Some set-theoretic reduction principles

Michael Bärtschi and Gerhard Jäger

Abstract In this article we study several reduction principles in the context of Simpson's set theory ATR_0^S and Kripke-Platek set theory KP (with infinity). Since ATR_0^S is the set-theoretic version of ATR_0 there is a direct link to second order arithmetic and the results for reductions over ATR_0^S are as expected and more or less straightforward. However, over KP we obtain several interesting new results and are lead to some open questions.

Dedicated to Peter Schroeder-Heister

1 Introduction

Peter Schroeder-Heister has been interested in the foundations of inference for many decades, most prominently in connection with a program that he baptized "proof-theoretic semantics". Though not related to this program in the strict sense, the work presented here is in direct connection to a talk given by the second author at the *Third Tübingen Conference on Proof-Theoretic Semantics* in 2019. Proof theory is the conceptual link between foundational questions considered under the heading of proof-theoretic semantics and our research on subsystems of second order arithmetic and set theory.

About terminology: Let $\mathcal{K}$ be a class of formulas of second order arithmetic. What Simpson calls the $\mathcal{K}$ separation principle in second order arithmetic is the collection of all

Michael Bärtschi
Institute of Computer Science, University of Bern, Neubrückstrasse 10, 3012 Bern, Switzerland, e-mail: michael.baertschi@inf.unibe.ch

Gerhard Jäger
Institute of Computer Science, University of Bern, Neubrückstrasse 10, 3012 Bern, Switzerland, e-mail: gerhard.jaeger@inf.unibe.ch

$$\neg\exists i(\varphi[i] \land \psi[i]) \rightarrow \exists Z \forall i(\varphi[i] \rightarrow i \in Z \rightarrow \neg\psi[i]),$$

where $\varphi[i]$ and $\psi[i]$ are formulas from $\mathcal{K}$. In Simpson [16] various such separation principles have been studied. They play an interesting role in reverse mathematics and are equivalent – over a weak base theory – to certain comprehension principles. In particular, it is shown in [16] that ACA_0 plus Σ^1_1 separation (Σ^1_1-Sep) is equivalent to the famous theory ATR_0 of arithmetical transfinite recursion.

However, this form of separation must not be confused with separation in set theory. There, $\mathcal{K}$ separation for a class $\mathcal{K}$ of formulas of set theory consists of all assertions

$$\forall x \exists y(y = \{z \in x : \varphi[z]\})$$

with $\varphi[z]$ ranging over $\mathcal{K}$. In order to avoid this conflict of notation we decided to call "reduction" what Simpson calls "separation"; see Definition 3. Thus we can use the same terminology in second order arithmetic and set theory.

In this article we consider several reduction principles, analog to Simpson's separation principles, though in the context of his set theory ATR^S_0 and in the context of Kripke-Platek set theory KP (with infinity). Since ATR^S_0 is the set-theoretic version of ATR_0 the results for reductions over ATR^S_0 are as expected and more or less straightforward. However, over KP we obtain several interesting results and are lead to some open questions.

This article begins with a review of separation principles in second order arithmetic – now, of course, under the new term "reduction principles" – and some important equivalences to comprehension principles. Then we have a section in which some basics about the theory ATR^S_0 and its extension ATR^S are presented before we address Kripke-Platek set theory KP and its relationship to ATR^S_0. Finally, we turn to Σ_1 reduction (Σ_1-Red) and Π_1 reduction (Π_1-Red). The respective strengths of (Σ_1-Red) and (Π_1-Red) over ATR^S_0 and ATR^S are then determined, in general by making use of the quantifier theorem and the fact that ATR^S_0 is equivalent to ATR_0. Afterwards, we change the environment and study (Σ_1-Red) and (Π_1-Red) in the context of KP. We end with some general comments and open problems. This paper is a mix of a survey article and new technical work.

2 Well-known reduction principles in second order arithmetic

Let $\mathcal{L}_2$ be a standard language of second order arithmetic with countably infinite supplies of two distinct sorts of variables; we also have the constant symbols 0 and 1 and function symbols for addition and multiplication plus relation symbols for the equality and less relation on the natural numbers. The first order variables are called *number variables* and supposed to range over natural numbers. The second order variables are known as *set variables* and intended to range over all sets of natural numbers. The number terms and formulas of $\mathcal{L}_2$ are built up as usual. See, for example, Simpson [16].

We use the following categories of letters (possibly with subscripts) as metavariables:

- i, j, k for first order variables;
- X, Y, Z for second order variables;
- $\eta, \theta, \varphi, \psi$ for formulas.

A formula of $\mathcal{L}_2$ without bound set variables is called *arithmetical*. For $1 \leq n \in \mathbb{N}$, a formula φ is said to be Σ^1_n or Π^1_n iff it is of the form

$$\exists X_1 \forall X_2 \ldots X_n \theta \quad \text{or} \quad \forall X_1 \exists X_2 \ldots X_n \theta,$$

respectively, where θ is arithmetical.

Throughout this paper we work in classical logic with equality for the first sort. Equality for sets in $\mathcal{L}_2$ is defined by saying that two sets are identical iff they contain the same elements.

ACA_0 is the system of second order arithmetic whose non-logical axioms comprise the defining axioms for all primitive recursive functions and relations, the axiom schema of *arithmetical comprehension*

$$\exists X \forall i (i \in X \leftrightarrow \varphi[i])$$

for all arithmetical formulas $\varphi[i]$, and the *induction axiom*

$$\forall X (0 \in X \land \forall i (i \in X \to i{+}1 \in X) \to \forall i (i \in X)).$$

ACA_0 is known to be a conservative extension of Peano arithmetic PA. The theory ACA is obtained from ACA_0 by adding the *schema of induction*

$$\varphi[0] \land \forall i (\varphi[i] \to \varphi[i{+}1]) \to \forall i \varphi[i]$$

for all $\mathcal{L}_2$ formulas $\varphi[i]$. Below we will make use of several further axiom schemas:

- (Σ^1_1-AC) is the schema

  $$\forall i \exists X \varphi[i, X] \to \exists Y \forall i \varphi[i, (Y)_i]$$

  for arbitrary Σ^1_1 formulas $\varphi[i, X]$;
- (Δ^1_2-CA) is the schema

  $$\forall i (\varphi[i] \leftrightarrow \psi[i]) \to \exists X \forall i (i \in X \leftrightarrow \varphi[i])$$

  for all Σ^1_2 formulas $\varphi[i]$ and Π^1_2 formulas $\psi[i]$;
- (Π^1_2-CA) is the schema

  $$\exists X \forall i (i \in X \leftrightarrow \varphi[i])$$

  for all Π^1_2 formulas $\varphi[i]$.

To simplify the notation we write Δ^1_2-CA_0 for the theory $\mathsf{ACA}_0 + (\Delta^1_2\text{-}\mathsf{CA})$ and Π^1_2-CA_0 for $\mathsf{ACA}_0 + (\Pi^1_2\text{-}\mathsf{CA})$.

Additional notation is necessary to formulate the principles of *arithmetical transfinite recursion* (ATR) and *bar induction* (BI), and to this end we follow [16] as closely as possible.

Working in ACA_0, we code binary relations on the natural numbers $\mathbb{N}$ as subsets of $\mathbb{N}$ via the pairing function

$$(i, j) \ := \ (i + j)^2 + i.$$

A set X of natural numbers is said to be *reflexive* iff

$$\forall i, j((i, j) \in X \ \rightarrow \ ((i, i) \in X \ \wedge \ (j, j) \in X)).$$

If X is reflexive, then $\mathit{Field}[X]$ is defined to be the set $\{i : (i, i) \in X\}$, and we write

$$i \leq_X j \ := \ (i, j) \in X,$$

$$i <_X j \ := \ (i, j) \in X \ \wedge \ (j, i) \notin X.$$

Furthermore, if X is reflexive we say that X is *well-founded* iff every non-empty subset of $\mathit{Field}[X]$ has an X-minimal element.[1] We say that X is a *linear ordering* if it is a reflexive linear ordering of its field, i.e.,

$$\forall i, j, k(i \leq_X j \ \wedge \ j \leq_X k \ \rightarrow \ i \leq_X k),$$

$$\forall i, j(i \leq_X j \ \wedge \ j \leq_X i \ \rightarrow \ i = j),$$

$$(\forall i, j \in \mathit{Field}[X])(i \leq_X j \ \vee \ j \leq_X i).$$

We say that X is a *well-ordering* iff it is both well-founded and a linear ordering. Let $\mathit{WF}[X]$, $\mathit{LO}[X]$, and $\mathit{WO}[X]$ be formulas saying that X is, respectively, well-founded, a linear ordering, and a well-ordering.

Definition 1 Given an $\mathcal{L}_2$ formula $\varphi[i]$, let $\mathit{TI}[X, \varphi]$ be the formula

$$\forall j((\forall i <_X j)\varphi[i] \ \rightarrow \ \varphi[j]) \ \rightarrow \ \forall j \varphi[j].$$

The schema (BI) of *bar induction* consists of all formulas

$$\forall X(\mathit{WF}[X] \ \rightarrow \ \mathit{TI}[X, \varphi]),$$

where φ ranges over all $\mathcal{L}_2$ formulas.

Now let $\varphi[i, Y]$ be any formula with distinguished free number variable i and distinguished free set variable Y. Define $\mathcal{H}_\varphi[X, Y]$ to be the formula

$$\mathit{LO}[X] \ \wedge \ Y = \{(i, j) : j \in \mathit{Field}[X] \ \wedge \ \varphi[i, Y^j]\},$$

[1] This is equivalent over ACA_0 to the definition given in [16].

where $Y^j := \{(i,k) \in Y : k <_X j\}$. Intuitively, $\mathcal{H}_\varphi[X,Y]$ says that X is a linear ordering and Y is the result of iterating φ along X.

Definition 2 The schema (ATR) of *arithmetical transfinite recursion* comprises

$$\forall X(WO[X] \rightarrow \exists Y \mathcal{H}_\varphi[X,Y])$$

for all arithmetical formulas $\varphi[i,Y]$. Accordingly, we set

$$\mathsf{ATR}_0 := \mathsf{ACA}_0 + (\mathsf{ATR}) \quad \text{and} \quad \mathsf{ATR} := \mathsf{ACA} + (\mathsf{ATR}).$$

ACA_0 and ATR_0 belong to the “big five” in the Friedman-Simpson program of *reverse mathematics*:

$$\mathsf{RCA}_0 \subsetneq \mathsf{WKL}_0 \subsetneq \mathsf{ACA}_0 \subsetneq \mathsf{ATR}_0 \subsetneq \Pi^1_1\text{-}\mathsf{CA}_0.$$

For more about these theories and the program of reverse mathematics in general we refer to Simpson [16].

It is also known that the proof-theoretic ordinals of ATR_0 and ATR are the ordinals Γ_0 and Γ_{ε_0}, respectively. For these results cf., for example, Friedman, McAloon and Simpson [4] and Jäger [7, 9].

In the following the theory ATR_0 will play a major role. Arithmetical transfinite recursion is relevant here because of its remarkable equivalence to Π^1_1 reduction over ACA_0. This and related reduction principles are introduced now.

Definition 3 Let n be a natural number greater than 0.

1. Σ^1_n *reduction* (Σ^1_n-Red) is the schema consisting of all formulas

$$\forall i(\varphi[i] \rightarrow \psi[i]) \rightarrow \exists X \forall i(\varphi[i] \rightarrow i \in X \rightarrow \psi[i]),$$

where $\varphi[i]$ is a Π^1_n and $\psi[i]$ a Σ^1_n formula.

2. Π^1_n *reduction* (Π^1_n-Red) is the schema consisting of all formulas

$$\forall i(\varphi[i] \rightarrow \psi[i]) \rightarrow \exists X \forall i(\varphi[i] \rightarrow i \in X \rightarrow \psi[i]),$$

where $\varphi[i]$ is a Σ^1_n and $\psi[i]$ a Π^1_n formula.

As mentioned in the introduction, what we call Σ^1_n reduction [Π^1_n reduction] has been called Π^1_n separation [Σ^1_n separation] by Simpson. Modulo this renaming the following characterizations are known.

Theorem 1 (Buchholz-Schütte, Simpson)

(i) The theory $\mathsf{ACA}_0 + (\Sigma^1_1\text{-AC})$ *proves* (Σ^1_1-Red).
(ii) $\mathsf{ACA}_0 + (\Pi^1_1\text{-Red})$ *is equivalent to* ATR_0.
(iii) $\mathsf{ACA}_0 + (\Sigma^1_2\text{-Red})$ *is equivalent to* $\Delta^1_2\text{-}\mathsf{CA}_0$.
(iv) $\mathsf{ACA}_0 + (\Pi^1_2\text{-Red})$ *is equivalent to* $\Pi^1_2\text{-}\mathsf{CA}_0$.

Assertion (i) is an easy observation; (ii) is a fairly complicated and technically demanding result presented in Simpson [16]; (iii) has been proved in Buchholz and Schütte [3]; (iv) is mentioned in [16] as an exercise. The following theorem is an immediate consequence of [16], Corollary VII.2.19.

Theorem 2 (Simpson) *The theory* $\mathsf{ACA}_0 + (\mathsf{BI})$ *proves all instances of* (ATR).

3 Basic set theory BS_0 and Simpson's ATR_0^S

After these introductory remarks we now turn to the two set theories that are in the center of this work: Simpson's ATR_0^S and Kripke-Platek set theory KP (with infinity).

The set-theoretic language $\mathcal{L}_\in$ is a one-sorted first order language with two binary relation symbols $\in$ and $=$, countably many set variables, and the usual connectives and quantifiers of first order logic. Terms and formulas of $\mathcal{L}_\in$ are as usual.

We shall make use of the common set-theoretic terminology and employ the standard notational conventions. In addition, we use as metavariables (possibly with subscripts):

- $a, b, c, f, i, j, k, r, u, v, w, x, y, z$ for set-theoretic variables,
- $\eta, \theta, \varphi, \psi$ for formulas.

We also follow the general conventions in defining the $\Delta_0, \Sigma, \Pi, \Sigma_n$, and Π_n formulas of $\mathcal{L}_\in$ $(1 \le n \in \mathbb{N})$.

Basic set theory BS_0 is a theory in the language $\mathcal{L}_\in$ – based on classical first order logic with equality – whose non-logical axioms are the universal closures of the following formulas:

- Extensionality: $\forall x(x \in a \leftrightarrow x \in b) \to a = b$.
- Regularity: $a \neq \emptyset \to (\exists x \in a)(\forall y \in x)(y \notin a)$.
- Infinity: $\exists a(\emptyset \in a \wedge (\forall x, y \in a)(x \cup \{y\} \in a))$.
- Rudimentary closure: We have axioms that formalize that the universe is closed under a series of rudimentary set-theoretic operations; cf. Simpson [16], Definition VII.3.3. We also claim that every set is a subset of a transitive set.

It is set-theoretic folklore and mentioned, for example, in Simpson [16] that BS_0 proves Δ_0 separation.

Lemma 1 BS_0 *proves for all* Δ_0 *formulas* $\varphi[x]$ *and all sets a that*

$$\exists y \forall x(x \in y \leftrightarrow x \in a \wedge \varphi[x]). \qquad (\Delta_0\text{-}\mathsf{Sep})$$

The theory BS is the extension of BS_0 resulting from extending regularity from sets to arbitrary formulas $\varphi[x]$,

$$\exists x \varphi[x] \to \exists x(\varphi[x] \wedge (\forall y \in x)\neg\varphi[y]).$$

Thus $\in$-induction is available in BS for arbitray $\mathcal{L}_\in$ formulas.

Lemma 2 BS *proves, for any $\mathcal{L}_\in$ formula $\varphi[x]$,*

$$\forall x((\forall y \in x)\varphi[y] \rightarrow \varphi[x]) \rightarrow \forall x \varphi[x]. \qquad (\mathcal{L}_\in\text{-}\mathsf{I}_\in)$$

Proof Aiming at the contrapositive, assume $\neg\varphi[a]$ for some a and let b be a transitive set such that $\{a\} \subseteq b$. Now set

$$\psi[x] := x \in b \wedge \neg\varphi[x].$$

Clearly, $\exists x \psi[x]$ and, therefore, the schema of regularity for formulas gives us an x satisfying

$$\psi[x] \wedge (\forall y \in x)\neg\psi[y],$$

i.e.,

$$x \in b \wedge \neg\varphi[x] \wedge (\forall y \in x)(y \notin b \vee \varphi[y]).$$

Since b is transitive, this can be simplified to

$$\neg\varphi[x] \wedge (\forall y \in x)\varphi[y],$$

and we have the desired statement. □

Moreover, we shall employ the common set-theoretic terminology and the standard notational conventions, for example:

$$Tran[a] := (\forall x \in a)(\forall y \in x)(y \in a),$$

$$Ord[a] := Tran[a] \wedge (\forall x \in a)Tran[x],$$

$$Succ[a] := Ord[a] \wedge (\exists x \in a)(a = x \cup \{x\}),$$

$$FinOrd[a] := \begin{cases} Ord[a] \wedge (a = \varnothing \vee Succ[a]) \wedge \\ (\forall x \in a)(x = \varnothing \vee Succ[x]). \end{cases}$$

In addition, we let ω be the collection of all finite ordinals and observe that it forms a set in BS_0.

There is a natural translation of $\mathcal{L}_2$ into $\mathcal{L}_\in$: The number variables of $\mathcal{L}_2$ are interpreted in $\mathcal{L}_\in$ as ranging over ω and the set variables of $\mathcal{L}_2$ are interpreted in $\mathcal{L}_\in$ as ranging over the subsets of ω. This means that

- first order quantifiers $\exists i$ and $\forall i$ of $\mathcal{L}_2$ translate into $(\exists i \in \omega)$ and $(\forall i \in \omega)$,
- second order quantifiers $\exists X$ and $\forall X$ of $\mathcal{L}_2$ translate into $(\exists x \subseteq \omega)$ and $(\forall x \subseteq \omega)$.

Then one has to verify that BS_0 proves the existence of set-theoretic functions on ω that correspond to the number-theoretic addition and multiplication. The number-theoretic less and equality relation go over into $<$ and $=$ on ω.

Clearly, each axiom of ACA_0 becomes a theorem of BS_0 under this translation. When working in $\mathcal{L}_\in$ we shall from now on identify $\mathcal{L}_2$ formulas with their translations into $\mathcal{L}_\in$.

Simpson's ATR_0^S is obtained from BS_0 by adding the axiom of countability (C) and the axiom (Beta). In the definition below we write $Tran[a]$ to express that the set a is transitive and $Inj[f, a, \omega]$ to state that f is an injective function from a to ω.

Definition 4 A set a is called *hereditarily countable* iff there exist a transitive superset x of a and an injection from x to ω,

$$HC[a] \ := \ \exists x, f(a \subseteq x \ \land \ Tran[x] \ \land \ Inj[f, x, \omega]).$$

The *axiom of countability* (C) claims that all sets are hereditarily countable,

$$(\mathsf{C}) \ := \ \forall x HC[x].$$

In the formulation of the axiom (Beta) we write $Dom[f, a]$ to express that f is a function with domain a. We write $\langle x, y\rangle$ for the ordered pair of x and y and $a \times b$ for the Cartesian product of a and b. Also, $r \subseteq a \times a$ is called *well-founded on a* iff every non-empty subset of a has an r-minimal element,

$$Wf[a, r] \ := \ (\forall b \subseteq a)(b \neq \varnothing \ \to \ (\exists x \in b)(\forall y \in b)(\langle y, x\rangle \notin r)).$$

Definition 5 The *axiom* (Beta) is the universal closure of the formula

$$\begin{aligned} &Wf[a, r] \ \to \\ &\qquad \exists f(Dom[f, a] \ \land \ (\forall x \in a)(f(x) = \{f(y) : y \in a \ \land \ \langle y, x\rangle \in r\})). \end{aligned}$$

This function f is said to be the *collapsing function* for r on a.

The axiom (Beta) has the effect of making the Π_1 predicate $Wf[a, r]$ a Δ_1 predicate since the existence of a collapsing function for r on a obviously implies the well-foundedness of r on a.

With these definitions the stage is set to introduce the theory ATR_0^S and its extension ATR^S:

$$\mathsf{ATR}_0^S \ := \ \mathsf{BS}_0 + (\mathsf{Beta}) + (\mathsf{C}) \quad \text{and} \quad \mathsf{ATR}^S \ := \ \mathsf{BS} + (\mathsf{Beta}) + (\mathsf{C}).$$

Below we write $\mathsf{ATR}_0^S \backslash (\mathsf{C})$ for the subsystem of ATR_0^S without the axiom of countability and $\mathsf{ATR}^S \backslash (\mathsf{C})$ for ATR^S without (C).

As we have mentioned above, there is a natural translation of $\mathcal{L}_2$ into $\mathcal{L}_\in$. The converse direction, i.e., the translation of $\mathcal{L}_\in$ into $\mathcal{L}_2$, is more complicated. In a nutshell: The sets of $\mathcal{L}_\in$ are represented by so-called *suitable trees*. Any suitable tree T is a well-founded subset of $\mathbb{N}^{<\mathbb{N}}$, and if $\langle n\rangle \in T$ then $T^{\langle n\rangle}$ is the subtree $\{\sigma : \langle n\rangle * \sigma \in T\}$ of T, with σ ranging over (the codes of) finite sequences. Elementhood is then coded by defining

$$S \in^* T \ := \ \exists n(\langle n\rangle \in T \ \land \ S \simeq T^{\langle n\rangle}),$$

where $S \simeq T^{\langle n \rangle}$ says that there exists a specific tree isomorphism between S and $T^{\langle n \rangle}$. For all details concerning this representation we refer to Simpson [16]. There it is also described how to each formula φ of $\mathcal{L}_\in$ we associate a formula $|\varphi|$ of $\mathcal{L}_2$.

The following two results of Simpson [16] make it clear that ATR_0^S is the set-theoretic variant of ATR_0.

Theorem 3 (Simpson)

(i) Every axiom of ATR_0 *is a theorem of* ATR_0^S.
(ii) If φ *is an axiom of* ATR_0^S*, then* $|\varphi|$ *is a theorem of* ATR_0.

This theorem can be easily extended to ATR^S. On the side of second order arithmetic, arithmetical transfinite recursion simply has to be replaced by bar induction.

Theorem 4 *(i) Every axiom of* $\mathsf{ACA}_0 + (\mathsf{BI})$ *is a theorem of* $\mathsf{ATR}^S \backslash (\mathsf{C})$.
(ii) If φ *is an axiom of* ATR^S*, then* $|\varphi|$ *is a theorem of* $\mathsf{ACA}_0 + (\mathsf{BI})$.

Proof It is a classic result that all instances of (BI) can be proved by means of (Beta) and $\in$-induction ($\mathcal{L}_\in$-$\mathsf{I}_\in$); see, for example, Jäger [10]. For (ii) we refer to Simpson [16], Theorem VII.3.34 and Exercise VII.3.38. □

Corollary 1 *If we write* $|T|$ *for the proof-theoretic ordinal of the theory* T*, then we have:*

(i) $|\mathsf{ATR}_0| = |\mathsf{ATR}_0^S| = \Gamma_0$.
(ii) $|\mathsf{ATR}| = \Gamma_{\varepsilon_0}$.
(iii) $|\mathsf{ATR}^S| = \Psi(\varepsilon_{\Omega+1})$ *(Bachmann-Howard ordinal).*

We end this introductory section by stating the so-called *quantifier theorem* which relates formulas of $\mathcal{L}_2$ with formulas of $\mathcal{L}_\in$. The formulation below is from Simpson [16]. Its first part also follows from corresponding results in Jäger [6, 10].

Theorem 5 (Quantifier theorem / Jäger, Simpson) *Let n be any natural number.*

(i) Each Σ^1_{n+2} *formula of* $\mathcal{L}_2$ *is equivalent – provably in* $\mathsf{ATR}_0^S \backslash (\mathsf{C})$ *– to a* Σ_{n+1} *formula of* $\mathcal{L}_\in$.
(ii) If φ *is a* Σ_n *formula of* $\mathcal{L}_\in$*, then* $|\varphi|$ *is equivalent – provably in* ATR_0 *– to a* Σ^1_{n+1} *formula of* $\mathcal{L}_2$.

In many set theories – Kripke-Platek set theory is a typical example – we do not have to make a big difference between Δ_0 and Δ_1 formulas. That this is not the case for ATR_0^S is an immediate consequence of the quantifier theorem. Recall from Lemma 1 that ATR_0^S proves Δ_0 separation. However:

Corollary 2 *There are instances of* Δ_1 *separation that are not provable in* ATR^S.

Proof In view of the quantifier theorem, $\mathsf{ATR}^S + (\Delta_1\text{-}\mathsf{Sep})$ comprises the theory $\Delta^1_2\text{-}\mathsf{CA}_0$, whose proof-theoretic ordinal is much greater than the Bachmann-Howard ordinal. Therefore, $(\Delta^1_2\text{-}\mathsf{CA})$ cannot be provable in ATR^S. □

4 Kripke-Platek set theory KP and its relationship to ATR_0^S

Kripke-Platek set theory KP (with infinity) is one of the best studied subsystems of Zermelo-Fraenkel set theory ZF. The transitive models of KP are called *admissible sets*, and $L_{\omega_1^{CK}}$, where ω_1^{CK} denotes the first non-recursive ordinal, is the least standard model of KP. Kripke-Platek set theory and admissible sets play an important role in generalized recursion theory, definability theory and, of course, in proof theory.

Kripke-Platek set theory KP is obtained from BS by adding the schema of Δ_0 collection, i.e.,

$$(\forall x \in a)\exists y \varphi[x, y] \ \rightarrow\ \exists z(\forall x \in a)(\exists y \in z)\varphi[x, y] \qquad (\Delta_0\text{-}\mathsf{Col})$$

for all Δ_0 formulas $\varphi[x, y]$;

$$\mathsf{KP} \ := \ \mathsf{BS} + (\Delta_0\text{-}\mathsf{Col}).^2$$

KP_0 is the subsystem of KP where regularity is restricted to sets (as in BS_0), i.e.,

$$\mathsf{KP}_0 \ := \ \mathsf{BS}_0 + (\Delta_0\text{-}\mathsf{Col}).$$

Thus, obviously, KP_0 + (Beta) and KP + (Beta) are the same theories as $\mathsf{ATR}_0^S\backslash(\mathsf{C})$ + $(\Delta_0\text{-}\mathsf{Col})$ and $\mathsf{ATR}^S\backslash(\mathsf{C})$ + $(\Delta_0\text{-}\mathsf{Col})$, respectively.

Below we list a series of known result that indicate that the relationship between ATR_0^S and KP is quite intricate. They are mostly taken from Simpson [15, 16] and Jäger [8, 10].

Theorem 6 (Overview)

(i) The proof-theoretic ordinal of KP *is the Bachmann-Howard ordinal* $\Psi(\varepsilon_{\Omega+1})$. KP *is proof-theoretically equivalent to* ATR^S *and to the theory of non-iterated inductive definitions* ID_1.
(ii) KP_0 + (Beta) *is proof-theoretically equivalent to* $\Delta^1_2\text{-}\mathsf{CA}_0$.
(iii) KP + (Beta) *is proof-theoretically equivalent to* $\Delta^1_2\text{-}\mathsf{CA}_0$ + (BI).
(iv) Any well-founded model of ATR_0^S *of height* ω_1^{CK} *is not a model of* KP.
(v) Any well-founded model of KP *of height* ω_1^{CK} *is not a model of* ATR_0^S.

From that we immediately deduce that ATR_0^S and KP are not compatible in the sense that

$$\mathsf{KP} \not\subseteq \mathsf{ATR}_0^S \quad \text{and} \quad \mathsf{ATR}_0^S \not\subseteq \mathsf{KP}.$$

We even have $\mathsf{KP}_0 \not\subseteq \mathsf{ATR}^S$. Otherwise, ATR^S would comprise KP + (Beta), and the ordinal of this theory is greater than the Bachmann-Howard ordinal.

---

[2] Traditionally, KP is defined to be the set theory whose non-logical axioms consist of Extensionality, Pairing, Union, Infinity, $(\mathcal{L}_\in\text{-}\mathsf{I}_\in)$, $(\Delta_0\text{-}\mathsf{Sep})$, and $(\Delta_0\text{-}\mathsf{Col})$; but both definitions are equivalent.

5 Reduction axioms in set theory

The next step is to add reduction principles, similar to those of second order arithmetic in Definition 3, to our set theories.

Definition 6 1. Σ_1 *reduction* (Σ_1-Red) is the schema consisting of all formulas

$$(\forall x \in a)(\varphi[x] \to \psi[x]) \;\to\; (\exists y \subseteq a)(\forall x \in a)(\varphi[x] \;\to\; x \in y \;\to\; \psi[x]),$$

where $\varphi[x]$ is a Π_1 and $\psi[x]$ a Σ_1 formula.

2. Π_1 *reduction* (Π_1-Red) is the schema consisting of all formulas

$$(\forall x \in a)(\varphi[x] \to \psi[x]) \;\to\; (\exists y \subseteq a)(\forall x \in a)(\varphi[x] \;\to\; x \in y \;\to\; \psi[x]),$$

where $\varphi[x]$ is a Σ_1 and $\psi[x]$ a Π_1 formula.

Our aim is to analyze the strengths of (Σ_1-Red) and (Π_1-Red) in the context of Simpsons ATR_0^S and Kripke-Platek set theory. We begin with ATR_0^S and ATR^S where the situation is clear.

5.1 ATR_0^S and ATR^S plus (Σ_1-Red) and (Π_1-Red)

Theorem 3 and Theorem 4, in combination with the quantifier theorem, immediately give us

$$\begin{aligned}
\mathsf{ATR}_0 + (\Sigma^1_2\text{-}\mathsf{Red}) &\subseteq \mathsf{ATR}_0^S + (\Sigma_1\text{-}\mathsf{Red}),\\
\mathsf{ACA}_0 + (\mathsf{BI}) + (\Sigma^1_2\text{-}\mathsf{Red}) &\subseteq \mathsf{ATR}^S + (\Sigma_1\text{-}\mathsf{Red}),\\
\mathsf{ATR}_0 + (\Pi^1_2\text{-}\mathsf{Red}) &\subseteq \mathsf{ATR}_0^S + (\Pi_1\text{-}\mathsf{Red}),\\
\mathsf{ACA}_0 + (\mathsf{BI}) + (\Pi^1_2\text{-}\mathsf{Red}) &\subseteq \mathsf{ATR}^S + (\Pi_1\text{-}\mathsf{Red}).
\end{aligned}$$

Therefore, together with Theorem 1 we have the following lower bound results for (Σ_1-Red) and (Π_1-Red).

Theorem 7 *(i)* $\Delta^1_2\text{-}\mathsf{CA}_0 \subseteq \mathsf{ACA}_0 + (\Sigma^1_2\text{-}\mathsf{Red}) \subseteq \mathsf{ATR}_0^S + (\Sigma_1\text{-}\mathsf{Red})$.
(ii) $\Delta^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}) \subseteq \mathsf{ACA}_0 + (\Sigma^1_2\text{-}\mathsf{Red}) + (\mathsf{BI}) \subseteq \mathsf{ATR}^S + (\Sigma_1\text{-}\mathsf{Red})$.
(iii) $\Pi^1_2\text{-}\mathsf{CA}_0 \subseteq \mathsf{ACA}_0 + (\Pi^1_2\text{-}\mathsf{Red}) \subseteq \mathsf{ATR}_0^S + (\Pi_1\text{-}\mathsf{Red})$.
(iv) $\Pi^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}) \subseteq \mathsf{ACA}_0 + (\Pi^1_2\text{-}\mathsf{Red}) + (\mathsf{BI}) \subseteq \mathsf{ATR}^S + (\Pi_1\text{-}\mathsf{Red})$.

Turning to the converse directions, we first make use of the quantifier theorem again and observe that for every (closed) instance φ of (Σ_1-Red) and (Π_1-Red) the corresponding $\mathcal{L}_2$ formula $|\varphi|$ is derivable in $\mathsf{ATR}_0 + (\Sigma^1_2\text{-}\mathsf{Red})$ and $\mathsf{ATR}_0 + (\Pi^1_2\text{-}\mathsf{Red})$, respectively. Therefore, Theorem 3 and Theorem 4 yield for all sentences φ of $\mathcal{L}_\in$:

$$\mathsf{ATR}_0^S + (\Sigma_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \mathsf{ATR}_0 + (\Sigma^1_2\text{-}\mathsf{Red}) \vdash |\varphi|,$$
$$\mathsf{ATR}^S + (\Sigma_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \mathsf{ACA}_0 + (\Sigma^1_2\text{-}\mathsf{Red}) + (\mathsf{BI}) \vdash |\varphi|,$$
$$\mathsf{ATR}_0^S + (\Pi_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \mathsf{ATR}_0 + (\Pi^1_2\text{-}\mathsf{Red}) \vdash |\varphi|,$$
$$\mathsf{ATR}^S + (\Pi_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \mathsf{ACA}_0 + (\Pi^1_2\text{-}\mathsf{Red}) + (\mathsf{BI}) \vdash |\varphi|.$$

The following upper bounds for $(\Sigma_1\text{-}\mathsf{Red})$ and $(\Pi_1\text{-}\mathsf{Red})$ are straightforward consequences of Theorem 1.

Theorem 8 *We have for every sentence φ of $\mathcal{L}_\in$:*

(i) $\mathsf{ATR}_0^S + (\Sigma_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \Delta^1_2\text{-}\mathsf{CA}_0 \vdash |\varphi|$.
(ii) $\mathsf{ATR}^S + (\Sigma_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \Delta^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}) \vdash |\varphi|$.
(iii) $\mathsf{ATR}_0^S + (\Pi_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \Pi^1_2\text{-}\mathsf{CA}_0 \vdash |\varphi|$.
(iv) $\mathsf{ATR}^S + (\Pi_1\text{-}\mathsf{Red}) \vdash \varphi \quad\Rightarrow\quad \Pi^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}) \vdash |\varphi|$.

To sum up, we know that $(\Sigma_1\text{-}\mathsf{Red})$ and $(\Pi_1\text{-}\mathsf{Red})$ added to ATR_0^S and ATR^S lead to the following proof-theoretic equivalences:

$$\mathsf{ATR}_0^S + (\Sigma_1\text{-}\mathsf{Red}) \equiv \Delta^1_2\text{-}\mathsf{CA}_0,$$
$$\mathsf{ATR}^S + (\Sigma_1\text{-}\mathsf{Red}) \equiv \Delta^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}),$$
$$\mathsf{ATR}_0^S + (\Pi_1\text{-}\mathsf{Red}) \equiv \Pi^1_2\text{-}\mathsf{CA}_0,$$
$$\mathsf{ATR}^S + (\Pi_1\text{-}\mathsf{Red}) \equiv \Pi^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}).$$

So for $(\Sigma_1\text{-}\mathsf{Red})$ and $(\Pi_1\text{-}\mathsf{Red})$ the situation is clear as long as we stay in the context of ATR_0^S and its extension ATR^S. The picture is completely different when we move to Kripke-Platek set theory.

5.2 KP_0 and KP plus $(\Sigma_1\text{-}\mathsf{Red})$ and $(\Pi_1\text{-}\mathsf{Red})$

A first observation is that $(\Sigma_1\text{-}\mathsf{Red})$ is irrelevant for Kripke-Platek set theory; it is provable there.

Lemma 3 *Every instance of* $(\Sigma_1\text{-}\mathsf{Red})$ *is provable in* KP_0.

Proof Suppose that, for some Π_1 formula $\varphi[x]$ and Σ_1 formula $\psi[x]$,

$$(\forall x \in a)(\varphi[x] \rightarrow \psi[x]).$$

By Σ reflection there exists a set b such that

$$(\forall x \in a)(\varphi^b[x] \rightarrow \psi^b[x]).$$

Then $c := \{n \in a : \psi^b[n]\}$ is the set we need. □

A further obvious observation is that an upper bound for (Π_1-Red) is provided by Σ_1 separation, i.e., the schema

$$\exists y \forall x (x \in y \ \leftrightarrow \ x \in a \ \wedge \ \varphi[x]) \qquad (\Sigma_1\text{-Sep})$$

for all Σ_1 formulas $\varphi[x]$. To see why, take a Σ_1 formula $\varphi[x]$ and a Π_1 formula $\psi[x]$ such that

$$(\forall x \in a)(\varphi[x] \ \rightarrow \ \psi[x])$$

for some set a. By (Σ_1-Sep) we can define the set $b := \{x \in a : \varphi[x]\}$ which is an obvious witness for (Σ_1-Red). Thus we have the following upper bounds.

Theorem 9 *(i)* $\mathsf{KP}_0 + (\Pi_1\text{-Red}) \ \subseteq \ \mathsf{KP}_0 + (\Sigma_1\text{-Sep})$.
(ii) $\mathsf{KP} + (\Pi_1\text{-Red}) \ \subseteq \ \mathsf{KP} + (\Sigma_1\text{-Sep})$.

It is still an open question whether these bounds are sharp. However, we have some partial results.

Theorem 10 *We have for all formulas φ of $\mathcal{L}_2$:*

(i) $\Pi^1_2\text{-}\mathsf{CA}_0 \ \vdash \ \varphi \quad \Rightarrow \quad \mathsf{KP}_0 + (\mathsf{Beta}) + (\Pi_1\text{-Red}) \ \vdash \ \varphi.$
(ii) $\Pi^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}) \ \vdash \ \varphi \quad \Rightarrow \quad \mathsf{KP} + (\mathsf{Beta}) + (\Pi_1\text{-Red}) \ \vdash \ \varphi.$

Proof We first recall from Theorem 1 that $\Pi^1_2\text{-}\mathsf{CA}_0$ is equivalent to $\mathsf{ACA}_0 + (\Pi^1_2\text{-Red})$. From the first part of the quantifier theorem we deduce that, within KP_0 + (Beta), every Σ^1_2 formula of $\mathcal{L}_2$ is equivalent to a Σ_1 formula and every Π^1_2 formula of $\mathcal{L}_2$ is equivalent to a Π_1 formula. Therefore, every instance of (Π^1_2-Red) is provable in KP_0 + (Beta) + (Π_1-Red). The rest follows from Theorem 4. □

Now we recall from, for example, Rathjen [14] that KP_0 + (Σ_1-Sep) and KP + (Σ_1-Sep) prove the same $\mathcal{L}_2$ sentences as $\Pi^1_2\text{-}\mathsf{CA}_0$ and $\Pi^1_2\text{-}\mathsf{CA}_0$ + (BI), respectively. Moreover, by following Barwise [2] (with some small modifications), we can also show that KP_0 + (Σ_1-Sep) proves (Beta). Thus the following assertions are direct consequences of the previous two theorems.

Corollary 3 *(i) The theories* $\Pi^1_2\text{-}\mathsf{CA}_0$, KP_0 + (Beta) + (Π_1-Red), *and* KP_0 + (Σ_1-Sep) *prove the same $\mathcal{L}_2$ sentences.*
(ii) The theories $\Pi^1_2\text{-}\mathsf{CA}_0$ + (BI), KP + (Beta) + (Π_1-Red), *and* KP + (Σ_1-Sep) *prove the same $\mathcal{L}_2$ sentences.*

Shown schematically, we therefore have the following proof-theoretic equivalences:

$$\mathsf{KP}_0 + (\mathsf{Beta}) + (\Pi_1\text{-Red}) \equiv \mathsf{KP}_0 + (\Sigma_1\text{-Sep}) \equiv \Pi^1_2\text{-}\mathsf{CA}_0,$$
$$\mathsf{KP} + (\mathsf{Beta}) + (\Pi_1\text{-Red}) \equiv \mathsf{KP} + (\Sigma_1\text{-Sep}) \equiv \Pi^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}).$$

The axiom of constructibility states that every set belongs to some level L_α of Gödel's hierarchy of constructible sets:

$$\forall x \exists \alpha (x \in L_\alpha). \qquad (V{=}L)$$

In addition, we write $(a <_L b)$ to express that a is smaller than b according to the well-order $<_L$ of the constructible universe. It is well-known that $(a \in L_\alpha)$ and $(a <_L b)$ are Δ over KP. Moreover, $(\exists x <_L a)$ and $(\forall x <_L a)$ may be treated as bounded quantifiers. $(x \in L)$ is short for $\exists \alpha (x \in L_\alpha)$. For more on the constructible universe see, e.g., Barwise [2] or Kunen [12].

In Jäger and Steila [11] an interesting separation principle is introduced and studied. Call a quantifier *subset bounded* iff it ranges over the subsets of a given set. Then let $\exists^{\mathcal{P}}(\Delta_1)$ separation be the separation principle that, given any set a, allows the introduction of all subsets of a defined by a subset bounded Σ_1 formula over a Δ_1 matrix, i.e.,

$$\begin{aligned}&(\forall x \in a)(\forall y \subseteq a)(\varphi[x,y] \leftrightarrow \psi[x,y]) \rightarrow \\ &\qquad (\exists z \subseteq a)(\forall x \in a)(x \in z \leftrightarrow (\exists y \subseteq a)\varphi[x,y]),\end{aligned} \qquad (\exists^{\mathcal{P}}(\Delta_1)\text{-Sep})$$

for all Σ_1 formulas $\varphi[x,y]$ and Π_1 formulas $\psi[x,y]$.

The relationship between (Σ_1-Sep) and ($\exists^{\mathcal{P}}(\Delta_1)$-Sep) is interesting. It is easy to see that all instances of ($\exists^{\mathcal{P}}(\Delta_1)$-Sep) are provable in KP + (Σ_1-Sep). For the converse direction see the follwing theorem from Jäger and Steila [11].

Theorem 11 (Jäger and Steila)
All instances of (Σ_1-Sep) *are provable in* KP + $(V{=}L)$ + ($\exists^{\mathcal{P}}(\Delta_1)$-Sep).

Therefore, in order to show that the theory KP + $(V{=}L)$ + (Π_1-Red) contains KP + (Σ_1-Sep), it is sufficient to prove that it contains ($\exists^{\mathcal{P}}(\Delta_1)$-Sep).

Lemma 4 KP + $(V{=}L)$ + (Π_1-Red) *proves all instances of* ($\exists^{\mathcal{P}}(\Delta_1)$-Sep).

Proof Working in KP + $(V{=}L)$ + (Π_1-Red), let us assume that

$$(\forall x \in a)(\forall y \subseteq a)(\varphi[x,y] \leftrightarrow \psi[x,y])$$

for a Σ_1 formula $\varphi[x,y]$ and a Π_1 formula $\psi[x,y]$. We define

$$\widetilde{\varphi}[x,y] \;:=\; \varphi[x,y] \;\wedge\; (\forall z <_L y)(z \subseteq a \;\rightarrow\; \neg\psi[x,z]),$$

$$\widetilde{\psi}[x,y] \;:=\; \psi[x,y] \;\wedge\; (\forall z <_L y)(z \subseteq a \;\rightarrow\; \neg\varphi[x,z]).$$

Thus $\widetilde{\varphi}[x,y]$ is a Σ formula and $\widetilde{\psi}[x,y]$ is a Π formula, and we have:

(1) $(\forall x \in a)(\forall y \subseteq a)(\widetilde{\varphi}[x,y] \leftrightarrow \widetilde{\psi}[x,y])$.
(2) $\widetilde{\varphi}[x,y] \rightarrow \varphi[x,y]$.
(3) $y \subseteq a \wedge \varphi[x,y] \rightarrow (\exists z \subseteq a)\widetilde{\varphi}[x,z]$.
(4) $(\forall y, z \subseteq a)(\widetilde{\varphi}[x,y] \wedge \widetilde{\varphi}[x,z] \rightarrow y = z)$.

We also define

$$\mathfrak{A}[u] \;:=\; (\exists v, w \in a)(\exists y \subseteq a)(u = \langle v, w\rangle \;\wedge\; \widetilde{\varphi}[v, y] \;\wedge\; w \in y),$$

$$\mathfrak{B}[u] \;:=\; (\forall v, w \in a)(\forall y \subseteq a)(u = \langle v, w\rangle \;\wedge\; \widetilde{\varphi}[v, y] \;\rightarrow\; w \in y).$$

Hence $\mathfrak{A}[u]$ is equivalent to a Σ_1 formula and $\mathfrak{B}[u]$ to a Π_1 formula. In addition, because of (4) we have

$$(\forall u \in a \times a)(\mathfrak{A}[u] \;\rightarrow\; \mathfrak{B}[u]).$$

So by (Π_1-Red) there is a $b \subseteq a \times a$ such that

$$(\forall u \in a \times a)(\mathfrak{A}[u] \;\rightarrow\; u \in b \;\rightarrow\; \mathfrak{B}[u]).$$

Claim: For all $v \in a$,

$$(\exists y \subseteq a)\varphi[v, y] \;\leftrightarrow\; \widetilde{\varphi}[v, (b)_v], \tag{*}$$

where $(b)_v$ stands for the set $\{w \in a : \langle v, w\rangle \in b\}$.

Proof of the claim. The direction from right to left is obvious. To prove the converse, assume $(\exists y \subseteq a)\varphi[v, y]$. Then there is a $c \subseteq a$ with $\widetilde{\varphi}[v, c]$ according to (3), and for this c we have $c = (b)_v$. This can be seen as follows: For all $w \in a$,

$$w \in c \;\rightarrow\; \mathfrak{A}[\langle v, w\rangle] \;\rightarrow\; \langle v, w\rangle \in b \;\rightarrow\; w \in (b)_v,$$

$$w \in (b)_v \;\rightarrow\; \langle v, w\rangle \in b \;\rightarrow\; \mathfrak{B}[\langle v, w\rangle] \;\rightarrow\; w \in c.$$

Since $c = (b)_v$ and $\widetilde{\varphi}[v, c]$, we have $\widetilde{\varphi}[v, (b)_v]$ as desired, finishing the proof of the claim.

In view of (1) the set $\{v \in a : \widetilde{\varphi}[v, (b)_v]\}$ exists by Δ separation and is the set we need according to (*). □

The following is an immediate consequence of Theorem 9, Theorem 11, and the previous lemma.

Corollary 4 *We have for all $\mathcal{L}_\in$ formulas φ that*

$$\mathsf{KP} + (V{=}L) + (\Sigma_1\text{-}\mathsf{Sep}) \;\vdash\; \varphi \quad\Leftrightarrow\quad \mathsf{KP} + (V{=}L) + (\Pi_1\text{-}\mathsf{Red}) \;\vdash\; \varphi.$$

In order to get rid of the axiom ($V{=}L$) on the left-hand side of the previous implication, we show that L is an inner model of $\mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep})$.

Theorem 12 *If φ is the universal closure of an axiom of* $\mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep})$*, we have that*

$$\mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep}) \;\vdash\; \varphi^L.$$

Here φ^L is the result of restricting all unbounded quantifiers in φ to L.

Proof In view of Barwise [2] we only have to deal with the instances of (Σ_1-Sep). So let $\varphi[x, y, z]$ be a Δ_0 formula with all free variables indicated; we suppress mentioning additional parameters. Given elements $a, b \in L$ we have to show that

$$(\exists z \in L)(\forall x \in L)(x \in z \;\leftrightarrow\; x \in a \;\wedge\; \exists y(y \in L \;\wedge\; \varphi[x, y, b])).$$

By (Σ_1-Sep) there exists the set

$$c = \{x \in a : \exists y(y \in L \;\wedge\; \varphi[x, y, b])\}$$

and thus we have

$$(\forall x \in c)\exists \xi(\exists y \in L_\xi)\varphi[x, y, b].$$

By Σ collection there is an α such that

$$(\forall x \in c)(\exists \xi < \alpha)(\exists y \in L_\xi)\varphi[x, y, b]$$

and so, by the properties of the L-hierarchy,

$$(\forall x \in c)(\exists y \in L_\alpha)\varphi[x, y, b].$$

This implies that $c = \{x \in a : (\exists y \in L_\alpha)\varphi[x, y, b]\}$. Hence c is the required witness in L. □

This theorem implies that $\mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep}) + (V{=}L)$ is conservative over $\mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep})$ for formulas which are absolute w.r.t. $\mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep})$, in particular for all arithmetical formulas.

Corollary 5 *We have for all arithmetical formulas φ that*

$$\mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep}) \vdash \varphi \quad\Leftrightarrow\quad \mathsf{KP} + (V{=}L) + (\Pi_1\text{-}\mathsf{Red}) \vdash \varphi.$$

If we summarize our results, we have the following proof-theoretic equivalences:

$$\mathsf{KP} + (V{=}L) + (\Pi_1\text{-}\mathsf{Red}) \;\equiv\; \mathsf{KP} + (\Sigma_1\text{-}\mathsf{Sep}) \;\equiv\; \Pi^1_2\text{-}\mathsf{CA}_0 + (\mathsf{BI}).$$

6 Comments and Questions

This article did not discuss the theory $\mathsf{KP}_0 + (V{=}L) + (\Pi_1\text{-}\mathsf{Red})$. There is the question of how to deal with the constructible hierarchy in KP_0 and whether there is an analogue of Theorem 11 with KP replaced by KP_0. However, we are not sure whether this leads to something interesting.

Our real concern in the present context is the question of the strength of $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$. We know that ($\Pi_1$-Red) is not provable in KP. This can be seen as follows:

(i) In $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$ we can prove (Π^1_1-Red) and therefore, according to Theorem 1, $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$ contains ATR_0.

(ii) In view of a result in Avigad [1] we thus know that $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$ proves that every X-positive arithmetical formula $\varphi[x, X^+]$ has a fixed point (which is a set).

(iii) On the other hand, we also know by results due to Gregoriades [5] and Probst [13] that there are positive arithmetical formulas that do not have hyperarithmetical fixed points.

(iv) So we conclude that $L_{\omega_1^{CK}}$ is not a model of $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$, implying that KP does not prove $(\Pi_1\text{-}\mathsf{Red})$.

But is the proof-theoretic strength of $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$ greater than that of KP? As a preparatory step for the analysis of the proof-theoretic strength of $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$ it could be useful to check whether $\mathsf{KP} + (\Pi_1\text{-}\mathsf{Red})$ proves $(\Pi^1_1\text{-}\mathsf{CA})$ or $(\Delta^1_2\text{-}\mathsf{CA})$.

A different line of research is to look at reduction principles of the form as discussed above in theories of sets and classes. Some first results are known, but in general this field is wide open.

References

1. J. Avigad, *On the relationship between* ATR_0 *and* $\widehat{\mathsf{ID}}_{<\omega}$, The Journal of Symbolic Logic **62** (1996), no. 3, 768–779.
2. J. Barwise, *Admissible Sets and Structures*, Perspectives in Mathematical Logic, no. 7, Springer, 1975.
3. W. Buchholz and K. Schütte, *Proof Theory of Impredicative Subsystems of Analysis*, Studies in Proof Theory. Monographs, vol. 2, Bibliopolis, 1988.
4. H. Friedman, K. McAloon, and S. G. Simpson, *A finite combinatorial principle which is equivalent to the* 1*-consistency of predicative analysis*, Patras Logic Symposion (G. Metakides, ed.), Studies in Logic and the Foundations of Mathematics, vol. 109, Elsevier, 1982, pp. 197–230.
5. V. Gregoriades, *On a question of Jaeger's*, arXiv:1905.09609, 2019.
6. G. Jäger, *Die konstruktible Hierarchie als Hilfsmittel zur beweistheoretischen Untersuchung von Teilsystemen der Mengenlehre und Analysis*, Ph.D. thesis, University of Munich, 1979.
7. G. Jäger, *Theories for iterated jumps*, Notes, Oxford, 1980.
8. G. Jäger, *Zur Beweistheorie der Kripke-Platek-Mengenlehre über den natürlichen Zahlen*, Archive for Mathematical Logic **22** (1982), no. 3-4, 121–139.
9. G. Jäger, *The strength of admissibility without foundation*, The Journal of Symbolic Logic **49** (1984), no. 3, 867–879.
10. G. Jäger, *Theories for Admissible Sets. A Unifying Approach to Proof Theory*, Studies in Proof Theory. Lecture Notes, vol. 2, Bibliopolis, 1986.
11. G. Jäger and S. Steila, *About some fixed point axioms and related principles in Kripke-Platek environments*, The Journal of Symbolic Logic **83** (2018), no. 2, 642–668.
12. K. Kunen, *Set Theory. An Introduction to Independence Proofs*, Studies in Logic and the Foundations of Mathematics, vol. 102, North-Holland, 1980.
13. D. Probst, *Pseudo-Hierarchies in Admissible Set Theories without Foundation and Explicit Mathematics*, Ph.D. thesis, University of Bern, 2005.
14. M. Rathjen, *Explicit mathematics with the monotone fixed point principle. II: Models*, The Journal of Symbolic Logic **64** (1999), no. 2.
15. S. G. Simpson, *Set theoretic aspects of* ATR_0, Logic Colloquium '80 (D. van Dalen, D. Lascar, and T. J. Smiley, eds.), Studies in Logic and the Foundations of Mathematics, vol. 108, Elsevier, 1980, pp. 255–271.

16. S. G. Simpson, *Subsystems of second order arithmetic*, second ed., Perspectives in Logic, Association for Symbolic Logic and Cambridge University Press, 2009.