First order theories for nonmonotone inductive definitions: recursively inaccessible and Mahlo

Gerhard Jäger

Abstract

In this paper first order theories for nonmonotone inductive definitions are introduced, and a proof-theoretic analysis for such theories based on combined operator forms à la Richter with recursively inaccessible and Mahlo closure ordinals is given.

1 Introduction

Let Φ be an operator on the power set $P(\mathbb{N})$ of the natural numbers, i.e. a mapping from $P(\mathbb{N})$ to $P(\mathbb{N})$. Then Φ can be used to generate subsets I_Φ^σ of the natural numbers if we define

$$I_\Phi^\sigma := I_\Phi^{<\sigma} \cup \Phi(I_\Phi^{<\sigma}) \qquad \text{and} \qquad I_\Phi^{<\sigma} := \bigcup \{I_\Phi^\tau : \tau < \sigma\}$$

by transfinite recursion on the ordinals. Furthermore we let

$$I_\Phi^\infty := \bigcup \{I_\Phi^\tau : \tau \ an \ ordinal \ \}$$

be the set of natural numbers *inductively defined by* Φ. Obviously there exists a least ordinal ρ so that $I_\Phi^\rho = I_\Phi^{<\rho}$. We call this ordinal the *closure ordinal* of the inductive definition generated by Φ and know that I_Φ^∞ is identical to I_Φ^ρ. The sets I_Φ^σ are the *stages* of the inductive definition generated by Φ. If $\mathcal{K}$ is a class of operators, then the closure ordinal of $\mathcal{K}$, denoted by $cl(\mathcal{K})$, is the supremum of the closure ordinals of the inductive definitions generated by operators from $\mathcal{K}$.

A lot is known about inductive definitions. The situation is particularly well analyzed if Φ is monotone, i.e. if $S_1 \subset S_2$ implies $\Phi(S_1) \subset \Phi(S_2)$ for all sets of natural numbers S_1 and S_2. Good introductions into the recursion theory, definability theory and proof theory of (special classes of) monotone inductive definitions are, for example, provided by Moschovakis [17] and Buchholz, Feferman, Pohlers and Sieg [6].

There exist also important recursion-theoretic results about various classes of nonmonotone inductive definitions, see e.g. Aczel and Richter [2], Richter [26] and the

1

papers quoted in these articles. The proof theory of nonmonotone inductive definitions, on the other hand, has not been developed to the same degree.

In the first part of this article we introduce first order theories $\mathsf{FID}(\mathcal{K})$ which are tailored for representing arbitrary classes $\mathcal{K}$ of first order inductive definitions. Then we turn to (the theories of) several specific nonmonotone inductive definitions which are interesting in the context of recursively inaccessible and recursively Mahlo ordinals. Not surprisingly, such theories are closely related to the corresponding theories for iterated admissible sets.

We are particularly interested in the operator classes $[\Pi_1^0, \Pi_1^0]$, $[\Pi_1^0, \Pi_0^0]$, $[\mathrm{POS}, \Pi_1^0]$ and $[\mathrm{POS}, \mathrm{QF}]$ according to the notation of Richter [26]; they are also defined in Section 3 in full detail. The embeddings of the corresponding theories $\mathsf{FID}(\mathcal{K})$ and some of their subsystems (which are obtained by restricting the induction principles) into suitable systems of iterated admissible sets provide a perspicuous treatment of interesting nonmonotone inductive definitions and simplify several proofs described in the literature. Moreover, they also provide (sharp) upper bounds for the proof-theoretic strength of these theories for nonmonotone inductive definitions.

Their lower proof-theoretic bounds are not explicitly analyzed in this article. However, they follow immediately from connections between specific theories for nonmonotone inductive definitions and systems of Feferman's explicit mathematics (cf. e.g. Feferman [7, 8]) such as T_0 and some of its natural subsystems and extensions.

Actually, one of the main reasons for introducing theories for nonmonotone inductive definitions is the desire to provide a natural and powerful framework for defining canonical models of explicit mathematics. More in this direction can be found in Jäger and Studer [16] and Studer [27].

2   Theories for first order inductive definitions

In this section we introduce adequate first order theories for representing first order inductive definitions, no matter whether their definition clauses are are positive or not. Such theories can be easily obtained if we have ordinals for representing the stages of these inductive definitions at our disposal.

In the following we let $\mathcal{L}$ denote some standard language of first order arithmetic. $\mathcal{L}$ includes *number variables* $(x, y, z, x_0, y_0, z_0, \ldots)$ and symbols for all primitive recursive functions and relations. The *number terms* $(s, t, s_0, t_0, \ldots)$ of $\mathcal{L}$ are defined as usual. The *atomic formulas* of $\mathcal{L}$ are all expressions $R(s_1, \ldots, s_n)$ so that R is a relation symbol of $\mathcal{L}$. The *formulas* of $\mathcal{L}$ are generated from the atomic formulas of $\mathcal{L}$ by closing under negations, disjunctions, conjunctions and numerical quantifications; the remaining logical connectives are abbreviated as expected.

If X is a fresh unary[1] relation symbol, then $\mathcal{L}(X)$ is the extension of $\mathcal{L}$ by X. The

---

[1]For notational simplicity we restrict ourselves to a fresh unary relation symbol; however, it is obvious how all following arguments can be generalized to arbitrary arities.

definition of positive and negative occurrences of X within $\mathcal{L}(X)$ formulas A is as always. If the $\mathcal{L}(X)$ formula A has no negative occurrences of X, then one speaks of an X *positive* $\mathcal{L}(X)$ formula; the collection of all X positive $\mathcal{L}(X)$ formulas is denoted by POS.

The QF formulas of $\mathcal{L}(X)$ are the quantifier-free $\mathcal{L}(X)$ formulas. Numerical quantifiers are bounded if they occur in a context $\exists x(x <_N t \wedge \ldots)$ or $\forall x(x <_N t \rightarrow \ldots)$ for the primitive recursive less relation $<_N$ and a number term t not containing x; then we often write $(\exists x <_N t)(\ldots)$ and $(\forall x <_N t)(\ldots)$, respectively.[2] The Π_0^0 formulas of $\mathcal{L}(X)$ are the $\mathcal{L}(X)$ formulas whose quantifiers are bounded, the Π_1^0 formulas of $\mathcal{L}(X)$ comprise the Π_0^0 formulas of $\mathcal{L}(X)$ as well as all $\mathcal{L}(X)$ formulas of the form $\forall x A(x)$ so that $A(x)$ is a Π_0^0 formula of $\mathcal{L}(X)$.

An $\mathcal{L}(X)$ formula which contains at most x free is called a (unary) *operator form*, and we let $\mathcal{A}(X, x)$ range over such forms. Observe that operator forms are not required to be X positive. Sometimes we simply write POS, QF, Π_0^0 and Π_1^0 for the collections of all operator forms which are positive, QF, Π_0^0 and Π_1^0 formulas of $\mathcal{L}(X)$, respectively.

Later we will turn to specific operator forms, but the following theories can be formulated without imposing any syntactic restrictions. Hence let $\mathcal{K}$ be an arbitrary collection of operator forms. Then we extend $\mathcal{L}$ to a new first order language $\mathcal{L}_\mathcal{K}$ by adding a new sort of *ordinal variables* $(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, \ldots)$, a new binary relation symbol $<$ for the less relation on the ordinals and a binary relation symbol $P_\mathcal{A}$ for each unary operator form $\mathcal{A}(X, x)$ from $\mathcal{K}$. In addition, we write $P_\mathcal{A}^\alpha(s)$ for $P_\mathcal{A}(\alpha, s)$.

The atomic formulas of $\mathcal{L}_\mathcal{K}$ comprise the atomic formulas of $\mathcal{L}$ plus the expressions of the form $(\alpha < \beta)$ and $P_\mathcal{A}^\alpha(s)$ for all operator forms $\mathcal{A}(X, x)$ from $\mathcal{K}$. The formulas $(A, B, C, A_0, B_0, C_0, \ldots)$ of $\mathcal{L}_\mathcal{K}$ are inductively generated as follows:

1. Each atomic formula of $\mathcal{L}_\mathcal{K}$ is an $\mathcal{L}_\mathcal{K}$ formula.

2. If A and B are $\mathcal{L}_\mathcal{K}$ formulas, then $\neg A$, $(A \vee B)$ and $(A \wedge B)$ are $\mathcal{L}_\mathcal{K}$ formulas.

3. If A is an $\mathcal{L}_\mathcal{K}$ formula, then $\exists x A$ and $\forall x A$ are $\mathcal{L}_\mathcal{K}$ formulas.

4. If A is an $\mathcal{L}_\mathcal{K}$ formula, then $\exists \alpha A$, $\forall \alpha A$, $(\exists \alpha < \beta) A$ and $(\forall \alpha < \beta) A$ are $\mathcal{L}_\mathcal{K}$ formulas.

Quantifiers of the form $(Q\alpha < \beta)$ are called *bounded ordinal quantifier*. Furthermore, we write A^α to denote the $\mathcal{L}_\mathcal{K}$ formula which is obtained from A by replacing all unbounded ordinal quantifiers $Q\beta$ in A by the bounded ordinal quantifier $(Q\beta < \alpha)$. Additional abbreviations are

$$P_\mathcal{A}^{<\alpha}(s) := (\exists \beta < \alpha) P_\mathcal{A}^\beta(s) \qquad \text{and} \qquad P_\mathcal{A}^\infty(s) := \exists \alpha P_\mathcal{A}^\alpha(s).$$

---

[2] We write $<_N$ for the primitive recursive less relation since later a less relation on the ordinals will also be used.

For any $\mathcal{L}(X)$ formula $A(X)$ and $\mathcal{L}_\mathcal{K}$ formula $B(x)$, perhaps with other free variables, $A(\lambda x.B(x))$, or simply $A(B)$, denotes the result of substituting $B(s)$ for each occurrence of $X(s)$ in $A(X)$.

The $\Delta_0^\mathbb{O}$ formulas of $\mathcal{L}_\mathcal{K}$ are the $\mathcal{L}_\mathcal{K}$ formulas which do not contain unbounded ordinal quantifiers. An $\mathcal{L}_\mathcal{K}$ formula is called a $\Sigma^\mathbb{O}$ $[\Pi^\mathbb{O}]$ formula of $\mathcal{L}_\mathcal{K}$ if it does not contain positive [negative] occurrences of unbounded universal ordinal quantifiers and negative [positive] occurrences of unbounded existential ordinal quantifiers.

All theories which we will consider contain certain induction principles. In our present context we distinguish between induction on the natural numbers and induction on the ordinals. If $\mathcal{F}$ is a collection of $\mathcal{L}_\mathcal{K}$ formulas, then induction on the natural numbers with respect to $\mathcal{F}$ consists of all formulas

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

so that $A(x)$ belongs to $\mathcal{F}$. Induction on the ordinals with respect to $\mathcal{F}$, on the other hand, consists of all formulas

$$(\mathcal{F}\text{-I}_\mathbb{O}) \qquad\qquad \forall \alpha[(\forall \beta < \alpha)A(\beta) \to A(\alpha)] \ \to \ \forall \alpha A(\alpha)$$

where $A(\alpha)$ is in $\mathcal{F}$. For us the induction schemas $(\Delta_0^\mathbb{O}\text{-I}_\mathbb{N})$, $(\Delta_0^\mathbb{O}\text{-I}_\mathbb{O})$, $(\mathcal{L}_\mathcal{K}\text{-I}_\mathbb{N})$ and $(\mathcal{L}_\mathcal{K}\text{-I}_\mathbb{O})$ will be central; they provide induction on the natural numbers and ordinals for all $\Delta_0^\mathbb{O}$ formulas of $\mathcal{L}_\mathcal{K}$ and arbitrary $\mathcal{L}_\mathcal{K}$ formulas, respectively.

Now we are ready to present the theory $\mathsf{FID}(\mathcal{K})$ for the inductive definitions with definition clauses from $\mathcal{K}$. It is formulated in the language $\mathcal{L}_\mathcal{K}$, and its axioms can be divided into the following four groups.

I. Number-theoretic axioms. The axioms of Peano arithmetic PA with exception of complete induction on the natural numbers.

II. Linearity axioms. They state that the binary relation symbols $<$ provides a linear ordering of the ordinals.

III. Operator axioms. For all operator forms $\mathcal{A}(X, x)$ from $\mathcal{K}$ we have the following axioms:

$$(\mathsf{Op.1}) \qquad\qquad P_\mathcal{A}^\alpha(s) \ \leftrightarrow \ P_\mathcal{A}^{<\alpha}(s) \vee \mathcal{A}(P_\mathcal{A}^{<\alpha}, s),$$

$$(\mathsf{Op.2}) \qquad\qquad \mathcal{A}(P_\mathcal{A}^\infty, s) \ \to \ P_\mathcal{A}^\infty(s).$$

IV. Induction principles. These consist of the schemas $(\mathcal{L}_\mathcal{K}\text{-I}_\mathbb{N})$ and $(\mathcal{L}_\mathcal{K}\text{-I}_\mathbb{O})$ for full induction on the natural numbers and ordinals.

Next we introduce subsystems of $\mathsf{FID}(\mathcal{K})$ by weakening the principles of induction which are permitted. $\mathsf{FID}^\mathsf{w}(\mathcal{K})$ results from $\mathsf{FID}(\mathcal{K})$ by restricting induction on the ordinals to $(\Delta_0^\mathbb{O}\text{-I}_\mathbb{O})$, and $\mathsf{FID}^\mathsf{r}(\mathcal{K})$ is obtained from $\mathsf{FID}(\mathcal{K})$ by restricting induction

on the natural numbers and on the ordinals to $\Delta_0^{\mathbb{O}}$ formulas of $\mathcal{L}_\mathcal{K}$, i.e. to $(\Delta_0^{\mathbb{O}}\text{-}\mathsf{I}_\mathbb{N})$ and $(\Delta_0^{\mathbb{O}}\text{-}\mathsf{I}_\mathbb{O})$.

The operator axioms stated above are tailored according to the usual treatment of monotone or nonmonotone inductive definitions as described, for example, in Richter [26]. First one formalizes that the sets $P_\mathcal{A}^\alpha$ are the stages of the inductive definition generated by the operator form $\mathcal{A}(X, x)$; then one says that $P_\mathcal{A}^\infty$ is the set inductively defined by $\mathcal{A}(X, x)$.

The axioms (Op.2) are closure properties which implicitly require that there are sufficiently many ordinals in $\mathsf{FID}(\mathcal{K})$ and its subsystems so that the process of forming the stages of the inductive definitions with clauses from $\mathcal{K}$ comes to an end. If we put no restrictions on $\mathcal{K}$ this means asking for much (cf. e.g. Aczel and Richter [2] and Richter [26]). On the other hand, if $\mathcal{K}$ is the collection of all X positive operator forms $\mathcal{A}(X, x)$, then $\mathsf{FID}(\mathcal{K})$ is nothing but a variant of the well-known theory ID_1 (cf. e.g. Buchholz, Feferman, Pohlers and Sieg [6]) which explicitly mentions the stages of the inductive definitions.

3   Recursively inaccessible and Mahlo ordinals

Let Φ and Ψ be operators on the power set $P(\mathbb{N})$ of the natural numbers, i.e. mappings from $P(\mathbb{N})$ to $P(\mathbb{N})$. Following Richter [26], one can then define a new operator $[\Phi, \Psi]$ by setting for all subsets S of $\mathcal{N}$:

$$[\Phi, \Psi](S) := \begin{cases} \Phi(S), & \text{if } \Phi(S) \not\subset S, \\ \Psi(S), & \text{if } \Phi(S) \subset S. \end{cases}$$

Operators of this form are nonmonotone in general. In constructing the stages of the inductive definitions generated by $[\Phi, \Psi]$, one applies Φ until closure under Φ is reached; then there is one application of Ψ, and afterwards Φ is active again. This process is continued until one has closure under Φ and Ψ.

Now let $\mathcal{K}_1$ and $\mathcal{K}_2$ be two classes of unary operator forms. Then $[\mathcal{K}_1, \mathcal{K}_2]$ is the class of all operators $\mathcal{A}(X, x)$ defined as

$\mathcal{A}(X, x) :=$
$[\, \mathcal{A}_1(X, x) \wedge \neg \forall y (\mathcal{A}_1(X, y) \rightarrow X(y)) \,] \ \vee \ [\, \mathcal{A}_2(X, x) \wedge \forall y (\mathcal{A}_1(X, y) \rightarrow X(y)) \,]$

so that $\mathcal{A}_1(X, x)$ belongs to $\mathcal{K}_1$ and, additionally, $\mathcal{A}_2(X, x)$ to $\mathcal{K}_2$. This operator form $\mathcal{A}(X, x)$ is called a *combined operator form* with first component $\mathcal{A}_1(X, x)$ and second component $\mathcal{A}_2(X, x)$. Obviously, this definition follows the pattern of the combination of operators à la Richter.

Richter [26] studies the closure ordinals of several classes of combined operator forms, in particular those which have Π_1^0 definable components. For this case he exhibits the exact relationship to the first recursively Mahlo ordinal. According to Aczel [1] and Richter [26] we have:

(i) $cl([\Pi_0^0, \Pi_0^0]) = \omega^2$.

(ii) $cl([\Pi_1^0, \Pi_0^0]) =$ *the first recursively inaccessible ordinal.*

(iii) $cl([\Pi_1^0, \Pi_1^0]) =$ *the first recursively Mahlo ordinal.*

In connection with modeling Fefermans's theory T_0, also the class of operator forms $[\mathrm{POS}, \mathrm{QF}]$ is of interest. It is shown in Jäger and Studer [16] and Studer [27] that $\mathsf{FID}([\mathrm{POS}, \mathrm{QF}])$, $\mathsf{FID}^\mathsf{w}([\mathrm{POS}, \mathrm{QF}])$ and $\mathsf{FID}^\mathsf{r}([\mathrm{POS}, \mathrm{QF}])$ provide natural frameworks for interpreting T_0 and some of its restrictions.

Together with the results of the next section this implies that $\mathsf{FID}([\mathrm{POS}, \mathrm{QF}])$ is proof-theoretically equivalent to the theory of iterated admissible sets KPi, which formalizes a recursively inaccessible universe. Hence $\mathsf{FID}([\mathrm{POS}, \mathrm{QF}])$ is a further first order theory reflecting the idea of recursive inaccessibility.

4 Theories for admissible sets

Theories for (iterated) admissible sets are generally based on Kripke-Platek set theory KP, a famous subsystem of Zermelo-Fraenkel set theory ZF whose transitive standard models are the admissible sets. Prominent extension of KP are the theories KPi and KPm (cf. e.g. Jäger [13] and Rathjen [21]) which formalize that the respective universes of sets are recursively inaccessible and recursively Mahlo.

In this paper we are interested in theories for admissible sets since they provide a natural framework for dealing with several classes of nonmonotone inductive definitions in a very perspicuous way. For our purpose it is convenient, although not necessary, to work with Kripke-Platek set theories above the natural numbers as urelements. Then we have two forms of induction, namely induction on the natural numbers and $\in$ induction, which correspond exactly to induction on the naturals and induction on the ordinals in the theories for inductive definitions.

In the following we repeat more or less the formalization of theories for admissible sets as, for example, in Jäger [12, 13, 14] and refer to these publications for all unexplained notions, technical details and further reading. Accordingly, our theories for admissible sets are formulated in the extension $\mathcal{L}^* = \mathcal{L}(\in, \mathsf{N}, \mathsf{S}, \mathsf{Ad})$ of $\mathcal{L}$ by the membership relation symbol $\in$, the set constant N for the set of the natural numbers, the unary relation symbol S in order to express that an object is a set and the unary relation symbol Ad for stating that an object is an admissible set.

From now on we use x, y, z, f, g (possibly with subscripts) to range over the variables of $\mathcal{L}^*$. The terms $(a, b, c, a_0, b_0, c_0, \ldots)$ and formulas $(A, B, C, A_0, B_0, C_0, \ldots)$ of $\mathcal{L}^*$ as well as the Δ_0, Σ, Π, Σ_n and Π_n formulas of $\mathcal{L}^*$ are defined as usual. The notation $\vec{a}$ is shorthand for a finite string $a_1, \ldots, a_n$ whose length will be specified by the

context. Equality between objects is not represented by a primitive symbol but defined by

$$(a = b) := \begin{cases} (a \in \mathsf{N} \wedge b \in \mathsf{N} \wedge (a =_{\mathsf{N}} b)) \ \vee \\ (\mathsf{S}(a) \wedge \mathsf{S}(b) \wedge (\forall x \in a)(x \in b) \wedge (\forall x \in b)(x \in a)) \end{cases}$$

where $=_{\mathsf{N}}$ is the symbol for the primitive recursive equality on the natural numbers. The formula A^a is the result of replacing each unrestricted quantifier $\exists x(\ldots)$ and $\forall x(\ldots)$ in A by $(\exists x \in a)(\ldots)$ and $(\forall x \in a)(\ldots)$, respectively. In addition, we freely make use of all standard set-theoretic notations and write, for example $\mathsf{Tran}(a)$ for the Δ_0 formula saying that a is a transitive set.

In $\mathcal{L}^*$ one can easily formulate induction on the natural numbers and $\in$ induction, i.e. foundation. To this end let $\mathcal{F}$ be a collection of $\mathcal{L}^*$ formulas. As in the case of theories for inductive definitions we let induction on the natural numbers with respect to $\mathcal{F}$ consist of all formulas

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

so that $A(x)$ belongs to $\mathcal{F}$. Furthermore, $\in$ induction with respect to $\mathcal{F}$ consists of all formulas

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

with $A(x)$ in $\mathcal{F}$. In analogy to theories for inductive definitions we will confine ourselves to the induction schemas $(\Delta_0\text{-}\mathsf{I}_{\mathbb{N}})$, $(\Delta_0\text{-}\mathsf{I}_{\in})$, $(\mathcal{L}^*\text{-}\mathsf{I}_{\mathbb{N}})$ and $(\mathcal{L}^*\text{-}\mathsf{I}_{\in})$.

Now we introduce three main theories KPu, KPi and KPm for admissible sets which differ in strength of their set existence axioms. Their logical axioms comprise the usual axioms of classical first order logic with equality. The logical axioms of KPu can be divided into the following four groups.

I. Ontological axioms. We have for all terms a, b and $\vec{c}$ of $\mathcal{L}^*$, all function symbols h and relation symbols R of $\mathcal{L}$ and all axioms $A(\vec{x})$ of group III whose free variables belong to the list $\vec{x}$:

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

(2) $\vec{c} \in \mathsf{N} \to h(\vec{c}) \in \mathsf{N}$.

(3) $R(\vec{c}) \to \vec{c} \in \mathsf{N}$.

(4) $a \in b \to \mathsf{S}(b)$.

(5) $\mathsf{Ad}(a) \to (\mathsf{N} \in a \wedge \mathsf{Tran}(a))$.

(6) $\mathsf{Ad}(a) \to (\forall \vec{x} \in a)A^a(\vec{x})$.

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

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

III. Set-theoretic axioms. We have for all terms a and b and all Δ_0 formulas $A(x)$ and $B(x, y)$ of $\mathcal{L}^*$:

(Pair) $\exists x(a \in x \wedge b \in x)$.

(Transitive Hull) $\exists x(a \subset x \wedge \mathsf{Tran}(x))$.

(Δ_0 Separation) $\exists y(\mathsf{S}(y) \wedge y = \{x \in a : A(x)\})$.

(Δ_0 Collection) $(\forall x \in a)\exists y B(x, y) \rightarrow \exists z(\forall x \in a)(\exists y \in z)B(x, y)$.

IV. Induction axioms. These consist of the schemas $(\mathcal{L}^*\text{-}\mathsf{I}_{\mathbb{N}})$ and $(\mathcal{L}^*\text{-}\mathsf{I}_{\in})$ for full induction on the natural numbers and full $\in$ induction.

KPu corresponds to Barwise's theory KPU^+ described in [5] with PA as theory for the urelements. It says that its universe is an admissible set which contains the set of natural numbers as an element. In our axiomatization we include the axiom about the existence of transitive hulls rather the the more familiar axiom about the existence of union sets since we often work with very restricted forms of $\in$ induction.

The set theory KPi results from KPu by adding a further limit axiom which expresses that every set is contained in an admissible set,

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

Hence the transitive standard models of KPi are admissible limits of admissible sets, the so called recursively inaccessible sets. Finally, the theory KPm is KPu augmented by the schema of Π_2 reflection on the admissible sets,

(Mahlo) $\forall x \exists y A(x, y, \vec{a}) \rightarrow \exists z[\mathsf{Ad}(z) \wedge \vec{a} \in z \wedge (\forall x \in z)(\exists y \in z)A(x, y, \vec{a})]$

for all Δ_0 formulas $A(x, y, \vec{a})$ whose parameters belong to the list $x, y, \vec{a}$. The transitive standard models of KPm are the recursively Mahlo sets. Furthermore, it is easy to see that (Limit) follows from (Mahlo).

Let T be one of the systems KPu, KPi or KPm. Then T^{w} is obtained from T by restricting $(\mathcal{L}^*\text{-}\mathsf{I}_{\in})$ to $(\Delta_0\text{-}\mathsf{I}_{\in})$, and T^{r} is T with $(\mathcal{L}^*\text{-}\mathsf{I}_{\in})$ replaced by $(\Delta_0\text{-}\mathsf{I}_{\in})$ and $(\mathcal{L}^*\text{-}\mathsf{I}_{\mathbb{N}})$ replaced by $(\Delta_0\text{-}\mathsf{I}_{\mathbb{N}})$.

The proof-theoretic analysis of KPu, KPi and KPm and their just mentioned subsystems has been carried through several years ago and belongs to the general area *admissible proof theory*. Relevant articles and sources for further reading are, for example, Jäger [9, 10, 11, 13], Jäger and Pohlers [15], Pohlers [18, 19, 20] and Rathjen [21, 22, 23, 24]. Arai [3, 4] presents an alternative proof-theoretic approach to dealing with Mahlo universes.

5 Reductions to theories for admissible sets

Now we turn to some crucial connections between nonmonotone inductive definitions and admissible sets, and this analysis then leads to straightforward interpretations of several systems $\mathsf{FID}(\mathcal{K})$ into theories for admissible sets. We concentrate ourselves on classes of combined operator forms whose first component is from POS or Π_1^0 and whose second component is from QF, Π_0^0 or Π_1^0.

The ordinals of the theories for inductive definitions will be represented as the ordinals of the theories for admissible sets; the latter are defined in the language $\mathcal{L}^*$ by a Δ_0 formula $\mathsf{Ord}(x)$. We use α, β, γ (possibly with subscripts) to range over ordinals and write $\alpha < \beta$ for $\alpha \in \beta$.

Let $\mathcal{A}(X, x)$ be an arbitrary operator form. Then we want $\mathcal{H}_{\mathcal{A}}(\alpha, f)$ to express that f describes the iteration of $\mathcal{A}$ along α and set

$$
\mathcal{H}_{\mathcal{A}}(\alpha, f) \; := \; \left\{ \begin{array}{l} \mathsf{Fun}(f) \wedge \mathsf{dom}(f) = \alpha \; \wedge \\[2mm] (\forall \beta < \alpha) \, [\, f(\beta) = \{x \in \mathsf{N} : \mathcal{A}^{\mathsf{N}}(\bigcup_{\gamma < \beta} f(\gamma), x)\} \cup \bigcup_{\gamma < \beta} f(\gamma) \,] \,. \end{array} \right.
$$

This formula $\mathcal{H}_{\mathcal{A}}(\alpha, f)$ is used in a Σ definition of the stages of the inductive definition generated by $\mathcal{A}(X, x)$ which is given now. We are also interested in theories with $\in$ induction restricted to Δ_0 formulas so that Σ recursion over the full universe is not necessarily available. Locally in each admissible set, however, we can make use of Σ recursion, and this is sufficient for our inductive definitions.

$$
\begin{array}{llll}
I_{\mathcal{A}}^{\alpha}(x) & := & \exists f[\mathcal{H}_{\mathcal{A}}(\alpha+1, f) \wedge x \in f(\alpha)], & \qquad I_{\mathcal{A}}^{<\beta}(x) \;\; := \;\; (\exists \alpha < \beta) I_{\mathcal{A}}^{\alpha}(x), \\[2mm]
I_{\mathcal{A}}^{<a}(x) & := & (\exists \alpha \in a) I_{\mathcal{A}}^{\alpha}(x), & \qquad I_{\mathcal{A}}^{\infty}(x) \;\; := \;\; \exists \alpha \, I_{\mathcal{A}}^{\alpha}(x).
\end{array}
$$

The next lemma provides some auxiliary results which are independent of the specific form of the operator forms. Their proofs are straightforward and will be omitted. In the following we could often replace KPi^r by weaker theories; however, in the end we will be interested in KPi^r and some of its extensions so that there is no point in being more restrictive now.

Lemma 1 KPi^r *proves for all operator forms* $\mathcal{A}(X, x)$:

1. $\mathcal{H}_{\mathcal{A}}(\alpha, f) \wedge \mathcal{H}_{\mathcal{A}}(\beta, g) \wedge \alpha < \beta \; \rightarrow \; (\forall \gamma < \alpha)(f(\gamma) = g(\gamma))$.

2. $\mathsf{Ad}(a) \; \rightarrow \; (\forall \alpha \in a)(\exists f \in a)\mathcal{H}_{\mathcal{A}}(\alpha, f)$.

3. $\forall \alpha \exists f \mathcal{H}_{\mathcal{A}}(\alpha, f)$.

4. $\mathsf{Ad}(a) \wedge \alpha \in a \wedge b \in \mathsf{N} \; \rightarrow \; (I_{\mathcal{A}}^{\alpha}(b) \leftrightarrow (\exists f \in a)[\mathcal{H}_{\mathcal{A}}(\alpha+1, f) \wedge b \in f(\alpha)])$.

5. $a \in \mathsf{N} \; \rightarrow \; (I_{\mathcal{A}}^{\alpha}(a) \leftrightarrow (I_{\mathcal{A}}^{<\alpha}(a) \vee \mathcal{A}^{\mathsf{N}}(I_{\mathcal{A}}^{<\alpha}, a)))$.

After these preparatory observations concerning general operator forms, we now turn to the combined operator forms as introduced in Section 3. We begin with the following simple property.

Lemma 2 *Let $\mathcal{A}(X, x)$ be a combined operator form whose first component is the $\mathcal{L}(X)$ formula $\mathcal{B}(X, x)$. Then we have:*

$$\mathsf{KPi^r} \vdash a \in \mathsf{N} \wedge \mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<\alpha}, a) \ \rightarrow \ I_{\mathcal{A}}^{\alpha}(a).$$

PROOF. This assertion is shown by a distinction of cases which takes the specific behaviour of combined operator forms with respect to their components into account. We work informally in $\mathsf{KPi^r}$.

Case 1. $\neg(\forall y \in \mathsf{N})(\mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<\alpha}, y) \rightarrow I_{\mathcal{A}}^{<\alpha}(y))$. Then the first component $\mathcal{B}(X, x)$ of $\mathcal{A}(X, x)$ is active at stage α of the hierarchy generated by $\mathcal{A}^{\mathsf{N}}(X, x)$, yielding that $I_{\mathcal{A}}^{\alpha}(a)$.

Case 2. $(\forall y \in \mathsf{N})(\mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<\alpha}, y) \rightarrow I_{\mathcal{A}}^{<\alpha}(y))$. Then we are immediately done since $I_{\mathcal{A}}^{<\alpha}$ is contained in $I_{\mathcal{A}}^{\alpha}$. $\qquad\square$

Combined operator forms whose first components belong to POS have the expected closure properties with respect to their first components. They are listed in the following lemma.

Lemma 3 *Let $\mathcal{A}(X, x)$ be a combined operator form whose first component is the X positive $\mathcal{L}(X)$ formula $\mathcal{B}(X, x)$. Then we have:*

1. $\mathsf{KPi^r} \vdash \mathsf{Ad}(a) \wedge b \in \mathsf{N} \wedge \mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<a}, b) \ \rightarrow \ I_{\mathcal{A}}^{<a}(b).$

2. $\mathsf{KPi^r} \vdash b \in \mathsf{N} \wedge \mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{\infty}, b) \ \rightarrow \ I_{\mathcal{A}}^{\infty}(b).$

PROOF. In view of our assumption about the operator form $\mathcal{A}(X, x)$ we know that it is of the form

$$[\, \mathcal{B}(X, x) \wedge \neg \forall y(\mathcal{B}(X, y) \rightarrow X(y)) \,] \ \vee \ [\, \mathcal{C}(X, x) \wedge \forall y(\mathcal{B}(X, y) \rightarrow X(y)) \,]$$

for some $\mathcal{L}(X)$ formula $\mathcal{C}(X, x)$. Now we work informally in $\mathsf{KPi^r}$ and assume that we have $\mathsf{Ad}(a)$, $b \in \mathsf{N}$ and $\mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<a}, b)$. Because of Lemma 1 this implies that

$$\mathcal{B}^{\mathsf{N}}(\lambda y.C^a(y), b)$$

for $C(y)$ being the Σ formula $\exists \alpha I_{\mathcal{A}}^{\alpha}(y)$. Since $\mathcal{B}(X, x)$ is X positive we can apply Σ reflection within a and obtain, after some intermediate steps, that there exists an ordinal β in a so that $\mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, b)$. Hence it follows from the previous lemma that $I_{\mathcal{A}}^{\beta}(b)$, thus $I_{\mathcal{A}}^{<a}(b)$. This completes the proof of our first assertion; the second assertion can be treated analogously. $\qquad\square$

The next lemma is a preparatory step for dealing with the second components of operator forms belonging to operator classes of the form $[\mathcal{K}, \mathrm{QF}]$ or $[\mathcal{K}, \Pi_0^0]$. It is tailored for handling negative occurrences of X in QF and Π_0^0 formulas of $\mathcal{L}(X)$.

Lemma 4 *Let $\mathcal{A}(X, x)$ be an operator form, $B(X, \vec{z})$ a QF formula and $C(X, \vec{z})$ a Π_0^0 formula of $\mathcal{L}(X)$ and assume that all free variables of $B(X, \vec{z})$ and $C(X, \vec{z})$ are from the list $\vec{z}$. Then we have:*

1. $\mathsf{KPi^r} \vdash \vec{a} \in \mathsf{N} \wedge B^{\mathsf{N}}(I_{\mathcal{A}}^{\infty}, \vec{a}) \ \rightarrow \ \exists\alpha\forall\beta(\alpha \leq \beta \rightarrow B^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, \vec{a}))$.

2. $\mathsf{KPi^w} \vdash \vec{a} \in \mathsf{N} \wedge C^{\mathsf{N}}(I_{\mathcal{A}}^{\infty}, \vec{a}) \ \rightarrow \ \exists\alpha\forall\beta(\alpha \leq \beta \rightarrow C^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, \vec{a}))$.

PROOF. Without loss of generality we can assume that $B(X, \vec{z})$ and $C(X, \vec{z})$ are in negation normal form; i.e. all negation symbols are pushed inside the formula as far as possible, and consecutive negation symbols of even number are removed afterwards. We begin with proving the first assertion by induction on the length of the formula $B(X, \vec{z})$.

(i) If X does not occur in $B(X, \vec{z})$, then the assertion is trivial. If $B(X, \vec{z})$ is of the form $X(t)$ or $\neg X(t)$, then the assertion follows from the definition of the formulas $I_{\mathcal{A}}^{\infty}(t)$ and $I_{\mathcal{A}}^{<\beta}(t)$.

(ii) If $B(X, \vec{z})$ is a formula $(B_0(X, \vec{z}) \vee B_1(X, \vec{z}))$, then the assertion follows immediately from the induction hypothesis.

(iii) If $B(X, \vec{z})$ is a formula $(B_0(X, \vec{z}) \wedge B_1(X, \vec{z}))$, then the induction hypothesis implies that

$$B^{\mathsf{N}}(I_{\mathcal{A}}^{\infty}, \vec{a}) \ \rightarrow \ [\exists\alpha_0\forall\beta(\alpha_0 \leq \beta \rightarrow B_0^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, \vec{a})) \wedge \exists\alpha_1\forall\beta(\alpha_1 \leq \beta \rightarrow B_1^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, \vec{a}))].$$

Choosing the larger witness for the two existential quantifiers yields the assertion.

Since $B(X, \vec{z})$ is quantifier-free, the proof of the first assertion is completed. The second assertion is proved by following the same pattern but with two additional cases for dealing with the bounded numerical quantifiers.

(iv) If $C(X, \vec{z})$ begins with a bounded existential numerical quantifier, we obtain the assertion by a simple application of the induction hypothesis.

(v) If $C(X, \vec{z})$ is of the form $(\forall x <_N t)C_0(X, \vec{z}, x)$ for some number term t, the induction hypothesis implies

$$\mathsf{KPi^w} \vdash \vec{a} \in \mathsf{N} \wedge C^{\mathsf{N}}(I_{\mathcal{A}}^{\infty}, \vec{a}) \ \rightarrow \ (\forall x <_N t)\exists\alpha\forall\beta(\alpha \leq \beta \rightarrow C_0^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, \vec{a}, x)). \quad (1)$$

At this stage we insert an auxiliary consideration. We write $A(x, \alpha)$ for the formula $\forall\beta(\alpha \leq \beta \rightarrow C_0^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, \vec{a}, x)$ and obtain by complete induction on the natural numbers that

$$\mathsf{KPi^w} \vdash (\forall y \in \mathsf{N})[(\forall x <_N y)\exists\alpha A(x, \alpha) \rightarrow \exists\alpha(\forall x <_N y)A(x, \alpha)]. \quad (2)$$

From (1) and (2) we conclude

$$\mathsf{KPi^w} \vdash \vec{a} \in \mathsf{N} \wedge C^{\mathsf{N}}(I_{\mathcal{A}}^{\infty}, \vec{a}) \ \rightarrow \ \exists\alpha\forall\beta(\alpha \leq \beta \rightarrow (\forall x <_N t)C_0^{\mathsf{N}}(I_{\mathcal{A}}^{<\beta}, \vec{a}, x)).$$

Hence also bounded universal numerical quantifiers can be handled, and therefore the second assertion is proved as well. $\square$

11

The proof of the second assertion of the previous lemma indicates that a bit more than Δ_0 induction on the natural numbers is necessary for treating bounded numerical quantifiers. Thus it cannot be carried through in $\mathsf{KPi^r}$, but $\mathsf{KPi^w}$ is amply sufficient.

Theorem 5 *Let $\mathcal{A}(X,x)$ be an operator form from $[\mathrm{POS},\mathrm{QF}]$ and $\mathcal{B}(X,x)$ an operator form from $[\mathrm{POS},\Pi_0^0]$. Then we have:*

1. $\mathsf{KPi^r} \vdash a \in \mathsf{N} \wedge \mathcal{A}^\mathsf{N}(I_\mathcal{A}^\infty, a) \rightarrow I_\mathcal{A}^\infty(a)$.

2. $\mathsf{KPi^w} \vdash a \in \mathsf{N} \wedge \mathcal{B}^\mathsf{N}(I_\mathcal{B}^\infty, a) \rightarrow I_\mathcal{B}^\infty(a)$.

PROOF. Since the operator form $\mathcal{A}(X,x)$ belongs to the class $[\mathrm{POS},\mathrm{QF}]$, it has to be of the form

$$[\,\mathcal{C}(X,x) \wedge \neg\forall y(\mathcal{C}(X,y) \rightarrow X(y))\,] \vee [\,\mathcal{D}(X,x) \wedge \forall y(\mathcal{C}(X,y) \rightarrow X(y))\,]$$

with $\mathcal{C}(X,x)$ an X positive and $\mathcal{D}(X,x)$ a quantifier-free formula of $\mathcal{L}(X)$. Now we work informally in $\mathsf{KPi^r}$ and only have to show

$$\mathcal{C}^\mathsf{N}(I_\mathcal{A}^\infty, a) \rightarrow I_\mathcal{A}^\infty(a) \qquad \text{and} \qquad \mathcal{D}^\mathsf{N}(I_\mathcal{A}^\infty, a) \rightarrow I_\mathcal{A}^\infty(a)$$

for all $a \in \mathsf{N}$ in order to establish the first part of our theorem. However, the left implication follows immediately from Lemma 3. Now suppose that $\mathcal{D}^\mathsf{N}(I_\mathcal{A}^\infty, a)$. Then the first assertion of Lemma 4 implies that there exists an ordinal α so that

$$\forall\beta(\alpha \leq \beta \rightarrow \mathcal{D}^\mathsf{N}(I_\mathcal{A}^{<\beta}, a)).$$

We use the axiom (Limit) of $\mathsf{KPi^r}$ to find an admissible set b which contains α as an element. It follows that $\{x \in b : \mathsf{Ord}(x)\}$ is a set in $\mathsf{KPi^r}$, namely an ordinal γ so that $\alpha < \gamma$. Hence we have $\mathcal{D}^\mathsf{N}(I_\mathcal{A}^{<\gamma}, a)$; in addition, Lemma 3 yields

$$(\forall x \in \mathsf{N})(\mathcal{C}^\mathsf{N}(I_\mathcal{A}^{<\gamma}, x) \rightarrow I_\mathcal{A}^{<\gamma}(x)).$$

This implies that $\mathcal{D}^\mathsf{N}(X,x)$ has to be active in defining the γth stage $I_\mathcal{A}^\gamma$ of our hierarchy. Because of $\mathcal{D}^\mathsf{N}(I_\mathcal{A}^{<\gamma}, a)$, we therefore obtain $I_\mathcal{A}^\gamma(a)$ and consequently $I_\mathcal{A}^\infty(a)$.

The second assertion of this theorem is proved analogously; we only have to use the second assertion of Lemma 4 instead of its first. $\qquad\square$

Our next considerations refer to combined operator forms from the classes $[\Pi_1^0, \mathrm{QF}]$ and $[\Pi_1^0, \Pi_0^0]$. In contrast to the previous cases, we now have to deal with nonmonotonicity already in the first components of these operator forms. For this purpose we introduce a suitable form of Π_2 reflection in our theories for admissibles. Global Π_2 reflection is the schema

$(\Pi_2 \; \mathsf{GRef})$ $\forall x \exists y A(x,y,\vec{a}) \rightarrow \exists z[\mathsf{Tran}(z) \wedge \vec{a} \in z \wedge (\forall x \in z)(\exists y \in z)A(x,y,\vec{a})]$

for all Δ_0 formulas $A(x, y, \vec{a})$ whose parameters belong to the list $x, y, \vec{a}$. This is global Π_2 reflection since it refers to the full universe of sets. Of course it also makes sense to claim that each admissible is closed under Π_2 reflection. This we call the schema of local Π_2 reflection,

$(\Pi_2 \; \mathsf{LRef}) \quad \mathsf{Ad}(a) \; \rightarrow \; (\forall \vec{x} \in a) A^a(\vec{x})$

for all instances $A(\vec{x})$ of $(\Pi_2 \; \mathsf{GRef})$ whose free variables belong to the list $\vec{x}$. Finally, $(\Pi_2 \; \mathsf{Ref})$ is the schema comprising all instances of $(\Pi_2 \; \mathsf{GRef})$ and $(\Pi_2 \; \mathsf{LRef})$. Although very specific forms of Π_2 reflection would be sufficient for our later purposes, we decided to work with this general form since thus further notation can be avoided and the proof-theoretic bounds are not affected.

Namely, if T is the theory KPi or KPm, then it is not difficult to see that $\mathsf{T^w} + (\Pi_2 \; \mathsf{Ref})$ and $\mathsf{T^w}$ as well as $\mathsf{T} + (\Pi_2 \; \mathsf{Ref})$ and T are pairwise proof-theoretically equivalent. Moreover, by adapting an argument of Barwise [5], one can check that subforms of Π_2 reflection, for example Π_2 reflection for ordinals, are provable in $\mathsf{T^w}$. We also want to mention that $\mathsf{T^r} + (\Pi_2 \; \mathsf{Ref})$ is slightly stronger than $\mathsf{T^r}$. A detailed proof-theoretic analysis of Π_2 reflection in theories for admissible sets with weak forms of induction will be given elsewhere.

In the following we want to distinguish between the positive and negative occurrences of the relation symbol X in $\mathcal{L}(X)$ formulas. This can be conveniently achieved by choosing a fresh unary relation symbol Y and working in the extension $\mathcal{L}(X, Y)$ of $\mathcal{L}(X)$. Then for each $\mathcal{L}(X)$ formula $A(X)$ there exists an $\mathcal{L}(X, Y)$ formula $B(X, Y)$ which is positive in X and Y so that $A(X)$ is logically equivalent to $B(X, \lambda y. \neg X(y))$. The following lemma is needed for the proof of Lemma 7 below.

Lemma 6 *Let $\mathcal{A}(X, x)$ be an operator form and $B(X, Y, \vec{z})$ a Π_1^0 formula of $\mathcal{L}(X, Y)$ which is positive in X and Y; assume also that all free variables of $B(X, Y, \vec{z})$ are from the list $\vec{z}$. Then we have:*

$\mathsf{KPi^r} \vdash \vec{a} \in \mathsf{N} \wedge \gamma \neq 0 \wedge (\forall \alpha < \gamma)(\exists \beta < \gamma) B^\mathsf{N}(I_\mathcal{A}^\beta, \mathsf{N} \setminus I_\mathcal{A}^\alpha, \vec{a}) \; \rightarrow \; B^\mathsf{N}(I_\mathcal{A}^{<\gamma}, \mathsf{N} \setminus I_\mathcal{A}^{<\gamma}, \vec{a}).$

PROOF. By exploiting the X positivity of $B(X, Y, \vec{z})$, a simple persistency argument immediately shows that $\mathsf{KPi^r}$ proves the following implication:

$\vec{a} \in \mathsf{N} \wedge \gamma \neq 0 \wedge (\forall \alpha < \gamma)(\exists \beta < \gamma) B^\mathsf{N}(I_\mathcal{A}^\beta, \mathsf{N} \setminus I_\mathcal{A}^\alpha, \vec{a}) \; \rightarrow \; (\forall \alpha < \gamma) B^\mathsf{N}(I_\mathcal{A}^{<\gamma}, \mathsf{N} \setminus I_\mathcal{A}^\alpha, \vec{a}).$

Hence it only remains to be shown in $\mathsf{KPi^r}$ that $(\forall \alpha < \gamma) B^\mathsf{N}(I_\mathcal{A}^{<\gamma}, \mathsf{N} \setminus I_\mathcal{A}^\alpha, \vec{a})$ implies $B^\mathsf{N}(I_\mathcal{A}^{<\gamma}, \mathsf{N} \setminus I_\mathcal{A}^{<\gamma}, \vec{a})$ under the given assumptions. As in the proof of Lemma 4 we assume without loss of generality that $B(X, Y, \vec{z})$ is in negation normal form and proceed by induction on the length of $B(X, Y, \vec{z})$.

(i) If Y does not occur in $B(X, Y, \vec{z})$, then the assertion is trivial. If $B(X, Y, \vec{z})$ is of the form $Y(t)$, then the assertion follows from the definition of the formula $I_\mathcal{A}^{<\gamma}(t)$.

(ii) If $B(X, Y, \vec{z})$ is a disjunction, a conjunction or if it begins with a bounded or unbounded universal numerical quantifier, then the assertion follows immediately from the induction hypothesis and the Y positivity of $B(X, Y, \vec{z})$.

(iii) If $B(X, Y, \vec{z})$ is of the form $(\exists x <_N t) B_0(X, Y, \vec{z}, x)$ for some number term t, then the induction hypothesis implies that

$$\vec{a} \in \mathsf{N} \wedge \gamma \neq 0 \wedge (\exists x <_N t)(\forall \alpha < \gamma) B_0^{\mathsf{N}}((I_{\mathcal{A}}^{<\gamma}, \mathsf{N} \setminus I_{\mathcal{A}}^{\alpha}, \vec{a}, x) \ \rightarrow \ B^{\mathsf{N}}(I_{\mathcal{A}}^{<\gamma}, \mathsf{N} \setminus I_{\mathcal{A}}^{<\gamma}, \vec{a})$$

is provable in $\mathsf{KPi^r}$. Now we proceed similar to the proof of Lemma 4 and use complete induction on the natural numbers to exchange the universal and existential quantifiers in the left hand side of the previous formula. However, $(\Delta_0\text{-}\mathsf{I_N})$ is sufficient, and therefore also this argument can be carried through in $\mathsf{KPi^r}$. $\square$

Lemma 7 *Let $\mathcal{A}(X, x)$ be a combined operator form whose first component is the Π_1^0 formula $\mathcal{B}(X, x)$ of $\mathcal{L}(X)$. Then we have:*

1. $\mathsf{KPi^r} + (\Pi_2\ \mathsf{LRef}) \vdash \mathsf{Ad}(a) \wedge b \in \mathsf{N} \wedge \mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<a}, b) \ \rightarrow \ I_{\mathcal{A}}^{<a}(b)$.

2. $\mathsf{KPi^r} + (\Pi_2\ \mathsf{GRef}) \vdash b \in \mathsf{N} \wedge \mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{\infty}, b) \ \rightarrow \ I_{\mathcal{A}}^{\infty}(b)$.

PROOF. One only has to deal with the first assertion; the second can be proved analogously. Hence we work informally in $\mathsf{KPi^r}$ and assume that we have $\mathsf{Ad}(a)$, $b \in \mathsf{N}$ and $\mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<a}, b)$. First we choose a Π_1^0 formula $C(X, Y, x)$ of $\mathcal{L}(X, Y)$ which is positive in X and Y so that $\mathcal{B}(X, x)$ is logically equivalent to $C(X, \lambda y.\neg X(y), x)$. Then we have $C^{\mathsf{N}}(I_{\mathcal{A}}^{<a}, \mathsf{N} \setminus I_{\mathcal{A}}^{<a}, b)$ so that

$$(\forall \alpha \in a) C^{\mathsf{N}}(I_{\mathcal{A}}^{<a}, \mathsf{N} \setminus I_{\mathcal{A}}^{\alpha}, b)$$

follows because of persistency. Therefore Σ reflection within the admissible set a implies that

$$(\forall \alpha \in a)(\exists \beta \in a) C^{\mathsf{N}}(I_{\mathcal{A}}^{\beta}, \mathsf{N} \setminus I_{\mathcal{A}}^{\alpha}, b).$$

Now we are in the position to apply local Π_2 reflection within a and conclude that a contains an ordinal γ, which can be chosen to be greater than 0, so that

$$(\forall \alpha < \gamma)(\exists \beta < \gamma) C^{\mathsf{N}}(I_{\mathcal{A}}^{\beta}, \mathsf{N} \setminus I_{\mathcal{A}}^{\alpha}, b).$$

Thus it follows from Lemma 6 that $C^{\mathsf{N}}(I_{\mathcal{A}}^{<\gamma}, \mathsf{N} \setminus I_{\mathcal{A}}^{<\gamma}, b)$, i.e. $\mathcal{B}^{\mathsf{N}}(I_{\mathcal{A}}^{<\gamma}, b)$. In view of Lemma 2 this means that $I_{\mathcal{A}}^{\gamma}(b)$, hence $I_{\mathcal{A}}^{<a}(b)$ since γ is an element of a. $\square$

This lemma shows in which way Π_2 reflection can be used in order to establish the closure property of combined operator forms with respect to their first components if these are Π_1^0 formulas of $\mathcal{L}(X)$ with possibly negative occurrences of X. The treatment of operator forms from $[\Pi_1^0, \mathrm{QF}]$ and $[\Pi_1^0, \Pi_0^0]$ in the theories $\mathsf{KPi^r} + (\Pi_2\ \mathsf{Ref})$ and $\mathsf{KPi^w} + (\Pi_2\ \mathsf{Ref})$, respectively, is now straightforward.

Theorem 8 *Let $\mathcal{A}(X, x)$ be an operator form from $[\Pi_1^0, \mathrm{QF}]$ and $\mathcal{B}(X, x)$ an operator form from $[\Pi_1^0, \Pi_0^0]$. Then we have:*

1. $\mathsf{KPi^r} + (\Pi_2 \text{ Ref}) \vdash a \in \mathsf{N} \land \mathcal{A}^{\mathsf{N}}(I^{\infty}_{\mathcal{A}}, a) \rightarrow I^{\infty}_{\mathcal{A}}(a).$

2. $\mathsf{KPi^w} + (\Pi_2 \text{ Ref}) \vdash a \in \mathsf{N} \land \mathcal{B}^{\mathsf{N}}(I^{\infty}_{\mathcal{B}}, a) \rightarrow I^{\infty}_{\mathcal{B}}(a).$

PROOF. We obtain a proof of this theorem from our proof of Theorem 5 simply by using Lemma 7 instead of Lemma 3. □

So far this section was concerned with the analysis of nonmonotone inductive definitions for the operator classes $[\text{POS}, \text{QF}]$, $[\text{POS}, \Pi^0_0]$, $[\Pi^0_1, \text{QF}]$ and $[\Pi^0_1, \Pi^0_0]$. Now we make a comparatively big step and turn to the classes $[\text{POS}, \Pi^0_1]$ and $[\Pi^0_1, \Pi^0_1]$. On the side of the theories for admissibles this means that we have to work within theories for a recursively Mahlo universe. The following lemma is crucial for dealing with combined operator forms whose second component is from Π^0_1.

Lemma 9 *Let $\mathcal{A}(X, x)$ be a combined operator and $B(X, x)$ a Π^0_1 formula of $\mathcal{L}(X)$. Then we have:*

$$\mathsf{KPm^r} \vdash a \in \mathsf{N} \land B^{\mathsf{N}}(I^{\infty}_{\mathcal{A}}, a) \rightarrow \exists x[\mathsf{Ad}(x) \land B^{\mathsf{N}}(I^{<x}_{\mathcal{A}}, a)].$$

PROOF. We work informally in $\mathsf{KPm^r}$ and assume that $a \in \mathsf{N}$ and $B^{\mathsf{N}}(I^{\infty}_{\mathcal{A}}, a)$. As in the proof of Lemma 7 we then choose a Π^0_1 formula $C(X, Y, x)$ of $\mathcal{L}(X, Y)$ which is positive in X and Y so that $B(X, x)$ is logically equivalent to $C(X, \lambda y.\neg X(y), x)$. Then we have $C^{\mathsf{N}}(I^{\infty}_{\mathcal{A}}, \mathsf{N} \setminus I^{\infty}_{\mathcal{A}}, a)$ and obtain, similar to the proof of Lemma 7, by making use of persistency and Σ reflection that

$$\forall \alpha \exists \beta C^{\mathsf{N}}(I^{\beta}_{\mathcal{A}}, \mathsf{N} \setminus I^{\alpha}_{\mathcal{A}}, a).$$

Now the Mahlo axiom of $\mathsf{KPm^r}$ comes into action and provides an admissible set b so that

$$(\forall \alpha \in b)(\exists \beta \in b) C^{\mathsf{N}}(I^{\beta}_{\mathcal{A}}, \mathsf{N} \setminus I^{\alpha}_{\mathcal{A}}, a).$$

Essentially by Lemma 6 we thus obtain $C^{\mathsf{N}}(I^{<b}_{\mathcal{A}}, \mathsf{N} \setminus I^{<b}_{\mathcal{A}}, a)$, i.e. $B^{\mathsf{N}}(I^{<b}_{\mathcal{A}}, a)$. Hence our lemma is proved. □

With this lemma and earlier results about the closure properties concerning the first components of combined operators we can now treat all operators belonging to the operator classes $[\text{POS}, \Pi^0_1]$ and $[\Pi^0_1, \Pi^0_1]$. Local Π_2 reflection is used for dealing with first components which are Π^0_1.

Theorem 10 *Let $\mathcal{A}(X, x)$ be an operator form from $[\text{POS}, \Pi^0_1]$ and $\mathcal{B}(X, x)$ an operator form from $[\Pi^0_1, \Pi^0_1]$. Then we have:*

1. $\mathsf{KPm^r} \vdash a \in \mathsf{N} \land \mathcal{A}^{\mathsf{N}}(I^{\infty}_{\mathcal{A}}, a) \rightarrow I^{\infty}_{\mathcal{A}}(a).$

2. $\mathsf{KPm^r} + (\Pi_2 \text{ LRef}) \vdash a \in \mathsf{N} \land \mathcal{B}^{\mathsf{N}}(I^{\infty}_{\mathcal{B}}, a) \rightarrow I^{\infty}_{\mathcal{B}}(a).$

PROOF. The first assertion follows from Lemma 3 and Lemma 9. For proving the second assertion, we first recall that the operator form $\mathcal{B}(X, x)$ is of the form

$$[\, \mathcal{C}(X, x) \wedge \neg \forall y (\mathcal{C}(X, y) \rightarrow X(y)) \,] \ \vee \ [\, \mathcal{D}(X, x) \wedge \forall y (\mathcal{C}(X, y) \rightarrow X(y)) \,]$$

with $\mathcal{C}(X, x)$ and $\mathcal{D}(X, x)$ Π_1^0 formulas of $\mathcal{L}(X)$. Following the pattern of the proof of Theorem 5, we show

$$\mathcal{C}^{\mathsf{N}}(I_{\mathcal{B}}^{\infty}, a) \rightarrow I_{\mathcal{B}}^{\infty}(a) \qquad \text{and} \qquad \mathcal{D}^{\mathsf{N}}(I_{\mathcal{B}}^{\infty}, a) \rightarrow I_{\mathcal{B}}^{\infty}(a)$$

for all $a \in \mathsf{N}$, now working informally in $\mathsf{KPm^r} + (\Pi_2\ \mathsf{LRef})$. Because of Lemma 9, $\mathcal{C}^{\mathsf{N}}(I_{\mathcal{B}}^{\infty}, a)$ implies $\mathcal{C}^{\mathsf{N}}(I_{\mathcal{B}}^{<b}, a)$ for some admissible set b and thus $I_{\mathcal{B}}^{\infty}(a)$ by Lemma 2. Hence $I_{\mathcal{B}}^{\infty}$ is closed under $\mathcal{C}(X, x)$.

Next suppose that $\mathcal{D}^{\mathsf{N}}(I_{\mathcal{B}}^{\infty}, a)$. Then Lemma 9 yields $\mathcal{D}^{\mathsf{N}}(I_{\mathcal{B}}^{<c}, a)$ for some admissible set c. Now we make use of Lemma 7 to conclude that

$$(\forall x \in \mathsf{N})(\mathcal{C}^{\mathsf{N}}(I_{\mathcal{B}}^{<c}, x) \rightarrow I_{\mathcal{B}}^{<c}(x)),$$

i.e. $I_{\mathcal{B}}^{<c}$ is closed under $\mathcal{C}(X, x)$. Hence $\mathcal{D}(X, x)$ has to be active at stage $\gamma := c \cap \mathsf{Ord}$ of our inductive definition. Because of $\mathcal{D}^{\mathsf{N}}(I_{\mathcal{B}}^{<\gamma}, a)$, we therefore obtain $I_{\mathcal{B}}^{\gamma}(a)$ and consequently $I_{\mathcal{B}}^{\infty}(a)$. $\qquad\square$

6 Proof-theoretic results

The previous section provides the crucial material for embedding our first order theories for nonmonotone inductive definitions into suitable systems for iterated admissible sets. These embeddings together with some other results about modeling systems of explicit mathematics in theories for nonmonotone inductive definitions are sufficient for determining the exact proof-theoretic bounds of those.

Let $\mathcal{K}$ be an arbitrary class of operator forms. Then the interpretation of the language $\mathcal{L}_{\mathcal{K}}$ of the theory $\mathsf{FID}(\mathcal{K})$ into the language $\mathcal{L}^*$ should be obvious. Both, $\mathcal{L}_{\mathcal{K}}$ and $\mathcal{L}^*$ extend the language $\mathcal{L}$ of first order arithmetic; the ordinals of $\mathcal{L}_{\mathcal{K}}$ are interpreted as the ordinals of $\mathcal{L}^*$, the less relation on the ordinal of $\mathcal{L}_{\mathcal{K}}$ goes over into the less relation on the ordinals of $\mathcal{L}^*$ and the atomic formulas $P_{\mathcal{A}}^{\alpha}(r)$ of $\mathcal{L}_{\mathcal{K}}$ are translated into the formulas $I_{\mathcal{A}}^{\alpha}(r)$ of $\mathcal{L}^*$. Now we describe this interpretation in more detail

We first assume that the number and ordinal variables of $\mathcal{L}_{\mathcal{K}}$ are mapped into the variables of $\mathcal{L}^*$ so that no conflicts arise; if we want to be very precise, we write $\hat{u}$ and $\hat{\alpha}$ for the variables of $\mathcal{L}^*$ corresponding to the number variable u and ordinal variable α of $\mathcal{L}_{\mathcal{K}}$. Given an $\mathcal{L}_{\mathcal{K}}$ formula A, the $\mathcal{L}^*$ formula A' is then obtained as follows: (i) All variables are replaced by their translations into variables of $\mathcal{L}^*$; (ii) all $\mathcal{L}_{\mathcal{K}}$ formulas of the form $(\alpha < \beta)$ are replaced by the $\mathcal{L}^*$ formulas $(\hat{\alpha} < \hat{\beta})$; (iii) all $\mathcal{L}_{\mathcal{K}}$ formulas of the form $P_{\mathcal{A}}^{\alpha}(r)$ are replaced by the $\mathcal{L}^*$ formulas $I_{\mathcal{A}}^{\hat{\alpha}}(r)$; (iv) all numerical

quantifiers $\exists x(\ldots)$ and $\forall x(\ldots)$ are replaced by $(\exists \hat{x} \in \mathsf{N})(\ldots)$ and $(\forall \hat{x} \in \mathsf{N})(\ldots)$; (v) all ordinal quantifiers $\exists \alpha(\ldots)$ and $\forall \alpha(\ldots)$ are replaced by $\exists \hat{\alpha}(\mathsf{Ord}(\hat{\alpha}) \wedge \ldots)$ and $\forall \hat{\alpha}(\mathsf{Ord}(\hat{\alpha}) \to \ldots)$.

Finally let A be an $\mathcal{L}_{\mathcal{K}}$ formula whose free number and ordinal variables belong to the lists $x_1, \ldots, x_m$ and $\alpha_1, \ldots, \alpha_n$, respectively. Then we associate to A the $\mathcal{L}^*$ formula F_A,

$$F_A := (\hat{x}_1 \in \mathsf{N} \wedge \ldots \wedge \hat{x}_m \in \mathsf{N} \wedge \mathsf{Ord}(\hat{\alpha}_1) \wedge \ldots \wedge \mathsf{Ord}(\hat{\alpha}_n) \to A').$$

Considering Lemma 1 once more, it is easy to check that in $\mathsf{KPi}^{\mathsf{r}}$ the $\Delta_0^{\mathbb{O}}$ formulas of $\mathcal{L}$, modulo this translation, can be treated as Δ_0 formulas of $\mathcal{L}^*$, provided that we have an admissible bound of their ordinal parameters,

$$\mathsf{KPi}^{\mathsf{r}} \vdash \mathsf{Ad}(a) \wedge \hat{\alpha}_1 \in a \wedge \ldots \wedge \hat{\alpha}_n \in a \to (F_A \leftrightarrow F_A^a).$$

Based on this translation, we say that $\mathsf{FID}(\mathcal{K})$ is contained in the $\mathcal{L}^*$ theory T if we have $\mathsf{T} \vdash F_A$ for every axiom A of $\mathsf{FID}(\mathcal{K})$.

If T is one of the theories KPi or KPm, then it is evident by the considerations of the previous section that T^{r} proves F_A for all number-theoretic axioms, linearity axioms and $(\mathsf{Op.1})$ axioms A of $\mathsf{FID}(\mathcal{K})$. For dealing with the induction principles of our theories for inductive definitions, we only have to observe that (the translations of) all instances of $(\Delta_0^{\mathbb{O}}\text{-}\mathsf{I}_{\mathbb{N}})$ and $(\Delta_0^{\mathbb{O}}\text{-}\mathsf{I}_{\mathbb{O}})$ are available in T^{r}. For proving all instances of $(\mathcal{L}_{\mathcal{K}}\text{-}\mathsf{I}_{\mathbb{N}})$ and $(\mathcal{L}_{\mathcal{K}}\text{-}\mathsf{I}_{\mathbb{O}})$, on the other hand, the set theories require the schemes $(\mathcal{L}^*\text{-}\mathsf{I}_{\mathbb{N}})$ and $(\mathcal{L}^*\text{-}\mathsf{I}_{\in})$, respectively.

These observations make it clear that, with exception of the closure axioms $(\mathsf{Op.2})$, all axioms of the theories $\mathsf{FID}(\mathcal{K})$ and their restrictions create no problems when interpreted into theories of admissibles. But in view of Theorem 5 we also know how to deal with the closure properties of $[\mathrm{POS}, \mathrm{QF}]$ and $[\mathrm{POS}, \Pi_0^0]$ operator forms.

Theorem 11   *1.* $\mathsf{FID}^{\mathsf{r}}([\mathrm{POS}, \mathrm{QF}])$ *is contained in* $\mathsf{KPi}^{\mathsf{r}}$.

  2. $\mathsf{FID}^{\mathsf{w}}([\mathrm{POS}, \mathrm{QF}])$ *and* $\mathsf{FID}^{\mathsf{w}}([\mathrm{POS}, \Pi_0^0])$ *are contained in* $\mathsf{KPi}^{\mathsf{w}}$.

  3. $\mathsf{FID}([\mathrm{POS}, \mathrm{QF}])$ *and* $\mathsf{FID}([\mathrm{POS}, \Pi_0^0])$ *are contained in* KPi.

The closure properties of combined operator forms with first Π_1^0 components require some form of Π_2 reflection. According to Theorem 8, we thus have the following embedding theorem.

Theorem 12   *1.* $\mathsf{FID}^{\mathsf{r}}([\Pi_1^0, \mathrm{QF}])$ *is contained in* $\mathsf{KPi}^{\mathsf{r}} + (\Pi_2\ \mathsf{Ref})$.

  2. $\mathsf{FID}^{\mathsf{w}}([\Pi_1^0, \mathrm{QF}])$ *and* $\mathsf{FID}^{\mathsf{w}}([\Pi_1^0, \Pi_0^0])$ *are contained in* $\mathsf{KPi}^{\mathsf{w}} + (\Pi_2\ \mathsf{Ref})$.

  3. $\mathsf{FID}([\Pi_1^0, \mathrm{QF}])$ *and* $\mathsf{FID}([\Pi_1^0, \Pi_0^0])$ *are contained in* $\mathsf{KPi} + (\Pi_2\ \mathsf{Ref})$.

The theories $\mathsf{FID}^\mathsf{r}([\mathrm{POS}, \Pi_0^0])$ and $\mathsf{FID}^\mathsf{r}([\Pi_1^0, \Pi_0^0])$ are not mentioned in the previous two theorems. Of course they would be contained in KPi^w and $\mathsf{KPi}^\mathsf{w} + (\Pi_2\,\mathsf{Ref})$, respectively, but we can certainly do better. The problem is to take care of the second Π_0^0 components with weak forms of induction on the natural numbers (cf. Lemma 4). Exact bounds will be given in a forthcoming note.

The closure properties of the operator forms from $[\mathrm{POS}, \Pi_1^0]$ and $[\Pi_1^0, \Pi_1^0]$ are considered in Theorem 10 and require a recursively Mahlo universe. We first turn to the interpretation of the theory $\mathsf{FID}([\mathrm{POS}, \Pi_1^0])$ and its subsystems and look at $\mathsf{FID}([\Pi_1^0, \Pi_1^0])$ afterwards.

Theorem 13   *1.* $\mathsf{FID}^\mathsf{r}([\mathrm{POS}, \Pi_1^0])$ *is contained in* KPm^r.

   2. $\mathsf{FID}^\mathsf{w}([\mathrm{POS}, \Pi_1^0])$ *is contained in* KPm^w.

   3. $\mathsf{FID}([\mathrm{POS}, \Pi_1^0])$ *is contained in* KPm.

As before (cf. Theorem 11 and Theorem 12) we add Π_2 reflection in order to obtain closure provided that the first component is Π_1^0. Because of the Mahlo axiom, the global form of Π_2 reflection is not necessary, but the proof of Theorem 10 reveals at which stage its local version is applied.

Theorem 14   *1.* $\mathsf{FID}^\mathsf{r}([\Pi_1^0, \Pi_1^0])$ *is contained in* $\mathsf{KPm}^\mathsf{r} + (\Pi_2\,\mathsf{LRef})$.

   2. $\mathsf{FID}^\mathsf{w}([\Pi_1^0, \Pi_1^0])$ *is contained in* $\mathsf{KPm}^\mathsf{w} + (\Pi_2\,\mathsf{LRef})$.

   3. $\mathsf{FID}([\Pi_1^0, \Pi_1^0])$ *is contained in* $\mathsf{KPm} + (\Pi_2\,\mathsf{LRef})$.

The previous four theorems furnish us with upper proof-theoretic bounds for a series of first order theories for nonmonotone inductive definitions based on combined operator forms. In fact, all these upper bounds are sharp, and each first order theory for nonmonotone inductive definitions mentioned in these four theorems is proof-theoretically equivalent to the theory of admissible sets in which it is embedded.

This can be seen, for example, by carrying through well-ordering proofs in the first order theories for nonmonotone inductive definitions considered above and by exploiting the ordinal analysis of the corresponding theories of admissible sets. Here we follow a different rack and refer to Jäger and Studer [16]. In this article we use theories $\mathsf{FID}(\mathcal{K})$ together with their restricted versions $\mathsf{FID}^\mathsf{r}(\mathcal{K})$ and $\mathsf{FID}^\mathsf{w}(\mathcal{K})$ for suitable classes $\mathcal{K}$ of operator forms to obtain generic model constructions for systems of explicit mathematics.

Among other things it is shown in [16] and [27] that Feferman's famous theory T_0 can be modeled in $\mathsf{FID}([\mathrm{POS}, \mathrm{QF}])$ and, naturally, its obvious subsystems in $\mathsf{FID}^\mathsf{r}([\mathrm{POS}, \mathrm{QF}])$ and $\mathsf{FID}^\mathsf{w}([\mathrm{POS}, \mathrm{QF}])$. Because of well-known proof-theoretic results we immediately obtain the following theorem.

Theorem 15 *Each of the theories* $\mathsf{FID}([\mathrm{POS},\mathrm{QF}])$, $\mathsf{FID}([\mathrm{POS},\Pi_0^0])$, $\mathsf{FID}([\Pi_1^0,\mathrm{QF}])$ *and* $\mathsf{FID}([\Pi_1^0,\Pi_0^0])$ *is proof-theoretically equivalent to the theory* KPi *and therefore also to the theories* $\mathsf{KPi} + (\Pi_2\ \mathsf{Ref})$, $(\Delta_2^1\text{-}\mathsf{CA}) + (\mathsf{BI})$ *and* $\mathsf{T_0}$.

Here $(\Delta_2^1\text{-}\mathsf{CA}) + (\mathsf{BI})$ is the usual system of second order arithmetic with Δ_2^1 comprehension and bar induction. Analogous theorems can also be formulated for the restricted versions of these systems, but we omit their tiring formulations.

It is more interesting that, as shown in [16], the theory $\mathsf{FID}([\mathrm{POS},\Pi_1^0])$ is an adequate framework for generating a model of the theory $\mathsf{T_0(M)}$ which extends $\mathsf{T_0}$ by a Mahlo axiom tailored for explicit mathematics. A well-ordering proof for $\mathsf{T_0(M)}$ has not been published yet. However, it should yield that the proof-theoretic ordinals of KPm and $\mathsf{T_0(M)}$ agree. Then we also have the proof-theoretic equivalence of the theories $\mathsf{FID}([\mathrm{POS},\Pi_1^0])$, $\mathsf{FID}([\Pi_1^0,\Pi_1^0])$, KPm, $\mathsf{KPm} + (\Pi_2\ \mathsf{Ref})$ and $\mathsf{T_0(M)}$. Of course one has also the corresponding equivalences for the restricted systems.

References

[1] P. Aczel. An introduction to inductive definitions. In J. Barwise, editor, *Handbook of Mathematical Logic*. North-Holland, 1977.

[2] P. Aczel and W. Richter. Inductive definitions and reflecting properties of admissible ordinals. In J.E. Fenstad and P.G. Hinman, editors, *Generalized Recursion Theory*. North-Holland, 1974.

[3] T. Arai. Ordinal diagrams for recursively Mahlo universes. Submitted.

[4] T. Arai. Proof theory for theories of ordinals I: recursively Mahlo ordinals. Submitted.

[5] J. Barwise. *Admissible Sets and Structures*. Springer, 1975.

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

[7] S. Feferman. A language and axioms for explicit mathematics. In J.N. Crossley, editor, *Algebra and Logic*. Lecture Notes in Mathematics 450, Springer, 1975.

[8] S. Feferman. Constructive theories of functions and classes. In M. Boffa, D. van Dalen, and K. McAloon, editors, *Logic Colloquium '78*. North-Holland, 1979.

[9] G. Jäger. Beweistheorie von KPN. *Archiv für Mathematische Logik und Grundlagenforschung*, 20, 1980.

[10] G. Jäger. Zur Beweistheorie der Kripke-Platek-Mengenlehre über den natürlichen Zahlen. *Archiv für Mathematische Logik und Grundlagenforschung*, 22, 1982.

[11] G. Jäger. A well-ordering proof for Feferman's theory $\mathsf{T_0}$. *Archiv für Mathematische Logik und Grundlagenforschung*, 23, 1983.

[12] G. Jäger. The strength of admissibility without foundation. *Journal of Symbolic Logic*, 49, 1984.

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

[14] G. Jäger. Some proof-theoretic contributions to theories of sets. In The Paris Logic Group, editor, *Logic Colloquium '85*. North-Holland, 1987.

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

[16] G. Jäger and T. Studer. Extending the system $\mathsf{T_0}$ of explicit mathematics: the limit and Mahlo axioms. Submitted.

[17] Y.N. Moschovakis. *Elementary Induction on Abstract Structures*. North-Holland, 1974.

[18] W. Pohlers. Proof theory and ordinal analysis. *Archive for Mathematical Logic*, 30, 1991.

[19] W. Pohlers. A short course in ordinal analysis. In P. Aczel, H. Simmons, and S. Wainer, editors, *Proof Theory*. Cambridge University Press, 1992.

[20] W. Pohlers. Pure proof theory: Aims, methods and results. *Bulletin of Symbolic Logic*, 2, 1996.

[21] M. Rathjen. Proof-theoretic analysis of KPM. *Archive for Mathematical Logic*, 30, 1991.

[22] M. Rathjen. Fragments of Kripke-Platek set theory with infinity. In P. Aczel, H. Simmons, and S. Wainer, editors, *Proof Theory*. Cambridge University Press, 1992.

[23] M. Rathjen. Admissible proof theory and beyond. In D. Prawitz, B. Skyrms, and D. Westerståhl, editors, *Logic, Methodology and Philosophy of Science IX*. Elsevier Science B.V., 1994.

[24] M. Rathjen. Collapsing functions based on recursively large ordinals: A well-ordering proof for KPM. *Archive for Mathematical Logic*, 33, 1994.

[25] M. Rathjen. Proof theory of reflection. *Annals of Pure and Applied Logic*, 68, 1994.

[26] W. Richter. Recursively Mahlo ordinals and inductive definitions. In R.O. Gandy and C.E.M. Yates, editors, *Logic Colloquium '69.* North-Holland, 1971.

[27] T. Studer. Explicit mathematics: W-type, models. Diploma thesis, Institut für Informatik und angewandte Mathematik, Universität Bern, 1997.

Address

Gerhard Jäger, Institut für Informatik und angewandte Mathematik, Universität Bern, Neubrückstrasse 10, CH-3012 Bern, Switzerland, jaeger@iam.unibe.ch.