Cumulative Higher-Order Logic as a Foundation for Set Theory

Wolfgang Degen[1] Jan Johannsen[2,*]

[1] IMMD 1, Universität Erlangen-Nürnberg, Germany
email: degen@informatik.uni-erlangen.de

[2] Institut für Informatik, Ludwig-Maximilians-Universität München
email: jjohanns@informatik.uni-muenchen.de

Abstract

The systems K_α of transfinite cumulative types up to α are extended to systems K_α^∞ that include a natural infinitary inference rule, the so-called *limit rule*. For countable α a semantic completeness theorem for K_α^∞ is proved by the method of reduction trees, and it is shown that every model of K_α^∞ is equivalent to a cumulative hierarchy of sets. This is used to show that several axiomatic first-order set theories can be interpreted in K_α^∞, for suitable α.

Keywords: cumulative types, infinitary inference rule, logical foundations of set theory.
MSC: 03B15 03B30 03E30 03F25

1 Introduction

The idea of founding mathematics on a theory of types was first proposed by Russell [20] (foreshadowed already in [19]), and subsequently implemented by Whitehead and Russell [26]. The formal systems presented in these works were later simplified and cast into their modern shape by Ramsey [18]. Gödel [9] and Tarski [25] were the first to restrict the type structure to types of unary predicates denoted by natural numbers, where 0 denotes the type of individuals, and $n+1$ denotes the type of predicates of objects of type n.

Several authors have proposed to extend this type structure to transfinite ordinals, e.g. Gödel [10]. This extension naturally leads to the idea of cumulativity, i.e., the applicability of a predicate of some type τ to objects of all smaller types $\sigma < \tau$. Formal systems based on such a transfinite cumulative type structure were, to our knowledge, only investigated in detail by Bustamante [4], Kemeny [14], Bowen [2, 3] and the first author [5]. Systems with transfinite function types were studied by L'Abbé [15] and Andrews [1].

*Most of this research was done while the second author was employed at the chair for Theoretical Computer Science (Prof. K. Leeb), Universität Erlangen-Nürnberg.

The concept of a transfinite cumulatively leveled structure also arises in another traditional approach to foundations of mathematics, viz. axiomatic set theory as initiated by Zermelo [27]. The universe of pure sets described by first-order set theories like ZF is naturally stratified into the levels V_α of sets of rank at most α. Thus instead of the traditional first-order description, an approach to describe this structure in a cumulatively typed language suggests itself.

In order to axiomatize the cumulative universe of sets within the typed language, we introduce certain laws in form of inference rules. Besides the usual logical rules, which are adapted to the cumulative framework, we require a strong extensionality rule and an infinitary inference rule that connects types represented by limit ordinals with the types below. It reflects the property of the cumulative hierarchy of sets, where V_γ for a limit γ is the union of the V_ξ with $\xi < \gamma$.

From this typed axiomatization, large parts of traditional first-order set theory can be recovered: the cumulative language is flexible enough to define a *type-homogeneous* membership relation, by use of which the first-order language of set theory can be interpreted in the typed language. If the type used as the target of the interpretation is a sufficiently large limit, then a fragment stronger than Zermelo's set theory, i.e., ZF without replacement, can be deduced.

All of our results can be proved in relatively weak subsystems of ZFC that contain some choice principles and sufficiently many ordinals to denote the types of the systems in question. We have made no effort to calibrate these subsystems precisely.

The outline of the paper is as follows: In the first section we define the syntax of the pure systems K_α of transfinite cumulative types of order α and their extensions and prove some of their basic properties. We then show that these systems are not stronger than K_ω with extensionality. In order to obtain stronger systems we extend them to the systems K_α^∞ that include the infinitary *limit rule*. In the second section we define a semantics for the systems K_α^∞ for countable α, and prove a completeness theorem by the method of reduction trees. We then show that every model of K_α^∞ is equivalent to a cumulative hierarchy of sets of length α. This is used in the third section to show that fragments of first-order set theory can be deduced in K_α^∞. Unfortunately, K_α^∞ also proves set-theoretic statements that are false, i.e., inconsistent with ZF. We finally give some sufficient criteria for a set-theoretic sentence provable in K_α^∞ to be consistent with ZF. The paper concludes with an appendix that corrects an error in the first author's book [5]. Some of the results of this paper were previously announced without proofs in [6].

2 Basic Definitions

2.1 The Pure Cumulative Systems K_α

Let α be an ordinal. We shall first define the formal system K_α of pure cumulative logic of order α.

Each ordinal $\tau < \alpha$ is a type. When using Greek letters as type symbols in the following we always assume that they are ordinals less than α.

For each type τ, there are countably many free variables a_i^τ, $i \in \mathbb{N}$ of type τ, and countably many bound variables x_j^τ, $j \in \mathbb{N}$ of type τ. These genuine variables of the system will hardly ever appear in the text, instead we use meta-variables $a^\tau, b^\tau, \ldots$ ranging over the free variables and $x^\tau, y^\tau, \ldots$ ranging over the bound variables.

Terms with their type and formulae are defined by simultaneous induction: Each free variable of type τ is a term of type τ. If $\tau < \sigma$ are types and s^σ and t^τ terms of the respective types, then $s^\sigma(t^\tau)$ is an atomic formula, also called a *predication*. If $\sigma > \tau + 1$, we call $s^\sigma(t^\tau)$ a *cumulative* predication.

Formulae are built up from atomic formulae by propositional connectives and quantifiers over variables of arbitrary type: If $A[a^\tau]$ is a formula in which the bound variable x^τ does not occur, then $\forall x^\tau\, A[x^\tau]$ and $\exists x^\tau\, A[x^\tau]$ are formulae. Finally, if $\tau + 1 < \alpha$, then $\big(\lambda x^\tau.\, A[x^\tau]\big)$ is a term of type $\tau + 1$.

The term $\big(\lambda x^\tau.\, A[x^\tau]\big)$ is intended to denote the set of those elements of type τ that satisfy the formula $A[a^\tau]$. A formula of the form $\big(\lambda x^\tau.\, A[x^\tau]\big)(s^\sigma)$ is called an *abstraction formula*, and if the formula $A[s^\sigma]$ is well-formed, it is called the *converse* of this abstraction formula. Note that the cumulative predication allows us to have abstraction formulae that have no well-formed converse, e.g. $\big(\lambda x^1.\, x^1(a^0)\big)(b^0)$.

The system K_α is formulated in a Gentzen-style sequent calculus for classical logic with the usual propositional rules, the cut rule (which is indispenable, as we shall see below), and with special rules for the introduction of quantifiers, where the type of the quantified variable can be shifted as follows: the right universal quantification rule is

$$\forall\text{:right} \qquad \frac{\Gamma \Longrightarrow \Delta\,,\ A[a^\sigma]}{\Gamma \Longrightarrow \Delta\,,\ \forall x^\tau\, A[x^\tau]}\,,$$

where $\tau \leq \sigma$, and the free variable a^σ does not occur in the lower sequent, whereas the left universal quantification rule is

$$\forall\text{:left} \qquad \frac{A[t^\sigma]\,,\ \Gamma \Longrightarrow \Delta}{\forall x^\tau\, A[x^\tau]\,,\ \Gamma \Longrightarrow \Delta}\,,$$

where t^σ is an arbitrary term of type σ and $\tau \geq \sigma$. The existential quantification rules $\exists$:right and $\exists$:left are defined dually. Furthermore there are the left and right abstraction rules

$$\lambda\text{:left} \qquad \frac{A[t^\sigma]\,,\ \Gamma \Longrightarrow \Delta}{\big(\lambda x^\tau.\, A[x^\tau]\big)(t^\sigma)\,,\ \Gamma \Longrightarrow \Delta}$$

$$\lambda\text{:right} \qquad \frac{\Gamma \Longrightarrow \Delta\,,\ A[t^\sigma]}{\Gamma \Longrightarrow \Delta\,,\ \big(\lambda x^\tau.\, A[x^\tau]\big)(t^\sigma)}\,,$$

where in both cases $\tau \geq \sigma$. Note that the type of the bound variable x^τ in all the above rules is further restricted by the implicit requirement that the

formulae in the lower sequent have to be well-formed. We use the following convention: if S denotes the sequent $\Gamma \Longrightarrow \Delta$, then $\hat{S}$ denotes the formula $\bigwedge \Gamma \to \bigvee \Delta$.

Let $\bar{K}_\alpha$ denote the restriction of K_α to the non-cumulative language, i.e., to those formulae where only predications of the form $t^{\tau+1}(s^\tau)$ occur. Thus $\bar{K}_\omega$ is a formulation of the usual, non-cumulative higher-order logic. An obvious question that we could not answer is the following.

Open Question. *Is K_ω conservative over $\bar{K}_\omega$?*

An important feature of the cumulative language of K_α is that the usual definition of Leibniz equality can be generalized to a type-heterogeneous equality. Let $\delta > \max(\sigma, \tau)$, then we define

$$a^\sigma =_\delta b^\tau :\leftrightarrow \forall z^\delta \left(z^\delta(a^\sigma) \to z^\delta(b^\tau)\right)$$

and in particular

$$a^\sigma = b^\tau :\leftrightarrow a^\sigma =_{\max(\sigma,\tau)+1} b^\tau \ .$$

For every $\max(\sigma, \tau) < \delta$, it is easily seen that $K_\alpha \vdash a^\sigma =_\delta b^\tau \leftrightarrow a^\sigma = b^\tau$, hence we shall not use the subscripted equality $=_\delta$ in the following. Furthermore it is easily shown that the defined equality has the usual properties of an equality:

Proposition 1. *$K_\alpha \vdash a^\tau = a^\tau$, and if both $A[a^\sigma]$ and $A[b^\tau]$ are well-formed, then $K_\alpha \vdash a^\sigma = b^\tau$, $A[a^\sigma] \Longrightarrow A[b^\tau]$.*

For $\sigma \geq \delta \geq \tau$, the following properties of equality, reflecting the cumulativity of the type structure, are easily proved in K_α:

$$\exists x^\sigma\, x^\sigma = a^\tau \quad \text{and} \quad a^\sigma = b^\tau \Longrightarrow \exists x^\delta\, (a^\sigma = x^\delta = b^\tau)$$

The formal system K_α^e is K_α extended by the extensionality rule

$$\text{Ext} \qquad \frac{\begin{array}{c} s_1^{\sigma+1}(a^\sigma)\,,\, \Gamma_1 \Longrightarrow \Delta_1\,,\, \exists x^\tau\, (s_2^{\tau+1}(x^\tau) \wedge a^\sigma = x^\tau) \\ s_2^{\tau+1}(b^\tau)\,,\, \Gamma_2 \Longrightarrow \Delta_2\,,\, s_1^{\sigma+1}(b^\tau) \end{array}}{t^\delta(s_1^{\sigma+1})\,,\, \Gamma_1\,,\, \Gamma_2 \Longrightarrow \Delta_1\,,\, \Delta_2\,,\, t^\delta(s_2^{\tau+1})}$$

where $\sigma \geq \tau$, $\delta > \sigma + 1$ and the variables a^σ, b^τ do not occur in the lower sequent, except possibly in the term t^δ.

Let $\bar{K}_\alpha^e$ be an extension of $\bar{K}_\alpha$ by an appropriate extensionality rule, e.g. a variant of Ext where $\sigma = \tau$, with the formula $\exists x^\tau\, (s_2^{\tau+1}(x^\tau) \wedge a^\sigma = x^\tau)$ in the first premise replaced by $s_2^{\tau+1}(a^\tau)$. Then the same question as above is also open for the systems with extensionality:

Open Question. *Is K_ω^e conservative over $\bar{K}_\omega^e$?*

A well-known fact already proved by Russell (see Takeuti [23] for a detailed proof) is that there is an interpretation of type theory within type theory without extensionality, i.e., of $\bar{K}_\omega^e$ in $\bar{K}_\omega$. We do not know whether the same is possible in the cumulative case.

Open Question. *Is there an interpretation of K^e_ω in K_ω?*

On the other hand, another obvious question *can* be answered in an unexpected way:

Theorem 2. *For $\alpha \geq 3$, K^e_α does not allow cut-elimination.*

Proof. Let $a^1 \subseteq b^1$ abbreviate $\forall x^0\, a^1(x^0) \to b^1(x^0)$. We shall show that the formula $A := \exists x^1\, x^1 \subseteq x^1 \wedge x^1 = a^0$ does not have a cut-free proof, although it is deduced by use of a cut as follows: we first deduce the sequent $b^1 = a^0 \Longrightarrow \exists x^1\, x^1 \subseteq x^1 \wedge x^1 = a^0$ in the obvious way. Likewise the formula $\exists x^1\, x^1 = a^0$ is easily proved. From these deduce the formula above by $\exists$:left and a cut.

Suppose there is a cut-free proof of A. The last inference can only be an $\exists : right$, whose premise is

$$\Longrightarrow \hat{A}\,,\ t^1_1 \subseteq t^1_1 \wedge t^1_1 = a^0$$

where $\hat{A}$ means that the formula A can be present or not. Note that due to the first conjunct, the type of t^1_1 has to be 1. Now this sequent can either be the conclusion of another $\exists$:right, or of an $\wedge$:right. Say there are k $\exists$:right inferences in the proof, where the terms that are quantified are $t^1_1, \ldots, t^1_k$.

For each of the auxiliary formulas $t^1_i \subseteq t^1_i \wedge t^1_i = a^0$ of these inferences, there are one or more $\wedge$:right inferences with it as the principal formula. We always write the premise containing the auxiliary formula $t^1_i \subseteq t^1_i$ as the left son, and the other one as the right son in the proof-tree. The left premise has an obvious cut-free proof, so we look only an the rightmost branch in the proof tree.

For each of the right auxiliary formulas $t^1_i = a^0$ of the $\wedge$:right inferences, there are one or more $\forall$:right inferences having it as principal formula, whose eigenvariables are $b^2_{i,1}, \ldots, b^2_{i,\ell_i}$. Finally, for each of the auxiliary formulas $b^2_{i,j}(t^1_i) \to b^2_{i,j}(a^0)$, there are one or more $\to$:right inferences having it as principal formula.

At some height in the rightmost branch of the proof tree, above all these inferences, there is a sequent of the following form:

$$b^2_{1,1}(t^1_1), \ldots, b^2_{1,\ell_1}(t^1_1), \ldots, b^2_{k,1}(t^1_k), \ldots, b^2_{k,\ell_k}(t^1_k) \\ \Longrightarrow b^2_{1,1}(a^0), \ldots, b^2_{1,\ell_1}(a^0), \ldots, b^2_{k,1}(a^0), \ldots, b^2_{k,\ell_k}(a^0)$$

Obviously, neither this sequent, nor any subsequent of it, can be the conclusion of any logical or extensionality inference. Hence it has no cut-free proof. $\square$

This is in sharp contrast to the non-cumulative systems $\bar{K}^e_\alpha$, for $\alpha \leq \omega$, for which a cut-elimination was proved by Takahashi [22] and Prawitz [17]. The proof also shows that there must be an error in Bowen [2], where a cut elimination theorem was proved for a formal system TT^θ of transfinite cumulative types. This system TT^θ is essentially an extension of K_θ by a second hierarchy of cumulative types over the base type of propositions. The same counterexample as in the proof of Theorem 2 shows that cut elimination does not hold for TT^θ.

Consider the following *substitution rule*:

$$\text{Subst} \qquad \frac{S[a^\sigma]}{S[t^\tau]}$$

where $\sigma \geq \tau$, and the free variable a^σ does not occur in the lower sequent, except possibly in the term t^τ. This rule is obviously derivable in K_α. Another interesting consequence of the proof of Theorem 2 is that this inference rule is not admissible in K_α without the cut rule: the counterexample $\exists x^1\, x^1 \subseteq x^1 \wedge x^1 = a^0$ is easily derived without cuts by use of the substitution rule. On the other hand, the cut rule cannot simply be replaced by the substitution rule, since the following sentence

$$\exists x^1 \left(x^1 \subseteq x^1 \wedge \exists y^0\, x^1 = y^0\right)$$

is provable in K_α, but cannot be derived without cuts even in presence of the substitution rule.

2.2 Membership and Extensionality

By use of type-heterogeneous equality, a membership relation between terms of arbitrary types can be defined. In particular, we can define a *type-homogeneous* membership relation between terms of equal type. This will allow us to translate set-theoretic formulae into the language of K_α. Let $\delta = \max(\sigma, \tau) + 1$, then we define

$$a^\sigma \in b^\tau :\leftrightarrow \exists x^\delta \left(x^\delta(a^\sigma) \wedge x^\delta = b^\tau\right).$$

It is easily seen that for $\sigma < \tau$ we have $a^\sigma \in b^\tau \leftrightarrow b^\tau(s^\sigma)$ in K_α, and we shall use this equivalence tacitly in the sequel.

For an ordinal ξ, let $\check{\xi}$ denote ξ if $\xi = 0$ or ξ is a limit, and ξ' if $\xi = \xi' + 1$. For $\tau \leq \sigma$, we define

$$\begin{aligned} a^\sigma \gtrsim b^\tau :\leftrightarrow\ & \forall x^{\check{\sigma}} \left(x^{\check{\sigma}} \in a^\sigma \to \exists y^{\check{\tau}} \left(y^{\check{\tau}} \in b^\tau \wedge x^{\check{\sigma}} = y^{\check{\tau}}\right)\right) \\ & \wedge \forall x^{\check{\tau}} \left(x^{\check{\tau}} \in b^\tau \to x^{\check{\tau}} \in a^\sigma\right) \end{aligned}$$

Proposition 3. *For $\sigma \leq \tau$, K_α^e proves the equivalence $a^\sigma \gtrsim b^\tau \leftrightarrow a^\sigma = b^\tau$.*

Proof. For the implication from left to right, we have to replace the $\in$ in the definition of $\gtrsim$ by predications in order to apply the extensionality rule Ext. This is trivial for successor types σ and τ.

For limit types σ, τ, we first derive $a^\sigma = b^\tau$ under the additional assumptions $a^{\sigma+1} = a^\sigma$ and $b^{\tau+1} = b^\tau$: replace a^σ by $a^{\sigma+1}$ and b^τ by $b^{\tau+1}$ in $a^\sigma \gtrsim b^\tau$ and then apply Ext and equality laws. Finally, we get rid of the assumptions by cuts, since $\exists x^{\sigma+1}\, (x^{\sigma+1} = a^\sigma)$ and $\exists x^{\tau+1}\, (x^{\tau+1} = b^\tau)$ are provable.

It is worth noting that the implication from right to left also requires the extensionality rule. Again we only treat the hard case where $\sigma > \tau$ are limits.

The most involved part is to prove the first conjunct of $a^\sigma \gtrsim b^\tau$ from $a^\sigma = b^\tau$, i.e., to derive the sequent

$$(*) \qquad a^\sigma = b^\tau, c^\sigma \in a^\sigma \implies \exists y^\tau\, (y^\tau \in b^\tau \wedge c^\sigma = b^\tau)\ .$$

Let $t^{\sigma+1}$ be the term $\big(\lambda z^\sigma.\, \exists y^\tau (y^\tau \in b^\tau \wedge z^\sigma = b^\tau)\big)$, then using the extensionality rule we can prove $t^{\sigma+1} = b^{\tau+1}$ under the assumption $b^{\tau+1} = b^\tau$, and equality laws yield $t^{\sigma+1} = b^\tau$. Now it is easy to prove $a^\sigma = b^\tau, c^\sigma \in a^\sigma \implies c^\sigma \in b^\tau$, hence using the equality $t^{\sigma+1} = b^\tau$ and λ-conversion yields the sequent $(*)$. □

Proposition 4. *Let $\delta > \max(\sigma, \tau)$, and let $a^\sigma \in_\delta b^\tau :\leftrightarrow \exists x^\delta\, (x^\delta(a^\sigma) \wedge x^\delta = b^\tau)$. Then K^e_α proves $a^\sigma \in_\delta b^\tau \leftrightarrow a^\sigma \in b^\tau$.*

Proof. The direction from right to left is easy. For the reverse direction, let $\gamma := \max(\sigma, \tau)$. We need a term $t^{\gamma+1}$ such that

$$c^\delta(a^\sigma) \wedge c^\delta = b^\tau \implies t^{\gamma+1}(a^\sigma) \wedge t^{\gamma+1} = b^\tau\ .$$

If we set $t^{\gamma+1} := \big(\lambda x^\gamma.\, c^\delta(x^\gamma)\big)$, then the above sequent can be proved by use of extensionality, i.e., Prop. 3. □

Proposition 4, as well as the corresponding equivalence for equality (i.e., $a^\sigma = b^\tau \leftrightarrow a^\sigma =_\delta b^\tau$), will from now on be used tacitly, e.g. we consider $a^\omega = b^5$, $a^\omega \in c^5 \implies b^5 \in c^5$ an instance of substitution, although the right-hand side should correctly be $b^5 \in_{\omega+1} c^5$.

The following proposition formalizes the intuition that an element of a set of type τ should have a type not bigger than τ.

Proposition 5. $K^e_\alpha \vdash a^\sigma \in b^\tau \implies \exists x^{\check{\tau}}\, a^\sigma = x^{\check{\tau}}$.

Proof. This is trivial if $\sigma < \tau$, so let $\sigma \geq \tau$. Now $a^\sigma \in b^\tau$ means that for some $c^{\sigma+1}$ we have $c^{\sigma+1}(a^\sigma)$ and $c^{\sigma+1} = b^\tau$. By Prop. 3 this implies $c^{\sigma+1} \gtrsim b^\tau$, and from this and $a^\sigma \in c^{\sigma+1}$ it follows that $\exists x^{\check{\tau}}\, a^\sigma = x^{\check{\tau}}$. □

2.3 The Compression Theorem

In this section we shall show that for every ordinal α, the system K^e_α is no stronger than the system K^e_ω of finite cumulative types with extensionality. In particular, it follows that no system K^e_α can prove an axiom of infinity and that for recursive ordinals α, the consistency of K^e_α can be proved constructively. Hence the extension of the type structure to the transfinite alone does not give us stronger systems.

Let S be a sequent in the language of K_α. A sequent S' is called a *typical variant* of S if there is an order-preserving function f from the types occurring in S into ON such that S' is obtained from S by replacing each free variable a^ξ_j in S by $\mathsf{a}^{f(\xi)}_j$ and each bound variable x^ξ_j in S by $\mathsf{x}^{f(\xi)}_j$.

Theorem 6. *Let S be a sequent provable in K^e_α. Then there is a typical variant $S^{(\omega)}$ of S belonging to the language of and provable in K^e_ω.*

Proof. Let $S_0, \ldots, S_n$ be a proof of S in K^e_α, such that $S_n = S$. Let $\alpha_0, \ldots, \alpha_m$ be the sequence of all types of (free or bound) variables appearing in one of the S_i, in their natural order. Construct a sequence of natural numbers $k_0, \ldots, k_m$ as follows:

$$k_0 := l_0 \qquad \text{where } \alpha_0 = \gamma + l_0 \text{ for some limit } \gamma$$

$$k_{i+1} := k_i + l_{i+1} \qquad \text{where } l_{i+1} = \begin{cases} c_i & \text{if } \alpha_{i+1} = \alpha_i + c_i \text{ for } c_i < \omega \\ c_i + 2 & \text{if } \alpha_{i+1} = \alpha_i + \nu_i \text{ for } \nu_i \geq \omega \text{ and} \\ & \nu_i = \gamma + c_i \text{ for some limit } \gamma \end{cases}$$

For each $i \leq n$, let $S_i^{(\omega)}$ be the typical variant of S_i induced by the function mapping α_i to k_i. Then an easy induction shows that $S_0^{(\omega)}, \ldots, S_n^{(\omega)}$ is a proof of $S_n^{(\omega)} =: S^{(\omega)}$ in K^e_ω. □

Corollary 7. *For each recursive ordinal $\alpha < \omega_1^{CK}$, the consistency of K^e_α can be proved in primitive recursive arithmetic PRA.*

Proof. For α a recursive ordinal, the syntax of K^e_α can be formalized in PRA. Then it is obvious that the proof of the Theorem 6 can also be formalized. Hence PRA proves that K^e_α and K^e_ω are equiconsistent.

Now the consistency of K^e_ω can be trivially proved in PRA, by considering the standard model $(V_\omega, \in)$, along the lines of Gentzen [8]. □

2.4 The Infinitary Systems K^∞_α

In order to get essentially stronger systems, we introduce a very natural infinitary inference rule, similar to the ω-rule in arithmetic. It is motivated by the intention that the set of objects of a limit type γ should be the union of the sets of objects of smaller type. We also add a further rule that is motivated by the desire to have only one object of type 0, viz. the empty set.

The limit rule is

$$\text{Lim} \qquad \frac{S[a^\xi] \quad \text{for all } \xi < \gamma}{S[a^\gamma]}$$

where γ is a limit ordinal and none of the variables a^ξ occur in the lower sequent. The null rule is

$$\text{Null} \qquad \frac{S[\emptyset^{\tau+1}]}{S[a^0]}$$

where $\emptyset^{\tau+1} := \big(\lambda x^\tau.\, x^\tau \neq x^\tau\big)$, and the variable a^0 does not occur in the upper sequent. The system K^e_α with the limit and null rule added is denoted K^∞_α.

Although the proofs in K^∞_α are infinite objects, the very general considerations of López-Escobar [16] show that for recursive ordinals $\alpha < \omega_1^{CK}$, it is sufficient to consider those proofs that can be recursively, i.e finitely presented.

Clearly, the null rule implies that every object type 0 is empty, and thus there is only one object of type 0, i.e., $\neg\exists x^\tau\, x^\tau \in a^0$ and $a^0 = b^0$ are provable.

As an example of the use of these rules, we show that a subset of a set of type τ is itself of type τ. For $\sigma \geq \tau$, let $a^\sigma \subseteq b^\tau$ abbreviate $\forall x^\sigma\, (x^\sigma \in a^\sigma \to x^\sigma \in b^\tau)$.

Proposition 8. $K_\alpha^\infty \vdash a^\sigma \subseteq b^\tau \implies \exists x^\tau\, a^\sigma = x^\tau$.

Proof. We have to distinguish three cases according to the nature of τ. For $\tau = 0$, one can easily deduce $a^\sigma \subseteq \emptyset^1 \implies a^\sigma \gtrsim \emptyset^1$, and an application of Null yields the claim.

For $\tau = \eta{+}1$, the idea is that $a \subseteq b$ means $a \cap b = a$. Let $t^\tau := \big(\lambda x^\eta.\, a^\sigma(x^\eta) \wedge b^\tau(x^\eta)\big)$, then we prove $a^\sigma \subseteq b^\tau \implies a^\sigma = t^\tau$. This is proved via extensionality from $a^\sigma \subseteq b^\tau \implies a^\sigma \gtrsim t^\tau$, which can be derived from Prop. 5.

For limit τ, we proceed inductively: Suppose that for each limit $\gamma < \tau$, we have $a^\sigma \subseteq b^\gamma \implies \exists x^\gamma\, a^\sigma = x^\gamma$. Then we also obtain $a^\sigma \subseteq b^\gamma \implies \exists x^\tau\, a^\sigma = x^\tau$ by lifting the type. Hence, together with the above cases, we obtain $a^\sigma \subseteq b^\xi \implies \exists x^\tau\, a^\sigma = x^\tau$ for each $\xi < \tau$, and an application of Lim yields the claim. □

A more liberal variant of the limit rule concerning several variables at once is actually derivable in the system.

Proposition 9. *For every $n \in \mathbb{N}$, the n-ary limit rule*

$$\frac{S[a_1^\xi, \dots, a_n^\xi] \;\; \textit{for all } \xi < \gamma}{S[a_1^\gamma, \dots, a_n^\gamma]}\,,$$

where γ is a limit and where for each $\tau \leq \gamma$, the variables $a_1^\tau, \dots, a_n^\tau$ are pairwise distinct, is a derived rule of K_α^∞.

Proof. Let the premises $S[a_1^\xi, \dots, a_n^\xi]$ for each $\xi < \gamma$ be given. To derive $S[a_1^\gamma, \dots, a_n^\gamma]$ by n applications of Lim, we need the premises $S[a_1^{\xi_1}, \dots, a_n^{\xi_n}]$ for each n-tuple $(\xi_1, \dots, \xi_n) \in \gamma^n$.

Let such an n-tuple $(\xi_1, \dots, \xi_n)$ be given, and let η be its maximum. Then from $S[a_1^\eta, \dots, a_n^\eta]$, we can derive $\forall x_1^\eta \dots \forall x_n^\eta\, \hat{S}[x_1^\eta, \dots, x_n^\eta]$. On the other hand, the sequent

$$\forall x_1^\eta \dots \forall x_n^\eta\, \hat{S}[x_1^\eta, \dots, x_n^\eta] \implies \hat{S}[a_1^{\xi_1}, \dots, a_n^{\xi_n}]$$

can be obtained by $\forall$:left with shifting of types. Hence a few cuts yield $S[a_1^{\xi_1}, \dots, a_n^{\xi_n}]$. □

Furthermore it suffices to have the premises of a limit inference only for cofinally many types:

Proposition 10. *Let γ be a limit, and let $\Gamma \subseteq \gamma$ be a* cofinal *subset of γ. Then the inference*

$$\frac{S[a^\xi] \;\; \textit{for all } \xi \in \Gamma}{S[a^\gamma]}$$

can be derived in K_α^∞ without use of the null rule.

Proof. Suppose we have the premises $S[a^\xi]$ for all $\xi \in \Gamma$. Consider first the case where $S[a^\eta]$ is well-formed for every $\eta < \gamma$. Then from $S[a^\xi]$ for some $\xi \in \Gamma$, we obtain $S[a^\eta]$ for every $\eta < \xi$ as in the proof of Prop. 9. Since Γ is cofinal in γ, this yields $S[a^\xi]$ for every $\xi < \gamma$, and Lim can be applied to obtain $S[a^\gamma]$.

If for some $\xi < \gamma$, the sequent $S[a^\xi]$ is not well-formed, then this can only be caused by some subformula $a^\gamma(t^\delta)$ in $S[a^\gamma]$. Replace every such subformula in $S[a^\gamma]$ by $t^\delta \in A^\gamma$, giving the equivalent sequent $S'[a^\gamma]$. Now $S'[a^\xi]$ is well-formed for every $\xi < \gamma$ and is obtained from $S'[a^\eta]$ for $\eta \in \Gamma$ as in the case above, and $S'[a^\eta]$ is obtained from the premise $S[a^\eta]$. Then Lim yields $S'[a^\gamma]$, from which we get $S[a^\gamma]$ again. □

This proposition shows that Prop. 8 for $\tau > 0$ can be proved without use of the null rule.

3 Semantics for K_α^∞

We will now introduce a semantics w.r.t. which the systems K_α for *countable* α and their extensions are correct and complete. The models for K_α are a cumulative variant of the "general models" of Henkin [11], they were first defined and the completeness theorem proved in the second author's thesis [13]. Special subclasses of this class of models will form the semantics for the various extensions of K_α. The restriction to countable types is due to the fact that our systems with their finitary syntax can only be complete for those. In order to get complete systems for uncountable types one would have to use a calculus with infinitely long sequents.

3.1 Models for K_α^∞

Let $\alpha < \aleph_1$. A *cumulative structure* $\mathfrak{M}$ of length α is a family of sets $\langle M_\xi \rangle_{\xi<\alpha}$ together with relations $\mathsf{P}_{\delta,\eta} \subseteq M_\delta \times M_\eta$ for $\delta < \eta < \alpha$ with $M_\delta \subseteq M_\eta$ for $\delta < \eta$ and

- $\forall x \in M_\delta\,,\ y \in M_\eta : x\ \mathsf{P}_{\delta,\eta}\ y \ \rightarrow\ \forall \xi < \eta\ (x \in M_\xi \rightarrow x\ \mathsf{P}_{\xi,\eta}\ y)$
- $\forall x \in M_\delta\,,\ y \in M_\eta : x\ \mathsf{P}_{\delta,\eta}\ y \ \rightarrow\ \forall \xi > \delta\ (x \in M_\xi \rightarrow x\ \mathsf{P}_{\delta,\xi}\ y)$

A cumulative structure $\mathfrak{M}$ is called a *Henkin-structure* if every assignment v can be extended to an evaluation $[\![t^\tau]\!]_v$ for arbitrary terms t^τ such that

$$\mathfrak{M}, v \models s^\sigma(t^\tau) \ \text{ iff } \ [\![t^\tau]\!]_v\ \mathsf{P}_{\tau,\sigma}\ [\![s^\sigma]\!]_v\ ,$$

and furthermore $[\![(\lambda x^\tau.\, F[x^\tau])]\!]_v \in M_{\tau+1}$ such that for all $a \in M_\tau$

$$a\ \mathsf{P}_{\tau,\tau+1}\ [\![(\lambda x^\tau.\, F[x^\tau])]\!]_v \ \text{ iff } \ \mathfrak{M}, v' \models F[a^\tau]\ ,$$

where v' is the assignment that differs from v only in that the variable a^τ (that must not occur in $F[x^\tau]$) gets the value a.

For $\gamma > \max(\delta, \eta)$, $x \in M_\delta$ and $y \in M_\eta$ we define

$$x \equiv_\gamma y \;\text{ iff }\; \forall z \in M_\gamma \, (x \;\mathsf{P}_{\delta,\gamma}\; z \;\leftrightarrow\; y \;\mathsf{P}_{\eta,\gamma}\; z) \; .$$

A cumulative structure is called *extensional*, if for every $x \in M_{\delta+1}$ and $y \in M_{\eta+1}$, $(\delta \leq \eta)$ the following holds: If $\forall z \in M_\delta (z \;\mathsf{P}_{\delta,\delta+1}\; x \;\leftrightarrow\; z \;\mathsf{P}_{\delta,\eta+1}\; y)$ and $\forall z \in M_\eta \, (z \;\mathsf{P}_{\eta,\eta+1}\; y \rightarrow \exists w \in M_\delta (w \;\mathsf{P}_{\delta,\delta+1}\; x \wedge z \equiv_{\eta+1} w))$, then $x \equiv_{\eta+2} y$.

Theorem 11. *Every cumulative Henkin structure is a model of K_α, and every extensional cumulative Henkin structure is a model of K^e_α.*

Proof. By an easy induction on the number of inferences in a proof. Note that the properties of a cumulative structure make the shifting of types in the quantifier inferences valid, and that the properties of a Henkin structure make the abstraction inferences valid. For the case with extensionality, note that the condition on a structure being extensional is just a semantic paraphrase of the extensionality rule. □

A cumulative structure is called *normal* if for every $x \in M_\gamma$ for some limit γ and every $y \in M_{\gamma+1}$ with $x \;\mathsf{P}_{\gamma,\gamma+1}\; y$ there is a z in some M_η, $\eta < \gamma$ such that $z \;\mathsf{P}_{\eta,\gamma+1}\; y$.

We call a cumulative structure *null-founded* if for every $x \in M_0$ and $y \in M_{\eta+2}$ with $x \;\mathsf{P}_{0,\eta+2}\; y$ there is a $z \in M_{\eta+1}$ with $\{ v \in M_\eta \, ; \, v \;\mathsf{P}_{\eta,\eta+1}\; z \} = \emptyset$ and $z \;\mathsf{P}_{\eta+1,\eta+2}\; y$.

Theorem 12. *Every normal cumulative Henkin structure is a model of K_α + Lim, and every extensional and null-founded cumulative Henkin structure is a model of K^e_α + Null.*

Proof. For the first part, let $\mathfrak{M}$ satisfy the premises of a limit inference $S[a^\xi]$ for every $\xi < \gamma$. Suppose there is an $a \in M_\gamma$ such that $\mathfrak{M}$ does not satisfy $S[a^\gamma]$. Since $\mathfrak{M}$ is a Henkin structure, this means that $a \;\mathsf{P}_{\gamma,\gamma+1}\; [\![(\lambda x^\gamma . \neg \hat{S}[x^\gamma])]\!]$. By the normality of $\mathfrak{M}$, there is $b \in M_\eta$ for some $\eta < \gamma$ such that $b \;\mathsf{P}_{\eta,\gamma+1}\; [\![(\lambda x^\gamma . \neg \hat{S}[x^\gamma])]\!]$, which means $\mathfrak{M} \models \neg \hat{S}[b^\eta]$, in contradiction to the premise $S[a^\eta]$.

For the second part, assume that $\mathfrak{M} \models S[\emptyset^{\eta+1}]$. Suppose that $\mathfrak{M}$ does not satisfy $S[a^0]$, so there is an $a \in M_0$ with $a \;\mathsf{P}_{0,\eta+2}\; [\![(\lambda x^{\eta+1} . \neg \hat{S}[x^{\eta+1}])]\!]$. Since $\mathfrak{M}$ is null-founded, there is $b \in M_{\eta+1}$ with $\{ x \in M_\eta \, ; \, x \;\mathsf{P}_{\eta,\eta+1}\; b \}$ empty and $b \;\mathsf{P}_{\eta+1,\eta+2}\; [\![(\lambda x^{\eta+1} . \neg \hat{S}[x^{\eta+1}])]\!]$. By extensionality, $b \equiv_{\eta+2} [\![\emptyset^{\eta+1}]\!]$, hence $[\![\emptyset^{\eta+1}]\!] \;\mathsf{P}_{\eta+1,\eta+2}\; [\![(\lambda x^{\eta+1} . \neg \hat{S}[x^{\eta+1}])]\!]$, which implies $\mathfrak{M} \models \neg \hat{S}[\emptyset^{\eta+1}]$, in contradiction to the assumption. □

In particular, every extensional, normal and null-founded cumulative Henkin structure of length α is a model of K^∞_α. The rest of this section is devoted to the proof of the converse of this result.

3.2 Reduction Trees

For every sequent S we define a *reduction tree* as e.g. in Takeuti [24]. For this section, we assume that the language is restricted in such a way as to contain only the logical symbols $\neg$, $\wedge$, $\forall$ and λ. The reduction tree for S is a tree labeled with sequents that can be infinitely branching, but in which every node has finite height, subject to the following conditions:

The root is labeled by the sequent S. Now let x be a node labeled by a sequent $\Gamma \Longrightarrow \Delta$, then the immediate successors of x and their labels are given as follows:

If Γ and Δ have at least one formula in common, then x has no immediate successors, i.e., x is a leaf of the tree. Otherwise we distinguish eleven cases according to the height $h(x) \bmod 11$, where $h(x)$ is the length of the unique path from x to the root.

If $h(x) \equiv 0, 1, 2, 3 \pmod{11}$, then the immediate successors of x are obtained by inverting $\neg$:left, $\neg$:right, $\wedge$:left and $\wedge$:right respectively, as usual (see e.g. [24]).

If $h(x) \equiv 4, 5 \pmod{11}$, the rules $\forall$:left and $\forall$:right are inverted as follows: For every type τ, let δ_τ be a surjective function from ω onto the set $\tau + 1$, and θ_τ a surjective function from ω onto $\alpha \setminus \tau$. Furthermore, let $\{\, t_i^\tau \,;\, i \in \omega \,\}$ be an enumeration of all terms of type τ.

Let $h(x) \equiv 4 \pmod{11}$, and let $\forall x_1^{\tau_1} A_1[x_1^{\tau_1}], \ldots, \forall x_n^{\tau_n} A_n[x_n^{\tau_n}]$ be all universally quantified formulae in Γ. Then x has exactly one immediate successor labeled by

$$\Pi_1, \ldots, \Pi_n, \Gamma \Longrightarrow \Delta \,,$$

where Π_k consists of all formulae $A_k[t_i^{\delta_{\tau_k}(j)}]$ for $i, j < h(x)$, as far as they are well-formed.

Let $h(x) \equiv 5 \pmod{11}$, and let now $\forall x_1^{\tau_1} A_1[x_1^{\tau_1}], \ldots, \forall x_n^{\tau_n} A_n[x_n^{\tau_n}]$ be all universally quantified formulae in Δ. Then x has exactly one immediate successor labeled by

$$\Gamma \Longrightarrow \Delta, \Phi_1, \ldots, \Phi_n \,,$$

where Φ_k consists of all formulae $A_k[a_{k,i}^{\theta_{\tau_k}(i)}]$ for $i < h(x)$, as far as they are well-formed, where the free variables $a_{k,i}^{\theta_{\tau_k}(i)}$ are all distinct and do not occur in $\Gamma \Longrightarrow \Delta$.

Let $h(x) \equiv 6 \pmod{11}$, and let $\big(\lambda x_1^{\tau_1} . A_1[x_1^{\tau_1}]\big)(t_1^{\sigma_1}), \ldots, \big(\lambda x_n^{\tau_n} . A_n[x_n^{\tau_n}]\big)(t_n^{\sigma_n})$ be all abstraction formulae in Γ. Then x has exactly one immediate successor with the label

$$A_{i_1}[t_{i_1}^{\sigma_{i_1}}] \,, \ldots, \, A_{i_m}[t_{i_m}^{\sigma_{i_m}}] \,, \Gamma \Longrightarrow \Delta \,,$$

where the $A_{i_j}[t_{i_j}^{\sigma_{i_j}}]$ are the converses of those of the above abstraction formulae that have well-formed converses. If $h(x) \equiv 7 \pmod{11}$, then the rule λ:right is inverted in the same manner.

If $h(x) \equiv 8 \pmod{11}$, then the extensionality rule is inverted as follows: Let $t_1^{\delta_1}, \ldots, t_m^{\delta_m}$ be all those terms such that there is a formula $t_i^{\delta_i}(s_1^{\sigma+1})$ in Γ and a formula $t_i^{\delta_i}(s_2^{\tau+1})$ in Δ for some terms of the resp. types, where $\sigma \geq \tau$. Let

$$t_1^{\delta_1}(s_{1,1}^{\sigma_1+1}) \,,\; t_1^{\delta_1}(s_{1,2}^{\tau_1+1}) \qquad \ldots \qquad t_m^{\delta_m}(s_{n,1}^{\sigma_n+1}) \,,\; t_m^{\delta_m}(s_{n,2}^{\tau_n+1})$$

be all such pairs of formulae in $\Gamma \Longrightarrow \Delta$, where n can be greater than m if one of the $t_j^{\delta_j}$ appears in more than one formula of the given form in Γ or Δ. Now we define the following formulae for $1 \leq i \leq n$:

$$\begin{array}{llll} A_{i,1} & :\equiv\; s_{i,1}^{\sigma_i+1}(a_i^{\sigma_i}) & B_{i,1} & :\equiv\; \exists x^{\tau_i}\,\big(s_{i,2}^{\tau_i+1}(x^{\tau_i}) \wedge a_i^{\sigma_i} = x^{\tau_i}\big) \\ A_{i,2} & :\equiv\; s_{i,2}^{\tau_1+1}(b_i^{\tau_i}) & B_{i,2} & :\equiv\; s_{i,1}^{\sigma_1+1}(b_i^{\tau_i}) \end{array}$$

where $a_i^{\sigma_i}$ and $b_i^{\tau_i}$ are free variables of the resp. types that are mutually distinct and do not occur in $\Gamma \Longrightarrow \Delta$. Then x has one immediate successor for each function $\pi : \{1, \ldots, n\} \to \{1, 2\}$ labeled by

$$A_{1,\pi(1)}, \ldots, A_{n,\pi(n)}, \Gamma \Longrightarrow \Delta, B_{n,\pi(n)}, \ldots, B_{1,\pi(1)} \,.$$

If $h(x) \equiv 9 \pmod{11}$, then the limit rule is inverted as follows:
Let $\{\, \ell(i) \,;\, i \in \omega \,\}$ be an enumeration of all free variables of limit types, and let $f(n) := n - \lfloor \sqrt{n} \rfloor^2$. Let $k := \frac{h(x)-9}{11}$, and let a^γ be the first variable in the sequence $\langle \ell(f(k+i)) \,;\, i \in \omega \rangle$ that occurs in $\Gamma \Longrightarrow \Delta$. Note that the sequence $\langle \ell(f(i)) \,;\, i \in \omega \rangle$ has the property that every free variable of some limit type appears in it infinitely often. Let n be the number of occurrences of a^γ in $\Gamma \Longrightarrow \Delta$, then there are $r := 2^n - 1$ systems of occurrences of a^γ in $\Gamma \Longrightarrow \Delta$. Let $\Gamma_j[A^\lambda] \Longrightarrow \Delta_j[A^\lambda]$ be the sequent $\Gamma \Longrightarrow \Delta$ where A^λ is indicated in the jth system of occurrences. Then x has one immediate successor for each $(\xi_1, \ldots, \xi_m) \in \gamma^m$ labeled by

$$\Gamma_{i_1}[a_1^{\xi_1}] \,, \ldots \,, \Gamma_{i_m}[a_m^{\xi_m}] \,, \Gamma \Longrightarrow \Delta \,, \Delta_{i_m}[a_m^{\xi_m}] \,, \ldots \,, \Delta_{i_1}[a_1^{\xi_1}] \,,$$

where the $i_1, \ldots, i_m$ are the indices of those systems of occurrences for which the sequent $\Gamma_{i_j}[a^\xi] \Longrightarrow \Delta_{i_j}[a^\xi]$ is well-formed for every $\xi < \gamma$, and all the variables $a_j^{\xi_j}$ are pairwise distinct and do not occur in $\Gamma \Longrightarrow \Delta$.
If $h(x) \equiv 10 \pmod{11}$, then the null rule is inverted as follows: Let σ be a surjective function from ω onto the set of successor ordinals τ with $1 \leq \tau < \alpha$, Let $k := \frac{h(x)-10}{11}$, and let a^0 be the first free variable in the sequence $\langle \mathsf{a}^0_{f(k+i)} \,;\, i \in \omega \rangle$ that occurs in $\Gamma \Longrightarrow \Delta$, such that $\Gamma = \tilde{\Gamma}[a^0]$ and $\Delta = \tilde{\Delta}[a^0]$. Then x has exactly one immediate successor labeled by

$$\tilde{\Gamma}[\emptyset^{\sigma(i_1)}] \,, \ldots \,, \tilde{\Gamma}[\emptyset^{\sigma(i_m)}] \,, \Gamma \Longrightarrow \Delta \,, \tilde{\Delta}[\emptyset^{\sigma(i_m)}] \,, \ldots \,, \tilde{\Delta}[\emptyset^{\sigma(i_1)}] \,,$$

where $i_1, \ldots, i_m$ are those values $1 \leq i_j \leq k$ for which $\tilde{\Gamma}[\emptyset^{\sigma(i_j)}]$ and $\tilde{\Delta}[\emptyset^{\sigma(i_j)}]$ are well-formed.

This completes the definition of the reduction tree for S.

Lemma 13. *If a sequent S has no cut-free proof in K_α^∞, then there is an infinite branch in the reduction tree for S.*

Proof. Suppose there is no infinite branch in the reduction tree for S. Then the reduction tree is well-founded, hence we can show by induction that there is a cut-free proof of every sequent appearing as a label in the tree, in particular of S.

If x is a leaf with label $\Gamma \Longrightarrow \Delta$, then there is a formula A that is contained in both Γ and Δ. Hence $\Gamma \Longrightarrow \Delta$ can be deduced from the axiom $A \Longrightarrow A$ by weakenings and structural inferences. This provides the induction basis.

Now suppose that x is an interior node, and there are cut-free proofs of the labels of x's immediate successors. Then a close inspection of the defining clauses of the reduction tree shows that the label of x can be deduced by several applications of the inference rule inverted at the height of x, and possibly some structural inferences. □

3.3 The Completeness Theorem

Now we are going to prove the completeness of K_α^∞ w.r.t. the semantics defined above. First we need the following fact that was first observed by Schütte [21] in the context of finite type theory. Let Id denote the formula $\forall x^0\, x^0 = x^0$.

Lemma 14. *A sequent $\Gamma \Longrightarrow \Delta$ is provable in K_α^∞ iff there is a cut-free proof of $Id, \Gamma \Longrightarrow \Delta$.*

Proof. If there is a cut-free proof of $Id, \Gamma \Longrightarrow \Delta$, then by a cut with $\Longrightarrow Id$ we can derive $\Gamma \Longrightarrow \Delta$. The reverse direction is proved by induction on the number of inferences in a proof of $\Gamma \Longrightarrow \Delta$.

The only interesting case is that of the last inference in the proof being a cut

$$\frac{\Gamma_1 \Longrightarrow \Delta_1\,,\, A \qquad A\,,\, \Gamma_2 \Longrightarrow \Delta_2}{\Gamma_1\,,\, \Gamma_2 \Longrightarrow \Delta_1\,,\, \Delta_2}\,.$$

By the induction hypothesis there are cut-free proofs of $Id, \Gamma_1 \Longrightarrow \Delta_1, A$ and $A, Id, \Gamma_2 \Longrightarrow \Delta_2$. Now use abstraction inferences to replace A by the formula $\big(\lambda x^0.\, A\big)(a^0)$, and then apply $\rightarrow$:left to deduce the sequent

$$\big(\lambda x^0.\, A\big)(a^0) \rightarrow \big(\lambda x^0.\, A\big)(a^0)\,,\, Id\,,\, \Gamma_1\,,\, \Gamma_2 \Longrightarrow \Delta_1\,,\, \Delta_2\,,$$

from which we get $Id, \Gamma_1, \Gamma_2 \Longrightarrow \Delta_1, \Delta_2$ by two applications of $\forall$:left and structural inferences. □

This lemma can also be interpreted in the following way: In K_α^∞ (and already in K_α) the cut rule can be replaced by the rule of inference

$$\frac{Id, \Gamma \Longrightarrow \Delta}{\Gamma \Longrightarrow \Delta}\,.$$

In fact, the system in [21] had a rule similar to this instead of the cut rule. We can also replace this rule by the more liberal rule

$$\frac{t^\tau = t^\tau, \Gamma \implies \Delta}{\Gamma \implies \Delta} \quad \text{or even} \quad \frac{a^\sigma = t^\tau, \Gamma \implies \Delta}{\Gamma \implies \Delta}$$

where t^τ is an arbitrary term, and in the latter rule $\sigma \geq \tau$ and the free variable a^σ does not occur in either t^τ or the lower sequent. It is not hard to see that these rules are deducible from each other without cuts.

Hence the systems K_α and their extensions can be presented in a formally cut-free way, in the sense that the cut rule is redundant in the system extended by any of the three rules above. But this cut-elimination property is a merely formal one, since the cuts are just hidden in other inferences, and no useful consequences can be derived from it.

Theorem 15. *If the sequent S is not provable in K_α^∞, then there is an extensional, normal and null-founded cumulative Henkin structure that does not satisfy $\hat{S}$.*

Proof. Suppose $\Gamma \implies \Delta$ is not provable, then by Lemma 14 there is no cut-free proof of $Id, \Gamma \implies \Delta$, hence by Lemma 13 there is an infinite path Z in the reduction tree for $Id, \Gamma \implies \Delta$.

We say that a formula A appears left (right) in Z, if A is a formula in the antecedent (succedent) of some label in Z. We shall use the following

Fact: Every formula A in the language of K_α appears either left or right in Z, but not both.

Proof. First, A cannot appear in Z on both sides, since then it would also appear on both sides in one sequent, and Z would be finite. Now, since Id appears left in Z, by the definition of the reduction tree, the formula $t^1(a^0) \wedge \neg t^1(a^0)$ must appear right in Z, for *every* term t^1 of type 1 and every free variable a^0 of type 0. Hence either $t^1(a^0)$ or $\neg t^1(a^0)$ must appear right in Z, and in the latter case, $t^1(a^0)$ appears left in Z. In particular, $\big(\lambda x^0.\, A\big)(a^0)$ must appear either left or right in Z, and hence its converse, which is A, must appear either left or right in Z. □

We now define an equivalence relation on terms by $t^\tau \equiv s^\sigma$ if the formula $t^\tau = s^\sigma$ appears left in Z. It is easily shown by use of the properties of the reduction tree that this is indeed an equivalence. E.g. the symmetry of $\equiv$ can be seen as follows: Suppose $s^\sigma = t^\tau$ appears left in Z, then by construction of the reduction tree, either $u^{\eta+1}(s^\sigma)$ appears left or $u^{\eta+1}(t^\tau)$ appears right in Z, for every term $u^{\eta+1}$ of type $\eta+1$, where $\eta = \max(\sigma, \tau)$. In particular, this holds for $u^{\eta+1} = \big(\lambda x^\eta.\, x^\eta = s^\sigma\big)$, but for this term, $u^{\eta+1}(s^\sigma)$ cannot appear right in Z, hence $u^{\eta+1}(t^\tau)$ appears left in Z, and by construction, $t^\tau = s^\sigma$ appears left in Z.

We denote the equivalence class of a term t^τ by $\big(t^\tau\big)^\equiv$. For $\eta < \alpha$, define the set $T_\eta := \big\{\, \big(t^\xi\big)^\equiv \,;\, \xi \leq \eta \,\big\}$, the set of equivalence classes of terms of type at most

η. For $s \in T_\delta$ and $t \in T_\eta$ with $\delta < \eta$ we define

$$s \; \mathsf{P}_{\delta,\eta} \; t \;:\equiv\; \text{for some } s^\sigma \in s \text{ and } t^\tau \in t \text{ with } \sigma < \tau \text{ the formula } t^\tau(s^\sigma) \text{ appears left in } Z.$$

Using the properties of the reduction tree, one can see that this definition is independent of the choice of representatives s^σ and t^τ. It is obvious from the definition that $\mathfrak{T} := \langle T_\xi \rangle_{\xi<\alpha}$ together with the so defined relations $\mathsf{P}_{\delta,\eta}$ forms a cumulative structure of length α.

Claim: Every class $x \in T_\eta$ contains some free variable c^η of type η.

Proof. Let $x = \left(t^\xi\right)^{\equiv}$, then the formula $\exists x^\eta \, x^\eta = t^\xi$ must appear left in Z, since otherwise $\forall x^\eta \, x^\eta \neq t^\xi$ would appear left in Z. By the construction of the reduction tree, then also $t^\xi \neq t^\xi$ would appear left in Z, but since $t^\xi = t^\xi$ also appears left, Z would be finite.

Therefore $\forall x^\eta \, x^\eta \neq t^\xi$ appears right, hence for some free variable c^η, $c^\eta \neq t^\tau$ appears right and hence $c^\eta = t^\xi$ appears left in Z. $\square$

To show that $\mathfrak{T}$ is a Henkin structure, we have to define the value $[\![t^\tau]\!]_v$ for any assignment v and term t^τ. Let t^τ be $\tilde{t}^\tau[a_1^{\tau_1}, \dots, a_n^{\tau_n}]$, where all free variables in t^τ are those indicated. Now for each $i \leq n$, choose a term $s_i^{\xi_i} \in v(a_i^{\tau_i})$ such that $\tilde{t}^\tau[s_1^{\xi_1}, \dots, s_n^{\xi_n}]$ is well-formed, and let

$$[\![t^\tau]\!]_v := \left(\tilde{t}^\tau[s_1^{\xi_1}, \dots, s_n^{\xi_n}]\right)^{\equiv} .$$

We shall show in the following that this definition is independent of the choice of the terms $s_i^{\xi_i}$, that it makes $\mathfrak{T}$ into a Henkin structure, and that for a formula $A[a_1^{\tau_1}, \dots, a_n^{\tau_n}]$ with only the free variables indicated,

$$\mathfrak{T}, v \models A[a_1^{\tau_1}, \dots, a_n^{\tau_n}] \text{ iff for some } s_i^{\xi_i} \in v(a_i^{\tau_i}), \text{ the formula } A[s_1^{\xi_1}, \dots, s_n^{\xi_n}] \text{ is well-formed and appears left in } Z.$$

All these are proved by simultaneous induction on the formation of terms and formulae.

The induction base is trivial for free variables, and for atomic formulae it follows directly from the definition of the relations $\mathsf{P}_{\delta,\eta}$. The inductive step is easy for negations and conjunctions.

Now let $A[a_1^{\tau_1}, \dots, a_n^{\tau_n}]$ be $\forall x^\tau \, B[x^\tau, a_1^{\tau_1}, \dots, a_n^{\tau_n}]$. By definition we have $\mathfrak{T}, v \models A[a_1^{\tau_1}, \dots, a_n^{\tau_n}]$ iff for every assignment v' with $v'(a_i^{\tau_i}) = v(a_i^{\tau_i})$ for all $i \leq n$,

$$\mathfrak{T}, v' \models B[b^\tau, a_1^{\tau_1}, \dots, a_n^{\tau_n}] .$$

By the above claim, there is a free variable $c^\tau \in v'(b^\tau)$, hence by the induction hypothesis this is equivalent to

$$B[c^\tau, s_1^{\xi_1}, \dots, s_n^{\xi_n}] \text{ appears left in } Z.$$

But then $A[s_1^{\xi_1}, \dots, s_n^{\xi_n}]$ cannot appear right in Z, since by construction of the reduction tree then $B[c^\tau, s_1^{\xi_1}, \dots, s_n^{\xi_n}]$ would also appear right in Z, thus Z would be finite. Hence $A[s_1^{\xi_1}, \dots, s_n^{\xi_n}]$ appears left in Z. Conversely, if $A[s_1^{\xi_1}, \dots, s_n^{\xi_n}]$ appears left in Z, then by construction $B[c^\tau, s_1^{\xi_1}, \dots, s_n^{\xi_n}]$ also appears left in Z.

Now consider the term $\big(\lambda x^\tau . B[x^\tau, a_1^{\tau_1}, \dots, a_n^{\tau_n}]\big)$, then by definition,

$$[\![\big(\lambda x^\tau . B[x^\tau, a_1^{\tau_1}, \dots, a_n^{\tau_n}]\big)]\!]_v = \big(\big(\lambda x^\tau . B[x^\tau, s_1^{\xi_1}, \dots, s_n^{\xi_n}]\big)\big)^{\equiv} .$$

Let $x \in T_\tau$, then we have to verify that

$$x \;\mathsf{P}_{\tau,\tau+1}\; \big(\big(\lambda x^\tau . B[x^\tau, s_1^{\xi_1}, \dots, s_n^{\xi_n}]\big)\big)^{\equiv} \;\text{ iff }\; \mathfrak{T}, v' \models B[b^\tau, a_1^{\tau_1}, \dots, a_n^{\tau_n}] ,$$

where $v'(a_i^{\tau_i}) = v(a_i^{\tau_i})$ for all $i \leq n$ and $v'(b^\tau) = x$. Let $s^\xi \in x$, then the left hand side is equivalent to $\big(\lambda x^\tau . B[x^\tau, s_1^{\xi_1}, \dots, s_n^{\xi_n}]\big)(s^\xi)$ appearing left in Z, and the right hand side is equivalent to $B[s^\xi, s_1^{\xi_1}, \dots, s_n^{\xi_n}]$ appearing left in Z. The latter two are equivalent by the construction of the reduction tree. Finally, it is easily verified that the value of $[\![\big(\lambda x^\tau . B[x^\tau, a_1^{\tau_1}, \dots, a_n^{\tau_n}]\big)]\!]_v$ is independent of the choice of the term $s_i^{\xi_i}$. This finishes the proof that $\mathfrak{T}$ is a Henkin structure.

Let v_0 be the assignment with $v_0(a^\tau) := \big(a^\tau\big)^{\equiv}$ for every free variable a^τ. Then by the above reasoning $\mathfrak{T}, v_0 \models \neg\hat{S}$, hence $\mathfrak{T} \not\models \hat{S}$. It remains to show that $\mathfrak{T}$ is extensional, normal and null-founded.

Let $\delta \leq \eta$ and $x \in T_{\delta+1}$, $y \in T_{\eta+1}$ such that not $x \equiv_{\eta+2}$, i.e., there is $z \in T_{\eta+2}$ with $x \;\mathsf{P}_{\delta+1,\eta+1}\; z$ but not $y \;\mathsf{P}_{\eta+1,\eta+2}\; z$. Let $t_1^{\delta+1} \in x$, $t_2^{\eta+1} \in y$ and $s^{\eta+2} \in z$ such that $s^{\eta+2}(t_1^{\delta+1})$ appears left in Z and $s^{\eta+2}(t_2^{\eta+1})$ appears right in Z. By the construction of the reduction tree there are two possibilities:

- $t_1^{\delta+1}(b^\delta)$ appears left and $t_2^{\eta+1}(b^\delta)$ appears right in Z, for some free variable b^δ. Hence if we set $u := \big(b^\delta\big)^{\equiv}$, we have a $u \in T_\delta$ with $u \;\mathsf{P}_{\delta,\delta+1}\; x$ but not $u \;\mathsf{P}_{\delta,\eta+1}\; y$.
- $t_2^{\eta+1}(a^\eta)$ appears left and $\exists x^\delta \; \big(t_1^{\delta+1}(x^\delta) \wedge a^\eta = x^\delta\big)$ appears right in Z. If we set $u = \big(a^\eta\big)^{\equiv}$, then we have $u \;\mathsf{P}_{\eta,\eta+1}\; y$, and the construction of the reduction tree guarantees that $u \equiv_{\eta+1} w$ and $w \;\mathsf{P}_{\delta,\delta+1}\; x$ cannot both hold for any $w \in T_\delta$.

Hence $\mathfrak{T}$ is extensional. Now let $\gamma < \alpha$ be a limit ordinal, and let $x \in T_\gamma$ and $y \in T_{\gamma+1}$ with $x \;\mathsf{P}_{\gamma,\gamma+1}\; y$. To show $\mathfrak{T}$ is normal, we have to find a $z \in T_\xi$ for some $\xi < \gamma$ with $z \;\mathsf{P}_{\xi,\gamma+1}\; y$.

By the above claim, there must be a free variable $c^\gamma \in x$. Let $t^{\gamma+1} \in y$ with $t^{\gamma+1}(c^\gamma)$ appearing left in Z. Then by the construction of the reduction tree, there must be a free variable c^τ of some type $\tau < \gamma$ such that $t^{\gamma+1}(c^\tau)$ also appears left in Z. We set $z := \big(c^\tau\big)^{\equiv}$, so we have $z \in T_\tau$ and $z \;\mathsf{P}_{\tau,\gamma+1}\; y$.

Finally, let $x \in T_0$ and $y \in T_{\eta+2}$ with $x \;\mathsf{P}_{0,\eta+2}\; y$, and let c^0 be a free variable with $c^0 \in x$ and $t^{\eta+2} \in y$ with $t^{\eta+2}(c^0)$ appearing left in Z. By construction of

the reduction tree, $t^{\eta+2}(\emptyset^{\eta+1})$ also appears left in Z, so we can set $z := \left(\emptyset^{\eta+1}\right)^{\equiv}$ and have $z \ \mathsf{P}_{\eta+1,\eta+2} \ y$. Furthermore, the construction of the reduction tree guarantees that $u \ \mathsf{P}_{\eta,\eta+1} \ z$ cannot hold for any $u \in T_\eta$. Hence $\mathfrak{T}$ is null-founded. $\square$

3.4 Set-Theoretic Models

A *cumulative hierarchy* of length α is a family of sets $\mathbb{H}_\alpha = \langle H_\xi \rangle_{\xi<\alpha}$ such that $H_0 = \{\emptyset\}$, $H_\xi \subseteq H_{\xi+1} \subseteq 2^{H_\xi}$ for each ξ with $\xi + 1 < \alpha$ and $H_\gamma = \bigcup_{\xi<\gamma} H_\xi$ for every limit ordinal $\gamma < \alpha$.

A cumulative hierarchy $\mathbb{H}_\alpha$ can be regarded as a cumulative structure by setting $\mathsf{P}_{\delta,\eta} := \in$ for every $\delta < \eta < \alpha$. By definition, every cumulative hierarchy is extensional, normal and null-founded.

A special case is the *full* cumulative hierarchy $\mathbb{V}_\alpha = \langle V_\xi \rangle_{\xi<\alpha}$, where $V_{\xi+1} := 2^{V_\xi}$ for each ξ with $\xi + 1 < \alpha$. It is obvious that $\mathbb{V}_\alpha$ is a Henkin structure, hence $\mathbb{V}_\alpha \models K_\alpha^\infty$. We are going to show that in fact every model of K_α^∞ is equivalent to some cumulative hierarchy. Hence when considering models of K_α^∞, it suffices to look at cumulative hierarchies that are Henkin structures. This will be crucial below when we prove the (non-)derivability of translations of set-theoretic sentences in K_α^∞ semantically.

Theorem 16. *Let $\mathfrak{M} \models K_\alpha^\infty$, then there is a cumulative hierarchy $\mathbb{H}_\alpha$ such that for every sentence A in the language of K_α, $\mathfrak{M} \models A$ iff $\mathbb{H}_\alpha \models A$.*

Proof. To avoid notational complexities, we shall first treat the case where α is a limit ordinal and hence there is no largest type. Let $\mathfrak{M}$ be a cumulative Henkin structure of length Δ. We first construct a quotient structure where Leibniz equality is interpreted as identity. For $x \in \bigcup \mathfrak{M}$ let

$$\begin{aligned} [x]_\eta &:= \begin{cases} \{\, y \in M_\eta \,;\, y \equiv_{\eta+1} x \,\} & \text{if } x \in M_\eta \\ \emptyset & \text{otherwise} \end{cases} \\ [x] &:= \bigcup_{\eta<\alpha} [x]_\eta \end{aligned}$$

and then set $\bar{M}_\eta := \left\{\, [x] \,;\, x \in M_\alpha \,\right\}$ and let $\bar{\mathfrak{M}}$ be the cumulative structure $\langle \bar{M}_\eta \rangle_{\eta<\alpha}$ with the relations

$$[x] \ \bar{\mathsf{P}}_{\delta,\eta} \ [y] \ := \ x \ \mathsf{P}_{\delta,\eta} \ y$$

for $x \in M_\delta$ and $y \in M_\eta$. One can easily show that this definition is independent of the choice of representatives.

For an assignment v in $\mathfrak{M}$, let $\bar{v}$ be the assignment in $\bar{\mathfrak{M}}$ with $\bar{v}(a^\tau) = \left[v(a^\tau)\right]$ for each variable a^τ. Clearly, every assignment in $\bar{\mathfrak{M}}$ is of the form $\bar{v}$ for some assignment v in $\mathfrak{M}$. We extend the assignment $\bar{v}$ to arbitrary terms by setting

$$[\![t^\tau]\!]_{\bar{v}} := \left[\, [\![t^\tau]\!]_v \,\right] .$$

By a straightforward simultaneous induction on the formation of terms and formulae, one shows that this definition makes $\bar{\mathfrak{M}}$ a Henkin structure, and that for every formula A

$$\bar{\mathfrak{M}}, \bar{v} \models A \text{ iff } \mathfrak{M}, v \models A\,.$$

Hence to prove the theorem, it remains to show that $\bar{\mathfrak{M}}$ is isomorphic to some cumulative hierarchy. We define a hierarchy $\mathbb{H}_\alpha$ together with an isomorphism ι as follows:

Since $K_\alpha^\infty \vdash a^0 = b^0$, we have $|\bar{M}_0| = 1$. Let x_0 be the unique element of $\bar{M}_0$, and define $\iota(x_0) := \emptyset$ and $H_0 := \{\iota(x_0)\} = \{\emptyset\}$. Now let $x \in \bar{M}_{\eta+1}$, then we set

$$\iota(x) := \left\{\, \iota(y)\,;\; y \in \bar{M}_\eta \text{ and } y\ \bar{\mathsf{P}}_{\eta,\eta+1}\ x \,\right\}$$

and $H_{\eta+1} := \left\{\, \iota(x)\,;\; x \in \bar{M}_{\eta+1} \,\right\}$. We need to show that ι restricted to $\bar{M}_{\eta+1}$ is injective. So let $\iota(x) = \iota(y)$, thus for all $z \in \bar{M}_\eta$, $z\ \bar{\mathsf{P}}_{\eta,\eta+1}\ x$ iff $z\ \bar{\mathsf{P}}_{\eta,\eta+1}\ y$, hence by extensionality $x \equiv_{\eta+2} y$, and hence by definition of $\bar{\mathfrak{M}}$ we have $x = y$.

Finally, since $\bar{\mathfrak{M}}$ is normal, $\bar{M}_\gamma = \bigcup_{\xi<\gamma} \bar{M}_\xi$ for every limit γ, which is seen as follows: Let $x \in \bar{M}_\gamma$, let $v(a^\gamma) = x$ and consider $t := [\![(\lambda x^\gamma . x^\gamma = a^\gamma)]\!]_v$. Then $y\ \mathsf{P}_{\gamma,\gamma+1}\ t$ holds only if $y = x$. By normality, there is $y' \in \bar{M}_\xi$ for some $\xi < \gamma$ with $y'\ \mathsf{P}_{\xi,\gamma+1}\ t$, hence $x = y'$ and so $x \in \bar{M}_\xi$. So ι is already defined on $\bar{M}_\gamma$, and we can set

$$H_\gamma := \left\{\, \iota(x)\,;\; x \in \bar{M}_\gamma \,\right\} = \bigcup_{\xi<\gamma} H_\xi\,.$$

The so defined $\mathbb{H}_\alpha := \langle H_\xi \rangle_{\xi<\alpha}$ is obviously a cumulative hierarchy isomorphic to $\bar{\mathfrak{M}}$.

For the case where α is a successor ordinal, say $\alpha = \beta + 1$, we have to find an equivalence relation $\equiv_{(\beta)}$ replacing $\equiv_{\beta+1}$ in the definition of $[x]_\beta$. Since we are only dealing with extensional structures, we use a version of extensional equivalence. If β is itself a successor $\beta' + 1$, then we define

$$x \equiv_{(\beta)} y \text{ iff } \forall z \in M_{\beta'}\ \ z\ \mathsf{P}_{\beta',\beta}\ x \leftrightarrow z\ \mathsf{P}_{\beta',\beta}\ y$$

and if β is a limit ordinal, we set

$$x \equiv_{(\beta)} y \text{ iff } \forall \xi < \beta\ \forall z \in M_\xi\ \ z\ \mathsf{P}_{\xi,\beta}\ x \leftrightarrow z\ \mathsf{P}_{\xi,\beta}\ y$$

If we now define $[x]_\beta$ as $\left\{\, y \in M_\beta\,;\; y \equiv_{(\beta)} x \,\right\}$, then the proof can proceed as before to yield the theorem in the general case. $\square$

4 Set Theory in K_α^∞

Let $L_\in$ denote the first order language of set theory, with equality and the membership relation $\in$ as the only non-logical symbols. A formula of $L_\in$ is called *bounded* if all quantifiers in it are of the form $\forall x \in y$ or $\exists x \in y$. The main

property of bounded formulae is their absoluteness, i.e., if M is a transitive class and φ is bounded, then φ and φ^M are equivalent.

Zermelo set theory Z is the theory in $L_\in$ given by the axioms of extensionality, the separation scheme, pairing, union, powerset, infinity and regularity, i.e., ZF minus the replacement scheme.

4.1 Embedding Set Theory

For a formula φ of $L_\in$ and a type τ, let the formula $\varphi^{(\tau)}$, called the τ-translation of φ, be obtained by replacing every variable in φ by a variable of type τ, and then interpreting $\in$ and $=$ as the defined relations in the language of K_α. For this to work, we have to assume $\alpha \geq \tau + 3$, since the definition of $a^\tau \in b^\tau$ involves variables of the types $\tau + 1$ and $\tau + 2$.

We shall show that K_α^∞ can prove the τ-translations of the theorems of Zermelo set theory, as well as the complete $L_\in$-theory of the hereditarily finite sets, for suitable types τ. The latter shows that the limit rule for type ω is as least as strong as the formally similar ω-rule.

Below, we use the following convention throughout: If we name an element $a \in H_\tau$ in some cumulative hierarchy $\mathbb{H}_\alpha$, and in the same context we speak of a formula or term containing a free variable denoted a^τ, then it is understood that the value a is assigned to this variable. The same applies to other names.

Lemma 17. *Let $\mathbb{H}_\alpha$ be a cumulative hierarchy of sets, and let $\alpha \geq \tau + 3$. Then for each formula $\varphi(\bar{x})$ of $L_\in$ and $\bar{a} \in H_\tau$, $\mathbb{H}_\alpha \models \varphi(\bar{a})^{(\tau)}$ iff $H_\tau \models \varphi(\bar{a})$.*

Proof. Since the τ-translation preserves the logical structure, it suffices to show the claim for atomic formulae of $L_\in$. Let $a, b \in H_\tau$.

Clearly, if $a = b$, then $\mathbb{H}_\alpha \models a = b$. If $a \neq b$, then by extensionality there is a c such that w.l.o.g. $c \in a$ but $c \notin b$. Then for some $\eta < \tau$, $c \in H_\eta$. Now the set $\{ x \in H_\tau \,;\, c \in x \}$ is an element of $H_{\tau+1}$, since it can be defined by $\big(\lambda x^\tau .\, x^\tau(c^\eta)\big)$, and it serves as a witness that $\mathbb{H}_\alpha \not\models a = b$.

Now let $a \in b$. Then, as $b \in H_{\tau+1}$, it witnesses that $\mathbb{H}_\alpha \models a \in b$. Vice versa, if $\mathbb{H}_\alpha \models a \in b$, then there must be $c \in H_{\tau+1}$ such that $c \in a$ and $\mathbb{H}_\alpha \models c = b$. As in the above it follows that $c = b$, hence $a \in b$. □

It is well-known that if $\gamma \geq \omega + \omega$ is a limit, then $V_\gamma \models Z$. We shall now show that in fact K_α proves the γ-translation of Z for such γ, whenever $\alpha \geq \gamma + 3$.

Theorem 18. *Let $\gamma \geq \omega + \omega$ be a limit and $\alpha \geq \gamma + 3$. Then $K_\alpha^\infty \vdash \varphi^{(\gamma)}$ for each theorem φ of Z.*

Proof. Let $\mathbb{H}_\alpha \models K_\alpha^\infty$. By Lemma 17 it suffices to show $H_\gamma \models Z$. The proof follows the lines of the usual proofs that V_γ satisfies Z for a limit $\gamma \geq \omega + \omega$.

We first show that H_γ is transitive. Let $a \in H_\gamma$, then there is $\eta < \gamma$ with $a \in H_{\eta+1}$, hence $a \subseteq H_\eta \subseteq H_\gamma$. Since H_γ is transitive, it satisfies extensionality.

Let $a \in H_\gamma$ and $\varphi(x)$ formula of $L_\in$. The set $s := \{\, x \in a \,;\, \varphi(x) \,\}$ is a subset of H_γ by transitivity. Now $a \in H_\eta$ for some $\eta < \gamma$, and by Prop. 5, every element of a is also in H_η, so $s \subseteq H_\eta$. Now s can be described by the term $\big(\lambda x^\eta.\, x^\eta \in a^\eta \wedge \varphi^{(\gamma)}(x^\eta)\big)$, hence $s \in H_{\eta+1}$ and thus $s \in H_\gamma$, so γ satisfies separation.

Let $a, b \in H_\gamma$, then there is $\eta < \gamma$ such that $\{a, b\} \subseteq H_\eta$. Then $\{a, b\} \in H_{\eta+1}$, as it is described by the term $\big(\lambda x^\eta.\, x^\eta = a^\eta \vee x^\eta = b^\eta\big)$. Hence $\{a, b\} \in H_\gamma$, so H_γ satisfies the pairing axiom.

Let $a \in H_\gamma$, then there is $\eta < \gamma$ with $a \in H_{\eta+2}$, hence $a \subseteq H_{\eta+1}$. But then $\bigcup a \subseteq H_\eta$, and so $\bigcup a \in H_{\eta+1}$, as it is described by the term $\big(\lambda x^\eta.\, \exists y^{\eta+1} a^{\eta+2}(y^{\eta+1}) \wedge y^{\eta+1}(x^\eta)\big)$. Hence $\bigcup a \in H_\gamma$, so H_γ satisfies the union axiom.

In order to see that H_γ satisfies the powerset axiom, we have to show that for each $a \in H_\gamma$, the set $p := \{\, x \in H_\gamma \,;\, x \subseteq a \,\}$ is an element of H_γ. Now there is $\eta < \gamma$ such that $a \in H_{\eta+1}$. For $b \in H_\gamma$ with $b \subseteq a$, Prop. 8 then yields $b \in H_{\eta+1}$ also. Hence $p = \{\, x \in H_{\eta+1} \,;\, x \subseteq a \,\}$, and this can be described by the term $\big(\lambda x^{\eta+1}.\, \forall y^\eta \, \big(x^{\eta+1}(y^\eta) \to a^{\eta+1}(y^\eta)\big)\big)$. Therefore $p \in H_{\eta+2}$ and thus $p \in H_\gamma$.

Consider the set H_ω: we have $\emptyset \in H_\omega$ and $\forall x \in H_\omega \; x \cup \{x\} \in H_\omega$, since for $a \in H_n$, $a \cup \{a\} \in H_{n+1}$ as it is described by the term $\big(\lambda x^n.\, x^n \in a^n \vee x^n = a^n\big)$. Furthermore $H_\omega \in H_{\omega+1}$, as it is described by the term $\big(\lambda x^\omega.\, x^\omega = x^\omega\big)$, and hence $H_\omega \in H_\gamma$. This shows that H_γ satisfies the axiom of infinity.

Let $\emptyset \neq a \in H_\gamma$, let η be the least ordinal such that $a \cap H_\eta \neq \emptyset$, and let $b \in a \cap H_\eta$. Then $a \cap b = \emptyset$, since either $\eta = 0$, then $b = \emptyset$, or else $\eta = \eta' + 1$. In the latter case no element $c \in b$ can be an element of a, since by Lemma 5 then $c \in H_{\eta'}$, in contradiction to the minimality of η. Thus H_γ satisfies regularity, which completes the proof. $\square$

Purely syntactical derivations of the $(\omega + \omega)$-translations of the axioms of Z in $K^\infty_{\omega+\omega+3}$, as well as a syntactical proof of the following theorem, were given by the second author in [12].

Theorem 19. *Let $\gamma \geq \omega$ be a limit, and $\alpha \geq \gamma + 3$. Then for each sentence φ of $L_\in$, $K^\infty_\alpha \vdash \big(\varphi^{V_\omega}\big)^{(\gamma)}$ iff $V_\omega \models \varphi$. Moreover $K^\infty_\alpha \vdash \varphi^{(\omega)}$ iff $V_\omega \models \varphi$.*

Proof. If $K^\infty_\alpha \vdash \big(\varphi^{V_\omega}\big)^{(\gamma)}$, then $V_\gamma \models \varphi^{V_\omega}$, and $V_\omega \models \varphi$ follows by absoluteness, since φ^{V_ω} is a bounded formula for every φ.

For the other direction, we need that for each $\mathbb{H}_\alpha \models K^\infty_\alpha$, $H_\omega = V_\omega$. To see this, we need to show that for every $n \in \mathbb{N}$, $H_n = V_n$. By definition, $H_0 = V_0 = \{\emptyset\}$. Now, let $H_n = V_n$. Since H_n is finite, we can define every subset of H_n by a term of type $n + 1$ with parameters from H_n, hence $H_{n+1} = 2^{H_n} = V_{n+1}$.

Now suppose $K^\infty_\alpha \nvdash \big(\varphi^{V_\omega}\big)^{(\gamma)}$. Then there is a cumulative hierarchy $\mathbb{H}_\alpha$ such that $\mathbb{H}_\alpha \models K^\infty_\alpha + \neg\big(\varphi^{V_\omega}\big)^{(\gamma)}$, hence $H_\gamma \models \neg\varphi^{V_\omega}$. As H_γ is transitive, $H_\omega \models \neg\varphi$ follows by absoluteness, hence by the above $V_\omega \models \neg\varphi$.

The second part follows from $\mathbb{H}_\alpha \models (\forall x \; x \in V_\omega)^{(\omega)}$, which is equivalent to $H_\omega \models \forall x \; x \in V_\omega$ since $x \in V_\omega$ is a bounded formula. But $H_\omega \models \forall x \; x \in V_\omega$ is

true, since $H_\omega = V_\omega$. $\square$

Note that e.g. the consistency of Z (or even ZFC) can be expressed in the form φ^{V_ω} for some sentence φ. Hence there are set-theoretic sentences whose translations are provable in K_α^∞, but that are not provable in Z. A more genuinely set-theoretic example of this is given in the following.

Let TC denote the sentence saying that every set has a transitive superset. This theorem of ZF is unprovable without the axiom of replacement, a folklore result that we found in Drake [7].

Theorem 20. *TC is not provable in Z.*

Proof. Let $t(x) := \{x\}$, and let $t^n(x)$ denote the n-fold iteration of this operation. Let $Q := \{\, t^n(n) \,;\, n \in \omega \,\}$, and define $Q_0 := Q$ and $Q_{i+1} := \bigcup Q_i$ for $i \in \omega$. Then define a model M by

$$\begin{aligned} M_0 &:= \omega \cup Q \\ M_{i+1} &:= M_i \cup 2^{M_i} \cup Q_{i+1} \\ M &:= \bigcup_{i \in \omega} M_i \end{aligned}$$

It is not difficult to see that $M \models Z$. First show that M is a transitive set, by induction on i in the construction.

If $x \in M_0$, then either $x \in \omega$ and thus $x \subseteq M_0$, or $x \in Q$, but then $x = \{q\}$ with $q \in M_1$, and thus $\{q\} \subseteq M_2$.

Now if $x \in M_{i+1}$, then either $x \in M_i$, hence $x \subseteq M$ by the induction hypothesis, or $x \subseteq M_i$, or $x \in Q^{(i+1)}$. In the latter case, either $x \in \omega$ or $x = \{q\}$ for some $q \in M_{i+2}$, and in both cases we have $x \subseteq M_{i+2}$.

Since M is transitive, it satisfies extensionality and regularity, and since $\omega \in M$, it also satisfies the axiom of infinity. It can be easily verified, in a manner similar to the proof of Theorem 18, that M satisfies the pairing, separation and powerset axioms. To show that M satisfies the union axiom, one has to show the following stronger claim by induction on i, which is routine:

Claim: If $x \in M_i$ or $x \subseteq M_i$, then $\bigcup x \in M$.

Now consider $Q \in M$. If M contained some transitive superset of Q, then since M satisfies separation, the set $\tilde{Q} := \{\, t^m(n) \,;\, m \leq n \in \omega \,\}$ would be in M. But for every $n \in \omega$, the set $t^{n+1}(2n+2) \notin M_n$, hence $\tilde{Q} \notin M$. $\square$

On the other hand, TC is quite easily provable in K_α^∞, thus is an example of a combinatorial set-theoretic statement witnessing that K_α^∞ is properly stronger than Z.

Theorem 21. *For every limit γ and $\alpha \geq \gamma + 3$, $K_\alpha^\infty \vdash TC^{(\gamma)}$.*

Proof. Let $a^\tau \in^* b^\tau$ be the formula

$$\forall x^{\tau+1} \left[x^{\tau+1}(b^\tau) \wedge \forall y^\tau, z^\tau \left(x^{\tau+1}(z^\tau) \wedge y^\tau \in z^\tau \to x^{\tau+1}(y^\tau) \right) \to x^{\tau+1}(a^\tau) \right. .$$

It is easily seen that $\in^*$ is a transitive super-relation of $\in$, i.e., we can prove in K_α^∞ that $a^\tau \in b^\tau \to a^\tau \in^* b^\tau$ and $a^\tau \in b^\tau \wedge b^\tau \in^* c^\tau \to a^\tau \in^* c^\tau$.

Let $\mathbb{H}_\alpha \models K_\alpha^\infty$, we show that $H_\gamma \models TC$. So let $a \in H_\gamma$, then $a \in H_\xi$ for some $\xi < \gamma$. Then the term $\big(\lambda x^\xi.\, x^\xi \in^* a^\xi\big)$ describes a transitive superset of a. Hence H_γ contains a transitive superset of each $a \in H_\gamma$. □

On the other hand, we have the following undesired property.

Theorem 22. *Let δ be a countable ordinal definable by a bounded formula in the language $L_\in$, and let $\gamma \leq \delta$ be a limit ordinal. Then for $\alpha \geq \gamma + 3$, $K_\alpha^\infty \vdash \big(\neg\exists x\, x = \delta\big)^{(\gamma)}$.*

Proof. Suppose not, then $\big(\exists x\, x = \delta\big)^{(\gamma)}$ is consistent with K_α^∞, hence there is a cumulative hierarchy $\mathbb{H}_\alpha$ such that $\mathbb{H}_\alpha \models \big(\exists x\, x = \delta\big)^{(\gamma)}$. This means that $H_\gamma \models \exists x\, x = \delta$, and by absoluteness we have $\delta \in H_\gamma$, which is impossible since every element of H_γ is of rank less than γ. □

Thus K_α^∞ proves set-theoretic statements that are "false", i.e., inconsistent with ZF. This raises the problem of determining which set-theoretic theorems of K_α^∞ are consistent with ZF, or at least give sufficient criteria for consistency. This problem is dealt with in the next section.

4.2 Criteria for Consistency

For this section, we fix a limit γ and $\alpha \geq \gamma + 3$, and let φ be a sentence of $L_\in$ such that $K_\alpha^\infty \vdash \varphi^{(\gamma)}$. We call φ *persistent*, if for every limit $\gamma' > \gamma$ there is an $\alpha' \geq \gamma' + 3$ such that $K_{\alpha'}^\infty \vdash \varphi^{(\gamma')}$.

Theorem 23. *If φ is persistent, then $ZF + \varphi$ is consistent.*

Proof. Since $\mathbb{V}_{\alpha'} \models K_{\alpha'}^\infty$, by the persistence of φ we have $V_{\gamma'} \models \varphi$ for all limits $\gamma' > \gamma$. Now suppose $ZF \vdash \neg\varphi$, then by the reflection principle of ZF there are arbitrary large limit ordinals δ such that $V_\delta \models \neg\varphi$, in contradiction to the above. □

Hence we are looking for criteria for persistence. One sufficient criterion turns out to be provability in K_α^∞ without any applications of the null rule.

For ordinals $\gamma < \delta$, let $\delta - \gamma$ denote the unique ordinal ξ with $\delta = \gamma + \xi$. For a sequent S, let $\boldsymbol{\ell}(S)$ denote the ordered sequence $\langle \gamma_1, \ldots, \gamma_n \rangle$ of those limit ordinals such that a (free or bound) variable of type $\gamma_i + k$ for some $k \in \omega$ occurs in S. We say that $\gamma_1, \ldots, \gamma_n$ are the limits occurring in S. Then if $\boldsymbol{\delta}$ is an ordered sequence $\langle \delta_1, \ldots, \delta_n \rangle$ of limit ordinals of the same length, we denote by $S(\boldsymbol{\delta})$ the typical variant (cf. section 2.3) of S induced by the function mapping $\gamma_i + k$ to $\delta_i + k$ for each $i \leq n$ and each $k < \omega$.

Theorem 24. *Let S with $\boldsymbol{\ell}(S) = \langle \gamma_1, \ldots, \gamma_n \rangle$ be provable in K_α^∞ without use of the null rule, and let $\boldsymbol{\delta} = \langle \delta_1, \ldots, \delta_n \rangle$ be an ordered sequence of limits with*

(1) $\delta_i \geq \gamma_i$ *for every* $i \leq n$,

(2) $\delta_{i+1} - \delta_i \geq \gamma_{i+1} - \gamma_i$ *and*

(3) if $\gamma_{i+1} = \gamma_i + \omega$, *then* $\delta_{i+1} = \delta_i + \omega$, *for every* $i < n$.

Then there is $\alpha' > \alpha$ *such that* $K^{\infty}_{\alpha'}$ *proves* $S(\boldsymbol{\delta})$ *without use of the null rule.*

Proof. By induction on the length of a proof S. Obviously, if S is an axiom, then also $\mathrm{S}(\boldsymbol{\delta})$ is an axiom, which gives the induction base.

For the inductive step, consider first the case where the last inference is a weakening or a $\wedge$:left whose premise is S'. then $\boldsymbol{\ell}(S')$ is a subsequence of $\boldsymbol{\ell}(S)$. Let $\boldsymbol{\delta}'$ be the subsequence of $\boldsymbol{\delta}$ corresponding to $\boldsymbol{\ell}(S')$, then by the induction hypothesis there is a proof of $S'(\boldsymbol{\delta}')$, and an inference of the same kind yields $S(\boldsymbol{\delta})$.

The case where the last inference is a negation inference or a $\wedge$:right is easy, since the types occurring in the premises are the same as those occurring in S. Hence the claim follows directly from the induction hypothesis by application of an inference of the same kind.

Now let the last inference be

$$\forall\text{:right} \qquad \frac{\Gamma \Longrightarrow \Delta\,,\ A[a^{\sigma}]}{\Gamma \Longrightarrow \Delta\,,\ \forall x^{\tau}\, A[x^{\tau}]}\,,$$

where $\tau = \gamma_i + k$ for some $k \in \omega$, and $\sigma = \eta + m$ for a limit η and $m \in \omega$. Let S' denote the premise of this inference. Then there are two possibilities: either $\eta = \gamma_j$ for some $j \geq i$, then $\boldsymbol{\ell}(S')$ is a subsequence of $\boldsymbol{\ell}(S)$, hence the argument works as in the case of a weakening above. Otherwise $\gamma_j < \eta < \gamma_{j+1}$ for some $j \geq i$, or $\eta \geq \gamma_n$. Then the condition (2) guarantees that there is η' with $\delta_j < \eta' < \delta_{j+1}$ or $\eta' > \delta_n$ respectively. Let $\boldsymbol{\delta}'$ denote the sequence $\boldsymbol{\delta}$ with η' inserted, and with δ_i omitted if γ_i does not occur in S'. Then by the induction hypothesis, there is a proof of $S'(\boldsymbol{\delta}')$, and a $\forall$:right yields $S(\boldsymbol{\delta})$.

Essentially the same argument applies when the last inference is a $\forall$:left, the difference is that in place of η there can be several limits $\eta_1, \ldots, \eta_r$ that occur in the term t^{σ} in the premise, but not in S. Then again condition (2) guarantees that there are suitable limits $\eta'_1, \ldots, \eta'_r$ to supplement the sequence $\boldsymbol{\delta}$. This argument also applies to the case where the last inference is a cut, where $\eta_1, \ldots, \eta_r$ are those limits occurring in the cut formula but not in S.

For the abstraction and extensionality inferences, the limits occurring in the premises again form a subsequence of $\boldsymbol{\ell}(S)$, hence the argument for a weakening applies. It remains to treat the case where the last inference is an application of Lim.

Let the conclusion of the last inference be $S[a^{\gamma}_i]$, and denote $S(\boldsymbol{\delta})$ by $\tilde{S}[a^{\delta_i}]$. Then we have to show that $\tilde{S}[a^{\xi}]$ is provable for every $\xi < \delta_i$.

Now let $\xi = \eta + k$ for a limit η and $k \in \omega$. Then if η is one of the δ_j for $j < i$, and thus $S[a^{\gamma_j + k}]$ is among the original premises, then the induction hypothesis yields the provability of $\tilde{S}[a^{\xi}]$.

Otherwise we have $\eta < \delta_0$ or $\delta_j < \eta < \delta_{j+1}$ for some $j < i$, and by condition (3) there is $\eta' < \gamma_1$ or $\gamma_j < \eta' < \gamma_{j+1}$ such that $S[a^{\eta'+k}]$ is among the original premises. Again the induction hypothesis yields the provability of $\tilde{S}[a^{\xi}]$. $\square$

Since for each set-theoretic sentence φ, the only limit occurring in $\varphi^{(\gamma)}$ is γ, this yields immediately:

Corollary 25. *If $\varphi^{(\gamma)}$ is provable in K_α^∞ without use of the null rule, then φ is persistent, and hence consistent with ZF.*

The criterion given by this corollary is only sufficient, but not necessary for the persistence of set-theoretic sentences. This can be seen by the following counterexample: Let Reg denote the axiom of regularity, which is clearly persistent.

Theorem 26. *$Reg^{(\gamma)}$ is not provable in K_α^∞ without the null rule.*

Proof. Work in set theory without regularity, and let Ω be a set with $\Omega = \{\Omega\}$. Let $H_0 := \{\emptyset, \Omega\}$, for ξ with $\xi + 1 < \alpha$ let $H_{\xi+1} := 2^{H_\xi}$, for limit γ let $H_\gamma := \bigcup_{\xi<\gamma} H_\xi$.

We show that $\mathbb{H}_\alpha := \big(H_\xi\big)_{\xi<\alpha}$ is a model of K_α^e plus the limit rule. Since H_0 is transitive, we have that for $\xi < \eta$, $H_\xi \subseteq H_\eta$. Clearly, $\mathbb{H}_\alpha$ is an extensional cumulative Henkin structure, and by the defining clause for H_γ, γ limit, it is also normal. Hence by Prop. 12 $\mathbb{H}_\alpha$ is a model of K_α^e plus the limit rule.

On the other hand, $H_\gamma \models \neg Reg$, hence $\mathbb{H}_\alpha \models \neg Reg^{(\gamma)}$. $\square$

The proofs in [12] show that in fact the regularity axiom is the only axiom of Z that requires the null rule to prove its γ-translation.

References

[1] Peter B. Andrews. *A Transfinite Type Theory with Type Variables.* North Holland, Amsterdam, 1965.

[2] Kenneth A. Bowen. Cut elimination in transfinite type theory. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 19:141–162, 1973.

[3] Kenneth A. Bowen. Systems of transfinite type theory based on intuitionistic and modal logics. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 20:355–372, 1974.

[4] Enrique Bustamante. *Transfinite Type Theory.* PhD thesis, Princeton, 1944.

[5] J. Wolfgang Degen. *Systeme der kumulativen Logik.* Philosophia Verlag, München, 1983.

[6] J. Wolfgang Degen. Two formal vindications of logicism. In Johannes Czermak, editor, *Philosophie der Mathematik*, volume 20 of *Schriftenreihe der Wittgenstein-Gesellschaft*, pages 243–250. Hölder-Pichler-Tempsky, Wien, 1993.

[7] Frank R. Drake. *Set Theory. An Introduction to Large Cardinals.* North Holland, Amsterdam, 1974.

[8] Gerhard Gentzen. Die Widerspruchsfreiheit der Stufenlogik. *Mathematische Zeitschrift*, 41:357–366, 1936.

[9] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.

[10] Kurt Gödel. Über die Länge von Beweisen. *Ergebnisse eines Mathematischen Kolloquiums*, 7:23–24, 1936.

[11] Leon Henkin. Completeness in the theory of types. *Journal of Symbolic Logic*, 15:81–91, 1950.

[12] Jan Johannsen. Kumulative Typenlogik mit einer infinitären Schlußregel. Student project, University of Erlangen, 1990.

[13] Jan Johannsen. Semantische Vollständigkeit halbformaler Systeme der transfiniten Typentheorie. Diploma thesis, University of Erlangen, 1991.

[14] John Kemeny. *Type Theory and Set Theory.* PhD thesis, Princeton, 1949.

[15] Maurice L'Abbé. Systems of transfinite type theory involving λ-conversion. *Journal of Symbolic Logic*, 18:209–224, 1953.

[16] E. G. K. López-Escobar. Infinite rules in finite systems. In A.I. Arruda, N.C.A. da Costa, and R. Chuaqui, editors, *Non-Classical Logics, Model Theory and Computability*, pages 75–97. North-Holland, Amsterdam, 1977.

[17] Dag Prawitz. Hauptsatz for higher-order logic. *Journal of Symbolic Logic*, 33:452–457, 1968.

[18] Frank P. Ramsey. The foundations of mathematics. *Proceedings of the London Mathematical Society, Ser. 2*, 25:338–384, 1925.

[19] Bertrand Russell. *The Principles of Mathematics.* Cambridge University Press, Cambridge, 1903.

[20] Bertrand Russell. Mathematical logic as based on the theory of types. *American Journal of Mathematics*, 30:222–262, 1908.

[21] Kurt Schütte. Syntactical and semantical properties of simple type theory. *Journal of Symbolic Logic*, 25:305–326, 1960.

[22] M. Takahashi. A proof of cut-elimination in simple type theory. *Journal of the Mathematical Society of Japan*, 19:399–410, 1967.

[23] Gaisi Takeuti. On a generalized logic calculus. *Japanese Journal of Mathematics*, 23:39–96, 1953.

[24] Gaisi Takeuti. *Proof Theory.* North-Holland, Amsterdam, 2nd edition, 1987.

[25] Alfred Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica, Lemberg*, 1:261–405, 1936.

[26] Alfred N. Whitehead and Bertrand Russell. *Principia Mathematica.* Cambridge University Press, Cambridge, 2nd edition, 1925-27.

[27] Ernst Zermelo. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen*, 65:261–281, 1908.

A Appendix: Reduction Systems

In the first author's book [5], the systems K^e_α were extended to the so-called reduction systems K^*_α, and it was shown that Zermelo set theory could be embedded into these: If $\tau \geq \gamma + 3$ for some limit ordinal $\gamma \geq \omega$, then $K^*_\alpha \vdash \varphi^{(\gamma)}$ for every theorem φ of Z. Unfortunately, it turned out later that the systems K^*_α are inconsistent.

In this appendix, after recalling the definition of K^*_α, we give a proof of this inconsistency. We then give a reformulation K^{**}_α of these systems, and we show that Zermelo set theory can be embedded into K^{**}_α in basically the same manner as shown in [5] for K^*_α. Finally, we show that K^{**}_α is a finitary subsystem of K^∞_α, showing that these reformulated systems are indeed consistent.

A.1 The Inconsistency of K^*_α

If $\sigma < \tau$, then we say that a^τ is *σ-small*, written $\downarrow_\sigma a^\tau$, if $\exists x^\sigma\, a^\tau = x^\sigma$. In particular, $a^{\tau+1}$ is *small*, written $\downarrow a^{\tau+1}$, if $\downarrow_\tau a^{\tau+1}$. Finally we define $a^{\tau+1} \prec b^{\tau+1}$ by $\exists x^\tau\, a^{\tau+1} = x^\tau \wedge b^{\tau+1}(x^\tau)$, and say that $a^{\tau+1}$ is *smaller than* $b^{\tau+1}$,

The relation $\prec$ is just another kind of type-homogeneous membership relation for successor types, and it is indeed equivalent to the usual one:

Proposition 27. $K^e_{\tau+4} \vdash a^{\tau+1} \prec b^{\tau+1} \leftrightarrow a^{\tau+1} \in b^{\tau+1}$

Proof. If $a^{\tau+1} \prec b^{\tau+1}$, then there is a $c^\tau = a^{\tau+1}$ with $b^{\tau+1}(c^\tau)$. Then let $d^{\tau+2} = b^{\tau+1}$, so we have $d^{\tau+2}(c^\tau)$ and thus $d^{\tau+2}(a^{\tau+1})$. But this means that $a^{\tau+1} \in b^{\tau+1}$.
If $a^{\tau+1} \in b^{\tau+1}$, by Prop. 5 there is $c^\tau = a^{\tau+1}$, and we have $c^\tau \in b^{\tau+1}$ which is equivalent to $b^{\tau+1}(c^\tau)$. This yields $a^{\tau+1} \prec b^{\tau+1}$. □

The central notion of the reduction systems is that of a *quasiuniverse* QU, which is defined by

$$QU(a^{\tau+1}) :\leftrightarrow \downarrow a^{\tau+1} \wedge \forall x^{\tau}\ (a^{\tau+1}(x^{\tau}) \rightarrow \exists y^{\tau+1} \subseteq a^{\tau+1}\ \ y^{\tau+1} = x^{\tau}) \wedge$$
$$\wedge\, \forall y^{\tau+1} \prec a^{\tau+1}\ \ \forall x^{\tau+1} \subseteq y^{\tau+1}\ \ x^{\tau+1} \prec a^{\tau+1}$$

Hence a quasiuniverse is a set of type $\tau + 1$ that is small, transitive and closed under subsets, i.e., it contains all subsets of its elements.

Now the system K^*_α is defined to be K^e_α together with the *separation rule*

$$\frac{t^{\tau+1}(a^{\tau}), \Gamma \implies \Delta, s^{\sigma+1}(a^{\tau})}{\downarrow_\eta s^{\sigma+1}, \Gamma \implies \Delta, \downarrow_\eta t^{\tau+1}}$$

where $\gamma \leq \tau \leq \sigma$ and the free variable a^{τ} does not occur in the conclusion, and the following two axioms on quasiuniverses for every limit γ with $\alpha \geq \gamma + 3$:

$$(I) \qquad \implies \exists x^{\gamma+1}\ \ QU(x^{\gamma+1}) \wedge x^{\gamma+1}(a^{\gamma})$$

$$(II) \qquad QU(a^{\gamma+1}), b^{\gamma+1} \prec a^{\gamma+1} \implies \exists x^{\gamma+1} \prec a^{\gamma+1}\ \ QU(x^{\gamma+1}) \wedge b^{\gamma+1} \subseteq x^{\gamma+1}$$

The separation rule immediately yields the following property:

$$a^{\gamma+1} \subseteq b^{\gamma+1} \wedge \downarrow b^{\gamma+1} \rightarrow \downarrow a^{\gamma+1}$$

Theorem 28. *For $\alpha \geq \omega + 3$, K^*_α is inconsistent.*

Proof. Let γ be a limit ordinal such that $\alpha \geq \gamma + 3$. As $\emptyset^{\gamma+1} \subseteq u^{\gamma+1}$ for every quasiuniverse $u^{\gamma+1}$, $\emptyset^{\gamma+1}$ is small, hence by axiom (I) there is a quasiuniverse $u_1^{\gamma+1}$ with $\emptyset^{\gamma+1} \prec u_1^{\gamma+1}$. Let $\{\emptyset\}^{\gamma+1} := \big(\lambda x^{\gamma}.\, x^{\gamma} = \emptyset^{\gamma+1}\big)$, then we have $\{\emptyset\}^{\gamma+1} \subseteq u_1^{\gamma+1}$, hence $\{\emptyset\}^{\gamma+1}$ is small and so there is a quasiuniverse $u_2^{\gamma+1}$ with $\{\emptyset\}^{\gamma+1} \prec u_2^{\gamma+1}$. Now $\{\{\emptyset\}\}^{\gamma+1} := \big(\lambda x^{\gamma}.\, x^{\gamma} = \{\emptyset\}^{\gamma+1}\big) \subseteq u_2^{\gamma+1}$, which again gives us that $\{\{\emptyset\}\}^{\gamma+1}$ is small, and hence $\{\{\emptyset\}\}^{\gamma+1} \prec u_3^{\gamma+1}$ for some quasiuniverse $u_3^{\gamma+1}$.

Let $s^{\gamma+1} := \big(\lambda x^{\gamma}.\, (x^{\gamma} = \emptyset^{\gamma+1} \vee x^{\gamma} = \{\emptyset\}^{\gamma+1} \vee x^{\gamma} = \{\{\emptyset\}\}^{\gamma+1})\big)$. Since $u_3^{\gamma+1}$, being a quasiuniverse, is transitive, we have $s^{\gamma+1} \subseteq u_3^{\gamma+1}$, and so $s^{\gamma+1}$ is small. Furthermore, it is easily seen and proved in K^*_α that $s^{\gamma+1}$ is transitive and closed under subsets, hence $s^{\gamma+1}$ is a quasiuniverse.

Now consider $\{\{\emptyset\}\}^{\gamma+1} \prec s^{\gamma+1}$. By axiom (II), there should be a quasiuniverse $t^{\gamma+1}$ with $t^{\gamma+1} \prec s^{\gamma+1}$ and $\{\{\emptyset\}\}^{\gamma+1} \subseteq t^{\gamma+1}$. But the only element $t^{\gamma+1}$ of $s^{\gamma+1}$ with $\{\{\emptyset\}\}^{\gamma+1} \subseteq t^{\gamma+1}$ is $\{\{\emptyset\}\}^{\gamma+1}$ itself, and it is easily seen and proved in K^*_α that $\{\{\emptyset\}\}^{\gamma+1}$ is not a quasiuniverse, as it is not transitive. Hence we get a contradiction to axiom (II). $\square$

A.2 A Remedy

As we saw, the main cause of the inconsistency of K^*_α is axiom (II): the requirement that every element of a quasiuniverse $u^{\gamma+1}$ is a subset of some quasiuniverse in $u^{\gamma+1}$. If we analyze the embedding of Zermelo set theory as done

in [5], we see that the only essential uses of axiom (II) are in the derivation of the translations of the pairing and regularity axioms.

Now, the use of (II) for the pairing axiom could be avoided if axiom (I) were strengthened to

$$(I') \qquad \Longrightarrow \exists x^{\gamma+1}\ QU(x^{\gamma+1}) \wedge x^{\gamma+1}(a^\gamma) \wedge x^{\gamma+1}(b^\gamma)$$

Furthermore, the nesting depth of applications of axiom (II) in the derivation of regularity is only two. Hence a remedy is to define the notion of a *super-quasiuniverse* SQU as a quasiuniverse $s^{\gamma+1}$ with the property that every element of $s^{\gamma+1}$ is a subset of some quasiuniverse in $s^{\gamma+1}$

$$\begin{aligned} SQU(a^{\gamma+1}) :\leftrightarrow\ & QU(a^{\gamma+1}) \wedge \\ & \wedge \forall x^{\gamma+1} \prec a^{\gamma+1}\ \exists y^{\gamma+1} \prec a^{\gamma+1}\ QU(y^{\gamma+1}) \wedge x^{\gamma+1} \subseteq y^{\gamma+1} \end{aligned}$$

and then iterating this once again to define the notion of an *ultra-quasiuniverse* UQU by

$$\begin{aligned} UQU(a^{\gamma+1}) :\leftrightarrow\ & SQU(a^{\gamma+1}) \wedge \\ & \wedge \forall x^{\gamma+1} \prec a^{\gamma+1}\ \exists y^{\gamma+1} \prec a^{\gamma+1}\ SQU(y^{\gamma+1}) \wedge x^{\gamma+1} \subseteq y^{\gamma+1} \end{aligned}$$

Now we define K_α^{**} as K_α^e extended by the separation rule and the following axiom on ultra-quasiuniverses

$$(I^*) \qquad \Longrightarrow \exists x^{\gamma+1}\ UQU(x^{\gamma+1}) \wedge x^{\gamma+1}(a^\gamma) \wedge x^{\gamma+1}(b^\gamma)$$

for every limit γ with $\alpha \geq \gamma + 3$. Then the γ-translations of the axioms of Z can be derived in K_α^{**} exactly as they are derived in K_α^* in [5], one only has to replace some occurrences of the notion of quasiuniverse by ultra- or super-quasiuniverse in such a way that uses of axiom (II) can be eliminated in favor of the respective definitions.

Finally, we show that K_α^{**} is a subsystem of K_α^∞. As K_α^∞ is consistent, having the standard model $\mathbb{V}_\alpha$, this implies that K_α^{**} is also consistent, hence we have succeeded in giving a remedied version of the reduction systems of [5].

Theorem 29. *Every theorem of K_α^{**} is provable in K_α^∞.*

Proof. It suffices to show that (a) the separation rule is an admissible inference rule in K_α^∞, and (b) the axiom (I^*) can be derived in K_α^∞. Part (a) is quite easy: from the premise $t^{\tau+1}(a^\tau), \Gamma \Longrightarrow \Delta, s^{\sigma+1}(a^\tau)$ we obtain $\Gamma \Longrightarrow \Delta, t^{\tau+1} \subseteq s^{\sigma+1}$. Then by use of Prop. 8 and some cuts the conclusion $\downarrow_\eta s^{\sigma+1}, \Gamma \Longrightarrow \Delta, \downarrow_\eta t^{\tau+1}$ is obtained.

Now for (b): For each type η, the universe $v^{\eta+1} := \big(\lambda x^\eta.\, x^\eta = x^\eta\big)$ is easily seen to be a quasiuniverse. Furthermore, every set $a^{\eta+1}$ is a subset of $v^{\eta+1}$, hence $v^{\eta+2}$, defined analogously, is a super-quasiuniverse. By the same argument, $v^{\eta+3}$ is an ultra-quasiuniverse. Now by abstraction, we obtain $v^{\eta+3}(a^\eta)$, hence existential quantification yields

$$\exists x^{\gamma+1}\, UQU(x^{\gamma+1}) \wedge x^{\gamma+1}(a^\eta) \wedge x^{\gamma+1}(b^\eta)$$

for every type $\eta < \gamma$, where γ is a limit with $\alpha \geq \gamma + 3$. Hence the binary limit rule, which is admissible by Prop. 9, yields axiom (I^*). $\square$