Relativized ordinal analysis: The case of Power Kripke-Platek set theory

Michael Rathjen

Department of Pure Mathematics, University of Leeds, Leeds LS2 9JT, England
rathjen@maths.leeds.ac.uk

Abstract. The paper relativizes the method of ordinal analysis developed for Kripke-Platek set theory to theories which have the power set axiom. We show that it is possible to use this technique to extract information about Power Kripke-Platek set theory, $\mathbf{KP}(\mathcal{P})$.

Key words: Power Kripke-Platek set theory, ordinal analysis, ordinal representation systems, proof-theoretic strength, power-admissible set
MSC Primary 03F15, 03F05, 03F35; Secondary 03F03, 03-03.

1 Introduction

Ordinal analyses of ever stronger theories have been obtained over the last 20 years (cf. [**1, 2, 3, 19, 20, 23, 24**]). The strongest theories for which proof-theoretic ordinals have been determined are subsystems of second order arithmetic with comprehension restricted to Π^1_2-comprehension (or even Δ^1_3-comprehension) (cf. [**26, 27, 28**]). Thus it appears that it is currently impossible to furnish an ordinal analysis of any set theory which has the power set axiom among its axioms as such a theory would dwarf the strength of second order arithmetic. Notwithstanding the foregoing, the current paper relativizes the techniques of ordinal analysis developed for Kripke-Platek set theory, **KP**, to obtain very useful information about Power Kripke-Platek set theory, $\mathbf{KP}(\mathcal{P})$, crystallizing in a bound for the transfinite iterations of the power set operation that are provable in the latter theory.

Technically we draw on tools that have been developed more than 30 years ago. With the work of Jäger and Pohlers (see [**13, 14**]) the forum of ordinal analysis switched from subsystems of second-order arithmetic to set theory, shaping what is called *admissible proof theory*, after the standard models of **KP**. We also draw on the framework of operator controlled derivations developed by Buchholz [**22**] that allows one to express the uniformity of infinite derivations and to carry out their bookkeeping in an elegant way.

The results and techniques of this paper have important applications. The characterization of the strength of $\mathbf{KP}(\mathcal{P})$ in terms of the von Neumann hierarchy is used in [**31**, Theorem 1.1] to calibrate the strength of the calculus of construction with one type universe (which is an intuitionistic type theory). Another application is made in connection with the so-called *existence property*, **EP**, that intuitionistic set theories may or may not have. Full intuitionistic Zermelo-Fraenkel set theory, **IZF**, does not have the existence property, where **IZF** is formulated with Collection (cf. [**12**]). By contrast, an ordinal analysis of intuitionistic $\mathbf{KP}(\mathcal{P})$ similar to the one given in this paper together

[1]Research supported by Templeton Foundation Grant #13152, the CRM Infinity Project, and by U.K. EPSRC grant No. EP/G029520/1. The author is also grateful for a Mercator Professorship of the German Science Foundation (DFG) which allowed him to work on some of the ideas of this article.

with results from [**30**] can be utilized to show that **IZF** with only bounded separation has the **EP**.

2 Power Kripke-Platek set theory

A particularly interesting (classical) subtheory of **ZF** is Kripke-Platek set theory, **KP**. Its standard models are called *admissible sets*. One of the reasons that this is an important theory is that a great deal of set theory requires only the axioms of **KP**. An even more important reason is that admissible sets have been a major source of interaction between model theory, recursion theory and set theory (cf. [**5**]). **KP** arises from **ZF** by completely omitting the power set axiom and restricting separation and collection to bounded formulae. These alterations are suggested by the informal notion of 'predicative'.

To be more precise, quantifiers of the forms $\forall x \in a$, $\exists x \in a$ are called *bounded*. *Bounded* or Δ_0*-formulae* are the formulae wherein all quantifiers are bounded. The axioms of **KP** consist of *Extensionality, Pair, Union, Infinity, Bounded Separation*

$$\exists x \, \forall u \, [u \in x \leftrightarrow (u \in a \, \wedge \, A(u))]$$

for all bounded formulae $A(u)$, *Bounded Collection*

$$\forall x \in a \, \exists y \, G(x,y) \, \rightarrow \, \exists z \, \forall x \in a \, \exists y \in z \, G(x,y)$$

for all bounded formulae $G(x,y)$, and *Set Induction*

$$\forall x \, [(\forall y \in x \, C(y)) \rightarrow C(x)] \, \rightarrow \, \forall x \, C(x)$$

for all formulae $C(x)$.

A transitive set A such that $(A, \in)$ is a model of **KP** is called an *admissible set*. Of particular interest are the models of **KP** formed by segments of Gödel's *constructible hierarchy* **L**. The constructible hierarchy is obtained by iterating the definable powerset operation through the ordinals

$$\begin{aligned}
\mathbf{L}_0 &= \emptyset, \\
\mathbf{L}_\lambda &= \bigcup \{\mathbf{L}_\beta : \beta < \lambda\} \ \lambda \text{ limit} \\
\mathbf{L}_{\beta+1} &= \big\{X : \ X \subseteq \mathbf{L}_\beta; \ X \text{ definable over } \ \langle \mathbf{L}_\beta, \in \rangle \big\}.
\end{aligned}$$

So any element of $\mathbf{L}$ of level α is definable from elements of $\mathbf{L}$ with levels $< \alpha$ and the parameter $\mathbf{L}_\alpha$. An ordinal α is *admissible* if the structure $(\mathbf{L}_\alpha, \in)$ is a model of **KP**.

If the power set operation is considered as a definite operation, but the universe of all sets is regarded as an indefinite totality, we are led to systems of set theory having Power Set as an axiom but only Bounded Separation axioms and intuitionistic logic for reasoning about the universe at large. The study of subsystems of **ZF** formulated in intuitionistic logic with Bounded Separation but containing the Power Set axiom was apparently initiated by Pozsgay [**17, 18**] and then pursued more systematically by Tharp [**33**], Friedman [**10**] and Wolf [**35**]. These systems are actually semi-intuitionistic as they contain the law of excluded middle for bounded formulae.

In the classical context, weak subsystems of **ZF** with Bounded Separation and Power Set have been studied by Thiele [**34**], Friedman [**11**] and more recently at great length by Mathias [**16**]. Mac Lane has singled out and championed a particular fragment of **ZF**, especially in his book *Form and Function* [**15**]. *Mac Lane Set Theory*, christened **MAC** in [**16**], comprises the axioms of Extensionality, Null Set, Pairing, Union, Infinity, Power Set, Bounded Separation, Foundation, and Choice. **MAC** is naturally related to systems derived from topos-theoretic notions and, moreover, to type theories.

Definition 2.1 We use subset bounded quantifiers $\exists x \subseteq y \ \ldots$ and $\forall x \subseteq y \ \ldots$ as abbreviations for $\exists x(x \subseteq y \wedge \ldots)$ and $\forall x(x \subseteq y \rightarrow \ldots)$, respectively.

The $\Delta_0^{\mathcal{P}}$ formulae are the smallest class of formulae containing the atomic formulae closed under $\wedge, \vee, \rightarrow, \neg$ and the quantifiers

$$\forall x \in a,\ \exists x \in a,\ \forall x \subseteq a,\ \exists x \subseteq a.$$

Definition 2.2 $\mathbf{KP}(\mathcal{P})$ has the same language as **ZF**. Its axioms are the following: Extensionality, Pairing, Union, Infinity, Powerset, $\Delta_0^{\mathcal{P}}$-Separation and $\Delta_0^{\mathcal{P}}$-Collection.

The transitive models of $\mathbf{KP}(\mathcal{P})$ have been termed **power admissible** sets in [**11**].

Remark 2.3 Alternatively, $\mathbf{KP}(\mathcal{P})$ can be obtained from **KP** by adding a function symbol $\mathcal{P}$ for the powerset function as a primitive symbol to the language and the axiom

$$\forall y\, [y \in \mathcal{P}(x) \leftrightarrow y \subseteq x]$$

and extending the schemes of Δ_0 Separation and Collection to the Δ_0 formulae of this new language.

Lemma 2.4 $\mathbf{KP}(\mathcal{P})$ *is* **not** *the same theory as* $\mathbf{KP} + \mathbf{Pow}$. *Indeed,* $\mathbf{KP} + \mathbf{Pow}$ *is a much weaker theory than* $\mathbf{KP}(\mathcal{P})$ *in which one cannot prove the existence of* $V_{\omega+\omega}$.

Proof: Note that in the presence of full Separation and Infinity there is no difference between our system **KP** and Mathias's [**16**] **KP**. It follows from [**16**, Theorem 14] that $\mathbf{Z} + \mathbf{KP} + \mathbf{AC}$ is conservative over $\mathbf{Z} + \mathbf{AC}$ for stratifiable sentences. **Z** and $\mathbf{Z} + \mathbf{AC}$ are of the same proof-theoretic strength as the constructible hierarchy can be simulated in **Z**; a stronger statement is given in [**16**, Theorem 16]. As a result, **Z** and $\mathbf{Z} + \mathbf{KP}$ are of the same strength. As $\mathbf{KP} + \mathbf{Pow}$ is a subtheory of $\mathbf{Z} + \mathbf{KP}$, we have that $\mathbf{KP} + \mathbf{Pow}$ is not stronger than **Z**. If $\mathbf{KP} + \mathbf{Pow}$ could prove the existence of $V_{\omega+\omega}$ it would prove the consistency of **Z**. On the other hand $\mathbf{KP}(\mathcal{P})$ prove the existence of V_α for every ordinal α and hence proves the existence of arbitrarily large transitive models of **Z**. $\Box$

Remark 2.5 Our system $\mathbf{KP}(\mathcal{P})$ is not quite the same as the theory $\mathbf{KP}^{\mathcal{P}}$ in Mathias' paper [**16**, 6.10]. The difference between $\mathbf{KP}(\mathcal{P})$ and $\mathbf{KP}^{\mathcal{P}}$ is that in the latter system set induction only holds for $\Sigma_1^{\mathcal{P}}$ formulae, or what amounts to the same, $\Pi_1^{\mathcal{P}}$ foundation ($A \neq \emptyset \rightarrow \exists x \in A\ x \cap A = \emptyset$ for $\Pi_1^{\mathcal{P}}$ classes A).

Friedman [**11**] includes only Set Foundation in his formulation of a formal system $\mathbf{PAdm}^s$ appropriate to the concept of recursion in the power set operation $\mathcal{P}$.

3 A Tait-style formalization of $\mathbf{KP}(\mathcal{P})$

For technical reasons we shall use a Tait–style sequent calculus version of $\mathbf{KP}(\mathcal{P})$ in which finite sets of formulae can be derived. In addition, formulae have to be in negation normal form (cf. [**32**]). The language consists of: free variables $a_0, a_1, \cdots$, bound variables $x_0, x_1, \cdots$; the predicate symbol $\in$; the logical symbols $\neg, \vee, \wedge, \forall, \exists$. One peculiarity will be that we treat bounded quantifiers and subset bounded quantifiers as quantifiers in their own right.

We will use $a, b, c, \cdots,\ x, y, z, \cdots,\ A, B, C, \cdots$ as metavariables whose domains are the domain of the free variables, bound variables, formulae, respectively.

The *atomic formulae* are those of the form $(a\in b), \neg(a\in b)$.

The *formulae* are defined inductively as follows:

(i) Atomic formulae are formulae.

(ii) If A and B are formulae, then so are $(A \wedge B)$ and $(A \vee B)$.

(iii) If $A(b)$ is a formula in which x does not occur, then $\forall x A(x)$, $\exists x A(x)$, $(\forall x \in a)A(x)$, $(\exists x \in a)A(x)$, $(\forall x \subseteq a)A(x)$, and $(\exists x \subseteq a)A(x)$ are formulae.

The quantifiers $\exists x, \forall x$ will be called *unrestricted*, whereas the other quantifiers will be referred to as *restricted quantifiers.* A $\Delta_0^{\mathcal{P}}$*–formula* is a formula which contains no unrestricted quantifiers. The Δ_0–formulae are those $\Delta_0^{\mathcal{P}}$-formulae that do not contain subset bounded quantifiers.

The *negation* $\neg A$ of a formula A is defined to be the formula obtained from A by (i) putting $\neg$ in front of any atomic formula, (ii) replacing $\wedge, \vee, \forall x, \exists x, (\forall x \in a), (\exists x \in a), (\forall x \subseteq a), (\exists x \subseteq a)$ by $\vee, \wedge, \exists x, \forall x, (\exists x \in a), (\forall x \in a), (\exists x \subseteq a), (\forall x \subseteq a)$, respectively, and (iii) dropping double negations. $A \to B$ stands for $\neg A \vee B$.

$\vec{a}, \vec{b}, \vec{c}, \cdots$ and $\vec{x}, \vec{y}, \vec{z}, \cdots$ will be used to denote finite sequences of free and bound variables, respectively.

We use $F[a_1, \cdots, a_n]$ (by contrast with $F(a_1, \cdots, a_n)$) to denote a formula the free variables of which are among $a_1, \cdots, a_n$. We will write

$a = \{x \in b : G(x)\}$ for $(\forall x \in a)[x \in b \wedge G(x)] \wedge (\forall x \in b)[G(x) \to x \in a]$.

$a = b$ stands for $(\forall x \in a)(x \in b) \wedge (\forall x \in b)(x \in a)$. $a \subseteq b$ stands for $(\forall x \in a)(x \in b)$. However, as part of a subset bounded quantifier $(\forall x \subseteq a)$ or $(\exists x \subseteq b)$, $\subseteq$ is considered to be a primitive symbol.

Definition 3.1 The sequent-style version of $\mathbf{KP}(\mathcal{P})$ derives finite sets of formulae denoted by $\Gamma, \Delta, \Theta, \Xi, \cdots$. The intended meaning of Γ is the disjunction of all formulae of Γ. We use the notation Γ, A for $\Gamma \cup \{A\}$, and Γ, Ξ for $\Gamma \cup \Xi$.

The *axioms of* $\mathbf{KP}(\mathcal{P})$ are the following:

Logical axioms:	$\Gamma, A, \neg A$ for every $\Delta_0^{\mathcal{P}}$–formula A.
Extensionality:	$\Gamma, a = b \wedge B(a) \to B(b)$ for every $\Delta_0^{\mathcal{P}}$-formula $B(a)$.
Pair:	$\Gamma, \exists x[a \in x \wedge b \in x]$
Union:	$\Gamma, \exists x(\forall y \in a)(\forall z \in y)(z \in x)$
$\Delta_0^{\mathcal{P}}$ *–Separation:*	$\Gamma, \exists y(y = \{x \in a : G(x)\})$ for every $\Delta_0^{\mathcal{P}}$–formula $G(b)$.
Set Induction:	$\Gamma, \forall u\, [(\forall x \in u)\, G(x) \to G(x)] \to \forall u\, G(u)$ for every formula $G(b)$.
Infinity:	$\Gamma, \exists x\, [(\exists y \in x)\, y \in x \wedge (\forall y \in x)(\exists z \in x)\, y \in z]$.
Power Set:	$\Gamma, \exists z\, (\forall x \subseteq a) x \in z$.

The *logical rules of inference* are:

$(\wedge)$	$\vdash \Gamma, A$ and $\vdash \Gamma, B$	$\Rightarrow$	$\vdash \Gamma, A \wedge B$
$(\vee)$	$\vdash \Gamma, A_i$ for $i \in \{0, 1\}$	$\Rightarrow$	$\vdash \Gamma, A_0 \vee A_1$
$(b\forall)$	$\vdash \Gamma, a \in b \to F(a)$	$\Rightarrow$	$\vdash \Gamma, (\forall x \in b)F(x)$
$(pb\forall)$	$\vdash \Gamma, a \subseteq b \to F(a)$	$\Rightarrow$	$\vdash \Gamma, (\forall x \subseteq b)F(x)$
$(\forall)$	$\vdash \Gamma, F(a)$	$\Rightarrow$	$\vdash \Gamma, \forall x F(x)$
$(b\exists)$	$\vdash \Gamma, a \in b \wedge F(a)$	$\Rightarrow$	$\vdash \Gamma, (\exists x \in b)F(x)$
$(pb\exists)$	$\vdash \Gamma, a \subseteq b \wedge F(a)$	$\Rightarrow$	$\vdash \Gamma, (\exists x \subseteq b)F(x)$
$(\exists)$	$\vdash \Gamma, F(a)$	$\Rightarrow$	$\vdash \Gamma, \exists x F(x)$
(Cut)	$\vdash \Gamma, A$ and $\vdash \Gamma, \neg A$	$\Rightarrow$	$\vdash \Gamma$.

In the foregoing rules $F(a)$ is an arbitrary formula. Of course, it is demanded that in $(b\forall)$, $(pb\forall)$ and $(\forall)$ the free variable a is not to occur in the conclusion; a is called the *eigenvariable* of that inference.

The *non–logical rule of inference* is:

$(\Delta_0^{\mathcal{P}}$–COLLR) $\quad \vdash \Gamma, (\forall x \in a)\exists y H(x,y) \quad \Rightarrow \quad \vdash \Gamma, \exists z(\forall x \in a)(\exists y \in z) H(x,y)$

for every $\Delta_0^{\mathcal{P}}$–formula $H(b,c)$.

We shall conceive of axioms as inferences with an empty set of premisses. The *minor formulae* (m.f.) of an inference are those formulae which are rendered prominently in its premises. The *principal formulae* (p.f.) of an inference are the formulae rendered prominently in its conclusion. (Cut) has no p.f. So any inference has the form

$$(*) \quad \text{For all } i < k \ \vdash \Gamma, \Xi_i \quad \Rightarrow \quad \vdash \Gamma, \Xi$$

$(0 \leq k \leq 2)$, where Ξ consists of the p.f. and Ξ_i is the set of m.f. in the i–th premise. The formulae in Γ are called *side formulae* (s.f.) of $(*)$.

Derivations are defined inductively, as usual. $\mathcal{D}$,$\mathcal{D}$',$\mathcal{D}_0, \cdots$ range as syntactic variables over derivations. All this is completely standard, and we refer to [**32**] for notions like "*length of a derivation* $\mathcal{D}$" (abbreviated by $\mid \mathcal{D} \mid$), "*last inference of* $\mathcal{D}$", "*direct subderivation of* $\mathcal{D}$". We write $\mathcal{D} \vdash \Gamma$ if $\mathcal{D}$ is a derivation of Γ.

4 A representation system for the Bachmann-Howard ordinal

Definition 4.1 Let Ω be a "big" ordinal, e.g. $\Omega = \aleph_1$ or ω_1^{ck}. By recursion on α we define sets $C^{\Omega}(\alpha,\beta)$ and the ordinal $\psi_\Omega(\alpha)$ as follows:

$$(1) \qquad C^{\Omega}(\alpha,\beta) \quad = \quad \begin{cases} \textit{closure of } \beta \cup \{0,\Omega\} \\ \textit{under:} \\ +, (\xi \mapsto \omega^{\xi}) \\ (\xi \longmapsto \psi_\Omega(\xi))_{\xi<\alpha} \end{cases}$$

$$(2) \qquad \psi_\Omega(\alpha) \quad \simeq \quad \min\{\rho < \Omega : \ C^{\Omega}(\alpha,\rho) \cap \Omega = \rho\}.$$

It can be shown that $\psi_\Omega(\alpha)$ is always defined and that

$$\psi_\Omega(\alpha) < \Omega.$$

In the case of Ω being ω_1^{ck}, this follows from [**22**]. Moreover,

$$[\psi_\Omega(\alpha),\Omega) \ \cap \ C^{\Omega}(\alpha,\psi_\Omega(\alpha)) \ = \ \emptyset\,.$$

Thus the order-type of the ordinals below Ω which belong to the set $C^{\Omega}(\alpha,\psi_\Omega(\alpha))$ is $\psi_\Omega(\alpha)$. $\psi_\Omega(\alpha)$ is also a countable ordinal. In more pictorial terms, $\psi_\Omega(\alpha)$ is the α^{th} collapse of Ω.

Let $\varepsilon_{\Omega+1}$ be the least ordinal $\alpha > \Omega$ such that $\omega^\alpha = \alpha$. The set of ordinals $C^{\Omega}(\varepsilon_{\Omega+1},0)$ gives rise to an elementary computable ordinal representation system (cf. [**13, 7, 22, 25**]). In what follows, $C^{\Omega}(\varepsilon_{\Omega+1},0)$ will sometimes be denoted by $\mathcal{T}(\Omega)$.

In point of fact,

$$C^{\Omega}(\varepsilon_{\Omega+1},0) \ \cap \ \Omega = \psi_\Omega(\varepsilon_{\Omega+1}).$$

The ordinal $\psi_\Omega(\varepsilon_{\Omega+1})$ is known as the **Bachmann-Howard ordinal**. Its relation to **KP** is that it is the proof-theoretic ordinal of this theory as was shown by Jäger [**13**]. Moreover it is the smallest ordinal such that $L_{\psi_\Omega(\varepsilon_{\Omega+1})}$ is a Π_2-model of **KP** (see [**21**, Theorem 2.1] or [**29**, theorem 4.3]), i.e., whenever **KP** proves a Π_2 sentence C of set theory, then $L_{\psi_\Omega(\varepsilon_{\Omega+1})} \models C$.

5 The infinitary proof system $RS_\Omega^{\mathcal{P}}$

Henceforth all ordinals will be assumed to belong to $C^\Omega(\varepsilon_{\Omega+1}, 0)$.

The problem of "naming" sets will be solved by building a formal von Neumann hierarchy using the ordinals $< \Omega$ belonging to this set (i.e., ordinals $< \psi_\Omega(\varepsilon_{\Omega+1})$).

Definition 5.1 We define the $RS_\Omega^{\mathcal{P}}$–terms. To each $RS_\Omega^{\mathcal{P}}$–term t we also assign its *level*, $|t|$.

1. For each $\alpha < \Omega$, $\mathbb{V}_\alpha$ is an $RS_\Omega^{\mathcal{P}}$–term with $|\mathbb{V}_\alpha| = \alpha$.
2. For each $\alpha < \Omega$, we have infinitely many free variables $a_1^\alpha, a_2^\alpha, a_3^\alpha, \ldots$ which are $RS_\Omega^{\mathcal{P}}$–terms with $|a_i^\alpha| = \alpha$.
3. If $F(x, \vec{y})$ is a $\Delta_0^{\mathcal{P}}$ formula (whose free variables are exactly those indicated) and $\vec{s} \equiv s_1, \cdots, s_n$ are $RS_\Omega^{\mathcal{P}}$–terms, then the formal expression

$$\{x \in \mathbb{V}_\alpha \mid F(x, \vec{s})\}$$

is an $RS_\Omega^{\mathcal{P}}$–term with $|\{x \in \mathbb{V}_\alpha \mid F(x, \vec{s})\}| = \alpha$.

The $RS_\Omega^{\mathcal{P}}$*–formulae* are the expressions of the form $F(s_1, \ldots, s_n)$, where $F[a_1, \ldots, a_n]$ is a formula of $\mathbf{KP}(\mathcal{P})$ and $s_1, \ldots, s_n$ are $RS_\Omega^{\mathcal{P}}$-terms. We set

$$|F(s_1, \ldots, s_n)| \quad = \quad \{|s_1|, \ldots, |s_n|\}.$$

If $F[a_1, \ldots, a_n]$ is in $\Delta_0^{\mathcal{P}}$, $F(s_1, \ldots, s_n)$ is also called a $\Delta_0^{\mathcal{P}}$ formula (of $RS_\Omega^{\mathcal{P}}$).

As in the case of the Tait-style version of $\mathbf{KP}(\mathcal{P})$, we let $\neg A$ be the formula which arises from A by (i) putting $\neg$ in front of each atomic formula, (ii) replacing $\wedge, \vee, (\forall x \in s), (\exists x \in s), (\forall x \subseteq s), (\exists x \subseteq s), \forall x, \exists x$ by $\vee, \wedge, (\exists x \in s), (\forall x \in s), (\exists x \subseteq s), (\forall x \subseteq s), \exists x, \forall x$, respectively, and (iii) dropping double negations. $A \to B$ stands for $\neg A \vee B$.

Convention: In the sequel, $RS_\Omega^{\mathcal{P}}$–formulae will simply be referred to as formulae. The same usage applies to $RS_\Omega^{\mathcal{P}}$–terms.

We denote by upper case Greek letters $\Gamma, \Delta, \Lambda, \ldots$ finite sets of $RS_\Omega^{\mathcal{P}}$–formulae. The intended meaning of $\Gamma = \{A_1, \cdots, A_n\}$ is the disjunction $A_1 \vee \cdots \vee A_n$. Γ, Ξ stands for $\Gamma \cup \Xi$ and Γ, A stands for $\Gamma \cup \{A\}$.

Definition 5.2 The *axioms* of $RS_\Omega^{\mathcal{P}}$ are:

(A1) $\Gamma, A, \neg A$ for A in $\Delta_0^{\mathcal{P}}$.

(A2) $\Gamma, t = t$.

(A3) $\Gamma, s_1 \neq t_1, \ldots, s_n \neq t_n, \neg A(s_1, \ldots, s_n), A(t_1, \ldots, t_n)$
for $A(s_1, \ldots, s_n)$ in $\Delta_0^{\mathcal{P}}$.

(A4) $\Gamma,\ s \in \mathbb{V}_\alpha$ if $|\,s\,| < \alpha$.

(A5) $\Gamma,\ s \subseteq \mathbb{V}_\alpha$ if $|\,s\,| \leq \alpha$.

(A6) $\Gamma, t \notin \{x \in \mathbb{V}_\alpha \mid F(x,\vec{s}\,)\}, F(t,\vec{s}\,)$
whenever $F(t,\vec{s}\,)$ is $\Delta_0^{\mathcal{P}}$ and $|\,t\,| < \alpha$.

(A7) $\Gamma, \neg F(t,\vec{s}\,), t \in \{x \in \mathbb{V}_\alpha \mid F(x,\vec{s}\,)\}$
whenever $F(t,\vec{s}\,)$ is $\Delta_0^{\mathcal{P}}$ and $|\,t\,| < \alpha$.

The *inference rules* of $RS_\Omega^{\mathcal{P}}$ are:

$$(\wedge) \qquad \frac{\Gamma, A \quad \Gamma, A'}{\Gamma, A \wedge A'}$$

$$(\vee) \qquad \frac{\Gamma, A_i}{\Gamma, A_0 \vee A_1} \quad \text{if } i = 0 \text{ or } i = 1$$

$$(b\forall)_\infty \qquad \frac{\Gamma,\ s{\in}t \to F(s) \ \text{ for all } |\,s\,| < |\,t\,|}{\Gamma,\ (\forall x{\in}t)F(x)}$$

$$(b\exists) \qquad \frac{\Gamma,\ s{\in}t \wedge F(s)}{\Gamma, (\exists x{\in}t)F(x)} \quad \text{if } |\,s\,| < |\,t\,|$$

$$(pb\forall)_\infty \qquad \frac{\Gamma, s \subseteq t \to F(s) \ \text{ for all } |\,s\,| \leq |\,t\,|}{\Gamma, (\forall x \subseteq t)F(x)}$$

$$(pb\exists) \qquad \frac{\Gamma, s \subseteq t \wedge F(s)}{\Gamma, (\exists x \subseteq t)F(x)} \quad \text{if } |\,s\,| \leq |\,t\,|$$

$$(\forall)_\infty \qquad \frac{\Gamma,\ F(s) \ \text{ for all } s}{\Gamma, \forall x\, F(x)}$$

$$(\exists) \qquad \frac{\Gamma,\ F(s)}{\Gamma, \exists x\, F(x)}$$

$$(\notin)_\infty \qquad \frac{\Gamma, r \in t \to r \neq s \ \text{ for all } |\,r\,| < |\,t\,|}{\Gamma, s{\notin}t}$$

$$(\in) \qquad \frac{\Gamma,\ r{\in}t \ \wedge \ r = s}{\Gamma, s{\in}t} \quad \text{if } |\,r\,| < |\,t\,|$$

$$(\not\subseteq)_\infty \qquad \frac{\Gamma, r \subseteq t \to r \neq s \ \text{ for all } |\,r\,| \leq |\,t\,|}{\Gamma, s{\not\subseteq}t}$$

$$(\subseteq) \qquad \frac{\Gamma,\ s = r \ \wedge \ r \subseteq t}{\Gamma, s \subseteq t} \quad \text{if } |\,r\,| \leq |\,s\,|$$

$$\text{(Cut)} \qquad \frac{\Gamma, A \qquad \Gamma, \neg A}{\Gamma}$$

$$(\Sigma^{\mathcal{P}}\text{-}Ref) \qquad \frac{\Gamma, A}{\Gamma, \exists z\, A^z} \quad \text{if } A \text{ is a } \Sigma^{\mathcal{P}}\text{-formula,}$$

where a formula is said to be in $\Sigma^{\mathcal{P}}$ if all its unbounded quantifiers are existential. A^z results from A by restricting all unbounded quantifiers to z.

5.1 $\mathcal{H}$–controlled derivations

In general in $RS_{\Omega}^{\mathcal{P}}$ we cannot remove cuts that have $\Delta_0^{\mathcal{P}}$ cut formulae. What's more, the rule $(\Sigma^{\mathcal{P}}\text{-}Ref)$ poses an obstacle to removing cuts involving $\Sigma_1^{\mathcal{P}}$ formulae. notwithstanding that, it will turn out that cuts of a complexity higher than $\Delta_0^{\mathcal{P}}$ can be removed from derivations of $\Sigma^{\mathcal{P}}$ formulae if they are of a very uniform kind.

For the presentation of infinitary proofs we draw on [**7**]. Buchholz developed a very elegant and flexible setting for describing uniformity in infinitary proofs, called operator controlled derivations.

Definition 5.3 *Let*

$$P(ON) = \{X : X \textit{ is a set of ordinals}\}.$$

A class function

$$\mathcal{H} : P(ON) \to P(ON)$$

will be called operator if $\mathcal{H}$ is a closure operator, i.e monotone, inclusive and idempotent, and satisfies the following conditions for all $X \in P(ON)$:

(1) $0 \in \mathcal{H}(X)$ *and* $\Omega \in \mathcal{H}(X)$.

(2) *If α has Cantor normal form $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$, then*

$$\alpha \in \mathcal{H}(X) \iff \alpha_1, ..., \alpha_n \in \mathcal{H}(X).$$

The latter ensures that $\mathcal{H}(X)$ will be closed under $+$ and $\sigma \mapsto \omega^{\sigma}$, and decomposition of its members into additive and multiplicative components.

For a sequent $\Gamma = \{A_1, \ldots, A_n\}$ we define

$$|\,\Gamma\,| \;:=\; |\,A_1\,| \cup \ldots \cup |\,A_n\,|\,.$$

If s is an $RS_{\Omega}^{\mathcal{P}}$-term, the operator $\mathcal{H}[s]$ is defined by

$$\mathcal{H}[s](X) \;=\; \mathcal{H}(X \cup \{|\,s\,|\}).$$

Likewise, if $\mathfrak{X}$ is a formula or a sequent we define

$$\mathcal{H}[\mathfrak{X}](X) \;=\; \mathcal{H}(X \cup |\,\mathfrak{X}\,|).$$

If $\mathfrak{Y}_i$ is a term, or a formula, or a sequent for all $1 \leq i \leq n$, we let $\mathcal{H}[\mathfrak{Y}_1, \mathfrak{Y}_2] = \mathcal{H}[\mathfrak{Y}_1])[\mathfrak{Y}_2]$, $\mathcal{H}[\mathfrak{Y}_1, \mathfrak{Y}_2, \mathfrak{Y}_3] = \mathcal{H}[\mathfrak{Y}_1, \mathfrak{Y}_2])[\mathfrak{Y}_3]$, etc.

Lemma 5.4 *Let $\mathcal{H}$ be an operator. Let s be a term and $\mathfrak{X}$ be a formula or a sequent.*

(i) $\forall X, X' \in P(On)[X' \subseteq X \Longrightarrow \mathcal{H}(X') \subseteq \mathcal{H}(X)]$.

(ii) $\mathcal{H}[s]$ *and* $\mathcal{H}[\mathfrak{X}]$ *are operators.*

(iii) $|\,\mathfrak{X}\,| \subseteq \mathcal{H}[\emptyset] \implies \mathcal{H}[\mathfrak{X}] = \mathcal{H}$.

(iv) $|\,s\,| \in \mathcal{H}[\emptyset] \implies \mathcal{H}[s] = \mathcal{H}$.

Since we also want to keep track of the complexity of cuts appearing in derivations, we endow each formula with an ordinal rank.

Definition 5.5 The *rank* of a formula is determined as follows.

(1) $rk(s{\in}t) := rk(s{\notin}t) := max\{|\,s\,|+1, |\,t\,|+1\}$.

(2) $rk((\exists x{\in}t)F(x)) := rk((\forall x{\in}t)F(x)) := max\{|\,t\,|, rk(F(\mathbb{V}_0))+2\}$.

(3) $rk((\exists x \subseteq t)F(x)) := rk((\forall x \subseteq t)F(x)) := max\{|\,t\,|, rk(F(\mathbb{V}_0))+2\}$.

(4) $rk(\exists x\, F(x)) := rk(\forall x\, F(x)) := max\{\Omega, rk(F(\mathbb{V}_0))+2\}$.

(5) $rk(A \wedge B) := rk(A \vee B) := max\{rk(A), rk(B)\}+1$.

Note that for a $\Delta_0^{\mathcal{P}}$ formula A we have $rk(A) < \Omega$.
There is plenty of leeway in designing the actual rank of a formula.

Definition 5.6 Let $\mathcal{H}$ be an operator and let Λ be a finite set of $RS_\Omega^{\mathcal{P}}$–formulae. $\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Lambda$ is defined by recursion on α.

If Λ is an **axiom** and $|\,\Lambda\,| \cup \{\alpha\} \subseteq \mathcal{H}(\emptyset)$, then $\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Lambda$.

Moreover, we have inductive clauses pertaining to the inference rules of $RS_\Omega^{\mathcal{P}}$, which come with the additional requirement that $|\,\Lambda\,| \cup \{\alpha\} \subseteq \mathcal{H}(\emptyset)$, where Λ is the sequent of the conclusion. We shall not repeat this requirement below. The clauses are the following:

$(\wedge)$
$$\frac{\mathcal{H} \stackrel{\alpha_0}{\underset{\rho}{\vert}} \Gamma, A_0 \qquad \mathcal{H} \stackrel{\alpha_0}{\underset{\rho}{\vert}} \Gamma, A_1}{\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Gamma, A_0 \wedge A_1} \qquad \alpha_0 < \alpha$$

$(\vee)$
$$\frac{\mathcal{H} \stackrel{\alpha_0}{\underset{\rho}{\vert}} \Lambda, A_i}{\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Gamma, A_0 \vee A_1} \qquad \begin{array}{r}\alpha_0 < \alpha\\ i \in \{0,1\}\end{array}$$

$(b\forall)_\infty$
$$\frac{\mathcal{H}[s] \stackrel{\alpha_s}{\underset{\rho}{\vert}} \Gamma, s{\in}t \to F(s) \text{ for all } |\,s\,| < |\,t\,|}{\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Gamma, (\forall x{\in}t)F(x)} \qquad |\,s\,| \le \alpha_s < \alpha$$

$(b\exists)$
$$\frac{\mathcal{H} \stackrel{\alpha_0}{\underset{\rho}{\vert}} \Gamma, s{\in}t \wedge F(s)}{\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Gamma, (\exists x{\in}t)F(x)} \qquad \begin{array}{r}\alpha_0 < \alpha\\ |\,s\,| < |\,t\,|\\ |\,s\,| < \alpha\end{array}$$

$(pb\forall)_\infty$
$$\frac{\mathcal{H}[s] \stackrel{\alpha_s}{\underset{\rho}{\vert}} \Gamma, s \subseteq t \to F(s) \text{ for all } |\,s\,| \le |\,t\,|}{\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Gamma, (\forall x \subseteq t)F(x)} \qquad |\,s\,| \le \alpha_s < \alpha$$

$(pb\exists)$
$$\frac{\mathcal{H} \stackrel{\alpha_0}{\underset{\rho}{\vert}} \Gamma, s \subseteq t \wedge F(s)}{\mathcal{H} \stackrel{\alpha}{\underset{\rho}{\vert}} \Gamma, (\exists x \subseteq t)F(x)} \qquad \begin{array}{r}\alpha_0 < \alpha\\ |\,s\,| \le |\,t\,|\\ |\,s\,| < \alpha\end{array}$$

$$(\forall)_\infty \qquad \frac{\mathcal{H}[s] \vdash^{\alpha_s}_{\rho} \Gamma, F(s) \text{ for all } s}{\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, (\forall x F(x)} \qquad |\,s\,| \leq \alpha_s + 1 < \alpha$$

$$(\exists) \qquad \frac{\mathcal{H} \vdash^{\alpha_0}_{\rho} \Gamma, F(s)}{\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, \exists x F(x)} \qquad \begin{array}{r} \alpha_0 < \alpha \\ |\,s\,| < \alpha \end{array}$$

$$(\notin)_\infty \qquad \frac{\mathcal{H}[r] \vdash^{\alpha_r}_{\rho} \Gamma, r{\in}t \rightarrow r \neq s \text{ for all } |\,r\,| < |\,t\,|}{\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, s \notin t} \qquad |\,r\,| \leq \alpha_r < \alpha$$

$$(\in) \qquad \frac{\mathcal{H} \vdash^{\alpha_0}_{\rho} \Gamma, r{\in}t \wedge r = s}{\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, s{\in}t} \qquad \begin{array}{r} \alpha_0 < \alpha \\ |\,r\,| < |\,t\,| \\ |\,r\,| < \alpha \end{array}$$

$$(\not\subseteq)_\infty \qquad \frac{\mathcal{H}[r] \vdash^{\alpha_r}_{\rho} \Gamma, r \subseteq t \rightarrow r \neq s \text{ for all } |\,r\,| \leq |\,t\,|}{\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, s \not\subseteq t} \qquad |\,r\,| \leq \alpha_r < \alpha$$

$$(\subseteq) \qquad \frac{\mathcal{H} \vdash^{\alpha_0}_{\rho} \Gamma, r \subseteq t \wedge r = s}{\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, s \subseteq t} \qquad \begin{array}{r} \alpha_0 < \alpha \\ |\,r\,| \leq |\,t\,| \\ |\,r\,| < \alpha \end{array}$$

$$(\text{Cut}) \qquad \frac{\mathcal{H} \vdash^{\alpha_0}_{\rho} \Lambda, B \qquad \mathcal{H} \vdash^{\alpha_0}_{\rho} \Lambda, \neg B}{\mathcal{H} \vdash^{\alpha}_{\rho} \Lambda} \qquad \begin{array}{r} \alpha_0 < \alpha \\ rk(B) < \rho \end{array}$$

$$(\Sigma^{\mathcal{P}}\text{-}Ref) \qquad \frac{\mathcal{H} \vdash^{\alpha_0}_{\rho} \Gamma, A}{\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, \exists z\, A^z} \qquad \begin{array}{r} \alpha_0, \Omega < \alpha \\ A{\in}\Sigma^{\mathcal{P}} \end{array}$$

Remark 5.7 Suppose $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma(s_1, \ldots, s_n)$ where $\Gamma(a_1, \ldots, a_n)$ is a sequent of $\mathbf{KP}(\mathcal{P})$ and $s_1, \ldots, s_n$ are $RS^{\mathcal{P}}_{\Omega}$-terms. Then we have that $|\,s_1\,|, \ldots, |\,s_n\,| \in \mathcal{H}(\emptyset)$. Standing in sharp contrast to the ordinal analysis of **KP** (cf. [**13, 7**]), however, the terms s_i may and often will contain subterms that the operator $\mathcal{H}$ does **not** control, that is, subterms t with $|\,t\,| \notin \mathcal{H}(\emptyset)$.

The following observation is easily established by induction on α.

Lemma 5.8 (Weakening)

$$\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma \;\wedge\; \alpha \leq \alpha' {\in} \mathcal{H} \;\wedge\; \rho \leq \rho' \;\wedge\; |\,\Lambda\,| \subseteq \mathcal{H}(\emptyset) \implies \mathcal{H} \vdash^{\alpha'}_{\rho'} \Gamma, \Lambda\,.$$

Lemma 5.9 (Inversion) (i) *If* $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, A \vee B$ *and* $rk(A{\vee}B) \geq \Omega$, *then* $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, A, B\,.$

(ii) *If* $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, A_0 \wedge A_1$, $i \in \{0, 1\}$ *and* $rk(A_0 \wedge A_1) \geq \Omega$, *then* $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, A_i\,.$

(iii) $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, \forall x\, F(x) \;\wedge\; \gamma{\in}\mathcal{H}(\emptyset) \;\wedge\; \gamma < \Omega \implies \mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, (\forall x{\in}\mathbb{V}_\gamma) F(x)\,.$

(iv) *If* $\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, (\forall x \in t)\, F(x)$ *and* $rk(F(\mathbb{V}_0)) \geq \Omega$, *then* $\mathcal{H}[s] \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, s{\in}t \to F(s)$ *for all* $|\, s\,| < |\, t\,|$.

(v) *If* $\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, (\forall x \subseteq t)\, F(x)$ *and* $rk(F(\mathbb{V}_0)) \geq \Omega$, *then* $\mathcal{H}[s] \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, s \subseteq t \to F(s)$ *for all* $|\, s\,| \leq |\, t\,|$.

Proof: All proofs are by induction on α. Note that if a formula C of $rk(C) \geq \Omega$ is active in a derivation then it must have been the principal formula of an inference.

We show (iii). Suppose that $\forall x\, F(x)$ was the principal formula of the last inference. Then we have $\mathcal{H}[s] \overset{\alpha_s}{\underset{\rho}{\vdash}} \Gamma, \forall x\, F(x), F(s)$ for all terms s, using weakening (Lemma 5.8) if $\forall x\, F(x)$ was not a side formula of the inference. Moreover, $|\, s\,| \leq \alpha_s + 1 < \alpha$ holds for all s. Inductively we have $\mathcal{H}[s] \overset{\alpha_s}{\underset{\rho}{\vdash}} \Gamma, (\forall x{\in}\mathbb{V}_\gamma) F(x), F(s)$ for all $|\, s\,| < \gamma$. As $\mathcal{H}[s] \overset{\alpha_s}{\underset{\rho}{\vdash}} \Gamma, (\forall x{\in}\mathbb{V}_\gamma) F(x), s \in \mathbb{V}_\gamma$ holds for $|\, s\,| < \gamma$ on account of being an axiom, we get $\mathcal{H}[s] \overset{\alpha_s+1}{\underset{\rho}{\vdash}} \Gamma, (\forall x{\in}\mathbb{V}_\gamma) F(x), s \in \mathbb{V}_\gamma \wedge F(s)$ for all $|\, s\,| < \gamma$ via an inference $(\wedge)$, and hence $\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, (\forall x{\in}\mathbb{V}_\gamma) F(x)$ via an inference $(b\forall)$.

If $\forall x\, F(x)$ is not the principal formula of the last inference, then the assertion follows by using the induction hypothesis to its premisses and re-applying the same inference.

□

6 Embedding

To connect $\mathbf{KP}(\mathcal{P})$ with the infinitary system $RS_\Omega^\mathcal{P}$ we show that $\mathbf{KP}(\mathcal{P})$ can be embedded into $RS_\Omega^\mathcal{P}$. Indeed, the finite $\mathbf{KP}(\mathcal{P})$-derivations give rise to very uniform infinitary derivations.

Definition 6.1 For $\Gamma = \{A_1, \ldots, A_n\}$ let

$$no(\Gamma) := \omega^{rk(A_1)} \# \cdots \# \omega^{rk(A_n)}.$$

We define

$$\Vdash \Gamma \quad :\Longleftrightarrow \quad \text{for all operators } \mathcal{H},\ \mathcal{H}[\Gamma] \overset{no(\Gamma)}{\underset{0}{\vdash}} \Gamma$$

and

$$\Vdash_\rho^\xi \Gamma \quad :\Longleftrightarrow \quad \text{for all operators } \mathcal{H},\ \mathcal{H}[\Gamma] \overset{no(\Gamma)\#\xi}{\underset{\rho}{\vdash}} \Gamma\,.$$

Lemma 6.2 *For all formulae* A,

$$\Vdash A, \neg A.$$

Proof: We proceed by induction on the syntactic complexity of A. For A in $\Delta_0^\mathcal{P}$ this is an axiom of $RS_\Omega^\mathcal{P}$. Suppose A is of the form $\forall x F(x)$. Let $\mathcal{H}$ be an arbitrary operator. Let $\alpha_s := no(\{F(s), \neg F(s)\})$ and $\alpha := no(\{\forall x F(x), \exists x \neg F(x)\})$. Inductively we have $\mathcal{H}[F(s)] \overset{\alpha_s}{\underset{0}{\vdash}} F(s), \neg F(s)$ for all terms s. Using an inference $(\exists)$ we get $\mathcal{H}[F(s)] \overset{no(\{F(s),\exists x \neg F(x)\})}{\underset{0}{\vdash}} F(s), \exists x \neg F(x)$. Hence, via an inference $(\forall)$, we arrive at $\mathcal{H}[\forall x F(x)] \overset{\alpha}{\underset{0}{\vdash}} \forall x F(x), \exists x \neg F(x)$, noting that $\mathcal{H}[F(s)\}] \subseteq (\mathcal{H}[\forall x \neg F(x)])[s]$.

The other cases are similar. □

Lemma 6.3 (Equality and Extensionality)

$$\Vdash s_1 \neq t_1, \ldots, s_n \neq t_n, \neg A(s_1, \ldots, s_n), A(t_1, \ldots, t_n).$$

Proof: We proceed by induction on the buildup of $A(\vec{s})$. If $A(\vec{s})$ is $\Delta_0^{\mathcal{P}}$ then this is an axiom.

Suppose $A(\vec{s})$ is a formula $\forall x\, F(x, \vec{s})$. Let $\vec{s} \neq \vec{t}$ stand for $s_1 \neq t_1, \ldots, s_n \neq t_n$. Let $\Gamma_r := \{\vec{s} \neq \vec{t}, \neg F(r, \vec{s}), F(r, \vec{t})\}$ and $\alpha_r := no(\Gamma_r)$. Let $\mathcal{H}$ be an arbitrary operator. Inductively we have

$$\mathcal{H}[\Gamma_r] \; \Big|\!\frac{\alpha_r}{0}\; \Gamma_r$$

for all terms r. Using an inference $(\exists)$ we obtain $\mathcal{H}[\tilde{\Gamma}_r] \; |\!\frac{\tilde{\alpha}_r}{0}\; \tilde{\Gamma}_r$ where

$$\tilde{\Gamma}_r := \{\vec{s} \neq \vec{t}, \exists x \neg F(x, \vec{s}), F(r, \vec{t})\}$$

and $\tilde{\alpha}_r := no(\tilde{\Gamma}_r)$, noting that $|\, r \,| < \Omega \leq no(\exists x \neg F(x, \vec{s}))$. Thus, using an inference $(\forall)$, we have

$$\mathcal{H}[\Gamma] \; \Big|\!\frac{no(\Gamma)}{0}\; \Gamma$$

where $\Gamma := \{\vec{s} \neq \vec{t}, \exists x \neg F(x, \vec{s}), \forall x F(x, \vec{t})\}$. In the latter we used the fact that $\mathcal{H}[\tilde{\Gamma}_r] \subseteq (\mathcal{H}[\Gamma])[r]$.

The other cases are similar. □

Lemma 6.4 (Set Induction)

$$\Vdash \forall x\, [(\forall y{\in}x) F(y) \to F(x)] \;\longrightarrow\; \forall x F(x).$$

Proof. Fix an operator $\mathcal{H}$. Let $A \equiv (\forall x\, [(\forall y{\in}x) F(y) \to F(x)]$. First, we show, by induction on $|\, s \,|$, that

$$(+) \qquad \mathcal{H}[A, s] \; \Big|\!\frac{\omega^{rk(A)} \# \omega^{|s|+1}}{0}\; \neg A, F(s)\,.$$

So assume that

$$\mathcal{H}[A, t] \; \Big|\!\frac{\omega^{rk(A)} \# \omega^{|t|+1}}{0}\; \neg A, F(t)$$

holds for all $|\, t \,| < |\, s \,|$. Using $(\vee)$, this yields

$$\mathcal{H}[A, s, t] \; \Big|\!\frac{\omega^{rk(A)} \# \omega^{|t|+1} + 1}{0}\; \neg A, t{\in}s \to F(t)$$

for all $|\, t \,| < |\, s \,|$, and hence

$$(1) \qquad \mathcal{H}[A, s] \; \Big|\!\frac{\omega^{rk(A)} \# \omega^{|s|} + 2}{0}\; \neg A, (\forall x{\in}s) F(x)$$

via $(\forall)_\infty$. Set $\eta_s := \omega^{rk(A)} \# \omega^{|s|} + 2$. By Lemma 6.2 we have $\mathcal{H}[A, s] \; |\!\frac{\eta_s}{0}\; \neg F(s), F(s)$. Therefore, using (1) and $(\wedge)$,

$$\mathcal{H}[A, s] \; \Big|\!\frac{\eta_s + 1}{0}\; \neg A, (\forall y{\in}s) F(y) \wedge \neg F(s), F(s)\,.$$

From the latter we obtain

$$\mathcal{H}[A, s] \; \Big|\!\frac{\eta_s + 2}{0}\; \neg A, \exists x\, [(\forall y{\in}x) F(y) \wedge \neg F(x)], F(s)$$

via $(\exists)$. This shows $(+)$.

Finally, $(+)$ enables us to deduce, via $(\forall)_\infty$, that

$$\mathcal{H}[A, s] \; \Big|\!\frac{\omega^{rk(A)} + \Omega}{0}\; \neg A, \forall x F(x)\,.$$

From this the assertion follows by applying ($\vee$) twice. $\square$

Lemma 6.5 (Infinity Axiom) *For any operator $\mathcal{H}$ we have*

$$\mathcal{H} \mathrel{|}\!\frac{\omega+2}{0}\, \exists x\, [(\exists y{\in}x)\, y \in x \;\wedge\; (\forall y \in x)(\exists z{\in}x)\, y \in z]\,.$$

Proof: Let s be a term with $|\,s\,| = n < \omega$. Then $\mathcal{H} \mathrel{|}\!\frac{0}{0}\, s{\in}\mathbb{V}_{n+1}$ and $\mathcal{H} \mathrel{|}\!\frac{0}{0}\, \mathbb{V}_{n+1} \in \mathbb{V}_\omega$ since these formulae are axioms. Via ($\wedge$) we deduce $\mathcal{H} \mathrel{|}\!\frac{1}{0}\, \mathbb{V}_{n+1} \in \mathbb{V}_\omega \wedge s{\in}\mathbb{V}_{n+1}$ and hence $\mathcal{H} \mathrel{|}\!\frac{n+2}{0}\, (\exists z \in \mathbb{V}_\omega) s{\in}z$, using ($b\exists$). An inference ($\vee$) yields

$$\mathcal{H} \mathrel{|}\!\frac{n+3}{0}\, s{\in}\mathbb{V}_\omega \to (\exists z \in \mathbb{V}_\omega) s{\in}z\,.$$

Since this holds for all terms s with $|\,s\,| < \omega$, we conclude that

(3) $$\mathcal{H} \mathrel{|}\!\frac{\omega}{0}\, (\forall y{\in}\mathbb{V}_\omega)(\exists z \in \mathbb{V}_\omega) y{\in}z\,.$$

Since $\mathbb{V}_0 \in \mathbb{V}_\omega$ is an axiom we have $\mathcal{H} \mathrel{|}\!\frac{1}{0}\, \mathbb{V}_0 \in \mathbb{V}_\omega \wedge \mathbb{V}_0 \in \mathbb{V}_\omega$ via ($\wedge$) and thus

(4) $$\mathcal{H} \mathrel{|}\!\frac{2}{0}\, (\exists z \in \mathbb{V}_\omega) z \in \mathbb{V}_\omega\,,$$

using ($b\exists$). Combining (3) and (4) we arrive at

$$\mathcal{H} \mathrel{|}\!\frac{\omega+1}{0}\, (\exists z \in \mathbb{V}_\omega) z \in \mathbb{V}_\omega \;\wedge\; (\forall y{\in}\mathbb{V}_\omega)(\exists z \in \mathbb{V}_\omega) y{\in}z\,.$$

Thus an inference ($b\exists$) furnishes us with

$$\mathcal{H} \mathrel{|}\!\frac{\omega+2}{0}\, \exists x\, [(\exists z \in x) z \in x \;\wedge\; (\forall y{\in}x)(\exists z{\in}x) y{\in}z]\,.$$

$\square$

Lemma 6.6 ($\Delta_0^{\mathcal{P}}$–Separation) *Let $A(a, b, c_1, \ldots, c_n)$ be a $\Delta_0^{\mathcal{P}}$–formula of $\mathcal{L}$ with all free variables among the exhibited. Let $r, s_1, \ldots, s_n$ be $RS_\Omega^{\mathcal{P}}$-terms. Let $\mathcal{H}$ be an arbitrary operator. Then:*

$$\mathcal{H}[r, \vec{s}\,] \mathrel{|}\!\frac{\alpha+8}{\rho}\, \exists y\, [(\forall x{\in}y)(x{\in}r \wedge A(x, r, \vec{s}\,) \;\wedge\; (\forall x{\in}r)(A(x, r, \vec{s}\,) \to x{\in}y)]\,,$$

where $\alpha = |\,r\,|$ and $\rho = \max\{|\,r\,|, |\,s_1\,|, \ldots, |\,s_n\,|\} + \omega$.

Proof: Define the $RS_\Omega^{\mathcal{P}}$-term p by

$$p := \{x \in \mathbb{V}_\alpha \mid x{\in}r \;\wedge\; A(x, r, \vec{s}\,)\}.$$

Then $|\,p\,| = \alpha$. Let $\tilde{\mathcal{H}} := \mathcal{H}[r, \vec{s}\,]$. We have $\tilde{\mathcal{H}}[t] \mathrel{|}\!\frac{0}{0}\, t \notin p, t \in r \wedge A(t, r, \vec{s}\,)$ for all $|\,t\,| < \alpha$ since this is an axiom. Hence $\tilde{\mathcal{H}}[t] \mathrel{|}\!\frac{2}{0}\, t{\in}p \to t \in r \wedge A(t, r, \vec{s}\,)$ using ($\vee$) twice, and therefore

(5) $$\tilde{\mathcal{H}} \mathrel{|}\!\frac{\alpha+3}{0}\, (\forall x{\in}p)(x \in r \;\wedge\; A(x, r, \vec{s}\,))$$

by applying $(b\forall)_\infty$. We also have $\tilde{\mathcal{H}}[t] \mathrel{|}\!\frac{0}{0}\, t \notin r, t{\in}r$ and $\tilde{\mathcal{H}}[t] \mathrel{|}\!\frac{0}{0}\, \neg A(t, r, \vec{s}\,), A(t, r, \vec{s}\,)$ as these sequents are axioms. Using ($\wedge$) and weakening (Lemma 5.8) we conclude that

(6) $$\tilde{\mathcal{H}}[t] \mathrel{|}\!\frac{1}{0}\, t \notin r, \neg A(t, r, \vec{s}\,), t{\in}r \;\wedge\; A(t, r, \vec{s}\,)\,.$$

Since $\tilde{\mathcal{H}}[t] \,\Big|\!\frac{0}{0}\, \neg(t{\in}r \,\wedge\, A(t,r,\vec{s}\,)), t{\in}p$ holds on account of being an axiom, a cut applied to (6) and the latter yields

$$\text{(7)} \qquad \tilde{\mathcal{H}}[t] \,\Big|\!\frac{1}{\rho}\, t \notin r, \neg A(t,r,\vec{s}\,), t{\in}p\,.$$

since $rk(t{\in}r \,\wedge\, A(t,r,\vec{s}\,)) < \rho$ holds for terms t with $|\,t\,| < \alpha$. Now use $(\vee)$ four times to arrive at

$$\text{(8)} \qquad \tilde{\mathcal{H}}[t] \,\Big|\!\frac{5}{\rho}\, t{\in}r \rightarrow (A(t,r,\vec{s}\,) \rightarrow t{\in}p)\,.$$

Applying $(b\forall)_\infty$ to (8) yields

$$\text{(9)} \qquad \tilde{\mathcal{H}} \,\Big|\!\frac{\alpha+6}{\rho}\, (\forall x \in r)(A(x,r,\vec{s}\,) \rightarrow x \in p)\,.$$

Combining (5) and (9) via $(\wedge)$ we have

$$\tilde{\mathcal{H}} \,\Big|\!\frac{\alpha+7}{\rho}\, (\forall x{\in}p)(x \in r \,\wedge\, A(x,r,\vec{s}\,)) \,\wedge\, (\forall x \in r)(A(x,r,\vec{s}\,) \rightarrow x{\in}p)\,.$$

Consequently, by means of $(b\exists)$,

$$\tilde{\mathcal{H}} \,\Big|\!\frac{\alpha+8}{\rho}\, \exists y[(\forall x{\in}y)(x \in r \,\wedge\, A(x,r,\vec{s}\,)) \,\wedge\, (\forall x \in r)(A(x,r,\vec{s}\,) \rightarrow x{\in}y)]\,.$$

□

Lemma 6.7 (Pair and Union) *For any operator $\mathcal{H}$ the following hold:*

(i) $\mathcal{H}[s,t] \,\Big|\!\frac{\alpha+1}{0}\, \exists z\,(s{\in}z \wedge t{\in}z)$ *where* $\alpha = \max(|\,s\,|, |\,t\,|) + 1$.

(ii) $\mathcal{H}[s] \,\Big|\!\frac{\beta+7}{0}\, \exists z\,(\forall y{\in}s)(\forall x{\in}y)(x{\in}z)$ *where* $\beta = |\,s\,|$.

Proof: (i): $s \in \mathbb{V}_\alpha$ and $t \in \mathbb{V}_\alpha$ are axioms. Thus $\mathcal{H}[s,t] \,\Big|\!\frac{1}{0}\, s \in \mathbb{V}_\alpha \,\wedge\, t \in \mathbb{V}_\alpha$, and hence $\mathcal{H}[s,t] \,\Big|\!\frac{\alpha+2}{0}\, \exists z\,(s{\in}z \wedge t{\in}z)$ by means of $(b\exists)$.

(ii): Let r and t be terms of levels $< \beta$. Since $r \in \mathbb{V}_\beta$ is an axiom, we have

$$\mathcal{H}[s] \,\Big|\!\frac{0}{0}\, t \notin s, r \notin t, r \in \mathbb{V}_\beta\ .$$

Thus we get

$$\mathcal{H}[s] \,\Big|\!\frac{2}{0}\, t \notin s, r{\in}t \rightarrow r \in \mathbb{V}_\beta$$

$$\mathcal{H}[s] \,\Big|\!\frac{\beta+3}{0}\, t \notin s, (\forall x{\in}t)x \in \mathbb{V}_\beta$$

$$\mathcal{H}[s] \,\Big|\!\frac{\beta+5}{0}\, t{\in}s \rightarrow (\forall x{\in}t)x \in \mathbb{V}_\beta$$

$$\mathcal{H}[s] \,\Big|\!\frac{\beta+6}{0}\, (\forall y{\in}s)(\forall x{\in}t)x \in \mathbb{V}_\beta$$

$$\mathcal{H}[s] \,\Big|\!\frac{\beta+7}{0}\, \exists z\,(\forall y{\in}s)(\forall x{\in}t)x{\in}z\ .$$

□

Lemma 6.8 (Power Set) *For any operator $\mathcal{H}$ the following holds:*

$$\mathcal{H}[s] \,\Big|\!\frac{\alpha+1}{0}\, \exists z\,(\forall x \subseteq s)\, x \in z\,,$$

where $\alpha = |\,s\,|$.

Proof: Let t be a term with $|\,t\,| \le \alpha$. Then $t \in \mathbb{V}_{\alpha+1}$ is an axiom. Whence, using $(\vee)$ (twice), $(pb\forall)_\infty$, and $(\exists)$, we have

$$\mathcal{H}[s] \,\big|\!\frac{0}{0}\, t \not\subseteq s, t \in \mathbb{V}_{\alpha+1}$$

$$\mathcal{H}[s] \,\big|\!\frac{2}{0}\, t \subseteq s \to t \in \mathbb{V}_{\alpha+1}$$

$$\mathcal{H}[s] \,\big|\!\frac{\alpha+3}{0}\, (\forall x \subseteq s) x \in \mathbb{V}_{\alpha+1}$$

$$\mathcal{H}[s] \,\big|\!\frac{\alpha+4}{0}\, \exists z\, (\forall x \subseteq s) x \in z\,.$$

□

Theorem 6.9 *If*

$$\mathbf{KP}(\mathcal{P}) \vdash \Gamma(a_1, \ldots, a_l)$$

then then there exist $m, n < \omega$ such that

$$\mathcal{H}[s_1, \ldots, s_l\,] \,\big|\!\frac{\omega^{\Omega+m}}{\Omega+n}\, \Gamma(s_1, \ldots, s_l)$$

holds for all $RS_\Omega^{\mathcal{P}}$-terms $s_1, \ldots, s_l$ and operators $\mathcal{H}$. Thus m and n depend only on the $\mathbf{KP}(\mathcal{P})$*-derivation of* $\Gamma(\vec{a})$.

Proof: One proceeds by induction on the length of the $\mathbf{KP}(\mathcal{P})$-derivation of $\Gamma(\vec{a}\,)$. Note that the rank of an $RS_\Omega^{\mathcal{P}}$-formula A is always $< \Omega + \omega$ and thus the norms of $RS_\Omega^{\mathcal{P}}$-sequents will always be $< \omega^{\Omega+\omega}$.

If $\Gamma(\vec{a}\,)$ is an axiom of $\mathbf{KP}(\mathcal{P})$ then the assertion follows from the earlier results of this section.

As an example of a rule we shall treat $(pb\exists)$. So suppose the last inference of our $\mathbf{KP}(\mathcal{P})$-derivation $\mathcal{D}$ was an instance of $(pb\exists)$. Then $\Gamma(\vec{a}\,)$ contains a formula of the form $(\exists x \subseteq a_i) \wedge F(x, \vec{a}\,)$ and there exists a shorter $\mathbf{KP}(\mathcal{P})$-derivation $\mathcal{D}_0$ whose end sequent is either of the form $\Gamma(\vec{a}\,), c \subseteq a_i \wedge F(c, \vec{a}\,)$ with c not occurring in $\Gamma(\vec{a}\,)$ or c is a_j for some $1 \le j \le l$. In the former case the induction hypothesis supplies us with $n_0, m_0 < \omega$ such that

(10) $$\mathcal{H}[\vec{s}\,] \,\big|\!\frac{\omega^{\Omega+m_0}}{\Omega+n_0}\, \Gamma(\vec{s}\,), \mathbb{V}_0 \subseteq s_i \wedge F(\mathbb{V}_0, \vec{s}\,)$$

holds for all terms $\vec{s}$. As $|\,\mathbb{V}_0\,| = 0 \le |\,s_i\,|$ we can apply an inference $(pb\exists)$ in the system $RS_\Omega^{\mathcal{P}}$, yielding

(11) $$\mathcal{H}[\vec{s}\,] \,\big|\!\frac{\omega^{\Omega+m_0+2}}{\Omega+n_0}\, \Gamma(\vec{s}\,), (\exists x \subseteq s_i) F(x, \vec{s}\,)$$

and thus $\mathcal{H}[\vec{s}\,] \,\big|\!\frac{\omega^{\Omega+m_0+2}}{\Omega+n_0}\, \Gamma(\vec{s}\,)$ as $(\exists x \subseteq s_i) F(x, \vec{s}\,)$ belongs to $\Gamma(\vec{s}\,)$.

Now let's turn to the case where c is a_j. Then, by the induction hypothesis, there are $n_0, m_0 < \omega$ such that

(12) $$\mathcal{H}[\vec{s}\,] \,\big|\!\frac{\omega^{\Omega+m_0}}{\Omega+n_0}\, \Gamma(\vec{s}\,), s_j \subseteq s_i \wedge F(s_j, \vec{s}\,)$$

holds for all terms $\vec{s}$. Owing to Lemma 6.3 we can find m_1, n_1 such that with $\rho := \omega^{\Omega+m_1}$ we have

$$\mathcal{H}[\vec{s}, r] \,\big|\!\frac{\rho}{\Omega+n_1}\, s_j \neq r, s_j \not\subseteq s_i, r \subseteq s_i$$

and $\mathcal{H}[\vec{s},r] \mid\!\frac{\rho}{\Omega+n_1} s_j \neq r, \neg F(s_j,\vec{s}\,), F(r,\vec{s}\,)$ hold for all $r,\vec{s}$. By applying weakening and $(\wedge)$ we thus get

$$\mathcal{H}[\vec{s},r] \mid\!\frac{\rho+1}{\Omega+n_1} r \not\subseteq s_i, s_j \neq r, \neg F(s_j,\vec{s}\,), r \subseteq s_i \wedge F(r,\vec{s}\,)$$

for all r with $|\,r\,| \leq |\,s_i\,|$. Now apply $(pb\exists)$, $(\vee)$ (twice), $(\not\subseteq)_\infty$, and $(\vee)$ (twice):

$$\mathcal{H}[\vec{s},r] \mid\!\frac{\rho+2}{\Omega+n_1} r \not\subseteq s_i, s_j \neq r, \neg F(s_j,\vec{s}\,), (\exists x \subseteq s_i)F(x,\vec{s}\,)$$

$$\mathcal{H}[\vec{s},r] \mid\!\frac{\rho+4}{\Omega+n_1} r \subseteq s_i \rightarrow s_j \neq r, \neg F(s_j,\vec{s}\,), (\exists x \subseteq s_i)F(x,\vec{s}\,)$$

$$\mathcal{H}[\vec{s}\,] \mid\!\frac{\rho+5}{\Omega+n_1} s_j \not\subseteq s_i, \neg F(s_j,\vec{s}\,), (\exists x \subseteq s_i)F(x,\vec{s}\,)$$

$$\mathcal{H}[\vec{s}\,] \mid\!\frac{\rho+7}{\Omega+n_1} \neg(s_j \subseteq s_i \wedge F(s_j,\vec{s}\,)), (\exists x \subseteq s_i)F(x,\vec{s}\,)\,. \tag{13}$$

Finally, by applying a cut to (12) and (13) we have

$$\mathcal{H}[\vec{s}\,] \mid\!\frac{\omega^{\Omega+m}}{\Omega+n} \Gamma(\vec{s}\,), (\exists x \subseteq s_i)F(x,\vec{s}\,)$$

i.e., $\mathcal{H}[\vec{s}\,] \mid\!\frac{\omega^{\Omega+m}}{\Omega+n} \Gamma(\vec{s}\,)$, where $m = \max(m_0,m_1)+1$ and n is chosen such that $n > n_0, n_1$ and $rk(s_j \subseteq s_i \wedge F(s_j,\vec{s}\,) < \Omega + n$ for all $\vec{s}$.

The case of the last inference being $(b\exists)$ is treated in the same vein as $(pb\exists)$. All the other inferences are straightforward as the desired assertion can be obtained immediately from the induction hypothesis applied to the premisses followed by the corresponding inference in $RS_\Omega^{\mathcal{P}}$. For example, in the case of the $(\Delta_0^{\mathcal{P}}$-COLLR) one inductively finds $m_0, n_0 < \omega$ such that

$$\mathcal{H}[\vec{s}\,] \mid\!\frac{\omega^{\Omega+m}}{\Omega+n} \Gamma_0(\vec{s}\,), (\forall x{\in}s_i)\exists y\, H(x,y,\vec{s}\,)$$

holds for all $\vec{s}$, where $H(x,y,\vec{a}\,)$ is $\Sigma^{\mathcal{P}}$. Using $(\Sigma^{\mathcal{P}}\text{-}Ref)$ one obtains

$$\mathcal{H}[\vec{s}\,] \mid\!\frac{\omega^{\Omega+m}}{\Omega+n} \Gamma_0(\vec{s}\,), \exists z(\forall x{\in}s_i)(\exists y \in z)\, H(x,y,\vec{s}\,)\,.$$

□

7 Cut elimination

The usual cut elimination procedure works as long as the cut formulae are not in $\Delta_0^{\mathcal{P}}$ and have not been introduced by an inference $(\Sigma^{\mathcal{P}}\text{-}Ref)$. As the principal formula of an inference $(\Sigma^{\mathcal{P}}\text{-}Ref)$ has rank Ω one gets the following result.

Theorem 7.1 (Cut elimination I)

$$\mathcal{H} \mid\!\frac{\alpha}{\Omega+n+1} \Gamma \quad \Rightarrow \quad \mathcal{H} \mid\!\frac{\omega_n(\alpha)}{\Omega+1} \Gamma$$

where $\omega_0(\beta) := \beta$ *and* $\omega_{k+1}(\beta) := \omega^{\omega_k(\beta)}$.

Proof: The proof is standard. For details see [**7**, Lemma 3.14]. □

Lemma 7.2 (Boundedness) *Let* A *be a* $\Sigma^{\mathcal{P}}$*-formula,* $\alpha \leq \beta < \Omega$*, and* $\beta \in \mathcal{H}(\emptyset)$*. If*

$$\mathcal{H} \mid\!\frac{\alpha}{\rho} \Gamma, A$$

then

$$\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, A^{\mathbb{V}_\beta}\,.$$

Proof: Note that the derivation contains no instances of ($\Sigma^{\mathcal{P}}$-*Ref*). The proof is by induction on α. For details see [**7**, Lemma 3.17]. □

The obstacle to pushing cut elimination further is exemplified by the following scenario:

$$\dfrac{\dfrac{\mathcal{H} \overset{\delta}{\underset{\Omega}{\vdash}} \Gamma, A}{\mathcal{H} \overset{\xi}{\underset{\Omega}{\vdash}} \Gamma, \exists z\, A^z}(\Sigma^{\mathcal{P}}\text{-}Ref) \qquad \dfrac{\cdots \mathcal{H}[s] \overset{\xi_s}{\underset{\Omega}{\vdash}} \Gamma, \neg A^s \cdots (s \in \mathcal{T})}{\mathcal{H} \overset{\xi}{\underset{\Omega}{\vdash}} \Gamma, \forall z\, \neg A^z}(\forall)}{\mathcal{H} \overset{\alpha}{\underset{\Omega+1}{\vdash}} \Gamma}(Cut)$$

Fortunately, it is possible to eliminate cuts in the above situation provided that the side formulae Γ are of complexity $\Sigma^{\mathcal{P}}$. The technique is known as "collapsing" of derivations.

If the length of a derivation of $\Sigma^{\mathcal{P}}$-formulae is $\geq \Omega$, then "collapsing" results in a shorter derivation, however, at the cost of a much more complicated controlling operator.

Definition 7.3

$$\mathcal{H}_\delta(X) = \bigcap\{C^\Omega(\alpha, \beta) :\ X \subseteq C^\Omega(\alpha,\beta)\ \wedge\ \delta < \alpha\}$$

Theorem 7.4 (Collapsing Theorem) *Let Γ be a set of $\Sigma^{\mathcal{P}}$-formulae such that $|\,\Gamma\,| \subseteq C^\Omega(\eta+1, \psi_\Omega(\eta+1))$. Also suppose that $\eta \in \mathcal{H}_\eta[\Gamma](\emptyset)$. Then we have*

$$\mathcal{H}_\eta[\Gamma] \overset{\alpha}{\underset{\Omega+1}{\vdash}} \Gamma \quad \Rightarrow \quad \mathcal{H}_{\hat{\alpha}}[\Gamma] \overset{\psi_\Omega(\hat{\alpha})}{\underset{\psi_\Omega(\hat{\alpha})}{\vdash}} \Gamma$$

where $\hat{\alpha} = \eta + \omega^{\Omega+\alpha}$.

Proof by induction on α. Suppose $\mathcal{H}_\eta[\Gamma] \overset{\alpha}{\underset{\Omega+1}{\vdash}} \Gamma$. We shall distinguish cases according to the last inference of $\mathcal{H}_\eta[\Gamma] \overset{\alpha}{\underset{\Omega+1}{\vdash}} \Gamma$. Firstly, note that $\eta \in \mathcal{H}_\eta[\Gamma](\emptyset)$ implies $\eta \in \mathcal{H}_{\hat{\alpha}}[\Gamma](\emptyset)$, and therefore

(14) $$\alpha \in \mathcal{H}_\eta[\Gamma](\emptyset) \quad \Rightarrow \quad \psi_\Omega(\hat{\alpha}) \in \mathcal{H}_{\hat{\alpha}}[\Gamma](\emptyset).$$

Case 0: Suppose Γ is an axiom. Then $\mathcal{H}_{\hat{\alpha}}[\Gamma] \overset{\psi_\Omega(\hat{\alpha})}{\underset{\psi_\Omega(\hat{\alpha})}{\vdash}} \Gamma$ follows immediately by (14).

Case 1: Suppose the last inference was $(pb\forall)_\infty$. Then there is an $A \in \Gamma$ of the form $(\forall x \subseteq t)F(x)$ and $\mathcal{H}_\eta[\Gamma][s] \overset{\alpha_s}{\underset{\Omega+1}{\vdash}} \Gamma, s \subseteq t \rightarrow F(s)$ and $\alpha_s < \alpha$ hold for all s with $|\,s\,| < |\,t\,|$. Since $|\,t\,| \in C^\Omega(\eta+1, \psi_\Omega(\eta+1) \cap \Omega$ we have $|\,t\,| < \psi_\Omega(\eta+1)$ and hence $|\,s\,| < \psi_\Omega(\eta+1)$ whenever $|\,s\,| < |\,t\,|$. As a result, $|\,s\,| \in C^\Omega(\eta+1, \psi_\Omega(\eta+1))$ holds for all $|\,s\,| < |\,t\,|$. Therefore, by the induction hypothesis,

(15) $$\mathcal{H}_{\hat{\alpha}_s}[\Gamma][s] \overset{\psi_\Omega(\hat{\alpha}_s)}{\underset{\psi_\Omega(\hat{\alpha}_s)}{\vdash}} \Gamma, s \subseteq t \rightarrow F(s)$$

for all $|\,s\,| < |\,t\,|$. Let $|\,s\,| < |\,t\,|$. Since $|\,s\,| < \psi_\Omega(\eta+1)$ one computes that $\psi_\Omega(\hat{\alpha}_s) < \psi_\Omega(\hat{\alpha})$. Therefore, an inference $(pb\forall)_\infty$ applied to (15) yields $\mathcal{H}_{\hat{\alpha}}[\Gamma] \overset{\psi_\Omega(\hat{\alpha})}{\underset{\psi_\Omega(\hat{\alpha})}{\vdash}} \Gamma$.

The cases were the last inference is an instance of $(b\forall)_\infty$, $(\notin)_\infty$, $(\not\subseteq)_\infty$, or $(\wedge)$ are dealt with in a similar manner.

Case 2: Suppose the last inference was $(\exists)$. Then there is a formula $A \in \Gamma$ of the form $\exists x\, F(x)$ such that $\mathcal{H}_\eta[\Gamma] \vdash^{\alpha_0}_{\Omega+1} \Gamma, F(s)$ holds for some term s and $\alpha_0 < \alpha$. The induction hypothesis yields

$$\mathcal{H}_{\hat{\alpha}_0}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha}_0)}_{\psi_\Omega(\hat{\alpha}_0)} \Gamma, F(s)\,.$$

Since $\alpha_0, |\, s\,| \in \mathcal{H}_\eta[\Gamma](\emptyset)$ and $|\,\Gamma\,| \subseteq C^\Omega(\eta+1, \psi_\Omega(\eta+1))$ we see that

$$\alpha_0, |\, s\,| \in C^\Omega(\eta+1, \psi_\Omega(\eta+1))\,.$$

Consequently we have $|\, s\,|, \psi_\Omega(\hat{\alpha}_0) < \psi_\Omega(\hat{\alpha})$. Thus, via $(\exists)$ we conclude that $\mathcal{H}_{\hat{\alpha}}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha})}_{\psi_\Omega(\hat{\alpha})} \Gamma$.

The cases were the last inference is an instance of $(b\exists)$, $(\in)$, $(\subseteq)$, or $(\vee)$ are dealt with in a similar manner.

Case 3: Suppose $\exists z\, A^z \in \Gamma$ and $\mathcal{H}_\eta[\Gamma] \vdash^{\alpha_0}_{\Omega+1} \Gamma, A$ with $\alpha_0 < \alpha$. This means that the last inference was ($\Sigma^{\mathcal{P}}$-*Ref*). Note that $|\, A\,| = |\, \exists z\, A^z\,|$, and hence $\mathcal{H}_\eta[\Gamma, A] = \mathcal{H}_\eta[\Gamma]$. The induction hypothesis therefore yields $\mathcal{H}_{\hat{\alpha}_0}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha}_0)}_{\psi_\Omega(\hat{\alpha}_0)} \Gamma, A$ and therefore, as A is a $\Sigma^{\mathcal{P}}$-formula, we get $\mathcal{H}_{\hat{\alpha}_0}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha}_0)}_{\psi_\Omega(\hat{\alpha}_0)} \Gamma, A^{\mathbb{V}_{\psi_\Omega(\hat{\alpha}_0)}}$ by Lemma 7.2. Since $\psi_\Omega(\hat{\alpha}_0) \in \mathcal{H}_{\hat{\alpha}}$ and $\psi_\Omega(\hat{\alpha}_0) < \psi_\Omega(\hat{\alpha})$, an inference $(\exists)$ yields $\mathcal{H}_{\hat{\alpha}}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha})}_{\psi_\Omega(\hat{\alpha})} \Gamma, \exists z\, A^z$, i.e. $\mathcal{H}_{\hat{\alpha}}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha})}_{\psi_\Omega(\hat{\alpha})} \Gamma$.

Case 4: Suppose the last inference is (Cut). Then

$$\mathcal{H}_\eta[\Gamma] \vdash^{\alpha_0}_{\Omega+1} \Gamma, A \quad \text{and} \quad \mathcal{H}_\eta[\Gamma] \vdash^{\alpha_0}_{\Omega+1} \Gamma, \neg A\,,$$

where $\alpha_0 < \alpha$ and A is a formula with $rk(A) \le \Omega$.

Since $|\, A\,| \subseteq \mathcal{H}_\eta[\Gamma](\emptyset)$ and $|\,\Gamma\,| \subseteq C^\Omega(\eta+1, \psi_\Omega(\eta+1))$, this implies

$$|\, A\,| \subseteq C^\Omega(\eta+1, \psi_\Omega(\eta+1))$$

and

$$\mathcal{H}_{\eta'}[\Gamma, A] = \mathcal{H}_{\eta'}[\Gamma]$$

for all $\eta' \ge \eta$.

Case 4.1: Suppose that $rk(A) < \Omega$. This implies $rk(A) \in C^\Omega(\eta+1, \psi_\Omega(\eta+1))$ and hence $rk(A) < \psi_\Omega(\eta+1) < \psi_\Omega(\hat{\alpha})$. Inductively we have

$$\mathcal{H}_{\hat{\alpha}_0}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha}_0)}_{\psi_\Omega(\hat{\alpha}_0)} \Gamma, A \ \text{and}\ \mathcal{H}_{\hat{\alpha}_0}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha}_0)}_{\psi_\Omega(\hat{\alpha}_0)} \Gamma, \neg A\,.$$

Thus $\mathcal{H}_{\hat{\alpha}} \vdash^{\psi_\Omega(\hat{\alpha})}_{\psi_\Omega(\hat{\alpha})} \Gamma$ by means of (Cut).

Case 4.2: Suppose that $rk(A) = \Omega$. Then A or $\neg A$ will be of the form $\exists z\, F(z)$ with $F(\mathbb{V}_0)$ being $\Delta^{\mathcal{P}}_0$. We may assume that the former is the case. Then the induction hypothesis applied to $\mathcal{H}_\eta[\Gamma] \vdash^{\alpha_0}_{\Omega+1} \Gamma, A$ yields $\mathcal{H}_{\hat{\alpha}_0}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha}_0)}_{\psi_\Omega(\hat{\alpha}_0)} \Gamma, A$. Since $\psi_\Omega(\hat{\alpha}_0) \in \mathcal{H}_{\hat{\alpha}_0}(\emptyset)$, we can apply the Boundedness Lemma 7.2, obtaining

$$\mathcal{H}_{\hat{\alpha}_0}[\Gamma] \vdash^{\psi_\Omega(\hat{\alpha}_0)}_{\psi_\Omega(\hat{\alpha}_0)} \Gamma, A^{\mathbb{V}_{\psi_\Omega(\hat{\alpha}_0)}}\,. \tag{16}$$

By applying inversion (Lemma 5.9(iii)) to $\mathcal{H}_{\hat{\alpha}_0}[\Gamma] \mid\frac{\alpha_0}{\Omega+1} \Gamma, \neg A$ we also get

$$(17) \qquad \mathcal{H}_{\hat{\alpha}_0}[\Gamma] \mid\frac{\alpha_0}{\Omega+1} \Gamma, \neg A^{\mathbb{V}_{\psi_\Omega(\hat{\alpha}_0)}} .$$

Observing that $\Gamma, \neg A^{\mathbb{V}_{\psi_\Omega(\hat{\alpha}_0)}}$ is a set of $\Sigma^{\mathcal{P}}$-formulae, we can apply the induction hypothesis to (17), yielding

$$(18) \qquad \mathcal{H}_{\alpha_1}[\Gamma] \mid\frac{\psi_\Omega(\alpha_1)}{\psi_\Omega(\alpha_1)} \Gamma, \neg A^{\mathbb{V}_{\psi_\Omega(\hat{\alpha}_0)}} ,$$

where $\alpha_1 = \hat{\alpha}_0 + \omega^{\Omega+\alpha_0} = \eta + \omega^{\Omega+\alpha_0} + \omega^{\Omega+\alpha_0} < \eta + \omega^{\Omega+\alpha} = \hat{\alpha}$. Moreover, we have $\psi_\Omega(\alpha_1) < \psi_\Omega(\hat{\alpha})$. Therefore (Cut) applied to (16) and (18) furnishes $\mathcal{H}_{\hat{\alpha}}[\Gamma] \mid\frac{\psi_\Omega(\hat{\alpha})}{\psi_\Omega(\hat{\alpha})} \Gamma$. □

Note that the Collapsing Theorem produces a derivation in which all instances of $(\Sigma^{\mathcal{P}}\text{-}Ref)$ have been removed.

Also note that we cannot eliminate cuts with $\Delta_0^{\mathcal{P}}$-formulae since we don't have predicative cut elimination as in the case **KP**.

Corollary 7.5 *Let A be a $\Sigma^{\mathcal{P}}$-sentence of* $\mathbf{KP}(\mathcal{P})$. *Suppose that* $\mathbf{KP}(\mathcal{P}) \vdash A$. *Then there exists an operator $\mathcal{H}$ and an ordinal $\rho < \psi_\Omega(\varepsilon_{\Omega+1})$ such that*

$$\mathcal{H} \mid\frac{\rho}{\rho} A .$$

Proof: Let $\mathcal{H}_0$ be defined as in Definition 7.3. By Theorem 6.9 we have

$$\mathcal{H}_0 \mid\frac{\omega^{\Omega+m}}{\Omega+m+1} A$$

for some $0 < m < \omega$. Applying ordinary cut elimination, Theorem 7.1, we get

$$\mathcal{H}_0 \mid\frac{\omega_m(\omega^{\Omega+m})}{\Omega+1} A .$$

Finally, using the Collapsing Theorem 7.4 we arrive at

$$\mathcal{H}_{\omega_{m+1}(\omega^{\Omega+m})} \mid\frac{\rho}{\rho} A$$

with $\rho := \psi_\Omega(\omega_{m+1}(\omega^{\Omega+m}))$. □

8 Soundness

For the main Theorem of this paper, we want to show that derivability in $RS_\Omega^{\mathcal{P}}$ entails truth. Since $RS_\Omega^{\mathcal{P}}$-formulae contain variables we need the notion of assignment. Let VAR be the set of free variables of $RS_\Omega^{\mathcal{P}}$. A variable assignment ℓ is a function

$$\ell : VAR \longrightarrow V_{\psi_\Omega(\varepsilon_{\Omega+1})}$$

satisfying $\ell(a^\alpha) \in V_{\alpha+1}$, where as per usual V_α denotes the α^{th} level of the von Neumann hierarchy.

ℓ can be canonically lifted to all $RS_\Omega^{\mathcal{P}}$-terms as follows:

$$\begin{aligned} \ell(\mathbb{V}_\alpha) &= V_\alpha \\ \ell(\{x \in \mathbb{V}_\alpha \mid F(x, s_1, \ldots, s_n)\}) &= \{x \in V_\alpha \mid F(x, \ell(s_1), \ldots, \ell(s_n))\} . \end{aligned}$$

Note that $\ell(s) \in V_{\psi_\Omega(\varepsilon_{\Omega+1})}$ holds for all $RS_\Omega^{\mathcal{P}}$-terms s. Moreover, $\ell(s) \in V_{|s|+1}$.

Theorem 8.1 (Soundness) *Let $\mathcal{H}$ be an operator and $\alpha, \rho < \psi_\Omega(\varepsilon_{\Omega+1})$. Let $\Gamma(s_1, \ldots, s_n)$ be a sequent consisting only of $\Sigma^{\mathcal{P}}$-formulae. Suppose*

$$\mathcal{H} \overset{\alpha}{\underset{\rho}{|}} \Gamma(s_1, \ldots, s_n) .$$

Then, for all variable assignments ℓ,

$$V_{\psi_\Omega(\varepsilon_{\Omega+1})} \models \Gamma(\ell(s_1), \ldots, \ell(s_n)) .$$

Proof: The proof proceeds by induction on α. Note that, owing to $\alpha, \rho < \Omega$, the proof tree pertaining to $\mathcal{H} \overset{\alpha}{\underset{\rho}{|}} \Gamma(s_1, \ldots, s_n)$ neither contains any instances of ($\Sigma^{\mathcal{P}}$-*Ref*) nor of $(\forall)_\infty$, and that all cuts are with $\Delta_0^{\mathcal{P}}$-formulae. The proof is straightforward as all the axioms of $RS_\Omega^{\mathcal{P}}$ are true under the interpretation and all other rules are truth preserving with respect to this interpretation. Observe that we make essential use of the free variables when showing the soundness of $(b\forall)_\infty$, $(pb\forall)_\infty$, $(\notin)_\infty$ and $(\not\subseteq)_\infty$. □

Combining Theorem 8.1 and Corollary 7.5 we have the following:

Theorem 8.2 *If A is a $\Sigma^{\mathcal{P}}$-sentence and*

$$\mathbf{KP}(\mathcal{P}) \vdash A$$

then

$$V_{\psi_\Omega(\varepsilon_{\Omega+1})} \models A.$$

The bound of this Corollary is actually sharp, that is, $\psi_\Omega(\varepsilon_{\Omega+1})$ is the first ordinal with that property. This follows immediately from [**21**, Theorem 4.9].

The previous result can be extended to $\Pi_2^{\mathcal{P}}$ sentences, basically by the same proof as for [**21**, Theorem 2.1].

Theorem 8.3 *Let A be a $\Pi_2^{\mathcal{P}}$-sentence. Then $\mathbf{KP}(\mathcal{P}) \vdash A$ implies $V_{\psi_\Omega(\varepsilon_{\Omega+1})} \models A$.*

Proof: Assume $\mathbf{KP}(\mathcal{P}) \vdash \forall u \exists w H(u, w)$ with $H(u, w)$ being $\Delta_0^{\mathcal{P}}$. Let $\sigma := \psi_\Omega(\varepsilon_{\Omega+1})$. Let $b \in V_\sigma$. We have to verify that $V_\sigma \models \exists w H(b, w)$. σ is a limit, so there is $\xi < \sigma$ such that $b \in V_\xi$. Since V_ξ does not satisfy all $\Sigma^{\mathcal{P}}$-sentences provable in $\mathbf{KP}(\mathcal{P})$, we have $\mathbf{KP}(\mathcal{P}) \vdash B$ and $V_\xi \models \neg B$ for some $\Sigma^{\mathcal{P}}$-sentence B. Since $\Sigma^{\mathcal{P}}$-reflection is provable in $\mathbf{KP}(\mathcal{P})$, we also get $\mathbf{KP}(\mathcal{P}) \vdash \exists \alpha \exists x (x = V_\alpha \wedge B^x)$. Then, using $\Delta_0^{\mathcal{P}}$–Collection, we obtain

$$\mathbf{KP}(\mathcal{P}) \vdash \exists z \exists \alpha \exists x [x = V_\alpha \wedge B^x \wedge (\forall u \in x)(\exists w \in z) H(u, w)] .$$

Since this formula is equivalent to a $\Sigma^{\mathcal{P}}$–formula in $\mathbf{KP}(\mathcal{P})$, we get

$$V_\sigma \models \exists \alpha \exists x [x = V_\alpha \wedge B^x \wedge (\forall u \in x) \exists w H(u, w)]$$

As the formula "$x = V_\alpha$" has the same meaning In V_σ as it has in V, there exists $\alpha < \sigma$ such that $V_\alpha \models B$ and $(\forall u \in V_\alpha)(\exists w \in V_\sigma) H(u, w)$. By the choice of B, this implies $\xi < \alpha$, hence $b \in V_\alpha$, thus $V_\sigma \models \exists w H(b, w)$. □

References

[1] T. Arai: *Ordinal Diagrams for PI3-Reflection.* J. Symb. Log. 65(3): 1375-1394 (2000)

[2] T. Arai: *Proof theory for theories of ordinals I: recursively Mahlo ordinals*, Annals of Pure and applied Logic 122 (2003) 1–85.

[3] T. Arai: *Proof theory for theories of ordinals II:* Π_3*-Reflection*, Annals of Pure and Applied Logic 129 (2004) 39–92.

[4] H. Bachmann: *Die Normalfunktionen und das Problem der ausgezeichneten Folgen von Ordinalzahlen.* Vierteljahresschrift Naturforsch. Ges. Zürich 95 (1950) 115–147.

[5] J Barwise: *Admissible Sets and Structures* (Springer, Berlin 1975).

[6] W. Buchholz: *Eine Erweiterung der Schnitteliminationsmethode*, Habilitationsschrift (München 1977).

[7] *A simplified version of local predicativity*, in: Aczel, Simmons, Wainer (eds.), Leeds Proof Theory 1991 (Cambridge University Press, Cambridge, 1993) 115–147.

[8] W. Buchholz, S. Feferman, W. Pohlers, W. Sieg: *Iterated inductive definitions and subsystems of analysis* (Springer, Berlin, 1981).

[9] W. Buchholz and K. Schütte: *Proof theory of impredicative subsystems of analysis.* (Bibliopolis, Naples, 1988).

[10] H. Friedman: *Some applications of Kleene's method for intuitionistic systems.* In: A. Mathias and H. Rogers (eds.): *Cambridge Summer School in Mathematical Logic*, volume 337 of *Lectures Notes in Mathematics* (Springer, Berlin, 1973) 113–170.

[11] H. Friedman: *Countable models of set theories*, in: A.R.D. Mathias, H. Rogers Jr. (Eds.), Cambridge Summer School in Mathematical Logic, August 1–21, 1971, Springer Lecture Notes in Mathematics, Vol. 337 (Springer, Berlin, 1973) 539-–573.

[12] H. Friedman, S. Ščedrov: *The lack of definable witnesses and provably recursive functions in intuitionistic set theory*, Advances in Mathematics 57 (1985) 1–13.

[13] G. Jäger: *Zur Beweistheorie der Kripke–Platek Mengenlehre über den natürlichen Zahlen*, Archiv f. Math. Logik 22 (1982) 121–139.

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

[15] S. Mac Lane: *Form and Function* (Springer, Berlin, 1992).

[16] A.R.D. Mathias: *The strength of Mac Lane set theory*, Annals of Pure and Applied Logic 110 (2001) 107–234.

[17] L. Pozsgay: *Liberal intuitionism as a basis for set theory*, in Axiomatic Set Theory, Proc. Symp. Pure Math. XIII, Part 1 (1971) 321-330.

[18] L. Pozsgay: *Semi-intuitionistic set theory*, Notre Dame J. of Formal Logic 13 (1972) 546-550.

[19] M. Rathjen: *Ordinal notations based on a weakly Mahlo cardinal*, Archive for Mathematical Logic 29 (1990) 249–263.

[20] M. Rathjen: *Proof-Theoretic Analysis of KPM*, Arch. Math. Logic 30 (1991) 377–403.

[21] M. Rathjen: *Fragments of Kripke–Platek set theory.* In: P. Aczel, S. Wainer, H. Simmons (eds.), Proof Theory (Cambridge University Press, 1992) 251–273.

[22] M. Rathjen: *How to develop proof–theoretic ordinal functions on the basis of admissible sets.* Mathematical Quarterly 39 (1993) 47–54.

[23] M. Rathjen: *Collapsing functions based on recursively large ordinals: A well–ordering proof for KPM.* Archive for Mathematical Logic 33 (1994) 35–55.

[24] M. Rathjen: *Proof theory of reflection.* Annals of Pure and Applied Logic 68 (1994) 181–224.

[25] M. Rathjen: *The realm of ordinal analysis.* S.B. Cooper and J.K. Truss (eds.): *Sets and Proofs.* (Cambridge University Press, 1999) 219–279.

[26] M. Rathjen: *Recent advances in ordinal analysis:* Π^1_2-*CA and related systems.* Bulletin of Symbolic Logic 1, 468–485 (1995).

[27] M. Rathjen: *An ordinal analysis of stability*, Archive for Mathematical Logic 44 (2005) 1 - 62.

[28] M. Rathjen: *An ordinal analysis of parameter-free* Π^1_2 *comprehension* Archive for Mathematical Logic 44 (2005) 263 - 362.

[29] M. Rathjen: *Theories and ordinals in proof theory.* Synthese 148 (2006) 719–743.

[30] M . Rathjen: *From the weak to the strong existence property,* Annals of Pure and Applied Logic (2012) doi:10.1016/j.apal.2012.01.012

[31] M. Rathjen: *Constructive Zermelo-Fraenkel Set Theory, Power Set, and the Calculus of Constructions*, 2010. To appear in: 2012, in: P. Dybjer, S. Lindström, E. Palmgren and G. Sundholm: *Epistemology versus Ontology: Essays on the Philosophy and Foundations of Mathematics in Honour of Per Martin-Löf*, (Logic, Epistemology and the Unity of Science Series, NewYork/Dordrecht: Springer Verlag.)

[32] H. Schwichtenberg: *Some applications of cut elimination*, in J. Barwise (ed.): *Handbook of Mathematical Logic* (North-Holland, Amsterdam, 1977) 868–895.

[33] L. Tharp: *A quasi-intuitionistic set theory*. Journal of Symbolic Logic 36 (1971) 456–460.

[34] E.J. Thiele: *Über endlich axiomatisierbare Teilsysteme der Zermelo-Fraenkel'schen Mengenlehre*, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 14 (1968) 39-58.

[35] R. S. Wolf: *Formally Intuitionistic Set Theories with Bounded Predicates Decidable*, PhD Thesis (Stanford University, 1974).