

Fixed point theories and dependent choice

Gerhard Jäger and Thomas Strahm

Abstract

In this paper we establish the proof-theoretic equivalence of (i) **ATR** and $\widehat{\text{ID}}_\omega$,
(ii) $\text{ATR}_0 + (\Sigma_1^1\text{-DC})$ and $\widehat{\text{ID}}_{<\omega^\omega}$, and (iii) $\text{ATR} + (\Sigma_1^1\text{-DC})$ and $\widehat{\text{ID}}_{<\varepsilon_0}$.

1 Introduction

The starting points for this paper are (i) Avigad's fixed point theory FP_0 (see Avigad [1]) and (ii) the recent work in metapredicativity about transfinitely iterated fixed point theories (see Jäger, Kahle, Setzer, Strahm [12]). Avigad showed that his FP_0 is equivalent to the famous theory ATR_0 introduced by Friedman and gave a proof-theoretic reduction of FP_0 to the union $\widehat{\text{ID}}_{<\omega}$ of the finitely iterated fixed point theories. In view of Feferman's proof of Hancock's conjecture in [5], this provides a new proof that ATR_0 (and FP_0) have proof-theoretic ordinal Γ_0 , the limiting number of predicative mathematics.

In ATR_0 and FP_0 complete induction on the natural numbers is restricted to sets. Adding the scheme of induction for arbitrary formulas yields a first step into metapredicativity. The corresponding theories are called **ATR** and **FP** and will be shown to be proof-theoretically equivalent to $\widehat{\text{ID}}_\omega$ and possess proof-theoretic ordinal Γ_{ε_0} .

According to, for example, Simpson [16] the Σ_1^1 axiom of choice can be proved in ATR_0 and FP_0 . This is not the case, however, for Σ_1^1 dependent choice and, as we will see, additional proof-theoretic strength is gained by adding $(\Sigma_1^1\text{-DC})$ to FP_0 and FP . More precisely, we will prove that there exist mutual interpretations of $\text{FP}_0 + (\Sigma_1^1\text{-DC})$ and $\widehat{\text{ID}}_{<\omega^\omega}$, and of $\text{FP} + (\Sigma_1^1\text{-DC})$ and $\widehat{\text{ID}}_{<\varepsilon_0}$. Using the characterization of the proof-theoretic ordinals of transfinitely iterated fixed point theories by means of a ternary φ function given in [12], this yields the proof-theoretic ordinals $\varphi 1\omega 0$ and $\varphi 1\varepsilon_0 0$ for $\text{FP}_0 + (\Sigma_1^1\text{-DC})$ and $\text{FP} + (\Pi_0^1\text{-DC})$, respectively.

It is proved in [3] that the subsystem of analysis based on arithmetic comprehension and restricted induction plus Σ_1^1 dependent choice, $(\Pi_0^1\text{-CA})_0 + (\Sigma_1^1\text{-DC})$, is proof-theoretically equivalent to the theory for iterated arithmetic comprehension $(\Pi_0^1\text{-CA})_{<\omega^\omega}$ and that the corresponding system with full formula induction, $(\Pi_0^1\text{-CA}) + (\Sigma_1^1\text{-DC})$, is equivalent to $(\Pi_0^1\text{-CA})_{<\varepsilon_0}$. Hence, the results of this paper can be seen as analogues of the latter equivalences in the sense that arithmetic comprehension is replaced by Avigad's fixed point axiom. Observe, however, that $(\Sigma_1^1\text{-DC})$ and $(\Sigma_1^1\text{-AC})$ have the same strength over $(\Pi_0^1\text{-CA})$, which is not the case if these principles are added to FP .

2 Second order fixed point theories

In this section we review the fixed point theories \mathbf{FP} and \mathbf{FP}_0 of Avigad [1] and define various dependent choice principles. In the rest of this paper we will then relate these fixed point theories with (and without) dependent choice to transfinitely iterated fixed point theories.

2.1 The theories \mathbf{FP} and \mathbf{FP}_0

As usual we let \mathcal{L}_2 denote the language of second order arithmetic. \mathcal{L}_2 includes *number variables* ($a, b, c, f, g, h, u, v, w, x, y, z, \dots$), *set variables* (U, V, W, X, Y, Z, \dots), symbols for all primitive recursive functions and relations, the symbol \in for elementhood between numbers and sets, as well as equality in both sorts. Furthermore, there is a symbol \sim for forming negative literals.¹

The *number terms* (r, s, t, \dots) of \mathcal{L}_2 are defined as usual; the *set terms* are just the set variables. Positive literals of \mathcal{L}_2 are all expressions ($s = t$), $R(s_1, \dots, s_n)$, ($s \in X$) and ($X = Y$) for R a symbol for an n -ary primitive recursive relation. The negative literals of \mathcal{L}_2 have the form $\sim E$ so that E a positive literal. We often write ($s \neq t$), ($s \notin X$) and ($X \neq Y$) instead of $\sim(s = t)$, $\sim(s \in X)$ and $\sim(X = Y)$, respectively. The formulas (A, B, C, \dots) of \mathcal{L}_2 are now generated from the positive and negative literals of \mathcal{L}_2 by closing against disjunction, conjunction, as well as existential and universal number and set quantification. The negation $\neg A$ of an \mathcal{L}_2 formula A is defined by making use of De Morgan's laws and the law of double negation. Moreover, the remaining logical connectives are abbreviated as usual.

An \mathcal{L}_2 formula is called *arithmetic* if it does not contain bound set variables (but possibly free set variables); we write Π_0^1 for the collection of these formulas. An arithmetic formula is said to be X -positive if it has no subformulas of the form ($t \notin X$) and ($X = Y$). An arithmetic X -positive \mathcal{L}_2 formula which contains at most X, Y, x, y free is called an *inductive operator form*; we let $\mathcal{A}(X, Y, x, y)$ range over such forms. Finally, \mathcal{L}_1 denotes the first order part of \mathcal{L}_2 , i.e., the sublanguage of \mathcal{L}_2 which does not contain any set variables.

The stage is now set in order to define Avigad's fixed point theory \mathbf{FP} . It is based on the usual axioms for the two sorted predicate calculus with equality in both sorts and extensionality for sets. The non-logical axioms of \mathbf{FP} comprise (i) defining axioms for all primitive recursive functions and relations, (ii) the induction scheme for arbitrary formulas of \mathcal{L}_2 , and (iii) the fixed point axioms

$$(\exists X)(\forall x)[x \in X \leftrightarrow \mathcal{A}(X, Y, x, y)]$$

for all inductive operator forms $\mathcal{A}(X, Y, x, y)$. Observe that this axiom entails arithmetic comprehension. \mathbf{FP}_0 is obtained from \mathbf{FP} by restricting complete induction on the natural numbers to sets.

¹This formulation of the language is chosen for the Tait-style reformulation of our systems in Section 5.

It is shown in Avigad [1] that FP_0 and FP are equivalent to Friedman's ATR_0 and ATR , respectively. It is established in Friedman et al. [7] and Jäger [10] that the proof-theoretic ordinal of ATR_0 is Γ_0 . Moreover, results of Friedman (cf. Simpson [15]) and Jäger [9] yield that Γ_{ε_0} is the proof-theoretic ordinal of ATR .

2.2 Dependent choice

Our main concern in the sequel is the study of our fixed point theories in the context of *arithmetic dependent choice*. Its formulation presupposes the usual way of coding an infinite sequence of sets into a single one: We assume that $\langle \cdot, \cdot \rangle$ is a primitive recursive pairing function with associated projections $(\cdot)_0$ and $(\cdot)_1$, and we write $s \in (X)_t$ instead of $\langle s, t \rangle \in X$.

Let \mathcal{K} be a collection of \mathcal{L}_2 formulas. Then we write $(\mathcal{K}\text{-DC})$ for the collection of all formulas

$$(\forall X)(\exists Y)A(X, Y) \rightarrow (\forall X)(\exists Z)[(Z)_0 = X \wedge (\forall u)A((Z)_u, (Z)_{u+1})] \quad (\mathcal{K}\text{-DC})$$

so that $A(X, Y)$ is a formula in the class \mathcal{K} . For us $(\Pi_0^1\text{-DC})$ will be most interesting. In the following proposition we show that this principle implies a (formally) more general form of dependent choice. In the literature one can find various formulations of dependent choice. Our formulation corresponds to the one e.g. in [2] with function variables replaced by set variables.

As usual we let Σ_1^1 be the collection of all \mathcal{L}_2 formulas $(\exists X)A(X)$ with $A(X)$ from Π_0^1 . Correspondingly, Π_1^1 is the collection of all formulas $(\forall X)B(X)$ with $B(X)$ from Π_0^1 . Then one has the following result, whose proof consists in a combination of folklore arguments and, therefore, will be omitted.

Proposition 1 *If $A(X, Y, u)$ is a Σ_1^1 formula, then $\text{FP}_0 + (\Pi_0^1\text{-DC})$ proves*

$$(\forall u)(\forall X)(\exists Y)A(X, Y, u) \rightarrow (\forall X)(\exists Z)[(Z)_0 = X \wedge (\forall u)A((Z)_u, (Z)_{u+1}, u)].$$

Of course the strength of FP_0 is not needed for this proposition, but we did not want to introduce further theories.

3 Transfinitely iterated fixed point theories

In this section we first present transfinitely iterated fixed point theories $\widehat{\text{ID}}_\alpha$ and $\widehat{\text{ID}}_{<\beta}$ for $\alpha < \varepsilon_0$ and $\beta \leq \varepsilon_0$. Of particular interest for us are the theories $\widehat{\text{ID}}_\omega$, $\widehat{\text{ID}}_{<\omega^\omega}$ and $\widehat{\text{ID}}_{<\varepsilon_0}$, whose relationship to second order fixed point theories with dependent choice will be studied. A general proof-theoretic treatment of the theories $\widehat{\text{ID}}_\alpha$ is given in Jäger, Kahle, Setzer and Strahm [12].

3.1 The theories $\widehat{\text{ID}}_\alpha$ for $\alpha < \varepsilon_0$

For the introduction of these theories we fix a canonical primitive recursive wellordering \prec of ordertype ε_0 . Without loss of generality we may assume that the field of \prec is the set of all natural numbers and 0 is the least element of \prec . Hence, each natural number codes an ordinal less than ε_0 . Moreover, there exist primitive recursive functions acting on these codes which correspond to the usual ordinal operations such as plus, times and exponentiation.

When working with formal first or second order theories, it is often convenient in order to simplify notation to use ordinals and ordinal operations instead of their codes and primitive recursive analogues. Then (for example) ω and $\omega + \omega$ stand for the natural numbers whose order type with respect to \prec are ω and $\omega + \omega$.

In order to formulate the theories $\widehat{\text{ID}}_\alpha$ we add to the first order language \mathcal{L}_1 a new unary relation symbol P^A for each inductive operator form $A(X, Y, x, y)$ and denote this new language by \mathcal{L}_{fix} . We write $P_s^A(t)$ for $P^A(\langle t, s \rangle)$ and $P_{\prec s}^A(t)$ for $t = \langle (t)_0, (t)_1 \rangle \wedge (t)_1 \prec s \wedge P^A(t)$.

If α is an ordinal less than ε_0 , then $\widehat{\text{ID}}_\alpha$ comprises the axioms of Peano arithmetic with the scheme of complete induction for all formulas of \mathcal{L}_{fix} as well as the fixed point axioms

$$(\forall a \prec \alpha)(\forall x)[P_a^A(x) \leftrightarrow A(P_a^A, P_{\prec a}^A, x, a)]$$

for all inductive operator forms $A(X, Y, x, y)$. We write $\widehat{\text{ID}}_{<\alpha}$ for the union of the theories $\widehat{\text{ID}}_\beta$ for $\beta < \alpha$. Of particular interest in the sequel are the theories $\widehat{\text{ID}}_\omega$, $\widehat{\text{ID}}_{<\omega^\omega}$, and $\widehat{\text{ID}}_{<\varepsilon_0}$.

3.2 Embeddings of $\widehat{\text{ID}}_{<\omega^\omega}$ and $\widehat{\text{ID}}_{<\varepsilon_0}$

In the sequel we are going to show that $\widehat{\text{ID}}_{<\omega^\omega}$ and $\widehat{\text{ID}}_{<\varepsilon_0}$ are contained in $\text{FP}_0 + (\Pi_0^1\text{-DC})$ and $\text{FP} + (\Pi_0^1\text{-DC})$, respectively. For that purpose is is convenient to assign a fundamental sequence $(\lambda[n] : n \in \mathbb{N})$ to each limit ordinal $\lambda < \varepsilon_0$ by setting for all natural numbers n :

- (i) If λ is of the form $\alpha + \omega^{\beta+1}$, then $\lambda[0] := 0$ and $\lambda[n+1] := \alpha + \omega^\beta \cdot n$.
- (ii) If λ is of the form $\alpha + \omega^\beta$ for a limit ordinal β , then $\lambda[0] := 0$ and $\lambda[n+1] := \alpha + \omega^{\beta[n]}$.

In addition, for each limit number $\lambda < \varepsilon_0$ and each natural number n we choose $\lambda^-[n]$ to denote the unique ordinal such that $\lambda[n] + \lambda^-[n] = \lambda[n+1]$. We let ℓ range over codes of limit ordinals less than ε_0 . Further, we set for all inductive operator forms $A(X, Y, x, y)$:

$$\begin{aligned} \text{Hier}_A(a, X) &:= (\forall b \prec a)(\forall x)[x \in (X)_b \leftrightarrow A((X)_b, (X)_{\prec b}, x, b)], \\ H_A(a, b, X, Y) &:= \text{Hier}_A(a, X) \rightarrow \text{Hier}_A(a+b, Y) \wedge (\forall c \prec a)[(X)_c = (Y)_c]. \end{aligned}$$

Here $(X)_{\prec b}$ denotes the set of all $s \in X$ so that $s = \langle (s)_0, (s)_1 \rangle \wedge (s)_1 \prec b$. The formula $\text{Hier}_{\mathcal{A}}(a, X)$ expresses that X represents a fixed point hierarchy with respect to \mathcal{A} below a . In the second formula, $a + b$ is supposed to denote ordinal addition in the sense of our notation system.

The next lemma is crucial for building fixed point hierarchies in $\text{FP}_0 + (\Pi_0^1\text{-DC})$ and $\text{FP} + (\Pi_0^1\text{-DC})$. Its proof illustrates the role of dependent choice in a perspicuous manner.

Lemma 2 $\text{FP}_0 + (\Pi_0^1\text{-DC})$ proves for all inductive operator forms $\mathcal{A}(X, Y, x, y)$:

$$(\forall u)(\forall X)(\exists Y)H_{\mathcal{A}}(a + \ell[u], \ell^-[u], X, Y) \rightarrow (\forall X)(\exists Y)H_{\mathcal{A}}(a, \ell, X, Y).$$

Proof. In the following we work informally in $\text{FP}_0 + (\Pi_0^1\text{-DC})$. We can readily apply dependent choice in the form of Proposition 1 to the premise of our assertion and yield

$$(\forall X)(\exists Z)[(Z)_0 = X \wedge (\forall u)H_{\mathcal{A}}(a + \ell[u], \ell^-[u], (Z)_u, (Z)_{u+1})]. \quad (1)$$

Now let a set X be given which satisfies $\text{Hier}_{\mathcal{A}}(a, X)$. Then (1) guarantees the existence of a choice set Z with $(Z)_0 = X$, so that an easy inductive argument shows

$$(\forall u)[\text{Hier}_{\mathcal{A}}(a + \ell[u], (Z)_u) \wedge (\forall b \prec a + \ell[u])((Z)_u)_b = ((Z)_{u+1})_b]. \quad (2)$$

If we take $Y := \{y : (\exists u)((y)_1 \prec a + \ell[u] \wedge y \in (Z)_u)\}$, then one immediately checks that $\text{Hier}_{\mathcal{A}}(a + \ell, Y)$. This finishes our argument. \square

Lemma 3 Let $\mathcal{A}(X, Y, x, y)$ be an inductive operator form and let $B(b)$ denote the \mathcal{L}_2 formula

$$(\forall a)(\forall X)(\exists Y)H_{\mathcal{A}}(a, b, X, Y).$$

Then $\text{FP}_0 + (\Pi_0^1\text{-DC})$ proves $B(\omega^k)$ for each natural number k , and $\text{FP} + (\Pi_0^1\text{-DC})$ proves $B(\alpha)$ for each ordinal α less than ε_0 .

Proof. Let us first sketch the argument for $\text{FP}_0 + (\Pi_0^1\text{-DC})$. We show by metamathematical induction on k that $\text{FP}_0 + (\Pi_0^1\text{-DC}) \vdash B(\omega^k)$. The case $k = 0$ follows by a simple application of the fixed point axiom. Let us now assume that $k > 0$; then our assertion holds by the previous lemma for $\ell = \omega^k$ and the induction hypothesis. The claim of our lemma for $\text{FP} + (\Pi_0^1\text{-DC})$ is easily handled by observing that in the presence of full induction, transfinite induction is available along initial segments of \prec for arbitrary \mathcal{L}_2 formulas. \square

From this lemma it is immediate that $\text{FP}_0 + (\Pi_0^1\text{-DC})$ proves $(\exists X)\text{Hier}_{\mathcal{A}}(\omega^k, X)$ for each natural number k ; correspondingly, $\text{FP} + (\Pi_0^1\text{-DC})$ proves $(\exists X)\text{Hier}_{\mathcal{A}}(\alpha, X)$ for each ordinal α less than ε_0 . Summing up, we have established an embedding of $\widehat{\text{ID}}_{<\omega^\omega}$ in $\text{FP}_0 + (\Pi_0^1\text{-DC})$ and $\widehat{\text{ID}}_{<\varepsilon_0}$ in $\text{FP} + (\Pi_0^1\text{-DC})$.

Theorem 4 $\widehat{\text{ID}}_{<\omega^\omega}$ can be embedded into $\text{FP}_0 + (\Pi_0^1\text{-DC})$, and $\widehat{\text{ID}}_{<\varepsilon_0}$ can be embedded into $\text{FP} + (\Pi_0^1\text{-DC})$; moreover, both embeddings preserve \mathcal{L}_1 theorems.

Later we will give an embedding of FP into $\widehat{\text{ID}}_\omega$. We do not know whether a direct interpretation of $\widehat{\text{ID}}_\omega$ in FP is possible. However, it is shown in [12] that the proof-theoretic ordinal of $\widehat{\text{ID}}_\omega$ is Γ_{ε_0} and, hence, $\widehat{\text{ID}}_\omega$ and FP are proof-theoretically equivalent according to our discussion in Section 2.1. In particular, FP and $\widehat{\text{ID}}_\omega$ prove the same arithmetic sentences.

4 Transfinitely iterated theories for self-reflecting truth

In this section we introduce transfinitely iterated theories for self-reflecting truth SRT_α and $\text{SRT}_{<\beta}$ for $\alpha < \varepsilon_0$ and $\beta \leq \varepsilon_0$. It will be immediate that SRT_α can be modeled in $\widehat{\text{ID}}_\alpha$. In the next section, theories for self-reflecting truth will be used in order to establish reductions of FP , $\text{FP}_0 + (\Pi_0^1\text{-DC})$, and $\text{FP} + (\Pi_0^1\text{-DC})$ to $\widehat{\text{ID}}_\omega$, $\widehat{\text{ID}}_{<\omega^\omega}$ and $\widehat{\text{ID}}_{<\varepsilon_0}$, respectively. Transfinitely iterated theories for self-reflecting truth have previously served as a convenient tool for carrying through wellordering proofs for $\widehat{\text{ID}}_\alpha$ in Jäger, Kahle, Setzer and Strahm [12].

The theories SRT_α are formulated in the language \mathcal{L}_{srt} , which extends \mathcal{L}_1 by two binary relation symbols T (for “true”) and F (for “false”). The terms of \mathcal{L}_{srt} are just the \mathcal{L}_1 terms; the formulas of \mathcal{L}_{srt} are given in a straightforward manner, taking into account the new atoms $T(s, t)$ (“ t is true on level s ”) and $F(s, t)$ (“ t is false on level s ”) as well as their complementations. In the following we often write $T_s(t)$ and $F_s(t)$ instead of $T(s, t)$ and $F(s, t)$, respectively.

If α is an ordinal less than ε_0 , then we obtain the sublanguage $\mathcal{L}_{\text{srt}}^\alpha$ of \mathcal{L}_{srt} by restricting atoms of the form $T_s(t)$ and $F_s(t)$ to closed terms s with value less than or equal to α (in the sense of our notation system). Hence, variables are not allowed in the first place of the relation symbols T and F in $\mathcal{L}_{\text{srt}}^\alpha$ formulas.

In order to describe transfinitely iterated truth theories below, we presuppose a standard Gödelization of the languages $\mathcal{L}_{\text{srt}}^\alpha$, uniformly in $\alpha < \varepsilon_0$. In particular, we have Gödelnumbers $\ulcorner t \urcorner$ and $\ulcorner A \urcorner$ for each \mathcal{L}_1 term t and each $\mathcal{L}_{\text{srt}}^\alpha$ formula A . Moreover, we will use the following \mathcal{L}_1 definable functions and predicates on Gödelnumbers: $\text{CTer}(x)$ (“ x is closed term of \mathcal{L}_1 ”); $\text{For}_n(f, a)$ (“ f is an $\mathcal{L}_{\text{srt}}^\alpha$ formula with at most n free variables”); $\text{Atm}(f, a)$ (“ f is a positive literal of $\mathcal{L}_{\text{srt}}^\alpha$ ”); $\text{num}(x)$ (“the x th numeral of \mathcal{L}_1 ”); $\text{val}(z)$ (“the value of the closed term z ”); $\text{neg}(f)$ (“the negation of the positive atom f ”); $\text{and}(f, g)$ (“the conjunction of f and g ”); $\text{or}(f, g)$ is defined analogously; $\text{all}(x, f)$ (“the universal quantification of f with respect to the x th variable of \mathcal{L}_1 ”); $\text{ext}(x, f)$ for existential quantification is defined analogously. In the following we write $\text{Sen}(f, a)$ instead of $\text{For}_0(f, a)$ and \dot{x} instead of $\text{num}(x)$. Finally, if R is an \mathcal{L}_1 relation symbol, $\text{CTer}(x)$ and $\text{CTer}(a)$, then expressions like $\ulcorner R(x) \urcorner$, $\ulcorner T_a(x) \urcorner$, and $\ulcorner F_a(x) \urcorner$ have their obvious meaning.

We conclude our tedious but standard preliminary remarks by adopting some conventions concerning substitution. If $For_n(f)$ and $CTer(x_1), \dots, CTer(x_n)$, then $f(x_1, \dots, x_n)$ denotes (the code of) the formula which is obtained from f by substituting the i th free variable of f by x_i ; in particular, $Sen(f(x_1, \dots, x_n))$. Similarly, if the free variables of A are among x_1, \dots, x_n , then $\lceil A(x_1, \dots, x_n) \rceil$ is an abbreviation for $\lceil A \rceil(x_1, \dots, x_n)$.

If α is an ordinal less than ε_0 , then the theory SRT_α for α times iterated self-reflecting truth comprises the axioms of Peano arithmetic with the scheme of complete induction for all formulas of \mathcal{L}_{srt} as well as the following axioms.

I. Atomic truth.

- (1) For each relation symbol R of \mathcal{L}_1 :

$$\begin{aligned} & CTer(x_1) \wedge \dots \wedge CTer(x_n) \wedge a \prec \alpha \rightarrow \\ & [T_a(\lceil R(x_1, \dots, x_n) \rceil) \leftrightarrow R(val(x_1), \dots, val(x_n))] \wedge \\ & [F_a(\lceil R(x_1, \dots, x_n) \rceil) \leftrightarrow \neg R(val(x_1), \dots, val(x_n))]. \end{aligned}$$

- (2) $CTer(x) \wedge CTer(b) \wedge val(b) \prec a \prec \alpha \rightarrow$
 $[T_a(\lceil T_b(x) \rceil) \leftrightarrow T_{val(b)}(val(x))] \wedge [T_a(\lceil F_b(x) \rceil) \leftrightarrow F_{val(b)}(val(x))] \wedge$
 $[F_a(\lceil T_b(x) \rceil) \leftrightarrow \neg T_{val(b)}(val(x))] \wedge [F_a(\lceil F_b(x) \rceil) \leftrightarrow \neg F_{val(b)}(val(x))].$

II. Composed truth.

- (3) $Atm(f, a) \wedge a \prec \alpha \rightarrow$

$$[T_a(neg(f)) \leftrightarrow F_a(f)] \wedge [F_a(neg(f)) \leftrightarrow T_a(f)],$$

- (4) $Sen(f, a) \wedge Sen(g, a) \wedge a \prec \alpha \rightarrow$

$$[T_a(and(f, g)) \leftrightarrow T_a(f) \wedge T_a(g)] \wedge [F_a(and(f, g)) \leftrightarrow F_a(f) \vee F_a(g)],$$

- (5) dual clauses for disjunction,

- (6) $Sen(all(v, f), a) \wedge a \prec \alpha \rightarrow$

$$[T_a(all(v, f)) \leftrightarrow (\forall x)T_a(f(\dot{x}))] \wedge [F_a(all(v, f)) \leftrightarrow (\exists x)F_a(f(\dot{x}))],$$

- (7) dual clauses for existential quantification.

III. Self reflecting truth.

- (8) $CTer(x) \wedge CTer(a) \wedge val(a) \prec \alpha \rightarrow$

$$\begin{aligned} & [T_{val(a)}(\lceil T_a(x) \rceil) \leftrightarrow T_{val(a)}(val(x))] \wedge [T_{val(a)}(\lceil F_a(x) \rceil) \leftrightarrow F_{val(a)}(val(x))] \wedge \\ & [F_{val(a)}(\lceil T_a(x) \rceil) \leftrightarrow F_{val(a)}(val(x))] \wedge [F_{val(a)}(\lceil F_a(x) \rceil) \leftrightarrow T_{val(a)}(val(x))]. \end{aligned}$$

This finishes the description of the theories SRT_α for α less than ε_0 . We write $\text{SRT}_{<\alpha}$ for the union of the theories SRT_β for $\beta < \alpha$. It is completely straightforward to model the theory SRT_α for α times iterated self-reflecting truth by means of a fixed point hierarchy of length α ; for a similar argument in the case of non-iterated truth theories, the curious reader is advised to consult Feferman [6]. Therefore, we can state the following proposition without proof.

Proposition 5 *We have for each ordinal α less than ε_0 that SRT_α can be embedded into $\widehat{\text{ID}}_\alpha$; moreover, \mathcal{L}_1 theorems are preserved under this embedding.*

Let us now turn to some crucial consequences of theories for self-reflecting truth. We first observe that T_a and F_a are complementary on sentences of level less than a . This is easily established by (formal) induction on the build up of such sentences.

Proposition 6 *We have for each ordinal α less than ε_0 that SRT_α proves*

$$(\forall a, b, f)[\text{Sen}(f, b) \wedge b \prec a \prec \alpha \rightarrow (T_a(f) \leftrightarrow \neg F_a(f))].$$

If a is a variable of \mathcal{L}_1 , then we call an \mathcal{L}_{srt} formula (T_a, F_a) *positive*, if it is built from \mathcal{L}_1 literals which do not contain a , atoms of the form $T_a(t), F_a(t)$ so that a does not occur in t , and by closing against conjunction, disjunction as well as quantification with respect to variables *different* from a . Hence, in a (T_a, F_a) positive formula only positive occurrences of T_a and F_a are allowed, and in addition, the variable a is used as an (unquantified) level variable only. If $A(a, x_1, \dots, x_n)$ is such a (T_a, F_a) -positive formula with level variable a and additional free variables among x_1, \dots, x_n , then there is associated uniformly in a valuation $\dot{a}, \dot{x}_1, \dots, \dot{x}_n$ for a, x_1, \dots, x_n a Gödelnumber $\Gamma A(\dot{a}, \dot{x}_1, \dots, \dot{x}_n)^\Gamma$ of the corresponding $\mathcal{L}_{\text{srt}}^a$ formula. The following proposition is characteristic for self reflective truth theories (cf. e.g. Cantini [4]) and is easily established by induction on the complexity of A .

Proposition 7 *Assume that $A(a, x_1, \dots, x_n)$ is a (T_a, F_a) positive formula with at most a, x_1, \dots, x_n free. Then we have for each ordinal α less than ε_0 that SRT_α proves*

$$(\forall a \prec \alpha)(\forall x_1, \dots, x_n)[T_a(\Gamma A(\dot{a}, \dot{x}_1, \dots, \dot{x}_n)^\Gamma) \leftrightarrow A(a, x_1, \dots, x_n)].$$

There is a natural notion of *ramified set* in our truth theoretic framework, namely, sets of natural numbers are understood as propositional functions. More formally, we define the notions “ f is a set of level a ”, $f \in S_a$, and “ x is an element of a set f of level a ”, $x \in_a f$, as follows.

$$\begin{aligned} f \in S_a &:= \text{For}_1(f, a) \wedge (\forall x)(T_a(f(\dot{x})) \leftrightarrow \neg F_a(f(\dot{x}))), \\ x \in_a f &:= T_a(f(\dot{x})). \end{aligned}$$

In the sequel we often write $x \in f$ instead of $x \in_a f$ if it is clear from the context that f is a set of level a . Moreover, if $A(X_1, \dots, X_n)$ is an arithmetic \mathcal{L}_2 formula

with its free set variables among X_1, \dots, X_n , then we write $A(f_1, \dots, f_n)$ for the \mathcal{L}_{srt} formula which is obtained from A by replacing each atom of the form $t \in X_i$ by $t \in f_i$ for $1 \leq i \leq n$.

We finish this section by stating a (uniform) version of the second recursion theorem, which is available in our truth theoretic framework. Its proof is standard and makes use of Proposition 7 and standard diagonalization techniques, cf. Cantini [4] for a similar argument.

Proposition 8 *Let $\mathcal{A}(X, Y, x, y)$ be an inductive operator form and α a limit ordinal less than ε_0 . Then there exists a primitive recursive function $\varphi_{\mathcal{A}}$, so that SRT_α proves*

$$(\forall a \prec \alpha)(\forall y)(\forall f \in S_a)[\varphi_{\mathcal{A}}(f, y) \in S_{a+1} \wedge (\forall x)(x \in \varphi_{\mathcal{A}}(f, y) \leftrightarrow \mathcal{A}(\varphi_{\mathcal{A}}(f, y), f, x, y))].$$

It is immediate that this proposition provides for an interpretation of the fixed point axiom in SRT_α and hence $\widehat{\text{ID}}_\alpha$ for limit ordinals α .

5 Reductions of FP , $\text{FP}_0 + (\Pi_0^1\text{-DC})$, and $\text{FP} + (\Pi_0^1\text{-DC})$

In this section we compute upper bounds of the theories FP , $\text{FP}_0 + (\Pi_0^1\text{-DC})$, and $\text{FP} + (\Pi_0^1\text{-DC})$ in terms of transfinitely iterated fixed point theories. In particular, we show that FP can be embedded into $\widehat{\text{ID}}_\omega$; moreover, we provide suitable partial cut elimination and asymmetric interpretation arguments, thus establishing conservativity with respect to arithmetic statements of $\text{FP}_0 + (\Pi_0^1\text{-DC})$ and $\text{FP} + (\Pi_0^1\text{-DC})$ over $\widehat{\text{ID}}_{<\omega^\omega}$ and $\widehat{\text{ID}}_{<\varepsilon_0}$, respectively.

5.1 Embedding FP into $\widehat{\text{ID}}_\omega$

We obtain a straightforward embedding of FP into SRT_ω by letting range the set variables of FP over sets of level $< \omega$ in the sense of SRT_ω . Equality between sets is interpreted as extensional equality according to this translation. Proposition 8 guarantees the translation of the fixed point axiom of FP and, hence, the embedding of FP into SRT_ω is established.

Theorem 9 *FP can be embedded into SRT_ω and, hence, $\widehat{\text{ID}}_\omega$; moreover, \mathcal{L}_1 theorems are preserved under this embedding.*

As we have mentioned at the end of Section 3.2, we do not know yet whether a direct converse embedding exists. Nevertheless, thanks to [9, 12], FP and $\widehat{\text{ID}}_\omega$ are proof-theoretically equivalent with proof-theoretic ordinal Γ_{ε_0} .

5.2 Reduction of $\text{FP}_0 + (\Pi_0^1\text{-DC})$ to $\text{SRT}_{<\omega^\omega}$

In the sequel we provide a Tait-style reformulation \mathbf{T} of $\text{FP}_0 + (\Pi_0^1\text{-DC})$, which allows us to prove partial cut elimination. Quasinormal derivations of \mathbf{T} are then reduced to $\text{SRT}_{<\omega^\omega}$ by means of an asymmetric interpretation argument.

We let Γ, Λ, \dots range over finite sets of \mathcal{L}_2 formulas; we often write (for example) Γ, A for the union of Γ and $\{A\}$. The Tait-calculus T is an extension of the classical Tait-calculus (cf. [14]) by the non-logical axioms of $\mathsf{FP}_0 + (\Pi_0^1\text{-DC})$. It comprises the following axioms and rules of inference.

I. **Axioms.** For all finite sets Γ of \mathcal{L}_2 formulas, all Σ_1^1 formulas A and all Σ_1^1 formulas B which are axioms of FP_0 :

$$\Gamma, \neg A, A \quad \text{and} \quad \Gamma, B.$$

II. **Propositional and quantifier rules.** These include the usual Tait-style inference rules for the propositional connectives and all sorts of quantifiers.

III. Π_0^1 **dependent choice.** For all finite sets Γ of \mathcal{L}_2 formulas, all Π_0^1 formulas $A(X, Y)$ and all set variables U :

$$\frac{\Gamma, (\forall X)(\exists Y)A(X, Y)}{\Gamma, (\exists Z)[(Z)_0 = U \wedge (\forall u)A((Z)_u, (Z)_{u+1})]}.$$

IV. **Cut rules.** For all finite sets Γ of \mathcal{L}_2 formulas and all \mathcal{L}_2 formulas A :

$$\frac{\Gamma, A \quad \Gamma, \neg A}{\Gamma}.$$

The formulas A and $\neg A$ are the cut formulas of this cut.

Observe that in our formulation the fixed point axioms of FP_0 are Σ_1^1 formulas so that these are the main formulas of axioms of T according to I. The notion $\mathsf{T} \vdash^n \Gamma$ is used to express that Γ is provable in T by a proof of depth less than or equal to n ; we write $\mathsf{T} \vdash_\star^n \Gamma$ if Γ is provable in T by a proof of depth less than or equal to n so that all its cut formulas are Σ_1^1 or Π_1^1 formulas. In addition, $\mathsf{T} \vdash \Gamma$ or $\mathsf{T} \vdash_\star \Gamma$ means that there exists a natural number n so that $\mathsf{T} \vdash^n \Gamma$ or $\mathsf{T} \vdash_\star^n \Gamma$, respectively.

One immediately observes that the main formulas of all non-logical axioms and rules of T are Σ_1^1 formulas. Hence, the following partial cut elimination theorem is a matter of routine.

Proposition 10 (Partial cut elimination for T) *We have for all finite sets Γ of \mathcal{L}_2 formulas:*

$$\mathsf{T} \vdash \Gamma \implies \mathsf{T} \vdash_\star \Gamma.$$

Furthermore, the axioms and rules of T are tailored so that $\mathsf{FP}_0 + (\Pi_0^1\text{-DC})$ can be easily embedded into T .

Proposition 11 *We have for all \mathcal{L}_2 formulas A :*

$$\mathsf{FP}_0 + (\Pi_0^1\text{-DC}) \vdash A \implies \mathsf{T} \vdash A.$$

Combining the previous two propositions yields the following corollary, which will be used for the proof of Theorem 13 below.

Corollary 12 *If the \mathcal{L}_2 formula A is provable in $\text{FP}_0 + (\Pi_0^1\text{-DC})$, then there exists a natural number n so that $\text{T} \vdash_*^n A$.*

In a next step we want to provide an *asymmetric interpretation* of T into $\text{SRT}_{<\omega^\omega}$. For that purpose, we need the following notation. Assume that $A(\vec{X})$ is an \mathcal{L}_2 formula with all its free set variables among \vec{X} . If the \vec{X} are replaced by sets \vec{x} in the sense of our truth theories, then we let $A(\vec{x})^{(a,b)}$ denote the \mathcal{L}_{srt} formula which is obtained from A by letting range the universal set quantifiers over S_a , and the existential set quantifiers over S_b . Similarly, if $\Gamma(\vec{X})$ is a finite set of \mathcal{L}_2 formulas, then $\Gamma(\vec{x})^{(a,b)}$ denotes the disjunction of the formulas $A(\vec{x})^{(a,b)}$ for A in Γ .

Theorem 13 (Asymmetric interpretation of T into $\text{SRT}_{<\omega^\omega}$) *Assume that Γ is a finite set of \mathcal{L}_2 formulas so that $\text{T} \vdash_*^n \Gamma$ for some natural number n . Then we have for all natural numbers $k > n$:*

$$\text{SRT}_{\omega^k} \vdash (\forall a \prec \omega^k)(\forall \vec{x} \in S_a)\Gamma(\vec{x})^{(a,a+\omega^n)}.$$

Proof. This theorem is proved by induction on n . In the following we tacitly use the usual persistency properties of our asymmetric interpretation. If Γ is a fixed point axiom, then the claim is immediate by Proposition 8; the remaining axioms of T are easily dealt with, too. If Γ is the conclusion of a propositional or quantifier rule, then the claim is immediate by the induction hypothesis. If Γ is the conclusion of a cut rule, then one proceeds as in similar asymmetric interpretations, cf. [8, 11, 13]. Hence, the only critical case comes up if Γ is the conclusion of an application of the dependent choice rule. Then there exists a natural number $n_0 < n$ and an Π_0^1 formula $A(X, Y)$ so that

$$\text{T} \vdash_*^{n_0} \Gamma, (\forall X)(\exists Y)A(X, Y). \quad (1)$$

Without loss of generality we may assume that $n_0 > 0$. Let us now choose a $k > n$ and work informally in SRT_{ω^k} . Then the induction hypothesis applied to (1) yields

$$(\forall a \prec \omega^k)(\forall \vec{x} \in S_a)[\Gamma(\vec{x})^{(a,a+\omega^{n_0})} \vee (\forall f \in S_a)(\exists g \in S_{a+\omega^{n_0}})A(f, g, \vec{x})]. \quad (2)$$

Let us now choose an arbitrary $a \prec \omega^k$ and sets $\vec{x} \in S_a$. We have that \vec{x} contains a parameter $y \in S_a$ which is the initial set in the conclusion of the dependent choice rule. If $\Gamma(\vec{x})^{(a,a+\omega^{n_0})}$ holds, then we are done by persistency. Otherwise we obtain from (2) that

$$(\forall f \in S_a)(\exists g \in S_{a+\omega^{n_0}})A(f, g, \vec{x}). \quad (3)$$

Depending on a and a natural number u we set $a(u) := a + (\omega^{n_0} \cdot u)$; further, we define a set $\varphi(u) \in S_{a(u)+1}$ by primitive recursion on u as follows. $\varphi(0)$ is just y and $\varphi(u+1)$ is the set in $S_{a(u+1)+1}$ which is given by the following formula $B(x)$, i.e. $\varphi(u+1)$ is the Gödelnumber of $B(x)$:

$$B(x) := (\exists z)[z = (\mu g \in S_{a(u+1)})A(\varphi(u), g) \wedge T_{a(u+1)}(z(\dot{x}))].$$

By making use of (3) one readily verifies that the following are true:

$$(\forall u)(\varphi(u) \in S_{a(u)+1}), \quad (4)$$

$$(\forall u)A(\varphi(u), \varphi(u+1), \vec{x}). \quad (5)$$

If we now choose h as the Gödelnumber of the formula $C(x)$ given by

$$C(x) := T_{a+\omega^n}(\varphi((x)_1)((x)_0)), \quad (6)$$

then we have by (4) that $h \in S_{a+\omega^n}$, and indeed we get

$$(h)_0 = y \wedge (\forall u)A((h)_u, (h)_{u+1}, \vec{x}). \quad (7)$$

Here equality and projection of sets are understood in the obvious way. But (7) readily yields $\Gamma(\vec{x})^{(a,a+\omega^n)}$, thus confirming our claim. This concludes the treatment of dependent choice and indeed finishes the proof of our asymmetric interpretation theorem. \square

Together with Theorem 4, Corollary 12 and Proposition 5 we thus have established the following corollary.

Corollary 14 *We have that $\text{FP}_0 + (\Pi_0^1\text{-DC})$ and $\widehat{\text{ID}}_{<\omega^\omega}$ prove the same arithmetic statements.*

This finishes our characterization of $\text{FP}_0 + (\Pi_0^1\text{-DC})$ in terms of transfinitely iterated fixed point theories. The argument for the theory $\text{FP} + (\Pi_0^1\text{-DC})$ is similar, and in the following let us briefly sketch the modifications which are necessary in the presence of full induction.

First of all, one provides an infinitary Tait-style reformulation T_∞ of $\text{FP} + (\Pi_0^1\text{-DC})$; T_∞ is essentially T plus ω rule. The latter rule is used in order to prove full formula induction as usual. T_∞ enjoys partial cut elimination in the same way as T ; the difference is that we now end up with quasinormal derivations of length less than ε_0 . Such derivations are then treated by means of an asymmetric interpretation as in Theorem 13 above, and the truth levels used are now bounded below ε_0 . Standard careful formalization of the whole procedure indeed yields conservativity with respect to arithmetic statements of $\text{FP} + (\Pi_0^1\text{-DC})$ over $\widehat{\text{ID}}_{<\varepsilon_0}$. Together with Theorem 4 we have thus established the following theorem.

Theorem 15 *We have that $\text{FP} + (\Pi_0^1\text{-DC})$ and $\widehat{\text{ID}}_{<\varepsilon_0}$ prove the same arithmetic statements.*

Our results also answer a question asked by Simpson in [15] concerning the proof-theoretic ordinals of the second order theories $\Sigma_1^1\text{-Tl}_0 + \Pi_1^1\text{-Tl}_0$ as well as $\Sigma_1^1\text{-Tl} + \Pi_1^1\text{-Tl}$. Without going into details we just repeat that $\Sigma_1^1\text{-Tl}_0$ is a system of second order arithmetic with arithmetic comprehension, induction on the natural numbers restricted to sets, and bar induction for Σ_1^1 formulas. $\Sigma_1^1\text{-Tl}$, on the other hand, is

$\Sigma_1^1\text{-Tl}_0$ plus full induction on the natural numbers. The systems $\Pi_1^1\text{-Tl}_0$ and $\Pi_1^1\text{-Tl}$ are defined analogously. For details and further notation cf. [15].

The proof-theoretic characterization of $\Pi_1^1\text{-Tl}_0$ and $\Pi_1^1\text{-Tl}$ is given in Simpson [15] by showing that the theories $(\Pi_0^1\text{-CA})_0 + (\Sigma_1^1\text{-DC})$ and $\Pi_1^1\text{-Tl}_0$ prove the same theorems. With respect to Σ_1^1 bar induction he shows that the theories $\Sigma_1^1\text{-Tl}_0$ and ATR_0 plus Σ_1^1 induction on the natural numbers prove the same theorems. It is left open, however, what the proof-theoretic ordinals of $\Sigma_1^1\text{-Tl}_0 + \Pi_1^1\text{-Tl}_0$ and $\Sigma_1^1\text{-Tl} + \Pi_1^1\text{-Tl}$ are. These ordinals are now immediately available by Simpson's results, the results of this paper and by [12].

Corollary 16 (i) *We have that $\Sigma_1^1\text{-Tl}_0 + \Pi_1^1\text{-Tl}_0$ and $\widehat{\text{ID}}_{<\omega^\omega}$ prove the same arithmetic statements.*

(ii) *We have that $\Sigma_1^1\text{-Tl} + \Pi_1^1\text{-Tl}$ and $\widehat{\text{ID}}_{<\varepsilon_0}$ prove the same arithmetic statements.*

6 Conclusion

In this paper we have related second order fixed point theories with (and without) dependent choice to transfinitely iterated first order fixed point theories. By Avigad's results [1] this actually also yields a correspondence between extensions of ATR with dependent choice and transfinitely iterated fixed point theories.

In Jäger, Kahle, Setzer and Strahm [12] a thorough ordinal analysis of the theories $\widehat{\text{ID}}_\alpha$ is carried through, and the proof-theoretic ordinals of transfinitely iterated fixed point theories are characterized by means of a ternary φ function (a straightforward generalization of the binary Veblen function). Hence, all the proof-theoretic ordinals of the theories studied here are available by [12].

The following theorem summarizes the results of this article, including the corresponding proof-theoretic ordinals. For completeness, we display the relevant theories without dependent choice as well.

Theorem 17 *We have the following proof-theoretic equivalences:*

- (i) $\text{FP}_0 \equiv \text{ATR}_0 \equiv \widehat{\text{ID}}_{<\omega}$.
- (ii) $\text{FP} \equiv \text{ATR} \equiv \widehat{\text{ID}}_\omega$.
- (iii) $\text{FP}_0 + (\Pi_0^1\text{-DC}) \equiv \text{ATR}_0 + (\Pi_0^1\text{-DC}) \equiv \widehat{\text{ID}}_{<\omega^\omega} \equiv \Sigma_1^1\text{-Tl}_0 + \Pi_1^1\text{-Tl}_0$.
- (iv) $\text{FP} + (\Pi_0^1\text{-DC}) \equiv \text{ATR} + (\Pi_0^1\text{-DC}) \equiv \widehat{\text{ID}}_{<\varepsilon_0} \equiv \Sigma_1^1\text{-Tl} + \Pi_1^1\text{-Tl}$.

The corresponding four proof-theoretic ordinals are Γ_0 , Γ_{ε_0} , $\varphi 1\omega 0$, and $\varphi 1\varepsilon_0 0$.

We mention that the above equivalences hold at least for arithmetic statements. Theories based on the second order fixed point axiom are even *logically* equivalent to their relatives based on arithmetic transfinite recursion, cf. [1].

References

- [1] AVIGAD, J. On the relationship between ATR_0 and $\widehat{ID}_{<\omega}$. *Journal of Symbolic Logic* 61, 3 (1996), 768–779.
- [2] BUCHHOLZ, W., FEFERMAN, S., POHLERS, W., AND SIEG, W. *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*, vol. 897 of *Lecture Notes in Mathematics*. Springer, Berlin, 1981.
- [3] CANTINI, A. On the relationship between choice and comprehension principles in second order arithmetic. *Journal of Symbolic Logic* 51 (1986), 360–373.
- [4] CANTINI, A. Notes on formal theories of truth. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 35 (1989), 97–130.
- [5] FEFERMAN, S. Iterated inductive fixed-point theories: application to Hancock’s conjecture. In *The Patras Symposium*, G. Metakides, Ed. North Holland, Amsterdam, 1982, pp. 171–196.
- [6] FEFERMAN, S. Reflecting on incompleteness. *Journal of Symbolic Logic* 56, 1 (1991), 1–49.
- [7] FRIEDMAN, H., MCALOON, K., AND SIMPSON, S. A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis. In *Patras Symposium*, G. Metakides, Ed. North Holland, Amsterdam, 1982, pp. 197–230.
- [8] JÄGER, G. Beweistheorie von KPN . *Archiv für mathematische Logik und Grundlagenforschung* 20 (1980), 53–64.
- [9] JÄGER, G. Theories for iterated jumps, 1980. Handwritten notes.
- [10] JÄGER, G. The strength of admissibility without foundation. *The Journal of Symbolic Logic* 49, 3 (1984), 867–879.
- [11] JÄGER, G. Fixed points in Peano arithmetic with ordinals. *Annals of Pure and Applied Logic* 60, 2 (1993), 119–132.
- [12] JÄGER, G., KAHLE, R., SETZER, A., AND STRAHM, T. The proof-theoretic analysis of transfinitely iterated fixed point theories. *Journal of Symbolic Logic*. To appear.
- [13] SCHÜTTE, K. *Proof Theory*. Springer, Berlin, 1977.
- [14] SCHWICHTENBERG, H. Proof theory: Some applications of cut-elimination. In *Handbook of Mathematical Logic*, J. Barwise, Ed. North Holland, Amsterdam, 1977, pp. 867–895.

- [15] SIMPSON, S. G. Σ_1^1 and Π_1^1 transfinite induction. In *Logic Colloquium '80*, D. van Dalen, D. Lascar, and J. Smiley, Eds. North Holland, Amsterdam, 1982, pp. 239–253.
- [16] SIMPSON, S. G. Subsystems of second order arithmetic. Tech. rep., Pennsylvania State University, 1986. Chapters II–V and VII.

Address

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

March 17, 1998