THE JOURNAL OF SYMBOLIC LOGIC
Volume 57, Number 3, Sept. 1992

ABOUT THE PROOF-THEORETIC ORDINALS OF WEAK FIXED POINT THEORIES

GERHARD JÄGER AND BARBARA PRIMO

Abstract. This paper presents several proof-theoretic results concerning weak fixed point theories over second order number theory with arithmetic comprehension and full or restricted induction on the natural numbers. It is also shown that there are natural second order theories which are proof-theoretically equivalent but have different proof-theoretic ordinals.

Fixed points and fixed point theories play an important role in many branches of mathematical logic and theoretical computer science. The spectrum ranges from the fixed point theorem in recursion theory to fixed point arguments in categorical logic and denotational semantics of high-level programming languages.

In proof theory, special emphasis has been put on formal systems for fixed points of (iterated) inductive definitions and their relationship to subsystems of analysis, set theory and constructive mathematics (cf. e.g. Buchholz, Feferman, Pohlers and Sieg [2], Feferman [3], and Jäger [7]). However, interpreted in the proper sense, also features of modern type theories can be studied in terms of fixed points, and many concepts in nonmonotonic reasoning (circumscription, completion of theories, etc.) are related to fixed point theories. To a certain extent even parts of logic programming are built upon fixed point constructions (cf. e.g. Lloyd [9] and Jäger and Stärk [8]).

The general purpose of this paper is to study several proof-theoretic aspects of the fixed point theories FP-ACA_0 and FP-ACA. They are formulated in the language of second order arithmetic, contain the axioms of primitive recursive arithmetic PRA, and comprise comprehension for arithmetic formulas. In addition, there has to be a fixed point for every positive arithmetic definition clause, and this fixed point can be proved to define a set. Both theories differ in the principles of complete induction which are available: In FP-ACA_0 complete induction on the natural numbers is restricted to sets, whereas FP-ACA contains complete induction for arbitrary formulas.

We will show that the proof-theoretic ordinals of FP-ACA_0 and FP-ACA are ε_0 and $\varphi 1\varepsilon_0$, respectively. This is remarkable since we will also see that FP-ACA_0 is proof-theoretically equivalent to the system (Σ^1_1-AC) of second order arithmetic,

Received March 8, 1991; revised October 28, 1991.

0022-4812/92/5703-0023/$02.20

which is known to have proof-theoretic ordinal $\varphi 1\varepsilon_0$. Hence we obtain the interesting consequence that FP-ACA$_0$ and (Σ^1_1-AC) are proof-theoretically equivalent theories of second order arithmetic but have different proof-theoretic ordinals.

In this paper we make use of the traditional approach to proof-theoretic ordinals. Hence we call an ordinal α *provable* in a second order theory Th if there exists a primitive recursive well-ordering $\prec$ on the natural numbers of order type α so that the well-foundedness of $\prec$ is provable in Th. Then the proof-theoretic ordinal of Th is defined to be the least ordinal which is not provable in Th. On the other hand, two theories Th$_1$ and Th$_2$ are called *proof-theoretically equivalent* if each can be reduced to the other.

In the literature one can also find different notions of proof-theoretic ordinal which are more or less directly based (often via consistency statements) on proof-theoretic reductions. An example is Feferman [3], where the proof-theoretic ordinal of a theory Th is defined as the least ordinal α such that the consistency statement for Th can be proved in HA + $TI(\prec_\alpha)$; here $TI(\prec_\alpha)$ is the scheme of transfinite induction up to α.

In most cases the different approaches to proof-theoretic ordinals do not really matter, and theories are proof-theoretically equivalent if they have the same proof-theoretic ordinals. Therefore the results of this paper are also interesting in the sense that they present natural subsystems of second order arithmetic which show that one must use the notions of proof-theoretic ordinal and proof-theoretic equivalence with some care. More on this can be found in §4, below.

In §1 we present some basic notions and give a precise definition of the theories FP-ACA$_0$ and FP-ACA. In §2 we introduce the infinitary system FP-ACA* which is needed for carrying through some proof-theoretic reductions. §3 is dedicated to collecting the proof-theoretic results about FP-ACA$_0$ and FP-ACA. The paper concludes with a general discussion of our results in view of related work. Several proofs of this paper are worked out in full detail in Primo [11] and only sketched in the following.

§1. The theories FP-ACA$_0$ and FP-ACA. Let L_2 be the usual language of second order arithmetic with number variables $x, y, z, x_1, y_1, z_1, \ldots$, set variables $X, Y, Z, X_1, Y_1, Z_1, \ldots$, the constant 0 and a symbol for each primitive recursive function and relation. The *terms* $s, t, s_1, t_1, \ldots$ and formulas $F, G, H, F_1, G_1, H_1, \ldots$ of L_2 are defined as usual; an L_2 formula is called an *arithmetic* L_2 formula if it does not contain bound set variables.

We write $\vec{X}$ and $\vec{x}$ for finite strings $X_1, \ldots X_n$ and $x_1, \ldots x_n$ of set and number variables. The notation $F[\vec{X}, \vec{x}]$ is used to indicate that all free variables of F come from the lists $\vec{X}$ and $\vec{x}$; $F(\vec{X}, \vec{x})$ may contain other free variables besides $\vec{X}$ and $\vec{x}$. The formula $F(G)$ denotes the formula that results from $F(X)$ if we replace each occurrence of $(t \in X)$ by $G(t)$. A formula is called *X-positive* if it does not contain negative occurrences of the set variable X. We write $\mathcal{N}$ for the usual set of the non-negative integers and Pow($\mathcal{N}$) for the power set of $\mathcal{N}$.

Each X-positive arithmetic formula $A[X, x]$ induces an operator Γ_A from Pow($\mathcal{N}$) to Pow($\mathcal{N}$) so that

$$\Gamma_A(S) := \{n \in \mathcal{N} : \mathcal{N} \models A[S, n]\}$$

for all $S \subset \mathcal{N}$. Since $A[X, x]$ is X-positive, Γ_A is a monotonic operator in the sense that $\Gamma_A(S_1) \subset \Gamma_A(S_2)$ if $S_1 \subset S_2$. For each such Γ_A we define the following subsets of $\mathcal{N}$ by transfinite recursion on the ordinals:

$$I_A^{<\alpha} := \bigcup_{\xi<\alpha} I_A^{\xi}, \qquad I_A^{\alpha} := \{n \in \mathcal{N} : \mathcal{N} \models A[I_A^{<\alpha}, n]\},$$

and we set $I_A := \bigcup_{\xi \in \mathrm{On}} I_A^{\xi}$.

We say that the set I_A is *inductively defined* by the X-positive arithmetic formula $A[X, x]$, and it is well known that I_A is the least fixed point of the operator Γ_A. For detailed information about inductive definitions we refer for example to Moschovakis [10].

The famous theory ID_1 provides a natural framework for studying the proof-theoretic aspects of least fixed points of inductive definitions in a first order extension of Peano arithmetic. A good source for more information about the proof theory of both the single and the iterated form of inductive definability is provided by Buchholz, Feferman, Pohlers and Sieg in [2].

In this paper we shift our interest from theories for least fixed points of inductive definitions to (weaker) formalisms which formalize the idea of arbitrary fixed points. For this end we extend our language L_2 to a language L_2^+ by adding a new unary relation symbol P_A for every X-positive arithmetic L_2 formula $A[X, x]$. According to our previous conventions, an L_2^+ formula is then called *arithmetical* if it contains no bounded set variables.

The strength of formal theories is often measured in terms of their *proof-theoretic ordinals*, whose definition is based on the following notions. Let $\prec$ be a binary primitive recursive relation and $R_\prec$ the corresponding relation symbol of L_2. In the following we write $(s \prec t)$ for $R_\prec(s, t)$, $(\forall x \prec t)F(x)$ for $(\forall x)(R_\prec(x, t) \to F(x))$, and $(\exists x \prec t)F(x)$ for $(\exists x)(R_\prec(x, t) \wedge F(x))$. The principle of *transfinite induction* for a formula F along $\prec$ is expressed in a second order context by the formula $TI(\prec, F)$; we call $\prec$ *well-founded* if we have transfinite induction along $\prec$ for all sets. Specifically,

$$TI(\prec, F) :\Leftrightarrow (\forall x)[(\forall y \prec x)F(y) \to F(x)] \to (\forall x)F(x),$$
$$WF(\prec) :\Leftrightarrow (\forall X)TI(\prec, X).$$

If Th is a theory in L_2 or an extension of L_2, then an ordinal α is called *provable in* Th if there exists a primitive recursive well-ordering $\prec$ of order type α so that $\mathrm{Th} \vdash \mathrm{WF}(\prec)$. The *proof-theoretic ordinal* of Th is often defined to be the least ordinal which is not provable in Th; it is denoted by $|\mathrm{Th}|$.

The theories in L_2^+ will be assumed to contain all axioms and rules of classical predicate calculus with equality in the first sort. The theory $\mathrm{FP\text{-}ACA}_0$ is given by the following additional axioms.

I. *Primitive recursion.* The usual axioms for 0, successor and the defining equations for the primitive recursive functions and relations.

II. *Axiom of induction on the natural numbers.*

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

III. *Arithmetic comprehension ACA.* For all arithmetic formulas $A(x)$ of L_2^+:

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

IV. *Fixed point axioms.* For all X positive arithmetic formulas $A[X, x]$ of L_2:

$$(\forall x)(A[P_A, x] \leftrightarrow P_A(x)).$$

Axiom IV formalizes that P_A represents a fixed point of the operator Γ_A. Hence we denote the relation symbols P_A as *fixed point constants.* FP-ACA is the extension of FP-ACA_0 which is obtained by replacing the axiom of induction on $\mathcal{N}$ by the corresponding *scheme of complete induction on the natural numbers*

$$F(0) \wedge (\forall x)(F(x) \to F(x')) \to (\forall x)F(x)$$

for arbitrary L_2^+ formulas. Finally, FP-$\mathrm{ACA}^{(1)}$ is the restriction of FP-ACA to the first order language L_1^+ which results from L_2^+ by deleting all formulas which contain set variables.

If Th is a theory containing FP-ACA_0, then we call the set

$$\mathrm{ProvPart}_{\mathrm{Th}}(A) := \{n \in \mathcal{N} : \mathrm{Th} \vdash P_A(n)\}$$

the *provable part* of the inductive definition (represented by) A with respect to Th. It is often the case that $\mathrm{ProvPart}_{\mathrm{Th}}(A) = I_A^{<\alpha}$ for some countable α.

The theory FP-ACA corresponds to the fixed point theory ID_1^- of Beeson [1], and FP-$\mathrm{ACA}^{(1)}$ agrees with the fixed point theory $\widehat{\mathrm{ID}}_1$ of Feferman [3]. Its relationship to FP-ACA_0 is characterized by the following theorem.

THEOREM 1. *FP-ACA_0 is a conservative extension of FP-$\mathrm{ACA}^{(1)}$ with respect to all L_1^+ formulas.*

It is obvious that FP-ACA_0 is an extension of FP-$\mathrm{ACA}^{(1)}$. The property of being conservative can be established by an easy model-theoretic argument: One shows that every (first order) model $\mathcal{M}$ of FP-$\mathrm{ACA}^{(1)}$ can be extended to a second order model $\mathcal{M}^{(2)}$ of FP-ACA_0 by choosing the collection of all first order definable subsets of the universe $|\mathcal{M}|$ of $\mathcal{M}$ as the range of the set variables of FP-ACA_0. Then we have

$$\mathcal{M} \models F \Leftrightarrow \mathcal{M}^{(2)} \models F$$

for all L_1^+ sentences F. Hence by completeness the provability of an L_1^+ sentence F in FP-ACA_0 implies its provability in FP-$\mathrm{ACA}^{(1)}$. A detailed proof of a result similar to Theorem 1 is given in Feferman [4].

§2. The infinitary system FP-ACA*. The proof-theoretic bounds of FP-ACA_0 and FP-ACA will be established by embedding these formal theories into the infinitary system FP-ACA* which admits the proof of a cut elimination and boundedness theorem. FP-ACA* is infinitary with respect to the ranks of its formulas and the length of its derivations and is related to the system RA* of Schütte [12].

We assume familiarity with a standard notation system for predicative mathematics. For all countable ordinals α we can define ordinal functions $\varphi\alpha$ by the following recursion: $\varphi 0\xi$ is ω^ξ; for $\alpha > 0$, $\varphi\alpha\xi$ is the ξth simultaneous fixed point of

all functions $\varphi\beta$ with $\beta < \alpha$. Then $\varphi 10$ is the ordinal ε_0, i.e. the least ξ such that $\omega^\xi = \xi$.

We now set up the infinitary system FP-ACA*. The language L^* of FP-ACA* starts out from the fragment of L_2^+ without free number variables. Hence the *number terms* of L^* are the closed number terms of L_2^+ and each number term s of L^* has a canonical value $V(s) \in \mathcal{N}$. Now let F be an atomic formula of L^* of the form $R_Z(s_1, \ldots, s_n)$, where R_Z is the relation symbol for the primitive recursive relation Z. Then F is called a *true* atomic formula if $Z(V(s_1), \ldots, V(s_n))$ is true; otherwise F is called *false*. Two atomic formulas of L^* are called *equivalent* if they differ in number terms with identical values only.

The *set terms* and *formulas* of L^* are generated by the following inductive definition:

1. Every set variable of L_2^+ is a set term of L^*.
2. Every atomic formula of L_2^+ without free number variables is an (atomic) L^* formula.
3. If S is a set term of L^* and s a number term of L^*, then $(s \in S)$ is an L^* formula.
4. If F is an L^* formula, then $\neg F$ is an L^* formula.
5. If F and G are L^* formulas and j is a binary junctor of L_2, then $(F\ j\ G)$ is an L^* formula.
6. If $F(x)$ is an L^* formula, then $(\forall x)F(x)$ and $(\exists x)F(x)$ are L^* formulas.
7. If $F(X)$ is an L^* formula, then $(\forall X)F(X)$ and $(\exists X)F(X)$ are L^* formulas.
8. If $F(x)$ is an L^* formula without bound set variables, then $\{x: F(x)\}$ is a set term of L^*.

The collection of all L^* formulas without set terms and without negative occurrences of fixed point constants P_A is denoted by $\mathcal{POS}$, and the collection of all L^* formulas without set terms and without positive occurrences of fixed point constants P_A by $\mathcal{NEG}$; a formula is called *positive* if it belongs to $\mathcal{POS}$ and *negative* if it belongs to $\mathcal{NEG}$. We write $\mathcal{PN}$ for the collection of positive and negative formulas, i.e. $\mathcal{PN} = \mathcal{POS} \cup \mathcal{NEG}$. It is obvious that we have $F \in \mathcal{POS}$ if and only if $\neg F \in \mathcal{NEG}$, and $G \in \mathcal{NEG}$ if and only if $\neg G \in \mathcal{POS}$.

The complexity of L^* formulas F is measured by their rank $\mathrm{rn}(F)$, which is inductively defined as follows.

1. If F belongs to $\mathcal{PN}$, then $\mathrm{rn}(F) := 0$.
2. For L^* formulas not in $\mathcal{PN}$ we define their rank according to the following rules (where Q stands for existential or universal quantifiers):

$$\begin{aligned} \mathrm{rn}(s \in X) &:= 1, \\ \mathrm{rn}(s \in \{x: F(x)\}) &:= \mathrm{rn}(F(0)) + 1, \\ \mathrm{rn}(\neg F) &:= \mathrm{rn}(F) + 1, \\ \mathrm{rn}(F\ j\ G) &:= \max(\mathrm{rn}(F), \mathrm{rn}(G)) + 1, \\ \mathrm{rn}((Qx)F(x)) &:= \mathrm{rn}(F(0)) + 1, \\ \mathrm{rn}((QX)F(X)) &:= \max(\omega, \mathrm{rn}(F(0)) + 1). \end{aligned}$$

This definition immediately yields that $0 \leq \mathrm{rn}(F) < \omega + \omega$ for all L^* formulas F; it is also clear that $\mathrm{rn}(F) = 0$ if and only if $F \in \mathcal{PN}$. If F does not belong to $\mathcal{PN}$, then the rank of each subformula of F is smaller than $\mathrm{rn}(F)$; in particular we have $\mathrm{rn}(F(S)) < \mathrm{rn}((QX)F(X))$ for all set terms S.

The proof system of FP-ACA* is an extension of the usual Gentzen sequent calculus LK (cf. e.g. Girard [6] and Takeuti [14]). The capital Greek letters $\Gamma, \Delta, \Sigma, \ldots$ denote finite sequences of L^* formulas, and *sequents* are formal expressions of the form $\Gamma \supset \Delta$.

The axioms and rules of FP-ACA* comprise (i) the axioms, (ii) the structural rules, (iii) the rules for the propositional connectives, and (iv) the cut rule of LK, and also include the following additional number-theoretic axioms and rules for quantification and fixed point constants.

Number-theoretic axioms. For all atomic formulas F_1 and F_2 which are equivalent, for all true atomic formulas G and all false atomic formulas H we choose as number-theoretic axioms:

$$F_1 \supset F_2, \qquad \supset G, \qquad H \supset.$$

Set-theoretic rules. For all L^* formulas $F(x)$ without bound set variables and all number terms s of L^* we have as set-theoretic rules:

$$\frac{\Gamma \supset \Delta, F(s)}{\Gamma \supset \Delta, s \in \{x: F(x)\}}, \qquad \frac{\Gamma, F(s) \supset \Delta}{\Gamma, s \in \{x: F(x)\} \supset \Delta}.$$

Quantifier rules. Formulated for universal quantifiers; the corresponding rules for existential quantifiers must also be included.

$$\frac{\Gamma \supset \Delta, F(s) \quad \text{for all } s}{\Gamma \supset \Delta, (\forall x)F(x)}, \qquad \frac{\Gamma, F(s) \supset \Delta \quad \text{for some } s}{\Gamma, (\forall x)F(x) \supset \Delta},$$

$$\frac{\Gamma \supset \Delta, F(S) \quad \text{for all } S}{\Gamma \supset \Delta, (\forall X)F(X)}, \qquad \frac{\Gamma, F(S) \supset \Delta \quad \text{for some } S}{\Gamma, (\forall X)F(X) \supset \Delta}.$$

Fixed point rules. For all X-positive arithmetic formulas $A[X, x]$ of L_2 and all number terms s of L^* we have as fixed point rules:

$$\frac{\Gamma \supset \Delta, A[P_A, s]}{\Gamma \supset \Delta, P_A(s)}, \qquad \frac{\Gamma \supset \Delta, P_A(s)}{\Gamma \supset \Delta, A[P_A, s]}.$$

The notion FP-ACA* $\vdash_\rho^\alpha \Gamma \supset \Delta$ is used to express that the sequent $\Gamma \supset \Delta$ is provable in FP-ACA* by a proof of length α so that all cut formulas have rank less than ρ; it is inductively defined as follows:

1. If $\Gamma \supset \Delta$ is an axiom of FP-ACA*, then we have FP-ACA* $\vdash_\rho^\alpha \Gamma \supset \Delta$ for all ordinals α and ρ.

2. If FP-ACA* $\vdash_\rho^{\alpha_i} \Gamma_i \supset \Delta_i$ and $\alpha_i < \alpha$ for every premise $\Gamma_i \supset \Delta_i$ of an FP-ACA* rule or a cut with a cut formula F so that $\mathrm{rn}(F) < \rho$, then we have FP-ACA* $\vdash_\rho^\alpha \Gamma \supset \Delta$.

Hence the sequent $\Gamma \supset \Delta$ is cut-free provable in FP-ACA* if there exists an ordinal α with FP-ACA* $\vdash_0^\alpha \Gamma \supset \Delta$. On the other hand, FP-ACA* $\vdash_1^\alpha \Gamma \supset \Delta$ means that $\Gamma \supset \Delta$ has a proof of length α so that all cut formulas belong to the collection $\mathscr{P}\mathscr{N}$. Because of the fixed point rules it is impossible to prove complete cut elimination for FP-ACA*. However, the principal formulas of the fixed point rules have rank 0. Therefore by applying the standard techniques of predicative proof theory as developed for example in Girard [6], Schütte [12], or Takeuti [14], one obtains the following weaker result.

THEOREM 2 (Cut elimination theorem). *For all sequents $\Gamma \supset \Delta$, all ordinals α and all ordinals $\rho > 0$:*

1. *FP-ACA* $\vdash^{\alpha}_{\rho+1} \Gamma \supset \Delta \Rightarrow$ FP-ACA* $\vdash^{\omega^{\alpha}}_{\rho} \Gamma \supset \Delta$;*
2. *FP-ACA* $\vdash^{\alpha}_{\omega} \Gamma \supset \Delta \Rightarrow$ FP-ACA* $\vdash^{\varphi 1 \alpha}_{1} \Gamma \supset \Delta$.*

It is an easy exercise to show that the formal theories FP-ACA, FP-ACA$_0$ and FP-ACA$^{(1)}$ can be embedded into FP-ACA*. Complete induction is proved in the infinitary system with the help of the infinitary rule for the universal number quantifier, arithmetic comprehension follows from the definition of set terms, and the existential rule for the set quantifiers and the fixed point axioms are a consequence of the fixed point rules.

THEOREM 3 (Embedding theorem). 1. *Let $F[\vec{X}, \vec{x}]$ be an L_2^+ formula which is provable in FP-ACA. Then there exist an $\alpha < \varepsilon_0$ and an $n < \omega$ so that for all set terms $\vec{S}$ and all number terms $\vec{s}$ of L^* we have*

$$FP\text{-}ACA^* \vdash^{\alpha}_{\omega+n} \supset F[\vec{S}, \vec{s}].$$

2. *Let $G[\vec{x}]$ be an L_1^+ formula which is provable in FP-ACA$_0$ and therefore—according to Theorem 1—also in FP-ACA$^{(1)}$. Then there exist an $\alpha < \varepsilon_0$ and an $n < \omega$ so that for all number terms $\vec{s}$ of L^* we have*

$$FP\text{-}ACA^* \vdash^{\alpha}_{n} \supset G[\vec{s}].$$

As preparation for the boundedness theorem we have to introduce some further notation. If F is a positive or negative L^* formula and α a countable ordinal, then $F\{\alpha\}$ denotes the result of substituting $I_A^{<\alpha}$ for every occurrence of P_A in F. Hence we have

$$(*) \qquad \mathcal{N} \models F\{\alpha\} \to F\{\beta\} \quad \text{and} \quad \mathcal{N} \models G\{\beta\} \to G\{\alpha\}$$

for every positive L^* formula F, every negative L^* formula G and all $\alpha \leq \beta$. For a finite sequence Γ of $\mathcal{P}\mathcal{N}$ formulas, we write $\Gamma^{\mathcal{P}}$ for the set of all positive formulas which occur in Γ and $\Gamma^{\mathcal{N}}$ for the set of all formulas in Γ which do not belong to $\Gamma^{\mathcal{P}}$. Hence every formula which occurs in Γ belongs to $\Gamma^{\mathcal{P}} \cup \Gamma^{\mathcal{N}}$. If $\Gamma \supset \Delta$ is a sequent of $\mathcal{P}\mathcal{N}$ formulas and if α and β are countable ordinals, then $\{\alpha\}(\Gamma \supset \Delta)\{\beta\}$ is defined to be the formula

$$\bigvee_{F \in \Gamma^{\mathcal{P}}} (\neg F)\{\alpha\} \vee \bigvee_{F \in \Gamma^{\mathcal{N}}} (\neg F)\{\beta\} \vee \bigvee_{F \in \Delta^{\mathcal{P}}} F\{\beta\} \vee \bigvee_{F \in \Delta^{\mathcal{N}}} F\{\alpha\}.$$

The following boundedness lemma and boundedness theorem are concerned with the provably accessible parts of the (least) fixed points I_A in the infinitary fixed point theory FP-ACA*. The boundedness lemma is based on an asymmetric interpretation of the positive and negative occurrences of the fixed point constants P_A in the sequents $\Gamma \supset \Delta$.

LEMMA 4 (Boundedness lemma). *Let Γ and Δ be finite sequences of $\mathcal{P}\mathcal{N}$ formulas. Then for all countable ordinals α and β we have*

$$FP\text{-}ACA^* \vdash^{\alpha}_{1} \Gamma \supset \Delta \Rightarrow \mathcal{N} \models \{\beta\}(\Gamma \supset \Delta)\{\beta + \omega^{\alpha}\}.$$

PROOF (by induction on α). If $\Gamma \supset \Delta$ is an axiom of FP-ACA*, then the assertion is trivial. Otherwise $\Gamma \supset \Delta$ is the conclusion of a derivation rule. We concentrate on the three critical cases and leave the rest to the reader.

1. $\Gamma \supset \Delta$ is the conclusion of the fixed point rule

$$\frac{\Gamma \supset \Sigma, A[P_A, s]}{\Gamma \supset \Sigma, P_A(s)}$$

where $\Delta = \Sigma, P_A(s)$. Then there exists a $\gamma < \alpha$ so that

$$\text{FP-ACA}^* \vdash_1^\gamma \Gamma \supset \Sigma, A[P_A, s],$$

and the induction hypothesis implies

$$\mathcal{N} \models \{\beta\}(\Gamma \supset \Sigma)\{\beta + \omega^\gamma\} \vee A[I_A^{<\beta+\omega^\gamma}, s].$$

By the definition of the sets I_A^ξ we may conclude that

$$\mathcal{N} \models \{\beta\}(\Gamma \supset \Sigma)\{\beta + \omega^\gamma\} \vee s \in I_A^{\beta+\omega^\gamma},$$

and the assertion follows by (*).

2. $\Gamma \supset \Delta$ is the conclusion of the fixed point rule

$$\frac{\Gamma \supset \Sigma, P_A(s)}{\Gamma \supset \Sigma, A[P_A, s]}$$

where $\Delta = \Sigma, A[P_A, s]$. Then there exists a $\gamma < \alpha$ so that

$$\text{FP-ACA}^* \vdash_1^\gamma \Gamma \supset \Sigma, P_A(s),$$

and the induction hypothesis implies

$$\mathcal{N} \models \{\beta\}(\Gamma \supset \Sigma)\{\beta + \omega^\gamma\} \vee s \in I_A^{<\beta+\omega^\gamma}.$$

We know that $\mathcal{N}$ is a model of $(s \in I_A^{<\xi} \to A[I_A^{<\xi}, s])$ for all ξ, so that we obtain

$$\mathcal{N} \models \{\beta\}(\Gamma \supset \Sigma)\{\beta + \omega^\gamma\} \vee A[I_A^{<\beta+\omega^\gamma}, s].$$

In view of (*) this yields the assertion.

3. $\Gamma \supset \Delta$ is the conclusion of a cut

$$\frac{\Gamma \supset \Delta, H \qquad \Gamma, H \supset \Delta}{\Gamma \supset \Delta}$$

with cut formula H. Then there exist $\gamma_1, \gamma_2 < \alpha$ so that

$$\text{FP-ACA}^* \vdash_1^{\gamma_1} \Gamma \supset \Delta, H, \tag{1}$$

$$\text{FP-ACA}^* \vdash_1^{\gamma_2} \Gamma, H \supset \Delta, \tag{2}$$

and $\text{rn}(H) < 1$. Hence H is an element of $\mathcal{PN}$, and by symmetry we can assume without loss of generality that H belongs to $\mathcal{POS}$. Therefore we can apply the induction hypothesis to (1) and (2) and obtain from (1)

$$\mathcal{N} \models \{\beta\}(\Gamma \supset \Delta)\{\delta\} \vee H\{\delta\} \tag{3}$$

for $\delta := \beta + \omega^{\gamma_1}$. On the other hand, if we replace β by δ, then the induction hypothesis applied to (2) gives

$$\mathcal{N} \models \{\delta\}(\Gamma \supset \Delta)\{\delta + \omega^{\gamma_2}\} \vee (\neg H)\{\delta\}. \tag{4}$$

Since $\delta + \omega^{\gamma_2} = \beta + \omega^{\gamma_1} + \omega^{\gamma_2} \leq \beta + \omega^{\alpha}$, we obtain from (3), (4) and (*) that

$$\mathcal{N} \models \{\beta\}(\Gamma \supset \Delta)\{\beta + \omega^{\alpha}\}.$$

This finishes the proof of case 3. In the remaining cases the assertion follows from the induction hypothesis immediately. □

A special case of this boundedness lemma for FP-ACA* is the boundedness theorem for FP-ACA*.

THEOREM 5 (Boundedness theorem). *For all fixed point constants P_A, the corresponding (least) fixed points I_A and all natural numbers n,*

$$FP\text{-}ACA^* \vdash_1^{\alpha} \supset P_A(n) \quad \Rightarrow \quad n \in I_A^{<\omega^{\alpha}}.$$

§3. Proof-theoretic bounds. We begin this section with a general observation about the provable part of inductive definitions in FP-ACA_0 and FP-ACA. This result is obtained by a combination of the embedding, cut elimination and boundedness theorem for FP-ACA*.

THEOREM 6. *Let $A[X, x]$ be an X-positive arithmetic L_2 formula. Then:*

1. $FP\text{-}ACA_0 \vdash P_A(n) \quad \Rightarrow \quad n \in I_A^{<\varepsilon_0}$.
2. $FP\text{-}ACA \vdash P_A(n) \quad \Rightarrow \quad n \in I_A^{<\varphi 1 \varepsilon_0}$.

PROOF. 1. Assume that FP-ACA_0 proves $P_A(n)$. According to the embedding theorem there exist $\alpha < \varepsilon_0$ and $n < \omega$ such that

$$\text{FP-ACA}^* \vdash_n^{\alpha} \supset P_A(n)$$

and hence by cut elimination (since ε_0 is closed under $\exp_{\omega}$)

$$\text{FP-ACA}^* \vdash_1^{\beta} \supset P_A(n)$$

for some $\beta < \varepsilon_0$. The boundedness theorem then yields the assertion.

2. If FP-ACA proves $P_A(n)$, then the embedding theorem implies that

$$\text{FP-ACA}^* \vdash_{\omega + n}^{\alpha} \supset P_A(n)$$

for some $\alpha < \varepsilon_0$ and $n < \omega$. As before we obtain by cut elimination that

$$\text{FP-ACA}^* \vdash_{\omega}^{\beta} \supset P_A(n)$$

for a suitable $\beta < \varepsilon_0$. Again by cut elimination we can now conclude that

$$\text{FP-ACA}^* \vdash_1^{\varphi 1 \beta} \supset P_A(n).$$

Since φ is monotonic in its second argument, the assertion follows from the boundedness theorem. □

Now the ground is prepared for determining the proof-theoretic ordinals of FP-ACA_0 and FP-ACA. Their lower bounds are obvious since FP-ACA_0 contains Peano arithmetic PA and FP-ACA contains the theory ACA of arithmetic comprehension. By standard results we therefore have that

(A) $$\varepsilon_0 \leq |\text{FP-ACA}_0|,$$

(B) $$\varphi 1 \varepsilon_0 \leq |\text{FP-ACA}|.$$

The converse directions are proved with the help of Theorem 6. For this purpose we introduce for every primitive recursive well-ordering $\prec$ the X-positive arith-

metic L_2 formula $A_\prec[X, x]$ defined by

$$A_\prec[X, x] :\Leftrightarrow (\forall y \prec x)(y \in X),$$

and we write $P_\prec$ for $P_{A_\prec}$ as well as $I_\prec$ for $I_{A_\prec}$. If the natural number n belongs to the field of $\prec$, then $|n|_\prec$ denotes the order type of n with respect to $\prec$; otherwise we put $|n|_\prec := 0$. It is easy to show by induction on α that

$$\text{(C)} \qquad n \in I^\alpha_\prec \quad \Rightarrow \quad |n|_\prec \leq \alpha$$

for all natural numbers n. The fixed point constants generate arithmetic formulas of L_2^+ and provably define sets in FP-ACA$_0$ and FP-ACA. Therefore we obtain from the fixed point axiom for $P_\prec$ that

$$\text{(D)} \qquad \text{FP-ACA}_0 \vdash WF(\prec) \to (\forall x) P_\prec(x)$$

for all primitive recursive well-orderings $\prec$. To sum up, we have the following result for FP-ACA$_0$ and FP-ACA.

THEOREM 7. 1. $|FP\text{-}ACA_0| = \varepsilon_0$.

2. $|FP\text{-}ACA| = \varphi 1 \varepsilon_0$.

PROOF. Let Th be the theory FP-ACA$_0$ or FP-ACA. If the ordinal α is provable in Th, then there exist a primitive recursive well-ordering $\prec$ and a natural number n such that

$$\text{Th} \vdash WF(\prec) \quad \text{and} \quad \alpha = |n|_\prec.$$

Now we apply (D) and obtain Th $\vdash P_\prec(n)$. Hence Theorem 6 yields $n \in I^{<\sigma}_\prec$ for $\sigma = \varepsilon_0$ if Th = FP-ACA$_0$ and $\sigma = \varphi 1 \varepsilon_0$ if Th = FP-ACA. It follows immediately from (C) that $\alpha = |n|_\prec < \sigma$. Together with (A) and (B) this gives $|\text{Th}| = \sigma$. □

§4. Discussion. Let Th_1 and Th_2 be two consistent theories so that $L_1 \subset L(\text{Th}_i)$ for $i = 1, 2$. In the literature one often calls Th_1 *proof-theoretically reducible to* Th_2, in symbols $\text{Th}_1 \leq \text{Th}_2$, if there is an effective method to transform every $L(\text{Th}_1)$ formula F into an $L(\text{Th}_2)$ formula F^* so that the following conditions are satisfied:

$$\text{(1)} \qquad \text{Th}_1 \vdash F \quad \Rightarrow \quad \text{Th}_2 \vdash F^*,$$

and

$$\text{(2)} \qquad \text{Th}_2 \vdash G \leftrightarrow G^*$$

for all $L(\text{Th}_1)$ sentences F and all L_1 sentences G. Then the *proof-theoretic equivalence* of Th_1 and Th_2 is defined by

$$\text{Th}_1 \equiv \text{Th}_2 :\Leftrightarrow \text{Th}_1 \leq \text{Th}_2 \text{ and } \text{Th}_2 \leq \text{Th}_1.$$

In most interesting cases—and especially if the theories are natural and proof-theoretically strong—one has for second order theories Th_1 and Th_2 that

$$|\text{Th}_1| = |\text{Th}_2| \Leftrightarrow \text{Th}_1 \equiv \text{Th}_2,$$

but in the general case there are counterexamples to both directions. So it is well known that

$$|\text{ACA}_0| = |\text{ACA}_0 + \text{Con}(\text{ACA}_0)|$$

where ACA_0 is the system of second order arithmetic with arithmetic comprehension and restricted induction[1] and $Con(ACA_0)$ is the consistency statement for ACA_0. Because of Gödel's result we also have that $ACA_0 \not\equiv ACA_0 + Con(ACA_0)$. A counterexample to the converse direction is provided, for example, by the theory $FP\text{-}ACA_0$.

Among other things, Feferman [3], following an idea due to Aczel, presents an interpretation of the fixed point theory $\widehat{ID}_1$ into the system $(\Sigma^1_1\text{-AC})$ of second order arithmetic with the Σ^1_1 axiom of choice, and because of Theorem 1 this interpretation immediately implies

$$FP\text{-}ACA_0 \leq FP\text{-}ACA^{(1)} = \widehat{ID}_1 \leq (\Sigma^1_1\text{-AC}).$$

But it also follows from Feferman's paper that $(\Sigma^1_1\text{-AC})$ can be proof-theoretically reduced to $\widehat{ID}_1$, so that

$$(\Sigma^1_1\text{-AC}) \leq \widehat{ID}_1 = FP\text{-}ACA^{(1)} \subset FP\text{-}ACA_0.$$

By Theorem 7 we also have $|FP\text{-}ACA_0| = \varepsilon_0$. Since $|(\Sigma^1_1\text{-AC})| = \varphi\varepsilon_0 0$ (cf. e.g. [2], [5], or [13]) and $\varphi\varepsilon_0 0 > \varepsilon_0$, we conclude that

$$FP\text{-}ACA_0 \equiv (\Sigma^1_1\text{-AC}) \quad \text{and} \quad |FP\text{-}ACA_0| < |(\Sigma^1_1\text{-AC})|.$$

Hence the second order theories $FP\text{-}ACA_0$ and $(\Sigma^1_1\text{-AC})$ are proof-theoretically equivalent although they have different proof-theoretic ordinals. The explanation of this unexpected relationship between proof-theoretic equivalence and proof-theoretic ordinals is as follows:

In the reduction of $(\Sigma^1_1\text{-AC})$ to $\widehat{ID}_1$ the fixed point axiom is used in order to define simultaneously: (i) a unary relation S and (ii) two binary relations $\in$ and $\notin$ for representing the (codes of the) sets, the element relation and the not-element relation.[2] In this reduction the universal second order quantifier in $WF(\prec)$ is replaced by a first order quantifier $(\forall x)(S(x) \to \cdots)$ so that the formula $WF(\prec)$ is not carried over from $(\Sigma^1_1\text{-AC})$ to $FP\text{-}ACA_0$. If $(\Sigma^1_1\text{-AC}) \vdash WF(\prec)$, then $TI(\prec, Z)$ is provably true in $FP\text{-}ACA_0$ for all sets Z coded by elements of S but not necessarily for all sets.

If a "free" relation symbol R is adjoined to L_2 and $(\Sigma^1_1\text{-AC}) \vdash TI(\prec, R)$, then we can tailor the reduction of $(\Sigma^1_1\text{-AC})$ to $\widehat{ID}_1$ so that $\widehat{ID}_1 \vdash TI(\prec, R)$ and $FP\text{-}ACA_0 \vdash TI(\prec, R)$, provided that we allow R to occur in the definition clause of the fixed point which is needed for the definition of S, $\in$ and $\notin$. Hence this R is a parameter in the definition clause of a fixed point axiom and cannot act as a free set variable. If we generalize our fixed point axioms to clauses with free set parameters, we obtain theories of greater proof-theoretic power.

Added in proof. With regard to Question 1.6, we have recently shown that if $\kappa > \omega$, then the equation is consistent for any $n \in \omega$.

[1] ACA_0 is a conservative extension of Peano arithmetic PA.

[2] Actually the precise argument is more complicated since the reduction of $(\Sigma^1_1\text{-AC})$ to $\widehat{ID}_1$ is via the intermediate theory $EM_0 + (J)$.

REFERENCES

[1] M. J. BEESON, ***Foundations of constructive mathematics***, Springer-Verlag, Berlin, 1985.

[2] W. BUCHHOLZ, S. FEFERMAN, W. POHLERS, and W. SIEG, ***Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies***, Lecture Notes in Mathematics, vol. 897, Springer-Verlag, Berlin, 1981.

[3] S. FEFERMAN, *Iterated inductive fixed-point theories: application to Hancock's conjecture*, ***Patras logic symposion*** (G. Metakides, editor), North-Holland, Amsterdam, 1982, pp. 171–196.

[4] ———, *Monotone inductive definitions*, ***The L. E. J. Brouwer centenary symposium*** (A. S. Troelstra and D. van Dalen, editors), North-Holland, Amsterdam, 1982, pp. 77–89.

[5] H. FRIEDMAN, *Iterated inductive definitions and* Σ^1_1*-AC*, ***Intuitionism and Proof Theory*** (A. Kino et al., editors), North-Holland, Amsterdam, 1970, pp. 435–442.

[6] J.-Y. GIRARD, ***Proof theory and logical complexity.*** Vol. I, Bibliopolis, Napoli, 1987.

[7] G. JÄGER, *Iterating admissibility in proof theory*, ***Proceedings of the Herbrand symposium/Logic Colloquium '81*** (J. Stern, editor), North-Holland, Amsterdam, 1982, pp. 137–146.

[8] G. JÄGER and R. F. STÄRK, *The defining power of stratified and hierarchical logic programs*, ***Proceedings of the triennial meeting of S. I. L. F. S.*** (Viareggio, 1990) (to appear).

[9] J. W. LLOYD, ***Foundations of logic programming***, 2nd ed., Springer-Verlag, Berlin, 1987.

[10] Y. N. MOSCHOVAKIS, ***Elementary induction on abstract structures***, North-Holland, Amsterdam, 1974.

[11] B. PRIMO, ***Überlegungen zur logischen Analyse von Fixpunkt-Theorien,*** Diplomarbeit, ETH, Zürich, 1988.

[12] K. SCHÜTTE, ***Proof theory***, Springer-Verlag, Berlin, 1977.

[13] W. W. TAIT, *Normal derivability in classical logic*, ***The syntax and semantics of infinitary languages*** (J. Barwise, editor), Lecture Notes in Mathematics, vol. 72, Springer-Verlag, Berlin, 1968, pp. 204–236.

[14] G. TAKEUTI, ***Proof theory***, 2nd ed., North-Holland, Amsterdam, 1987.

INSTITUT FÜR INFORMATIK UND ANGEWANDTE MATHEMATIK
UNIVERSITÄT BERN
3012 BERN, SWITZERLAND

E-mail: jaeger@iam.unibe.ch

MATHEMATICS DEPARTMENT
EIDGENÖSSISCHE TECHNISCHE HOCHSCHULE
8092 ZURICH, SWITZERLAND