).
[19] G. Jäger and W. Pohlers, Eine beweistheoretische Untersuchung von Δ^1_2-CA + (BI) und verwandter Systeme, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse, 1982.
[20] W. Pohlers, Proof Theory. An Introduction, Lecture Notes in Mathematics, Vol. 1407 (Springer, New York, 1989).
[21] M. Rathjen, Monotone inductive definitions in explicit mathematics, J. Symbolic Logic 61 (1996) 125–146.
[22] K. Schütte, Proof Theory, Grundlehren der mathematischen Wissenschaften, Vol 225 (Springer, Berlin, 1977).
[23] H. Schwichtenberg, Proof theory: some applications of cut-elimination, in: J. Barwise, ed., Handbook of Mathematical Logic (North-Holland, Amsterdam, 1977) 867–895.
[24] T. Setzer, Proof Theoretical Strength of Martin-Löf Type Theory with W-Type and One Universe, Ph.D. Thesis, Universität München, 1993.
[25] S.G. Simpson, Subsystems of second order arithmetic, Chs. II-V and VII, Tech. Report, Pennsylvania State University, 1986.
[26] Stanford Report, mimeographed, 1963.
[27] S. Takahashi, Monotone inductive definitions in a constructive theory of functions and classes, Ann. Pure Appl. Logic 42 (1989) 255–297.