Universes in Type Theory Part II – Autonomous Mahlo

Anton Setzer*

20 December 2009

Abstract

We introduce the autonomous Mahlo universe which is an extension of Martin-Löf type theory which we consider as predicatively justified and which has a strength which goes substantially beyond that of the Mahlo universe, which is before writing this paper the strongest predicatively justified published extension of Martin-Löf type theory. We conjecture it to have the same proof theoretic strength as Kripke-Platek set theory extended by one recursively autonomous Mahlo ordinal and finitely many admissibles above it. Here a recursively autonomous Mahlo universe ordinal is an ordinal κ which is recursively hyper$^{\alpha}$-Mahlo for all $\alpha < \kappa$. We introduce as well as intermediate steps the hyper-Mahlo and hyper$^{\alpha}$-Mahlo universes, and give meaning explanations for these theories as well as for the super and the Mahlo universe. We introduce a model for the autonomous Mahlo universe, and determine an upper bound for its proof theoretic strength, therefore establishing one half of the conjecture mentioned before. The autonomous Mahlo universe is the crucial intermediate step for understanding the Π_3-reflecting universe, which will be published in a successor of this article and which is even stronger and will slightly exceed the strength of Kripke-Platek set theory plus the principle of Π_3-reflection.

1 Introduction

This article is a step in a research programme of the author with the goal of introducing proof theoretically as strong as possible extensions of Martin-Löf type theory, which still can be regarded as predicatively justified. We have three main reasons for following such a research programme:

(1) We hope that this approach gives more insights into the development of ordinal theoretic proof theory. Results in the area of proof theory of

*Department of Computer Science, Swansea University, Singleton Park, Swansea SA2 8PP, UK, Email: a.g.setzer@swan.ac.uk, http://www.cs.swan.ac.uk/∼csetzer/, Tel: +44 1792 513368, Fax: +44 1792 295651. Supported by EPSRC grant EP/G033374/1.

impredicative theories are often regarded as very difficult to understand by non-specialists. The theories developed in this programme use crucial ideas from proof theory, while – as we hope – being much easier understandable by a more general audience. We hope this allows more researchers to understand the insights gained by recent proof theoretic developments.

(2) This research can be seen as part of a revised Hilbert's programme. The goal of Hilbert's original programme was to prove the consistency of theories for formalising mathematical proofs using finitary methods. By Gödel's second incompleteness theorem we know that such a programme cannot be carried out, except for very weak theories. In ordinal theoretic proof theory the consistency of theories of increasing strength is reduced to the well-foundedness of ordinal notation systems. Using strong extensions of Martin-Löf type theory one is able to prove the well-foundedness of such ordinal notation systems and therefore the consistency of the original theories. This provides a reduction of the consistency of classical theories (mainly fragments of second order logic and fragments of set theory) to constructive theories, which have a philosophical argument for the validity of their theorems. Therefore extensions of Martin-Löf type theory can be regarded as a substitute for Hilbert's finitary methods.

(3) We hope that by developing such extensions we will discover new data structures. One example where this program has been successful was the discovery of the data type of inductive-recursive definitions ([16, 17, 18, 19]), which was strongly influenced by the Mahlo universe, a first step in this proof theoretic programme. Variants of this data type can be used in the area of generic programming ([12, 15]) and it has similarities with (and possibly influenced) the development of generic extensions of Haskell (polytypic programming, e.g. [25]).

In the article [53], we have introduced type theories of strength Kripke-Platek set theory extended by one recursively inaccessible, one recursively hyperinaccessible and one recursively Mahlo ordinal, respectively, and finitely many admissibles above those ordinals. The type theories considered were Martin-Löf type theory with W-type and one universe, one super universe and one Mahlo universe, respectively. We have as well given basic model constructions in the corresponding extensions of Kripke-Platek set theory, in order to obtain an upper bound for their proof theoretic strength. (For the reader not familiar with Kripke-Platek set theory it suffices to understand these constructions as model constructions in which one uses as little strength as possible from the set theory, in which the models are developed.) Lower bounds for the proof theoretic strength of these theories have been shown in [42, 46, 49].

In this article we introduce a new universe construction into Martin-Löf type theory of expected strength Kripke-Platek set theory plus one recursively autonomous Mahlo ordinal and finitely many admissibles above it. Here a recursively autonomous Mahlo ordinal is an ordinal κ which is recursively hyper$^\kappa$-

Mahlo, or, equivalently, which is recursively hyper$^\alpha$-Mahlo for all $\alpha < \kappa$. (See Subsect. 5.2 for details.)

The step from the Mahlo universe to the autonomous Mahlo universe is a very natural step, in which we move from the Mahlo universe to the hyper-Mahlo universe, then to the hyper$^\alpha$-Mahlo universe and then to an autonomous iteration of such universes. It is an important intermediate step before introducing the Π_3-reflecting universe, in which the iteration of Mahlo degrees (the Mahlo degree of a hyper$^\alpha$-Mahlo universe would be α) is carried out even further. We plan to introduce the Π_3-reflecting universe in a follow up paper.

The step from Mahlo to Π_3-reflection follows recent developments in proof theory. After an analysis of $(\Delta^1_2 - \mathrm{CA}) + (\mathrm{BI})$ or, equivalently, Kripke Platek set theory plus one recursively inaccessible had been carried out, the next step in the development of proof theory was an analysis of Kripke-Platek set theory plus one Mahlo ordinal, carried out independently by Rathjen and Arai (see [33, 34, 36] and [3, 4, 8, 9]). Then both researchers analysed Kripke-Platek set theory plus one Π_3-reflecting ordinal ([35, 37] and [3, 4, 7, 10]), which was the crucial stepping stone before analysing theories of strength $(\Pi^1_2 - \mathrm{CA}) + (\mathrm{BI})$ and beyond [38, 39, 40] and [3, 5, 6].

Relationship to the article [53]. This article is a follow-up article of the article [53], in which we introduced type theories with one universe, one super universe, and one Mahlo universe. There, we also gave the basic model construction and a proof of its correctness. However, since the author hasn't published yet an article laying out the basic infrastructure for forming models of this kind, details about the basic setting (how to form a model for a type theory in Kripke-Platek set theory in general including terms and their reduction rules, and how to interpret the basic set constructions) have not been carried out yet. (They have been carried out in [42] for a type theory with one universe, however that setting, while based on the same principles, was not generic enough to be directly extensible.) The main goal of this article is to introduce the autonomous Mahlo universe and give a model construction with the same level of detail. Full details of the model will be introduced in a follow up article. This article might in fact be more accessible to a general audience than the more detailed model construction to be introduced later.

The model will be carried out in Kripke-Platek set theory plus one recursively autonomous Mahlo ordinal and finitely many admissibles above it. Therefore we obtain an upper bound for the proof-theoretic strength of this universe construction. In order to obtain a lower bound for its strength, we plan to extend our approach to well-ordering proofs for ordinal notation systems based on ordinal systems ([47, 50, 52]) to Kripke-Platek set theory extended by a recursively autonomous Mahlo set.

Content. The structure of this article is as follows: In Sect. 2 we repeat the basic notations used in this article. In Sect. 3 we briefly revisit the super universe and the Mahlo universe, and develop meaning explanations for these theories. We will discuss as well the relationship between recursive and inductive subuniverses in this setting. In Sect. 4 we introduce the steps towards the autonomous Mahlo universe as follows: After a short motivation we introduce in Subsect. 4.1 the hyper-Mahlo universe, in Subsect. 4.2 the hyper-n-Mahlo and the hyper$^\alpha$-Mahlo universes, and in Subsect. 4.3 the autonomous Mahlo universe itself. For each of these theories we will develop meaning explanations. We regard these meaning explanation as a predicative justification for the theories introduced. Finally, in Sect. 5 we introduce a model for the autonomous Mahlo universe and give the main proof ideas of its correctness. This way we determine an upper bound for its proof-theoretic strength.

Related work. G. Jäger and T. Strahm have introduced in [24] a Π_3-reflecting universe in the context of Feferman's systems of explicit mathematics, which doesn't require to define an autonomous Mahlo universe first. Because they rely on partial application, their setting is much simpler, but because of its heavy reliance on partial application it is not possible to mimic this in type theory. Their theory seems as well to introduce elements which have no canonical counterpart, so it is not as explicit as demanded in type theory. We will discuss details about this in the follow up paper on the Π_3-reflecting universe. Jäger and Strahm carried out as well a full proof theoretic analysis. Since they are working in a Meta-predicative setting (translated to type theory this means that the W-type is omitted), the proof theoretic ordinal they obtained is well below the Bachmann-Howard ordinal, and well below the strength of type theory with the W-type and no universes except for atom.

This work is heavily inspired by M. Rathjen's and T. Arai's proof theoretic analysis of Π_3-reflection ([35, 37] and [3, 4, 7, 10]) – the Π_3-reflecting universe will mimic the ordinal notation systems of that strength. Without their work it would have been difficult to discover the Π_3-reflecting universe.

Acknowledgements. This research benefited very much from intensive discussions with both T. Arai and M. Rathjen, whose research we deeply admire. It was as well inspired by intensive discussions with Peter Dybjer on inductive-recursive definitions. We would like as well to thank the anonymous referee of a previous version of this article for valuable comments, and my PhD student Frederik Forsberg for careful proof reading.

2 Notations

We will frequently make use of notations and some basic lemmata introduced in [53], but will repeat the most important ones briefly, so that the reader, who is not interested in all details, can read it without having to go through the article [53] first.

- We use mathematical style for application, so we write $f(a)$ for the application of f to a. We write $f(a, b)$ instead of $f(a)(b)$, similarly for longer sequences of applications.

- We will omit equality versions of rules. For instance when we write the rule for the Π-type

$$\frac{A : \mathrm{Set} \qquad x : A \Rightarrow B : \mathrm{Set}}{(\Pi x : A.B) : \mathrm{Set}}$$

we implicitly introduce as well the rule

$$\frac{A = A' : \mathrm{Set} \qquad x : A \Rightarrow B = B' : \mathrm{Set}}{(\Pi x : A.B) = (\Pi x : A'.B') : \mathrm{Set}}$$

The reason that we sometimes need such proper rules is that we for foundational reasons will use only the small logical framework (see below).

- When expressing equality rules, we will usually omit premises. For instance, a rule $\mathrm{T}(\widehat{\Pi}(a,b)) = \Pi x : \mathrm{T}(a).\mathrm{T}(b(x))$ needs to be read as

$$\frac{a : \mathrm{U} \qquad b : \mathrm{T}(a) \to \mathrm{U}}{\mathrm{T}(\widehat{\Pi}(a,b)) = \Pi x : \mathrm{T}(a).\mathrm{T}(b(x)) : \mathrm{Set}}$$

  where the premises are what is needed in order to derive $\mathrm{T}(\widehat{\Pi}(a,b)) : \mathrm{Set}$ and $\Pi x : \mathrm{T}(a).\mathrm{T}(b(x)) : \mathrm{Set}$.

- The **basic set constructions** are the rules for forming the finite sets N_n, the set of natural numbers N, the disjoint union $A + B$, the Π-sets $\Pi x : A.B$, the Sigma set $\Sigma x : A.B$, the W-sets $\mathrm{W} x : A.B$, and the intensional equality $\mathrm{Id}(A, a, b)$. Codes for these constructions as elements of a universe will usually have the form $\widehat{\mathrm{N}}_i$, $\widehat{\mathrm{N}}$, $a \widehat{+} b$ etc.. More details can be found in Sect. 2.3 of [53].

- The elements of N_n are denoted by A^n_i $(i = 0, \ldots, n-1)$. We write $\bot$ for N_0 (the empty set, which represents the false proposition), $\top$ for N_1 (which represents the true proposition), triv for A^1_0 (the *triv*ial proof of $\top$), Bool for N_2, ff for A^2_0, tt for A^2_1 (and when using universes write sometimes $\widehat{\bot}, \widehat{\top}, \widehat{\mathrm{Bool}}$ for $\widehat{\mathrm{N}}_0$, $\widehat{\mathrm{N}}_1$, $\widehat{\mathrm{N}}_2$, respectively).

- The rules for **atom** consist of the rules $x : \mathrm{N}_2 \Rightarrow \mathrm{atom}(x) : \mathrm{Set}$, $\mathrm{atom}(\mathrm{tt}) = \bot$, $\mathrm{atom}(\mathrm{ff}) = \top$. Therefore, atom translates a Boolean value into an *atomic* formula which is provable iff the Boolean value is true, therefore the name atom.

- The **small logical framework** consists of the dependent function type $(x : A) \to B$ and the dependent product $(x : A) \times B$, restricted to Set. λ-abstraction is denoted by $(x : A)b$ or shorter $(x)b$, and application by $f(a)$. The canonical elements of $(x : A) \times B$ are denoted by $\langle a, b \rangle$ and the projections of these elements are denoted by $\pi_0(c)$, $\pi_1(c)$. See Sect. 2.1 of [53] for details.

- The **full logical framework**, which is not part of any of the type theories in this article unless stated explicitly, adds a type level Type on top of set.

We have Set : Type and Type contains all elements of Set. Furthermore both Set and Type are closed under the dependent function type and dependent product. See Sect. 2.2 of [53] for details.

- The **type of families of sets** is defined (using the full logical framework) as $\mathrm{Fam}(\mathrm{Set}) := (X : \mathrm{Set}) \times (X \to \mathrm{Set})$. (See Def. 4.1 (b) of [53].) If we don't have the logical framework, then we write

  $$(U, T) : \mathrm{Fam}(\mathrm{Set})$$

  for the following rules:

  $$U : \mathrm{Set} \qquad \frac{a : U}{T(a) : \mathrm{Set}}$$

- An **operator on families of sets** is a function $h : \mathrm{Oper}(\mathrm{Set})$, where $\mathrm{Oper}(\mathrm{Set}) := \mathrm{Fam}(\mathrm{Set}) \to \mathrm{Fam}(\mathrm{Set})$. (See Def. 5.1 (a) of [53].) By uncurrying it and splitting it into two functions, such an h can be replaced by two functions

  $$\begin{array}{lll} f & : & (X : \mathrm{Set}, Y : X \to \mathrm{Set}) \to \mathrm{Set} \ , \\ g & : & (X : \mathrm{Set}, Y : X \to \mathrm{Set}, Z : f(X, Y)) \to \mathrm{Set} \ . \end{array}$$

  We write

  $$f, g : \mathrm{Oper}(\mathrm{Set})$$

  for f, g having the types just mentioned. We sometimes abbreviate f, g by $\vec{h}$ and write then

  $$\vec{h} = f, g : \mathrm{Oper}(\mathrm{Set}) \ .$$

- Assume $(U, T) : \mathrm{Fam}(\mathrm{Set})$. The **set of internal families of sets in** (U, T) is defined as $\mathrm{IFam}(U, T) := (x : U) \times (T(x) \to U)$. (See Def. 4.1 (b) of [53].)

- Assume $(U, T) : \mathrm{Fam}(\mathrm{Set})$. An **operator on families of sets internal to** (U, T) is a function $h : \mathrm{IOper}(U, T)$, where $\mathrm{IOper}(U, T) := \mathrm{IFam}(U, T) \to \mathrm{IFam}(U, T)$. (See Def. 5.1 (a) of [53].) Such functions can be split into two components and uncurried, then we obtain functions

  $$\begin{array}{lll} f & : & (a : U, b : T(a) \to U) \to U \ , \\ g & : & (a : U, b : T(a) \to U, c : T(f(a, b))) \to U \ . \end{array}$$

  We write

  $$f, g : \mathrm{IOper}(U, T)$$

  for f, g having the type just mentioned. We sometimes abbreviate f, g by $\vec{h}$ and write then

  $$\vec{h} = f, g : \mathrm{IOper}(U, T) \ .$$

- A **universe** is a pair (U, T), where $(\mathrm{U}, \mathrm{T}) : \mathrm{Fam}(\mathrm{Set})$, such that U is closed under the basic set constructions. For instance, closure under Π means that for $a : \mathrm{U}$, $b : \mathrm{T}(a) \to \mathrm{U}$ we have $\widehat{\Pi}(a, b) : \mathrm{U}$, and $\mathrm{T}(\widehat{\Pi}(a, b)) = \Pi x : \mathrm{T}(a).\mathrm{T}(b(x))$. For each universe different names for the constructor will be chosen. We will indicate the names if they are used later, otherwise leave those names implicit.

  So the rules for a universe are as follows:

  rules expressing $(\mathrm{U}, \mathrm{T}) : \mathrm{Fam}(\mathrm{Set})$,
  $\widehat{\Pi} : (a : \mathrm{U}, b : \mathrm{T}(a) \to \mathrm{U}) \to \mathrm{U}$,
  $\mathrm{T}(\widehat{\Pi}(a, b)) = \Pi x : \mathrm{T}(a).\mathrm{T}(b(x))$,
  similarly for the other basic set constructions.

- When defining nested universes, we consider both **recursive subuniverses** and **inductive subuniverses**.

  - A **recursive subuniverse** of a universe $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ is given by a set U together with a recursively defined function $\widehat{\mathrm{T}}_{\mathrm{U}} : \mathrm{U} \to \mathrm{V}$, which determines for every $a : \mathrm{U}$ the element in V it corresponds to. Then one defines $\mathrm{T}_{\mathrm{U}}(a) := \mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{T}}_{\mathrm{U}}(a))$. So for every element $a : \mathrm{U}$ an element of V corresponding to it needs to have been defined using other rules. We have always closure of $(\mathrm{U}, \widehat{\mathrm{T}}_{\mathrm{U}})$ under the basic set constructions, with constructors named in a usual way e.g. $\widehat{\mathrm{N}}_{\mathrm{U}}$, $\widehat{\Pi}_{\mathrm{U}}$ etc. We will indicate the names of those constructors, if they are used in the following, and indicate the name of T_{U} by saying "$(\mathrm{U}, \widehat{\mathrm{T}}_{\mathrm{U}})$ is a recursive subuniverse of $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ with full decoding T_{U}."

    So the rules expressing that $(\mathrm{U}, \widehat{\mathrm{T}}_{\mathrm{U}})$ is a recursive subuniverse of $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ with full decoding T_{U} are as follows (assuming that the constructors for the basic set constructions in V are denoted by $\widehat{\mathrm{N}}_{\mathrm{V}}$, $\widehat{\Pi}_{\mathrm{V}}$ etc.)

    $\mathrm{U} : \mathrm{Set}$, $\qquad \widehat{\mathrm{T}}_{\mathrm{U}} : \mathrm{U} \to \mathrm{V}$,
    $$\frac{x : \mathrm{U}}{\mathrm{T}_{\mathrm{U}}(x) : \mathrm{Set}} \qquad \mathrm{T}_{\mathrm{U}}(x) = \mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{T}}_{\mathrm{U}}(x)) \ ,$$
    $\widehat{\Pi}_{\mathrm{U}} : (a : \mathrm{U}, b : \mathrm{T}_{\mathrm{U}}(a) \to \mathrm{U}) \to \mathrm{U}$,
    $\widehat{\mathrm{T}}_{\mathrm{U}}(\widehat{\Pi}_{\mathrm{U}}(a, b)) = \widehat{\Pi}_{\mathrm{V}}(\widehat{\mathrm{T}}_{\mathrm{U}}(a), \widehat{\mathrm{T}}_{\mathrm{U}} \circ b)$,
    rules for the other basic set constructions.

  - In an **inductive subuniverse**, U is defined together with recursively defined function $\mathrm{T}_{\mathrm{U}} : \mathrm{U} \to \mathrm{Set}$. Then we define the embedding of U into V, namely $\widehat{\mathrm{T}}_{\mathrm{U}} : \mathrm{U} \to \mathrm{V}$, as a constructor of V with equality rule $\mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{T}}_{\mathrm{U}}(x)) = \mathrm{T}_{\mathrm{U}}(x)$. So when we define an element a of U, we don't need to have an element in V corresponding to it which is defined using other rules. All we need is to have a way of defining $\mathrm{T}_{\mathrm{U}}(a)$.

So the rules for $(\mathrm{U}, \mathrm{T}_{\mathrm{U}})$ being an inductive subuniverse of $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ with embedding $\widehat{\mathrm{T}}_{\mathrm{U}}$ are as follows:

Rules expressing that $(\mathrm{U}, \mathrm{T}_{\mathrm{U}})$ is a universe,

$$\widehat{\mathrm{T}}_{\mathrm{U}} : \mathrm{U} \to \mathrm{V} \ , \qquad \mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{T}}_{\mathrm{U}}(x)) = \mathrm{T}_{\mathrm{U}}(x) \ .$$

A universe $(\mathrm{U}, \mathrm{T}_{\mathrm{U}})$ can be an inductive subuniverse of several universes, in which case the rules expressing that $(\mathrm{U}, \mathrm{T}_{\mathrm{U}})$ is a universe occur only once.

  - A more detailed discussion on inductive vs. recursive subuniverses can be found in paragraph "Type theories with several universes" in Subsect. 3.1 of [53]. Note that inductive subuniverses result in doubling of codes, for instance $\widehat{\mathrm{T}}_{\mathrm{U}}(\widehat{\mathrm{N}}_{\mathrm{U}})$ and $\widehat{\mathrm{N}}_{\mathrm{V}}$ are two different codes for the natural numbers, although without universe elimination rules we cannot prove in type theory that they are different. This is the reason why we regard recursive subuniverses as more aesthetic, as long as they don't create a substantial overhead.

- By "the inductive subuniverse $(\mathrm{U}, \mathrm{T}_{\mathrm{U}})$ of $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ is represented in V by $\widehat{\mathrm{U}}$" we mean the following rules

$$\widehat{\mathrm{U}} : \mathrm{V} \ , \qquad \mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{U}}) = \mathrm{U} \ .$$

3 The Super Universe and the Mahlo Universe Revisited

In this section we repeat briefly the super universe and the Mahlo universe, which were discussed in more detail in [53].

3.1 Super Universe

The super universe was introduced by Palmgren (see e.g. [30], Sect. 3.) In our setting, a super universe $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ is a universe which contains additionally for every $\langle a, b\rangle : \mathrm{IFam}(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$, a universe containing $\langle a, b\rangle$. This means that for every $\langle a, b\rangle : \mathrm{IFam}(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$, there exists a recursive subuniverse $(\mathrm{U}_{a,b}, \widehat{\mathrm{T}}_{\mathrm{U},a,b})$ with full decoding $\mathrm{T}_{\mathrm{U},a,b}$, and a code $\widehat{\mathrm{U}}_{a,b} : \mathrm{V}$ for it. $\mathrm{U}_{a,b}$ is closed under a, b, which means that there are codes $\widehat{\mathrm{a}}_{a,b} : \mathrm{U}_{a,b}$ for a, and $\widehat{\mathrm{b}}_{a,b} : \mathrm{T}_{\mathrm{V}}(a) \to \mathrm{U}_{a,b}$ for b.

Definition 3.1 *The rules for the super universe are as follows:*

- *Rules for the basic set constructions and the small logical framework.*
- *Rules expressing that* $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ *is universe.*

Assume in the following $a : \mathrm{V}$, $b : \mathrm{T_V}(a) \to \mathrm{V}$. *Then we have the following rules:*

- *Rules expressing that* $(\mathrm{U}_{a,b}, \widehat{\mathrm{T}}_{\mathrm{U},a,b})$ *is a recursive subuniverse of* V *with full decoding* $\mathrm{T}_{\mathrm{U},a,b}$.
- *Closure of* $(\mathrm{U}_{a,b}, \mathrm{T}_{\mathrm{U},a,b})$ *under* a, b *as expressed by the following rules:*

$$\begin{array}{ll} \widehat{\mathrm{a}}_{a,b} : \mathrm{U}_{a,b} \ , & \widehat{\mathrm{T}}_{\mathrm{U},a,b}(\widehat{\mathrm{a}}_{a,b}) = a \ , \\ \widehat{\mathrm{b}}_{a,b} : \mathrm{T_V}(a) \to \mathrm{U}_{a,b} \ , & \widehat{\mathrm{T}}_{\mathrm{U},a,b}(\widehat{\mathrm{b}}_{a,b}(x)) = b(x) \ . \end{array}$$

- $\mathrm{U}_{a,b}$ *is represented in* V*:*

$$\widehat{\mathrm{U}}_{a,b} : \mathrm{V} \ , \qquad \mathrm{T}(\widehat{\mathrm{U}}_{a,b}) = \mathrm{U}_{a,b} \ .$$

See Subsect. 4.1 of [53] for details.

As I learned from Peter Hancock, who has worked a lot on families of sets (see e.g. [20]), the rules for $\widehat{\mathrm{T}}_{\mathrm{U},a,b}$, $\widehat{\mathrm{U}}_{a,b}$ express that the successor of $(\mathrm{U}_{a,b}, \mathrm{T}_{\mathrm{U},a,b})$ is a subfamily of $(\mathrm{V}, \mathrm{T_V})$, where the successor of a family of sets (A, B) is the family of sets $(A + \top, B')$ where $B'(\mathrm{inl}(a)) = B(a)$, $B'(\mathrm{inr}(\mathrm{triv})) = A$. Remember that $\top$ is another name for N_1.

We note the differences with **Palmgren's original formulation** in [30], Sect. 3, p. 194 - 195. He introduces first using the full logical framework a universe operator $\mathrm{U}(A, B)$, which forms for arbitrary $\langle A, B\rangle : \mathrm{Fam}(\mathrm{Set})$ a universe $\mathrm{U}(A, B)$ (so it is closed under the basic set constructions) containing codes for A and B, together with decoding $\mathrm{T}(A, B, x) : \mathrm{Set}$ for $x : \mathrm{U}(A, B)$. Then the superuniverse is a universe (V, S) containing for every $\langle a, b\rangle : \mathrm{IFam}(\mathrm{V}, \mathrm{S})$ with $A := \mathrm{S}(a)$ and $B(x) := \mathrm{S}(b(x))$ an inductive subuniverse $\mathrm{u}(a, b)$ representing $\mathrm{U}(A, B)$. So we have $\mathrm{u}(a, b) : \mathrm{V}$, $\mathrm{S}(\mathrm{u}(a, b)) = \mathrm{U}(A, B)$, and for $x : \mathrm{T}(a)$ we have $\mathrm{t}(a, b, x) : \mathrm{V}$, $\mathrm{S}(\mathrm{t}(a, b, x)) = \mathrm{T}(A, B, x)$. So u, t are additional constructors of V. In this article we want to avoid the full logical framework. Therefore, we have restricted the super universe operator to elements of $\mathrm{IFam}(\mathrm{V}, \mathrm{T})$. The reason for using a recursive subuniverse rather than an inductive one is essentially a matter of taste, we prefer to use recursive subuniverses as long as this does not lead to any complications.

We can say that without introducing $\mathrm{U}(A, B)$ we get an external super universe, similar to the external Mahlo universe introduced in the next subsection: Here Set plays the rôle of the super universe, since Set is closed under the universe operator. When introducing (V, S), Palmgren makes the step towards an internal super universe, i.e. of forming a universe which reflects the closure of Set under the universe operator.

Relationship to inductive-recursive definitions. $\mathrm{U}(A, B)$ is an instance of an inductive-recursive definition [14, 15, 16, 17, 18, 19] with parameters A, B. Then one can see that Palmgren's super universe is as well an instance of an inductive-recursive definition. The formulation of the super universe in Def. 3.1 is not directly an instance of an inductive-recursive definition. Note that the formulation of inductive-recursive definitions makes heavily use of the full logical framework, so it is no surprise that a definition which avoids it is no longer an instance of that particular schema. However we can still observe an inductive-recursive nature in it (and might consider it as an instance of a suitable extension of inductive-recursive definitions): V, $\mathrm{U}_{a,b}$ are defined simultaneously inductively, while T_{V}, $\widehat{\mathrm{T}}_{\mathrm{U},a,b}$ are defined recursively. The constructors for the inductive definitions are the constructors for the codes for the basic set constructions, $\widehat{\mathrm{a}}$, $\widehat{\mathrm{b}}$ and $\widehat{\mathrm{U}}$.

Meaning explanations for the super universe. We will give meaning explanations for our formulation of the super universe in order to justify its constructive validity. We are giving them on request of the referee of a previous version of this article. We are giving them quite reluctantly, since we lack a philosophical education in order to have a proper understanding of meaning theory. We hope that researchers with a proper background in this area will be able to transform them into proper meaning explanations. The basic concept of meaning explanations and meaning explanations for the basic set constructions are as in Martin-Löf's article and book [27, 26].

So it suffices to give meaning explanations for V and for $\mathrm{U}_{a,b}$ for each element a of V and each function b mapping elements of $\mathrm{T}_{\mathrm{V}}(a)$ to elements of V. These meaning explanations are given simultaneously. Whenever we introduce a canonical element c of V, we determine the set $\mathrm{T}_{\mathrm{V}}(c)$, and for each canonical element d of $\mathrm{U}_{a,b}$ we determine the element $\widehat{\mathrm{T}}_{\mathrm{U},a,b}(d)$ of V. This is done in such a way that if c, c' are equal canonical elements of V, we have that $\mathrm{T}_{\mathrm{V}}(c)$ and $\mathrm{T}_{\mathrm{V}}(c')$ are equal sets, and if c, c' are equal canonical elements of $\mathrm{U}_{a,b}$, then $\widehat{\mathrm{T}}_{\mathrm{U},a,b}(c)$ and $\widehat{\mathrm{T}}_{\mathrm{U},a,b}(c')$ are equal elements of V. A non-canonical element c of V is equal to a canonical element c' of V' and then we define $\mathrm{T}_{\mathrm{V}}(c)$ as $\mathrm{T}_{\mathrm{V}}(c')$. The same applies to $\widehat{\mathrm{T}}_{\mathrm{U},a,b}$ and all other recursive functions defined in this article.

The canonical elements of V are those given by the fact that V is closed under the basic set constructions, and closure under $\widehat{\mathrm{U}}$. Regarding closure under the basic set constructions, we only treat the case of Π: If c is an element of V and d a function mapping elements of $\mathrm{T}_{\mathrm{V}}(c)$ to V, then $\widehat{\Pi}_{\mathrm{V}}(c, d)$ is a canonical element of V, and $\mathrm{T}_{\mathrm{V}}(\widehat{\Pi}_{\mathrm{V}}(c, d))$ is the set $\Pi x : \mathrm{T}_{\mathrm{V}}(c).\mathrm{T}_{\mathrm{V}}(d(x))$. $\widehat{\Pi}_{\mathrm{V}}(c, d)$ and $\widehat{\Pi}_{\mathrm{V}}(c', d')$ are equal canonical elements of V, if c, c' are equal elements of V and for x in $\mathrm{T}_{\mathrm{V}}(c)$ $d(x)$ and $d'(x)$ are equal elements of V. One observes now that if $\widehat{\Pi}_{\mathrm{V}}(c, d)$ and $\widehat{\Pi}_{\mathrm{V}}(c', d')$ are equal canonical elements, then $\mathrm{T}_{\mathrm{V}}(\widehat{\Pi}_{\mathrm{V}}(c, d))$ and $\mathrm{T}_{\mathrm{V}}(\widehat{\Pi}_{\mathrm{V}}(c', d'))$ are equal sets: we assume that $\mathrm{T}_{\mathrm{V}}(c)$ and $\mathrm{T}_{\mathrm{V}}(c')$ are equal sets, and for x in $\mathrm{T}_{\mathrm{V}}(c)$ that $\mathrm{T}_{\mathrm{V}}(d(x))$ and $\mathrm{T}_{\mathrm{V}}(d'(x))$ are equal sets. Then $\mathrm{T}_{\mathrm{V}}(\widehat{\Pi}(c, d))$ and $\mathrm{T}_{\mathrm{V}}(\widehat{\Pi}(c', d'))$ are equal sets as well. By saying in future meaning explanations that a universe $\mathrm{V}, \mathrm{T}_{\mathrm{V}}$ is closed under the basic set constructions we mean the explanations given before (using appropriate names for the constructors).

The definition of when two canonical elements are equal and an argument why this implies that recursively defined functions applied to them give the same result have to be carried out for all future canonical elements, a task we usually leave to the reader.

Closure of V under $\widehat{\mathrm{U}}$ is given as follows: If a is an element of V, and b is a function mapping elements of $\mathrm{T_V}(a)$ to V, then $\widehat{\mathrm{U}}_{a,b}$ is a canonical element of V, and we define $\mathrm{T_V}(\widehat{\mathrm{U}}_{a,b}) = \mathrm{U}_{a,b}$, where $\mathrm{U}_{a,b}$ will be explained later, where we will only use the fact that V contains a and $b(x)$ for x in $\mathrm{T_V}(a)$, and that V is closed under the basic set constructions (so we don't use closure of V under $\widehat{\mathrm{U}}$). This concludes the meaning explanations for V.

For elements a of V and functions c mapping elements of $\mathrm{T_V}(a)$ to V we explain the meaning of $\mathrm{U}_{a,b}$. Since $\mathrm{U}_{a,b}$ is a recursive subuniverse of V, we define $\mathrm{T}_{\mathrm{U},a,b}(c)$ as $\mathrm{T_V}(\widehat{\mathrm{T}}_{\mathrm{U},a,b}(c))$.

$\mathrm{U}_{a,b}$ is closed under the basic set constructions, and again we explain only closure under Π: If c is an element of $\mathrm{U}_{a,b}$ and d is a function mapping elements of $\mathrm{T}_{\mathrm{U},a,b}(c)$ to $\mathrm{U}_{a,b}$, then $\widehat{\Pi}_{\mathrm{U},a,b}(c,d)$ is a canonical element of $\mathrm{U}_{a,b}$, and we define $\widehat{\mathrm{T}}_\mathrm{V}(\widehat{\Pi}_{\mathrm{U},a,b}(c,d))$ as $\widehat{\Pi}_\mathrm{V}(\widehat{\mathrm{T}}_\mathrm{V}(c),(x)\widehat{\mathrm{T}}_\mathrm{V}(d(x)))$, which one verifies is a (canonical) element of V. (Note that we have to define, when two canonical elements $\widehat{\Pi}_{\mathrm{U},a,b}(c,d)$ and $\widehat{\Pi}_{\mathrm{U},a,b}(c',d')$ are equal and see, why in this case $\widehat{\mathrm{T}}_\mathrm{V}(\widehat{\Pi}_{\mathrm{U},a,b}(c,d))$ and $\widehat{\mathrm{T}}_\mathrm{V}(\widehat{\Pi}_{\mathrm{U},a,b}(c',d'))$ are equal.) By saying in future that a recursive subuniverse $(\mathrm{U}_{a,b},\widehat{\mathrm{T}}_{a,b})$ of a universe $(\mathrm{V},\mathrm{T_V})$ is closed under the basic set constructions we mean the explanations given before (using appropriate names for the constructors).

Furthermore $\mathrm{U}_{a,b}$ is closed under $\widehat{\mathrm{a}}_{a,b}$, $\widehat{\mathrm{b}}_{a,b}$: $\widehat{\mathrm{a}}_{a,b}$ is a canonical element of $\mathrm{U}_{a,b}$, and we define $\widehat{\mathrm{T}}_{\mathrm{U},a,b}(\widehat{\mathrm{a}}_{a,b})$ to be a. If $x : \mathrm{T}_{\mathrm{U},a,b}(a)$, then $\widehat{\mathrm{b}}_{a,b}(x)$ is a canonical element of $\mathrm{U}_{a,b}$, and we define $\widehat{\mathrm{T}}_{\mathrm{U},a,b}(\widehat{\mathrm{b}}_{a,b}(x))$ to be $b(x)$.

3.2 The External Mahlo Universe

The super universe is a universe $(\mathrm{V},\mathrm{T_V})$ which is closed under one operator on families of sets, namely the universe operator U. We obtain the external Mahlo universe (see [18], Sect. 6) by assuming that for any operator on families of sets there exist a universe closed under it. As indicated in the introduction, we will split such an operator into two components and introduce the following abbreviation (see the introduction for what is meant by notations like $f, g : \mathrm{Oper}(U,T)$).

Definition 3.2 *(Assuming the full logical framework.) Assume*

$$\vec{h} = f, g : \mathrm{Oper}(\mathrm{Set})$$

and $\langle U, T\rangle : \mathrm{Fam}(\mathrm{Set})$.

By $\widehat{\mathrm{f}},\widehat{\mathrm{g}} : \mathrm{IOper}(U,T)$ *reflect* f, g *we mean the following rules:*

$$\widehat{\mathrm{f}},\widehat{\mathrm{g}} : \mathrm{IOper}(U,T)$$

(which form constructors of U*) and*

$$\begin{array}{rcl} T(\widehat{\mathrm{f}}(a,b)) & = & f(T(a), T\circ b) \ , \\ T(\widehat{\mathrm{g}}(a,b,c)) & = & g(T(a), T\circ b, c) \ . \end{array}$$

Definition 3.3 *The rules for the external Mahlo universe are as follows:*

- *We have the full logical framework and the rules for the basic set constructions.*

Assume in the following

$$\vec{h} = f, g : \mathrm{Oper}(\mathrm{Set}) \ .$$

Then we have the following:

- *We have*

$$\mathrm{U}_{\vec{h}} : \mathrm{Set} \ , \qquad \mathrm{T}_{\mathrm{U},\vec{h}} : \mathrm{U}_{\vec{h}} \to \mathrm{Set} \ .$$

- *We have the rules expressing that* $(\mathrm{U}_{\vec{h}}, \mathrm{T}_{\mathrm{U},\vec{h}})$ *is a universe.*
- *We have the rules expressing that*

$$\widehat{\mathrm{f}}, \widehat{\mathrm{g}} : \mathrm{IOper}(\mathrm{U}_{\vec{h}}, \mathrm{T}_{\vec{h}})$$

  reflect $\vec{h}$.

Relationship to inductive-recursive definitions and meaning explanations. The external Mahlo universe is directly an instance of an inductive-recursive definition with parameter $\vec{h}$. We are at the moment not able to give meaning explanations for the external Mahlo universe, since we don't have yet a sufficient understanding of meaning explanations for the full logical framework.

3.3 The (Internal) Mahlo Universe

The internal Mahlo universe is the Mahlo universe as originally formulated by the author in [48], and we usually refer to it as the Mahlo universe. The internal Mahlo universe avoids the logical framework, which means that we don't refer to Fam(Set). Instead we form a universe $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$, which is called the Mahlo universe, such that for any operator on families of sets internal to $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ there exists a recursive subuniverse in V closed under this operator. Note that in case of the external Mahlo universe, Set plays the rôle which is played by V in the internal Mahlo universe. We will below introduce as well the notion of an internal inductive Mahlo universe, where we replace the recursive subuniverse by an inductive one.

As for the external Mahlo universe, we will split these operators into its two components, and introduce the following abbreviation (see the introduction for what is meant by notations like $f, g : \mathrm{IOper}(U, T)$):

Definition 3.4 *Assume* $(U, T) : \mathrm{Fam}(\mathrm{Set})$*,*

$$\vec{h} = f, g : \mathrm{IOper}(U, T) \ .$$

Assume (U', T') *was declared to be an inductive subuniverse of* (U, T) *with embedding* $\widehat{T} : U \to U'$*, or* $(U', \widehat{T})$ *was declared to be a recursive subuniverse of* (U, T) *with full decoding* T'.

By

$$\widehat{\mathrm{f}}, \widehat{\mathrm{g}} : \mathrm{IOper}(U', T') \textit{ reflect } f, g \enspace ,$$

we mean the following rules:

$$\widehat{\mathrm{f}}, \widehat{\mathrm{g}} : \mathrm{IOper}(U', T')$$

(which form constructors of U) and

(a) *in case of (U', T') being a recursive subuniverse of (U, T)*

$$\begin{array}{rcl} \widehat{T}(\widehat{\mathrm{f}}(a, b)) & = & f(\widehat{T}(a), \widehat{T} \circ b) \enspace , \\ \widehat{T}(\widehat{\mathrm{g}}(a, b, c)) & = & g(\widehat{T}(a), \widehat{T} \circ b, c) \enspace ; \end{array}$$

(b) *and in case (U', T') being an inductive subuniverse of (U, T)*

$$\begin{array}{rcl} T'(\widehat{\mathrm{f}}(a, b)) & = & T(f(\widehat{T}(a), \widehat{T} \circ b)) \enspace , \\ T'(\widehat{\mathrm{g}}(a, b, c)) & = & T(g(\widehat{T}(a), \widehat{T} \circ b, c)) \enspace . \end{array}$$

Definition 3.5 *(a) The rules for the internal Mahlo universe (also called simply the Mahlo universe or the internal recursive Mahlo universe) are as follows:*

- *We have the rules of the small logical framework and of the basic set constructions.*
- *We have rules expressing that $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ is a universe.*

Assume in the following

$$\vec{h} = f, g : \mathrm{IOper}(\mathrm{V}, \mathrm{T}_{\mathrm{V}}) \enspace .$$

- *We have rules expressing that $(\mathrm{U}_{\vec{h}}, \widehat{\mathrm{T}}_{\mathrm{U}, \vec{h}})$ is a recursive subuniverse of $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ with full decoding $\mathrm{T}_{\mathrm{U}, \vec{h}}$.*
- *We have rules expressing*

$$\widehat{\mathrm{f}}_{\vec{h}}, \widehat{\mathrm{g}}_{\vec{h}} : \mathrm{IOper}(\mathrm{U}_{\vec{h}}, \mathrm{T}_{\mathrm{U}, \vec{h}}) \textit{ reflect } \vec{h} \enspace .$$

- *We have the rule that V contains a code for $\mathrm{U}_{\vec{h}}$.*

$$\widehat{\mathrm{U}}_{\vec{h}} : \mathrm{V} \enspace , \qquad \mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{U}}_{\vec{h}}) = \mathrm{U}_{\vec{h}} \enspace .$$

(b) *The rules for the internal inductive Mahlo universe are obtained by replacing in (a) "recursive subuniverse" by "inductive subuniverse".*

Inductive-recursive nature of the construction. The Mahlo universe and all universe constructions defined later in this article are no longer instances of inductive-recursive definitions, since $\widehat{\mathrm{U}}$ is a constructor which is no longer strictly positive (Note that $\widehat{\mathrm{U}} : \mathrm{IOper}(\mathrm{V}, \mathrm{T}_{\mathrm{V}}) \to \mathrm{V}$, where $\mathrm{IOper}(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ forms the two components of $\mathrm{IFam}(\mathrm{V}, \mathrm{T}_{\mathrm{V}}) \to \mathrm{IFam}(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$). However, we can still see some inductive-recursive nature in it, by some sets being defined inductively, and some functions recursively. In case of the Mahlo universe, V, $\mathrm{U}_{\vec{h}}$ are defined simultaneously inductively, while T_{V}, $\widehat{\mathrm{T}}_{\mathrm{U},\vec{h}}$ are defined recursively. The constructors for the inductively defined sets are the constructors for the codes for the basic set constructions, $\widehat{\mathrm{f}}$, $\widehat{\mathrm{g}}$ and $\widehat{\mathrm{U}}$.

Meaning explanations for the internal Mahlo universe. As for the super universe we give simultaneously the meaning of V and of $\mathrm{U}_{f,g}$. Here f, g are such that from the understanding of the meaning of the sets and elements introduced before we can conclude that for every element a of V and every element b mapping elements of $\mathrm{T}_{\mathrm{V}}(a)$ to V we have that $f(a, b)$ is in V, and for every element c of $\mathrm{T}_{\mathrm{V}}(f(a, b))$ we have $g(a, b, c)$ is in V. We say in the following "$(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ is closed under (f, g)" for the condition on f, g just stated (which includes the proviso "from the understanding of the meaning of the sets and elements introduced before we can conclude that").

As for the super universe, $\mathrm{T}_{\mathrm{V}}(c)$ and $\widehat{\mathrm{T}}_{\mathrm{U},f,g}(c)$ will be defined for any canonical element of V and $\mathrm{U}_{f,g}$, respectively, and then defined for arbitrary elements in the same way, and we define $\mathrm{T}_{\mathrm{U},f,g}(c)$ as $\mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{T}}_{\mathrm{U},f,g}(c))$.

We first explain V. $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ is closed under the basic set constructions. Assume now that $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ is closed under f, g (as stated before). Then $\widehat{\mathrm{U}}_{f,g}$ is a canonical element of V and we define $\mathrm{T}_{\mathrm{V}}(\widehat{\mathrm{U}}_{f,g}) = \mathrm{U}_{f,g}$, where $\mathrm{U}_{f,g}$ will be explained below while referring only to the closure of V under the basic set constructions and under f, g, not to the closure under $\widehat{\mathrm{U}}$. If f, f' and g, g' are equal functions of their respective types, than $\widehat{\mathrm{U}}_{f,g}$ and $\widehat{\mathrm{U}}_{f',g'}$ are equal canonical elements of V. Since equal elements of V are mapped by T_{V} to equal sets we get that in this case $\mathrm{U}_{f,g}$ and $\mathrm{U}_{f',g'}$ are equal since the elements introduced and $\widehat{\mathrm{T}}_{\mathrm{U},f,g}$ and $\widehat{\mathrm{T}}_{\mathrm{U},f',g'}$ applied to them give the same results in V. This concludes the meaning explanations for V.

We now explain $\mathrm{U}_{f,g}$. $(\mathrm{U}_{f,g}, \widehat{\mathrm{T}}_{f,g})$ is closed under the basic set constructions. $\mathrm{U}_{f,g}$ is closed under f, g using constructors $\widehat{\mathrm{f}}_{f,g}, \widehat{\mathrm{g}}_{f,g}$, which means the following (this notion will be used in future meaning explanations): Assume a is an element of $\mathrm{U}_{f,g}$, and b is a function, mapping elements x of $\mathrm{T}_{\mathrm{U},f,g}(a)$ to $\mathrm{U}_{f,g}$. Then $a' := \widehat{\mathrm{T}}_{\mathrm{U},f,g}(a)$ is an element of V and $b' := (x)\widehat{\mathrm{T}}_{\mathrm{U},f,g}(b(x))$ is a function mapping elements of $\mathrm{T}_{\mathrm{V}}(a')$ to V. Since we know that V is closed under f, g, we know that $f(a', b')$ is an element of V and for c an element of $\mathrm{T}_{\mathrm{V}}(f(a', b'))$ we have that $g(a', b', c)$ is an element of V. Now $\widehat{\mathrm{f}}_{f,g}(a, b)$ and for c in $\mathrm{T}_{\mathrm{U},f,g}(f(a, b))$ $\widehat{\mathrm{g}}_{f,g}(a, b, c)$ are canonical elements of $\mathrm{U}_{f,g}$, and we define $\widehat{\mathrm{T}}_{\mathrm{U},f,g}(\widehat{\mathrm{f}}_{f,g}(a, b)) = f(a', b')$ and $\widehat{\mathrm{T}}_{\mathrm{U},f,g}(\widehat{\mathrm{g}}_{f,g}(a, b, c)) = g(a', b', c)$. $\widehat{\mathrm{f}}_{f,g}(a, b)$ is equal to $\widehat{\mathrm{f}}_{f,g}(a', b')$, if a, a' are equal elements of $\mathrm{U}_{f,g}$ and b, b' are equal functions from $\mathrm{T}_{\mathrm{U},f,g}(a)$ to $\mathrm{U}_{f,g}$. Equality for elements $\widehat{\mathrm{g}}_{f,g}(a, b, c)$ is defined similarly. This concludes the meaning explanation for $\mathrm{U}_{f,g}$.

4 The Autonomous Mahlo Universe

We are now going to carry out steps towards the autonomous Mahlo universe. The first step is to iterate the Mahlo construction and to form in Subsect. 4.1 the hyper Mahlo universe and in Subsect. 4.2 the hyper^n and hyper$^\alpha$-Mahlo universes. Then we will describe how to iterate the degrees of Mahloness autonomously, and obtain in Subsect. 4.3 the full formulation of the autonomous Mahlo universe. In the followup paper the Π_3-reflecting universe will be obtained by iterating the formation of higher degrees of Mahloness even further.

4.1 The hyper-Mahlo universe.

The step from a Mahlo universe to a hyper-Mahlo universe is similar to the step from a super universe to a hyper-super-universe (see the beginning of Sect. 5.1 of [53]): A hyper-Mahlo universe is a universe $(\mathrm{U}_2, \mathrm{T}_2)$ such that for every $\vec{h}_2 : \mathrm{IOper}(\mathrm{U}_2, \mathrm{T}_2)$ there exists a subuniverse $\mathrm{U}_{1,\vec{h}_2}$ represented in U_2 such that $\mathrm{U}_{1,\vec{h}_2}$ is a Mahlo universe closed under $\vec{h}_2$. That $\mathrm{U}_{1,\vec{h}_2}$ with its decoding function $\mathrm{T}_{1,\vec{h}_2}$ is a Mahlo universe means that for every $\vec{h}_1 : \mathrm{IOper}(\mathrm{U}_{1,\vec{h}_2}, \mathrm{T}_{1,\vec{h}_2})$ there exists a subuniverse $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$ of $\mathrm{U}_{1,\vec{h}_2}$ closed under $\vec{h}_1$ and represented in $\mathrm{U}_{1,\vec{h}_2}$. If one wants to define $\mathrm{U}_{1,\vec{h}_2}$ as a recursive subuniverse of U_2, one has one small problem, namely that one needs to have a code for $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$ in U_2. But it is straightforward to introduce such a code.

As before we split operators on families of sets internal to a universe into its two components, and we obtain the following rules:

Definition 4.1 *The rules for the hyper-Mahlo universe are as follows:*

- *Rules for the basic set constructions and the small logical framework.*
- *Rules expressing that* $(\mathrm{U}_2, \mathrm{T}_2)$ *is universe.*

Assume in the following

$$\vec{h}_2 = f_2, g_2 : \mathrm{IOper}(\mathrm{U}_2, \mathrm{T}_2) \ .$$

- *We have rules expressing that* $(\mathrm{U}_{1,\vec{h}_2}, \widehat{\mathrm{T}}_{1,\vec{h}_2})$ *is a recursive subuniverse of* $(\mathrm{U}_2, \mathrm{T}_2)$ *with full decoding* $\mathrm{T}_{1,\vec{h}_2}$.
- *We have rules expressing that*

$$\widehat{\mathrm{f}}_{1,\vec{h}_2}, \widehat{\mathrm{g}}_{1,\vec{h}_2} : \mathrm{IOper}(\mathrm{U}_{1,\vec{h}_2}, \mathrm{T}_{1,\vec{h}_2}) \textit{ reflect } \vec{h}_2 \ .$$

- *We have the rule that* U_2 *contains a code for* $\mathrm{U}_{1,\vec{h}_2}$*:*

$$\widehat{\mathrm{U}}_{2,1,\vec{h}_2} : \mathrm{U}_2 \ , \qquad \mathrm{T}_2(\widehat{\mathrm{U}}_{2,1,\vec{h}_2}) = \mathrm{U}_{1,\vec{h}_2} \ .$$

Assume in the following in addition

$$\vec{h}_1 = f_1, g_1 : \mathrm{IOper}(\mathrm{U}_{1,\vec{h}_2}, \mathrm{T}_{1,\vec{h}_2}) \enspace .$$

- *We have rules expressing that* $(\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}, \widehat{\mathrm{T}}_{0,\vec{h}_2,\vec{h}_1})$ *is a recursive subuniverse of* $(\mathrm{U}_{1,\vec{h}_2}, \mathrm{T}_{1,\vec{h}_2})$ *with full decoding* $\mathrm{T}_{0,\vec{h}_2,\vec{h}_1}$*.*
- *We have rules expressing that*

$$\widehat{\mathrm{f}}_{0,\vec{h}_2,\vec{h}_1}, \widehat{\mathrm{g}}_{0,\vec{h}_2,\vec{h}_1} : \mathrm{IOper}(\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}, \mathrm{T}_{0,\vec{h}_2,\vec{h}_1}) \text{ reflect } \vec{h}_1 \enspace .$$

- *We have the rule that* U_2 *contains a code for* $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$*:*

$$\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1} : \mathrm{U}_2 \enspace , \qquad \mathrm{T}_2(\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1}) = \mathrm{U}_{0,\vec{h}_2,\vec{h}_1} \enspace .$$

- *We have the rule that* $\mathrm{U}_{1,\vec{h}_2}$ *contains a code for* $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$*:*

$$\widehat{\mathrm{U}}_{1,0,\vec{h}_2,\vec{h}_1} : \mathrm{U}_{1,\vec{h}_2} \enspace , \qquad \widehat{\mathrm{T}}_{1,\vec{h}_2}(\widehat{\mathrm{U}}_{1,0,\vec{h}_2,\vec{h}_1}) = \widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1} \enspace .$$

Inductive-recursive nature of the construction. In the above U_2, $\mathrm{U}_{1,\vec{h}_2}$, $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$ are defined simultaneously inductively, while T_2, $\widehat{\mathrm{T}}_{1,\vec{h}_2}$, $\widehat{\mathrm{T}}_{0,\vec{h}_2,\vec{h}_1}$ are defined recursively. The constructors for the inductively defined sets are the constructors for the codes for the basic set constructions, $\widehat{\mathrm{f}}_1$, $\widehat{\mathrm{g}}_1$, $\widehat{\mathrm{U}}_{2,1}$, $\widehat{\mathrm{f}}_0$, $\widehat{\mathrm{g}}_0$, $\widehat{\mathrm{U}}_{2,0}$, $\widehat{\mathrm{U}}_{1,0}$.

Meaning explanations for the hyper-Mahlo universe. As for the Mahlo universe we give simultaneously the meaning of

- U_2,
- of $\mathrm{U}_{1,\vec{h}_2}$ for every $\vec{h}_2 = f_2, g_2$ such that $(\mathrm{U}_2, \mathrm{T}_2)$ is closed under $\vec{h}_2$ (note that this was introduced as an abbreviation in the meaning explanations for the Mahlo universe with a special proviso),
- and of $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$ for every $\vec{h}_2$ as before and every $\vec{h}_1 = f_1, g_1$ such that $(\mathrm{U}_{1,\vec{h}_2}, \mathrm{T}_{1,\vec{h}_2})$ is closed under $\vec{h}_1$.

As for the super universe, $\mathrm{T}_2(a)$, $\widehat{\mathrm{T}}_{1,\vec{h}_2}(a)$, and $\widehat{\mathrm{T}}_{0,\vec{h}_2,\vec{h}_1}(a)$ will be defined for any canonical elements of U_2, $\mathrm{U}_{1,\vec{h}_2}$, $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$ respectively, and then defined for arbitrary elements in the same way. Furthermore $\mathrm{T}_{1,\vec{h}_2}$ and $\mathrm{T}_{0,\vec{h}_2,\vec{h}_1}$ are defined as before. In addition to the explanations given below, $(\mathrm{U}_2, \mathrm{T}_2)$, $(\mathrm{U}_{1,\vec{h}_2}, \widehat{\mathrm{T}}_{1,\vec{h}_2})$, $(\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}, \widehat{\mathrm{T}}_{0,\vec{h}_2,\vec{h}_1})$ will be closed under the basic set constructions.

We explain U_2: Assume U_2 is closed under $\vec{h}_2$. Then $\widehat{\mathrm{U}}_{2,1,\vec{h}_2}$ is a canonical element of U_2, and we define $\mathrm{T}_2(\widehat{\mathrm{U}}_{2,1,\vec{h}_2})$ to be $\mathrm{U}_{1,\vec{h}_2}$, which will be explained below while only referring to the closure of U_2 under the basic set constructions and under $\vec{h}_2$. Furthermore, if $\mathrm{U}_{1,\vec{h}_2}$ is closed under $\vec{h}_1$, then $\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1}$ is a canonical element of U_2 and we define $\mathrm{T}_2(\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1}) = \mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$, where $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$ will be explained below, while only referring to the closure of U_2 and $\mathrm{U}_{1,\vec{h}_2}$ under the basic set constructions and under $\vec{h}_2$ and $\vec{h}_1$ respectively. If $\vec{h}_2, \vec{h}'_2$ and $\vec{h}_1, \vec{h}'_1$ are equal functions of their respective types, then $\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1}$ and $\widehat{\mathrm{U}}_{2,0,\vec{h}'_2,\vec{h}'_1}$ are equal canonical elements of U_2, similar for $\widehat{\mathrm{U}}_{1,\vec{h}_2}$ and $\widehat{\mathrm{U}}_{1,\vec{h}'_2}$. We can observe, when looking at their explanations below, that they are mapped by T_2 to equal sets. This concludes the explanations of U_2.

We explain $\mathrm{U}_{1,\vec{h}_2}$: $(\mathrm{U}_{1,\vec{h}_2}, \widehat{\mathrm{T}}_{1,\vec{h}_2})$ is closed under $\vec{h}_2$ using constructors $\widehat{\mathrm{f}}_{1,\vec{h}_2}$, $\widehat{\mathrm{g}}_{1,\vec{h}_2}$. Furthermore, if $\mathrm{U}_{1,\vec{h}_2}$ is closed under $\vec{h}_1$, then $\widehat{\mathrm{U}}_{1,0,\vec{h}_2,\vec{h}_1}$ is a canonical element of $\mathrm{U}_{1,\vec{h}_2}$ and we define $\widehat{\mathrm{T}}_{1,\vec{h}_2}(\widehat{\mathrm{U}}_{1,0,\vec{h}_2,\vec{h}_1}) = \widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1}$. If $\vec{h}_1$, $\vec{h}'_1$ are equal functions of their types then $\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1}$ and $\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}'_1}$ are equal canonical elements of $\mathrm{U}_{1,\vec{h}_2}$ which are mapped by $\widehat{\mathrm{T}}_{1,\vec{h}_2}$ to equal elements of U_2.

We explain $\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}$: $(\mathrm{U}_{0,\vec{h}_2,\vec{h}_1}, \mathrm{T}_{0,\vec{h}_2,\vec{h}_1})$ is closed under $\vec{h}_1$ using constructors $\widehat{\mathrm{f}}_{2,\vec{h}_2,\vec{h}_1}$, $\widehat{\mathrm{g}}_{2,\vec{h}_2,\vec{h}_1}$, which finishes the explanation of it.

4.2 The hyper^n- and Hyper$^\alpha$-Mahlo universe.

The hyper-Mahlo universe can easily be generalised to a hyper^n-Mahlo universe, which consists of $n+1$ nested Mahlo universes (so the hyper^0-Mahlo universe is a Mahlo universe.)

We give now a generalisation to a hyper$^\alpha$-Mahlo universe. For this we need first to introduce a set the ordinals α are elements of. Assume some ordinal notation system, for which we have proved in some other type theory that it is well-founded (e.g. an ordinal notation system of order type ϵ_0). Let the set of notations be given by a Boolean valued function $\widehat{\mathrm{OT}} : \mathrm{N} \to \mathrm{Bool}$ which decides whether a natural number denotes an ordinal notation or not, and the less-than relation be given by a Boolean valued function $\widehat{\prec}' : \mathrm{N} \to \mathrm{N} \to \mathrm{Bool}$, which is written infix, which are definable in Peano Arithmetic, and therefore definable using the basic set constructions plus the rules for atom (needed in order to obtain Peano's fourth axiom).

In our type theory we use the rules for the basic set constructions, the small logical framework, and atom, and define using them $\widehat{\mathrm{OT}}$, $\widehat{\prec}'$. Then let $\mathrm{OT}(x) := \mathrm{atom}(\widehat{\mathrm{OT}}(x))$, $x \prec' y := \mathrm{atom}(x \mathrel{\widehat{\prec}'} y)$, and write $x \in \mathrm{OT}$ for $\mathrm{OT}(x)$. We define now the set of ordinals as $\mathrm{Ord} := (n : \mathrm{N}) \times (n \in \mathrm{OT})$, and $\langle n, p\rangle \prec \langle m, q\rangle := n \prec' m$.

A hyper$^\alpha$-universe is a universe $(\mathrm{U}^\alpha, \mathrm{T}^\alpha)$, such that for every $\beta \prec \alpha$ and $\vec{h} : \mathrm{IOper}(\mathrm{U}^\alpha, \mathrm{T}^\alpha)$ there exists a hyper$^\beta$-sub-universe of $(\mathrm{U}^\alpha, \mathrm{T}^\alpha)$, closed under $\vec{h}$ and represented in U^α. In order to formulate it one defines for $\alpha : \mathrm{Ord}$ a set of universes Univ_α of universes of Mahlo degree α, i.e. of hyper$^{-1+\alpha}$-Mahlo universes. (Univ_0 contains universes which are not Mahlo at all, Univ_1 contains the Mahlo-universes, which one might call hyper^0-Mahlo universes, Univ_2 contains hyper-Mahlo or hyper^1-Mahlo universes, etc.; so the "Mahlo degree" determines "how much hyper$^\alpha$-Mahlo" a universe is). For $u : \mathrm{Univ}_\alpha$ we have $(\mathrm{U}_{\alpha,u}, \mathrm{T}_{\mathrm{U},\alpha,u}) : \mathrm{Fam}(\mathrm{Set})$ (note that this is an abbreviation introduced in the introduction which does not refer to the logical framework).

When defining it, one faces the problem that it is possible, but cumbersome, to use recursive subuniverses. The reason is that we get chains of subuniverses: If $\alpha_0 \succ \alpha_1 \succ \cdots \succ \alpha_n$, we get chains of subuniverses $\mathrm{u}_i : \mathrm{U}_{\alpha_i}$ such that $(\mathrm{U}_{\alpha_{i+1}\mathrm{u}_{i+1}}, \mathrm{T}_{\mathrm{U},\alpha_{i+1},\mathrm{u}_{i+1}})$ is a subuniverse of $(\mathrm{U}_{\alpha_i\mathrm{u}_i}, \mathrm{T}_{\mathrm{U},\alpha_i,\mathrm{u}_i})$. When using recursive subuniverses, we need to have codes for U_i in every universe U_j for $j > i$ since the code for it will be recursively handed upwards until we reach the uppermost universe for which this code is mapped to the set U_i. This can be seen already for the hyper-Mahlo universe, where we first needed to introduce $\widehat{\mathrm{U}}_{2,0,\vec{h}_2,\vec{h}_1} : \mathrm{U}_2$ before we were able to define $\widehat{\mathrm{U}}_{1,0,\vec{h}_2,\vec{h}_1} : \mathrm{U}_{1,\vec{h}_2}$, otherwise we couldn't define an equality rule for $\widehat{\mathrm{T}}_{1,\vec{h}_2}(\widehat{\mathrm{U}}_{1,0,\vec{h}_2,\vec{h}_1})$. (See the end of Def. 4.1.) This is possible, but notationally complicated.[1]

It is much easier to work instead with inductive subuniverses. Then elements created in subuniverses are automatically passed on to higher universes containing it. We will follow this approach in the following.

Definition 4.2 *The rules for the hyper$^\alpha$-Mahlo universe are as follows:*

- *Rules for the basic set constructions, the small logical framework, and atom.*
- *The definition equations for* $\widehat{\mathrm{OT}}, \widehat{\prec}', \mathrm{OT}, \prec', \mathrm{Ord}, \prec$*, as given above. The order type of the ordinal notation system is supposed to be* $1 + \alpha$*.*
- *Rules expressing that* $(\mathrm{V}, \mathrm{T}_\mathrm{V})$ *is universe.*
- *Defining rules for* Univ_β*: Assume* $\beta : \mathrm{Ord}$*.*
  - $\mathrm{Univ}_\beta : \mathrm{Set}$.
  - *Rules expressing that for* $\beta : \mathrm{Ord}, u : \mathrm{Univ}_\beta$ *we have* $(\mathrm{U}_{\beta,u}, \mathrm{T}_{\mathrm{U},\beta,u})$ *is an inductive subuniverse of* $(\mathrm{V}, \mathrm{T}_\mathrm{V})$ *with embedding* $\widehat{\mathrm{T}}_{\mathrm{U},\beta,u}$*, which is represented in* V *by* $\widehat{\mathrm{U}}_{\mathrm{U},\beta,u}$*.*
- *Assume* $\beta : \mathrm{Ord}$*,*
$$\vec{h} = f, g : \mathrm{IOper}(\mathrm{V}, \mathrm{T}_\mathrm{V}) \ .$$
  - *We have*
$$\mathrm{v}_{\beta,\vec{h}} : \mathrm{Univ}_\beta \ .$$
    Let locally in this item only $\mathrm{v}_- := \mathrm{v}_{\beta,\vec{h}} : \mathrm{Univ}_\beta$*.*
  - *We have rules expressing that*
$$\widehat{\mathrm{f}}_{\mathrm{v},\beta,\vec{h}}, \widehat{\mathrm{g}}_{\mathrm{v},\beta,\vec{h}} : \mathrm{IOper}(\mathrm{U}_{\beta,\mathrm{v}_-}, \mathrm{T}_{\mathrm{U},\beta,\mathrm{v}_-}) \textit{ reflect } \vec{h} \ .$$

[1] In [53] we assumed that this is not possible because we expected to get infinite chains of universes. But in fact we never get such infinite chains: when going upwards in the Mahlo degrees, we will always in finitely many steps reach a top level. However the number of steps is different for universes having the same Mahlo degree.

- *Assume* $\beta : \mathrm{Ord}$*,* $u : \mathrm{Univ}_\beta$*,* $\gamma : \mathrm{Ord}$*,* $p : \gamma \prec \beta$*, and*

$$\vec{h} = f, g : \mathrm{IOper}(\mathrm{U}_{\beta,u}, \mathrm{T}_{\mathrm{U},\beta,u}) \enspace .$$

  - *We have*

$$\mathrm{u}_{\beta,u,\gamma,p,\vec{h}} : \mathrm{Univ}_\gamma \enspace .$$

    Let locally in this item only

$$\begin{array}{lcllll} \mathrm{U}_+ & := & \mathrm{U}_{\beta,u} & : & \mathrm{Set} \enspace , & \\ \mathrm{T}_+(a) & := & \mathrm{T}_{\mathrm{U},\beta,u}(a) & : & \mathrm{Set} & (\textit{where } a : \mathrm{U}_+) \enspace , \\ \mathrm{u}_- & := & \mathrm{u}_{\beta,u,\gamma,p,\vec{h}} & : & \mathrm{Univ}_\gamma \enspace , & \\ \mathrm{U}_- & := & \mathrm{U}_{\gamma,\mathrm{u}_-} & : & \mathrm{Set} \enspace , & \\ \mathrm{T}_-(a) & := & \mathrm{T}_{\mathrm{U},\gamma,\mathrm{u}_-}(a) & : & \mathrm{Set} & (\textit{where } a : \mathrm{U}_-) \enspace . \end{array}$$

  - $(\mathrm{U}_-, \mathrm{T}_-)$ *is an inductive subuniverse of* $(\mathrm{U}_+, \mathrm{T}_+)$ *with embedding* $\widehat{\mathrm{T}}_{\mathrm{u},\beta,u,\gamma,p,\vec{h}} : \mathrm{U}_- \to \mathrm{U}_+$*, which is represented in* U_+ *by* $\widehat{\mathrm{U}}_{\mathrm{u},\beta,u,\gamma,p,\vec{h}}$.
  - *We have rules expressing that*

$$\widehat{\mathrm{f}}_{\mathrm{u},\beta,u,\gamma,p,\vec{h}}, \widehat{\mathrm{g}}_{\mathrm{u},\beta,u,\gamma,p,\vec{h}} : \mathrm{IOper}(\mathrm{U}_-, \mathrm{T}_-) \textit{ reflect } \vec{h} \enspace .$$

Note that in case of $\alpha = 0$ we have that the ordinal notation system has one element and we obtain essentially the rules for the internal inductive Mahlo universe.

We leave it to the reader to model a hyper$^\alpha$-Mahlo universe in Kripke-Platek set theory plus one recursively hyper$^\alpha$-Mahlo ordinal and finitely many admissibles above it. (This model can as well be obtained by cutting down the model of the autonomous Mahlo universe below).

Inductive-recursive nature of the construction. In the above V, $\mathrm{U}_{\beta,u}$, Univ_β are defined simultaneously inductively, while T_V, $\mathrm{T}_{\mathrm{U},\beta,u}$ are defined recursively. The constructors for the inductively defined sets are the constructors for the codes for the basic set constructions, $\widehat{\mathrm{T}}_\mathrm{U}$, $\widehat{\mathrm{U}}_\mathrm{U}$, v, $\widehat{\mathrm{f}}_\mathrm{v}$, $\widehat{\mathrm{g}}_\mathrm{v}$, u, $\widehat{\mathrm{T}}_\mathrm{u}$, $\widehat{\mathrm{U}}_\mathrm{u}$, $\widehat{\mathrm{f}}_\mathrm{u}$, $\widehat{\mathrm{g}}_\mathrm{u}$.

Meaning explanations for the hyper$^\alpha$-Mahlo universe. For the hyper$^\alpha$-Mahlo universe we assume that we have already shown without using the Mahlo universe that OT is well-founded, and therefore validated the principle of transfinite induction over Ord. The meaning explanations for the basic set constructions, $\widehat{\mathrm{OT}}$, $\widehat{\prec}'$, OT, $\prec'$, Ord, $\prec$ are standard. We will now give simultaneously the meaning of

- V,
- of Univ_β for $\beta : \mathrm{Ord}$,
- and of $\mathrm{U}_{\beta,u}$ for $\beta : \mathrm{Ord}$ and $u : \mathrm{Univ}_\beta$.

This will be done in such a way that $\mathrm{U}_{\beta,u}$ will not refer to V or $\mathrm{U}_{\gamma,u}$ for $\beta \prec \gamma$ except for the closure properties known at the time of introducing u, and by not referring to $\mathrm{U}_{\beta,u'}$ for other u', except when defining it as equal to such a set because u and u' are equal elements of Univ_β. So we only assume, when introducing $\mathrm{U}_{\beta,u}$ a complete understanding of $\mathrm{U}_{\gamma,u'}$ for $\gamma \prec \beta$ for certain u' in Univ_γ.

Simultaneously recursively we define $\mathrm{T}_\mathrm{V}(x)$ and $\mathrm{T}_{\mathrm{U},\beta,u}(x)$ for elements x of V and $\mathrm{U}_{\beta,u}$, respectively. These operations will as before be first defined for canonical elements and then extended to arbitrary elements.

$(\mathrm{V}, \mathrm{T}_\mathrm{V})$ and $(\mathrm{U}_{\beta,u}, \mathrm{T}_{\beta,u})$ will in addition to what is said below be closed under the basic set constructions.

Meaning of V: If β is in Ord, u is in Univ_β, then $u' := \widehat{\mathrm{U}}_{\beta,u}$ is a canonical element of V and $\mathrm{T}_\mathrm{V}(u')$ is defined as $\mathrm{U}_{\beta,u}$. Furthermore, for a in $\mathrm{U}_{\beta,u}$ we have that $a' := \widehat{\mathrm{T}}_{\beta,u}(a)$ is a canonical element of V and we define $\mathrm{T}_\mathrm{V}(a') = \mathrm{T}_{\beta,u}(a)$.

Meaning of Univ_β: If $(\mathrm{V}, \mathrm{T}_\mathrm{V})$ is closed under f, g, then $\mathrm{v}_{\beta,f,g}$ is a canonical element of Univ_β. Assume that $(\mathrm{U}_{\gamma,u}, \mathrm{T}_{\gamma,u})$ is closed under f, g. Assume p is an element of $\beta \prec \gamma$. Then $\mathrm{u}_{\gamma,u,\beta,p,f,g}$ is a canonical element of Univ_β.

Meaning of $\mathrm{U}_{\beta,u}$: We have closure properties applying to all $u \in \mathrm{Univ}_\beta$: Assume that $(\mathrm{U}_{\beta,u}, \mathrm{T}_{\beta,u})$ is closed under f, g. Assume p is an element of $\gamma \prec \beta$. Let $u' := \mathrm{u}_{\beta,u,\gamma,p,f,g}$ which is a canonical element of Univ_β. Then $u'' := \widehat{\mathrm{U}}_{\mathrm{u},\beta,u,\gamma,p,f,g}$ is a canonical element of $\mathrm{U}_{\beta,u}$ with $\mathrm{T}_{\beta,u}(u'')$ defined as $\mathrm{U}_{\beta,u'}$. Furthermore for every a in $\mathrm{U}_{\gamma,u'}$ we have that $a' := \widehat{\mathrm{T}}_{\mathrm{u},\beta,u,\gamma,p,f,g}(a)$ is a canonical element of $\mathrm{U}_{\beta,u}$ and we define $\mathrm{T}_{\beta,u}(a')$ as $\mathrm{T}_{\gamma,u'}(a)$.

The additional closure properties for $u' := \mathrm{u}_{\beta,u,\gamma,p,f,g}$ as defined before are: $(\mathrm{U}_{\gamma,u'}, \mathrm{T}_{\gamma,u'})$ is closed under f, g via $\widehat{\mathrm{T}}' := \widehat{\mathrm{T}}_{\mathrm{u},\beta,u,\gamma,p,f,g}$ using constructors $\widehat{\mathrm{f}}' := \widehat{\mathrm{f}}_{\mathrm{u},\beta,u,\gamma,p,f,g}$ and $\widehat{\mathrm{g}}' := \widehat{\mathrm{g}}_{\mathrm{u},\beta,u,\gamma,p,f,g}$. This means that if a is an element of $\mathrm{U}_{\gamma,u'}$ and b a function mapping $\mathrm{T}_{\gamma,u'}(a)$ to $\mathrm{U}_{\gamma,u'}$, then $a' := \widehat{\mathrm{f}}'(a, b)$ is a canonical element of $\mathrm{U}_{\gamma,u'}$ with $\mathrm{T}_{\gamma,u'}(a') = A := \mathrm{T}_{\beta,u}(f(\widehat{\mathrm{T}}'(a), \widehat{\mathrm{T}}' \circ b))$. Furthermore, if c is an element of A, then $c' := \widehat{\mathrm{g}}'(a, b, c)$ is a canonical element of $\mathrm{U}_{\gamma,u'}$ with $\mathrm{T}_{\gamma,u'}(c') = \mathrm{T}_{\beta,u}(g(\widehat{\mathrm{T}}'(a), \widehat{\mathrm{T}}' \circ b, c))$.

The additional closure properties for $v' := \mathrm{v}_{\beta,f,g}$ as defined before are: $(\mathrm{U}_{\beta,v'}, \mathrm{T}_{\beta,v'})$ is closed under f, g via $\widehat{\mathrm{T}}_{\beta,v'}$ using constructors $\widehat{\mathrm{f}}_{\mathrm{v},\beta,f,g}, \widehat{\mathrm{g}}_{\mathrm{v},\beta,f,g}$.

4.3 The Autonomous Mahlo Universe

The next step is to define a universe $(\mathrm{V}, \mathrm{T}_\mathrm{V})$ which has strength $(\mathrm{KP}\omega + (\mathrm{AutMahlo}))^+$. Here $(\mathrm{KP}\omega + (\mathrm{AutMahlo}))^+$ means that there exists an ordinal κ which is recursively hyper$^\kappa$-Mahlo plus finitely many admissibles above it. We will call this universe the autonomous Mahlo universe.

(Recursively) inaccessible sets (and hyper$^\kappa$-Mahlo sets are inaccessible sets) correspond in type theory to universes, so we introduce in type theory a universe $(\mathrm{V}, \mathrm{T}_\mathrm{V})$ which correspond to an autonomous Mahlo set. Ordinals can be modelled in type theory by elements of a W-type. The set of ordinals in a universe $(\mathrm{V}, \mathrm{T}_\mathrm{V})$ are the elements of the W-type having as branching types all sets in V, which means that it can be represented by $\mathrm{W}x : \mathrm{V}.\mathrm{T_V}(x)$[2]. We define the set Deg of degrees preliminarily as this set.

So we need to formulate that $(\mathrm{V}, \mathrm{T_V})$ is hyper^d-Mahlo for any $d : \mathrm{Deg}$. In order to define this, we introduce for $d : \mathrm{Deg}$ the set Univ_d of codes for universes which are hyper^d-Mahlo. For each $u : \mathrm{Univ}_d$ we introduce the actual universe given by it as $\mathrm{U}_{d,u} : \mathrm{Set}$ and $\mathrm{T}_{d,u} : \mathrm{U}_{d,u} \to \mathrm{Set}$.

That the autonomous Mahlo ordinal κ is hyper$^\alpha$-Mahlo for any $\alpha < \kappa$ will be expressed by rules stating that for every $\vec{h} : \mathrm{IOper}(\mathrm{V}, \mathrm{T_V})$ and every degree $d : \mathrm{Deg}$ there exists a universe $u : \mathrm{Univ}_d$ s.t. $\mathrm{U}_{d,u}$ is closed under $\vec{h}$.

That a universe $u : \mathrm{Univ}_d$ is hyper^d-Mahlo will be expressed by rules stating that any $\vec{h} : \mathrm{IOper}(\mathrm{U}_{d,u}, \mathrm{T}_{d,u})$ can be reflected into smaller degrees d'. If $d = \mathrm{sup}(r, s)$, where $r : \mathrm{V}$ and $s : \mathrm{T_V}(a) \to \mathrm{Deg}$, then the degrees d' smaller than d are the degrees $d' := s(a)$ for $a : \mathrm{T_V}(a)$. Then we demand that any universe $u : \mathrm{Univ}_d$ is Mahlo reflecting in $\mathrm{Univ}_{d'}$ for any d' smaller than d: If $\vec{h} : \mathrm{IOper}(\mathrm{U}_{d,u}, \mathrm{T}_{d,u})$, then there exists an inductive subuniverse of $(\mathrm{U}_{d,u}, \mathrm{T}_{d,u})$ in $\mathrm{Univ}_{d'}$ which is closed under $\vec{h}$ and represented in $\mathrm{U}_{d,u}$.

(The reader might suggest that one should close "smaller" transitively, but that won't result in introducing new universes: in order for instance to reflect $\vec{h}$ as above into a $u'' : \mathrm{Univ}_{d''}$, where d'' is smaller than d' which is again smaller than d, we can first reflect $\vec{h}$ into some $u' : \mathrm{Univ}_{d'}$ and reflect the function reflecting $\vec{h}$ in $\mathrm{U}_{d',u'}$ into some $u'' : \mathrm{Univ}_{d''}$.)

One problem seems to be that Deg depends on all of $\mathrm{V}, \mathrm{T_V}$ – therefore Deg is only available once the construction of $\mathrm{V}, \mathrm{T_V}$ is complete. But when constructing V one needs to know Deg in order to be able to introduce new universes for appropriate Mahlo degrees.

However, one can easily see that for every $d : (\mathrm{W}x : \mathrm{V}.\mathrm{T_V}(x))$ there exists an $\langle a, b\rangle : \mathrm{IFam}(\mathrm{V}, \mathrm{T_V})$ such that d and its subtrees only refer to elements of V in $(b(x))_{x:a}$ (collect all branching degrees of d and its subtrees into one family of sets in V). More precisely, the situation is as follows: Define for $\langle a, b\rangle$ as above $f_{a,b} : (\mathrm{W}x : \mathrm{T_V}(a).\mathrm{T_V}(b(x))) \to \mathrm{W}x : \mathrm{V}.\mathrm{T_V}(x)$, $f_{a,b}(\mathrm{sup}(r, s)) = \mathrm{sup}(b(r), (y)f_{a,b}(s(y)))$. Then for every $d : \mathrm{Deg}$ there exist some a, b, d' s.t. $d = f_{a,b}(d')$ for some $\langle a, b\rangle$ and $d' : (\mathrm{W}x : \mathrm{T_V}(a).\mathrm{T_V}(b(x)))$.

Therefore the set of degrees depends only locally on V and one can construct Deg simultaneously while constructing V – whenever one constructs new elements of V one obtains new elements of Deg, which allow to construct new elements of V, namely subuniverses of V having the new degrees. The existence of an autonomous Mahlo universe means that we claim that this process eventually stops after transfinitely many steps. We distinguish now Deg, which is defined simultaneously with V from $\mathrm{W}x : \mathrm{V}.\mathrm{T_V}(x)$ (which can fully only be defined after the definition of $(\mathrm{V}, \mathrm{T_V})$ is complete) and give the constructor introducing elements of Deg the name deg rather than sup.

It turns out that we don't need any elimination rules for Deg, except for rules decomposing $\mathrm{deg}(r, s)$ into r and s. Full elimination rules are unnecessary, since there exists an obvious embedding $g : (\mathrm{W}x : \mathrm{V}.\mathrm{T_V}(x)) \to \mathrm{Deg}$. So if one

[2] Note that this type is the same as Peter Aczel's type V of iterative sets [1].

wanted to refer to the least set of degrees introduced like this, one could refer to $g(w)$ for $w : (\mathrm{W}x : \mathrm{V}.\mathrm{T}_{\mathrm{V}}(x))$ and use the elimination rule for $\mathrm{W}x : \mathrm{V}.\mathrm{T}_{\mathrm{V}}(x)$.

We therefore obtain now the following rules for the type theory with one autonomous Mahlo universe:

Definition 4.3 *The rules for the autonomous Mahlo universe are as follows:*

- *Rules for the basic set constructions and the small logical framework.*
- *Rules expressing that* $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ *is universe.*
- *Formation rule for* Deg*:*
$$\mathrm{Deg} : \mathrm{Set} \ ,$$
- *Introduction rule for* Deg *and elimination rules (referring to* $\mathrm{bdeg}, \mathrm{subdeg}$*):*
$$\begin{array}{ll} \mathrm{deg} : (a : \mathrm{V}, b : \mathrm{T}_{\mathrm{V}}(a) \to \mathrm{Deg}) \to \mathrm{Deg} \ , & \\ \mathrm{bdeg} : \mathrm{Deg} \to \mathrm{V} \ , & \mathrm{subdeg} : (d : \mathrm{Deg}, a : \mathrm{T}_{\mathrm{V}}(\mathrm{bdeg}(d))) \to \mathrm{Deg} \ . \\ \mathrm{bdeg}(\mathrm{deg}(a, b)) = a \ , & \mathrm{subdeg}(\mathrm{deg}(a, b), c) = b(c) \ . \end{array}$$
- *Defining rules for* Univ_d*: Assume* $d : \mathrm{Deg}$*.*
  - $\mathrm{Univ}_d : \mathrm{Set}$.
  - *Rules expressing that for* $d : \mathrm{Deg}$*,* $u : \mathrm{Univ}_d$ *we have* $(\mathrm{U}_{d,u}, \mathrm{T}_{\mathrm{U},d,u})$ *is an inductive subuniverse of* $(\mathrm{V}, \mathrm{T}_{\mathrm{V}})$ *with embedding* $\widehat{\mathrm{T}}_{\mathrm{U},d,u}$*, which is represented in* V *as* $\widehat{\mathrm{U}}_{\mathrm{U},d,u}$*.*
- *Assume* $d : \mathrm{Deg}$*,*
$$\vec{h} = f, g : \mathrm{IOper}(\mathrm{V}, \mathrm{T}_{\mathrm{V}}) \ .$$
  - *We have*
$$\mathrm{v}_{d,\vec{h}} : \mathrm{Univ}_d \ .$$
    Let locally in this item only $\mathrm{v}_{-} := \mathrm{v}_{d,\vec{h}} : \mathrm{Univ}_d$*.*
  - *We have rules expressing that*
$$\widehat{\mathrm{f}}_{\mathrm{v},d,\vec{h}}, \widehat{\mathrm{g}}_{\mathrm{v},d,\vec{h}} : \mathrm{IOper}(\mathrm{U}_{d,\mathrm{v}_{-}}, \mathrm{T}_{\mathrm{U},d,\mathrm{v}_{-}}) \ \textit{reflect} \ \vec{h} \ .$$
- *Assume now* $d : \mathrm{Deg}$*,* $u : \mathrm{Univ}_d$*,* $a : \mathrm{T}_{\mathrm{V}}(\mathrm{bdeg}(d))$*,*
$$\vec{h} = f, g : \mathrm{IOper}(\mathrm{U}_{d,u}, \mathrm{T}_{\mathrm{U},d,u}) \ .$$
  Let locally in this item only $\mathrm{d}_{-} := \mathrm{subdeg}(d, a) : \mathrm{Deg}$*.*
  - *We have*
$$\mathrm{u}_{d,u,a,\vec{h}} : \mathrm{Univ}_{\mathrm{d}_{-}} \ .$$

Let additionally locally in this item only

$$\begin{array}{lllll} \mathrm{U}_+ & := & \mathrm{U}_{d,u} & : & \mathrm{Set}\ , \\ \mathrm{T}_+(a) & := & \mathrm{T}_{\mathrm{U},d,u}(a) & : & \mathrm{Set} \qquad (\textit{where } a : \mathrm{U}_+)\ , \\ \mathrm{u}_- & := & \mathrm{u}_{d,u,a,\vec{h}} & : & \mathrm{Univ}_{\mathrm{d}_-}\ , \\ \mathrm{U}_- & := & \mathrm{U}_{\mathrm{d}_-,\mathrm{u}_-} & : & \mathrm{Set}\ , \\ \mathrm{T}_-(a) & := & \mathrm{T}_{\mathrm{U},\mathrm{d}_-,\mathrm{u}_-}(a) & : & \mathrm{Set} \qquad (\textit{where } a : \mathrm{U}_-)\ . \end{array}$$

- $(\mathrm{U}_-, \mathrm{T}_-)$ *is an inductive subuniverse of* $(\mathrm{U}_+, \mathrm{T}_+)$ *with embedding* $\widehat{\mathrm{T}}_{\mathrm{u},d,u,a,\vec{h}} : \mathrm{U}_- \to \mathrm{U}_+$*, which is represented in* U_+ *by* $\widehat{\mathrm{U}}_{\mathrm{u},d,u,a,\vec{h}}$.
- *We have rules expressing that*

$$\widehat{\mathrm{f}}_{\mathrm{u},d,u,a,\vec{h}}, \widehat{\mathrm{g}}_{\mathrm{u},d,u,a,\vec{h}} : \mathrm{IOper}(\mathrm{U}_-, \mathrm{T}_-) \textit{ reflect } \vec{h}\ .$$

Inductive-recursive nature of the construction. In the above V, Deg, Univ_d, $\mathrm{U}_{d,u}$ are defined simultaneously inductively, while T_V, bdeg, subdeg, $\mathrm{T}_{\mathrm{U},d,u}$ are defined recursively. The constructors for the inductively defined sets are the constructors for the codes for the basic set constructions, deg, $\widehat{\mathrm{T}}_\mathrm{U}$, $\widehat{\mathrm{U}}_\mathrm{U}$, v, $\widehat{\mathrm{f}}_\mathrm{v}$, $\widehat{\mathrm{g}}_\mathrm{v}$, u, $\widehat{\mathrm{T}}_\mathrm{u}$, $\widehat{\mathrm{U}}_\mathrm{u}$, $\widehat{\mathrm{f}}_\mathrm{u}$, $\widehat{\mathrm{g}}_\mathrm{u}$.

Meaning explanations for the autonomous Mahlo universe. These meaning explanations will serve as a predicative justification of its constructive validity. We give simultaneously the meaning of Deg, V, Univ_d (for d in Deg) and $\mathrm{U}_{d,u}$ (for d in Deg and u in Univ_d). We will as well define simultaneously sets $\mathrm{T}_\mathrm{V}(x)$ for x in V, and $\mathrm{T}_{d,u}(a)$ for a in $\mathrm{U}_{d,u}$, and elements $\mathrm{bdeg}(d)$ in V for d in Deg and $\mathrm{subdeg}(d, a)$ for d in Deg and a in $\mathrm{T}_\mathrm{V}(\mathrm{bdeg}(d))$. Here $\mathrm{U}_{d,u}$ will refer only to a complete understanding of $\mathrm{U}_{d',u'}$ for some u' in $\mathrm{Univ}_{d'}$ and degrees d' introduced before introducing d. Otherwise it will only refer to $\mathrm{U}_{d',u'}$ for other degrees d' and to V by using closure properties known when introducing u.

$(\mathrm{V}, \mathrm{T}_\mathrm{V})$ and $(\mathrm{U}_{d,u}, \mathrm{T}_{d,u})$ will in addition to what is said below be closed under the basic set constructions.

Explanation of V: If d is in Deg, u is in Univ_d, then $\widehat{\mathrm{U}}_{\mathrm{U},d,u}$ is a canonical element of V and we define $\mathrm{T}_\mathrm{V}(\widehat{\mathrm{U}}_{\mathrm{U},d,u})$ as $\mathrm{U}_{d,u}$. Furthermore, if a is in $\mathrm{U}_{d,u}$, then $\widehat{\mathrm{T}}_{\mathrm{U},d,u}(a)$ is a canonical element of V and we define $\mathrm{T}_\mathrm{V}(\widehat{\mathrm{T}}_{\mathrm{U},d,u}(a))$ as $\mathrm{T}_{\mathrm{U},d,u}(a)$.

Explanation of Deg: If a is in V, b is a function mapping elements of $\mathrm{T}_\mathrm{V}(a)$ to Deg, then $\mathrm{deg}(a, b)$ is a canonical element of Deg and we define $\mathrm{bdeg}(\mathrm{deg}(a, b)) = a$, $\mathrm{subdeg}(\mathrm{deg}(a, b), x) = b(x)$. Note that as usual we leave it to the reader to explain when two canonical elements of Deg are equal and that bdeg and subdeg applied to equal canonical elements result in equal elements of V and Deg, respectively.

Explanation of Univ_d: Assume $(\mathrm{V}, \mathrm{T}_\mathrm{V})$ is closed under f, g and d is an element of Deg. Then $\mathrm{v}_{d,f,g}$ is a canonical element of Univ_d. Assume d is in Deg, u in Univ_d. Assume $(\mathrm{U}_{d,u}, \mathrm{T}_{d,u})$ is closed under f, g, a is in $\mathrm{T}_\mathrm{V}(\mathrm{bdeg}(d))$, and let $d_- := \mathrm{subdeg}(d, a)$. Then we have $\mathrm{u}_{d,u,a,f,g}$ is a canonical element of Univ_{d_-}.

Explanation of $\mathrm{U}_{d,u}$: As for the hyper$^\alpha$-Mahlo universe, we first give an explanation of the closure properties of $\mathrm{U}_{d,u}$ which apply to all elements u of

Univ_d: If the conditions for forming $u_- := \mathrm{u}_{d,u,a,f,g}$ above are fulfilled, then $\widehat{\mathrm{U}}_{\mathrm{u},d,u,a,f,g}$ is a canonical element of $\mathrm{U}_{d,u}$ and we define $\mathrm{T}_{d,u}(\widehat{\mathrm{U}}_{\mathrm{u},d,u,a,f,g})$ as U_{d_-,u_-}. Furthermore for a in U_{d_-,u_-} we define $\widehat{\mathrm{T}}_{\mathrm{u},d,u,a,f,g}(a)$ as a canonical element of $\mathrm{U}_{d,u}$ and define $\mathrm{T}_{d,u}(\widehat{\mathrm{T}}_{\mathrm{u},d,u,a,f,g}(a))$ as $\mathrm{T}_{d_-,u_-}(a)$.

Now we define the additional closure properties for specific elements $\mathrm{U}_{d,u}$: In case of $u = \mathrm{v}_{d,f,g}$ for which the conditions for introducing it are fulfilled we have that $\mathrm{U}_{d,u}$ is closed under f, g via $\widehat{\mathrm{T}}_{\mathrm{U},d,u}$ using constructors $\widehat{\mathrm{f}}_{\mathrm{v},d,f,g}, \widehat{\mathrm{g}}_{\mathrm{v},d,f,g}$. In case of $u = \mathrm{u}_{d,u,a,f,g}$ which is an element of Univ_{d_-} with d_- as stated above we have that $\mathrm{U}_{d_-,u}$ is closed under f, g via $\widehat{\mathrm{T}}_{\mathrm{u},d,u,a,f,g}$ using constructors $\widehat{\mathrm{f}}_{\mathrm{u},d,u,a,f,g}, \widehat{\mathrm{g}}_{\mathrm{u},d,u,a,f,g}$.

5 A Model of the Autonomous Mahlo Universe

5.1 Notations

We will in the following make use of the notations and basic principles for developing models as introduced in the sections on model constructions of [53]. We repeat here the most important ones:

- Terms are interpreted as simplified terms of the language (we throw away some typing information which is not relevant in the model). For a term r let $[\![r]\!]_\rho$ be the result of substituting free variables x by $\rho(x)$ in r and omitting the information to be thrown away. Let $[\![\mathrm{Term}]\!]$ be the set of terms.

- A set A in an environment ρ is interpreted as a partial equivalence relation (PER; i.e. a symmetric and transitive relation; see Def. 2.6(c) of [53]) $[\![A]\!]_\rho$, where $\langle a, b\rangle \in [\![A]\!]_\rho$ means that a and b are equal elements of $[\![A]\!]_\rho$. $\mathrm{Dom}([\![A]\!]_\rho)$ is the underlying set of terms, i.e. since $[\![A]\!]_\rho$ is a PER, $\mathrm{Dom}([\![A]\!]_\rho) = \{a \mid \langle a, a\rangle \in [\![A]\!]_\rho\}$. (In [53] we used Flat instead of Dom – we change the notation on suggestion by the referee, in order to be closer to standard notations for PER models.)

- We will sometimes for simplicity treat $[\![A]\!]_\rho$ as if it were a set of terms rather than a set of pairs of terms, and therefore identify $[\![A]\!]_\rho$ with $\mathrm{Dom}([\![A]\!]_\rho)$.

- We have a reduction relation on terms corresponding to the elimination rules of the type theory. All interpretations of sets will be in addition to what is stated in the following be closed under reductions.

- We have operations on PERs corresponding to the operations of the standard set constructions and of the small logical framework, where we write $[\![\to]\!]$, $[\![\times]\!]$, $[\![+]\!]$ infix. Especially we write $(x \in A)\, [\![\to]\!]\, B(x)$ for the set of functions mapping $a \in \mathrm{Dom}(A)$ to $\mathrm{Dom}(B(a))$, or more precisely the corresponding partial equivalence relation. $(x \in A)\, [\![\times]\!]\, B(x)$ stands for the PER corresponding to the set of pairs $\langle a, b\rangle$ for $a \in \mathrm{Dom}(A)$ and $b \in \mathrm{Dom}(B(a))$. $A \,[\![+]\!]\, B$ is to be understood in a similar way.

- $[\![\,\mathrm{IFam}\,]\!](U,T) := (x \in U) \,[\![\times]\!]\, (T(x) \,[\![\rightarrow]\!]\, U)$, which is the PER corresponding to the set of pairs $\langle a, b\rangle$ such that $a \in \mathrm{Dom}(U)$ and $b \in \mathrm{Dom}(T(a) \,[\![\rightarrow]\!]\, U)$.
- $[\![\,\mathrm{IFamOper}\,]\!]_0(U,T) := (x \in U) \,[\![\rightarrow]\!]\, (T(x) \,[\![\rightarrow]\!]\, U) \,[\![\rightarrow]\!]\, U$,
  $[\![\,\mathrm{IFamOper}\,]\!]_1(U,T,f) :=$
  $(x \in U) \,[\![\rightarrow]\!]\, (y \in (T(x) \,[\![\rightarrow]\!]\, U)) \,[\![\rightarrow]\!]\, T(f(x,y)) \,[\![\rightarrow]\!]\, U$,
  $[\![\,\mathrm{IFamOper}\,]\!](U,T) := (f \in [\![\,\mathrm{IFamOper}\,]\!]_0(U,T)) \,[\![\times]\!]\, [\![\,\mathrm{IFamOper}\,]\!]_1(U,T,f)$.

  $[\![\,\mathrm{IFamOper}\,]\!]_0(U,T)$ and $[\![\,\mathrm{IFamOper}\,]\!]_1(U,T,f)$ form the two components of the set of functions $[\![\,\mathrm{IFam}\,]\!](U,T) \,[\![\rightarrow]\!]\, [\![\,\mathrm{IFam}\,]\!](U,T)$.
- The closure of X under reductions is defined as

$$\mathrm{Clos}_{\rightarrow}(X) := \{\langle s,t\rangle \mid \exists \langle s',t'\rangle \in X \mid s \longrightarrow^* s' \wedge t \longrightarrow^* t'\} \ .$$

5.2 Autonomous Mahlo and Hyper$^\alpha$-Mahlo Ordinals

In order to define what it means for an ordinal to be recursively autonomous hyper$^\alpha$-Mahlo we first introduce the notion of an recursively hyper$^\alpha$-Mahlo. We will in the following usually omit the adverb "recursively" (or put it in brackets) which has to be added to all occurrences of hyper$^\alpha$-Mahlo and autonomous Mahlo sets and ordinals.

Remember the definition of a Mahlo ordinal from [53]: A *(recursively) Mahlo set* is an admissible set $\mathrm{a_M}$ such that

$$\begin{array}{ll}(\mathrm{Mahlo}) & \forall \vec{z} \in \mathrm{a_M}.((\forall x \in \mathrm{a_M}.\exists y \in \mathrm{a_M}.\varphi(x,y,\vec{z})) \\ & \quad \rightarrow \exists u \in \mathrm{a_M}.\mathrm{Ad}(u) \wedge \vec{z} \in u \wedge (\forall x \in u.\exists y \in u.\varphi(x,y,\vec{z}))) \\ & \text{for any } \Delta_0\text{-formula } \varphi(x,y,\vec{z}) \text{ such that } \mathrm{FV}(\varphi(x,y,\vec{z})) \subseteq \{x,y,\vec{z}\} \quad .\end{array}$$

Here $\mathrm{Ad}(u)$ means that u is an admissible set. A *(recursively) Mahlo ordinal* is an ordinal M such that $\mathrm{L_M}$ is a (recursively) Mahlo set.

Recursively Mahlo ordinals can be defined purely recursion theoretically ([21] Definition VIII.6.7 p. 422) as the set of ordinals κ such that for any $F : \kappa \rightarrow \kappa$ which is κ-recursive in parameters there exist a recursively regular (or admissible) ordinal $\lambda < \kappa$ s.t. $\forall \pi < \lambda. F(\pi) < \lambda$.

Recursively hyper$^\alpha$-Mahlo ordinal are obtained by stating instead that for any $\beta < \alpha$ any F as above can be reflected into a recursively hyper$^\beta$-Mahlo ordinal κ'. Since we want recursively hyper$^\alpha$-Mahlo ordinals to be Mahlo, we need to allow as well reflection into admissible sets.

(Note that we only demand κ' to be recursively hyper$^\beta$-Mahlo for one $\beta < \alpha$, not for all $\beta < \alpha$ simultaneously, which would give a slightly different notion – with that different notion a recursively autonomous Mahlo κ is not recursively hyper$^\kappa$-Mahlo, but only recursively hyper$^\alpha$-Mahlo for all $\alpha < \kappa$).

Moving back to the definitions in Kripke-Platek set theory, we define for ordinals α the collection of *(recursively) hyper$^\alpha$-Mahlo sets* hyper$^\alpha$MahloSet by recursion on α as follows: $\mathrm{hyper}^{\alpha}\mathrm{MahloSet}$ is defined by having the following property:

$$
\begin{array}{ll}
(\mathrm{HyperMahlo})^{\alpha} & a \in \mathrm{hyper}^{\alpha}\mathrm{MahloSet} \rightarrow \\
& \quad \mathrm{Ad}(a) \wedge \\
& \quad (\forall \vec{z} \in a.(\forall x \in a.\exists y \in a.\varphi(x,y,\vec{z})) \\
& \qquad \rightarrow ((\forall \beta < \alpha.\exists u \in a.u \in \mathrm{hyper}^{\beta}\mathrm{MahloSet} \wedge \vec{z} \in u \\
& \qquad\qquad\qquad\qquad \wedge (\forall x \in u.\exists y \in u.\varphi(x,y,\vec{z}))) \\
& \qquad\quad \wedge \exists u \in a.\mathrm{Ad}(u) \wedge \vec{z} \in u \\
& \qquad\qquad\qquad \wedge (\forall x \in u.\exists y \in u.\varphi(x,y,\vec{z})))) \\
& \text{for any } \Delta_0\text{-formula } \varphi(x,y,\vec{z}) \text{ such that} \\
& \mathrm{FV}(\varphi(x,y,\vec{z})) \subseteq \{x,y,\vec{z}\} \ .
\end{array}
$$

So a hyper^{α}-Mahlo means that a is admissible and any Π_2-property $\forall x \in a.\exists y \in a.\varphi(x,y,\vec{z})$ can be reflected for any $\beta < \alpha$ into a hyper^{β}-Mahlo set in a and be reflected into an admissible set in a. Especially $\mathrm{hyper}^{0}\mathrm{MahloSet}$ are the (recursively) Mahlo sets. A *(recursively) Hyper$^{\alpha}$-Mahlo-ordinal* is an ordinal κ such that L_{κ} is a (recursively) Hyper^{α}-Mahlo set.

Now an ordinal κ is (recursively) autonomous Mahlo (we say in the following briefly autonomous Mahlo) if it is (recursively) hyper^{κ}-Mahlo.

The theory $(\mathrm{KP}\omega + (\mathrm{Aut} - \mathrm{Mahlo}))^{+}$ of Kripke Platek set theory plus one (recursively) autonomous Mahlo ordinal and finitely many admissibles above is defined as follows:

- The language is that of KPi as in [13], and in addition a binary predicate $u \in \mathrm{hyper}^{\beta}\mathrm{MahloSet}$ (depending on u, β), a constant AutMahloSet for an admissible which is autonomous Mahlo and for $n \in \mathbb{N}$ constants a_n for the nth admissible above AutMahlo.
- We have the axioms of KPi as in [13], p. 124, but excluding $(\Delta_0 - \mathrm{Col})$ and (Lim).
- We have axioms $(\mathrm{HyperMahlo})^{\alpha}$ (uniform in α, for any formula φ stated there).
- We have axioms

  $$\mathrm{Ad}(\mathrm{AutMahloSet}) \wedge \forall \alpha \in \mathrm{AutMahloSet}.\mathrm{AutMahloSet} \in \mathrm{hyper}^{\alpha}\mathrm{MahloSet} \ .$$

- We have for $n \in \mathbb{N}$ axioms

  $$\mathrm{Ad}(\mathrm{a}_n) \wedge \mathrm{AutMahloSet} \in \mathrm{a}_n \wedge \mathrm{a}_n \in \mathrm{a}_{n+1} \ .$$

5.3 The Basic Idea of the Model

The general idea for constructing models is as in Subsect. 3.2 of [53]: One defines by recursion on α $(\mathrm{V}^{\alpha}, \mathrm{T}^{\alpha}) \in [\![\mathrm{Fam}]\!](\mathrm{Set})$ fulfilling Assumption 3.2. of [53] by closing V^{α} under the basic set constructions and adding codes for universes (i.e. elements of Univ_d), whenever for some $\beta < \alpha$ we have that $(\mathrm{V}^\beta, \mathrm{T}^\beta)$ is suitable as an interpretation for this universe. The added universe is then interpreted as V^β for the minimal such β. Then $[\![\mathrm{V}]\!] := \mathrm{V}^{<\mathrm{AutM}}$ and $[\![\mathrm{T}(a)]\!]_\rho := \mathrm{T}^{<\mathrm{AutM}}([\![a]\!]_\rho)$ for a (recursively) autonomous Mahlo ordinal AutM.

Therefore the main task is to define when $(\mathrm{V}^\beta, \mathrm{T}^\beta)$ is sufficiently closed so that it can be used as an interpretation of an element of Univ_d.

One problem when dealing with the autonomous Mahlo universe is the fact that the set Deg seems to depend on the whole universe we are defining. But we have observed above that it only depends locally on it, so each element of Deg depends only on locally small set of elements of the autonomous Mahlo universe. So we will be able to define approximations Deg^α of $[\![\mathrm{Deg}]\!]$ simultaneously while defining V^α, T^α, and define $[\![\mathrm{Deg}]\!] := \mathrm{Deg}^{<\mathrm{AutM}} := \bigcup_{\alpha<\mathrm{AutM}} \mathrm{Deg}^\alpha$, and refer, when defining V^α only to elements of $\mathrm{Deg}^{<\alpha}$.

As for the Mahlo universe (defined in [53]) we have the problem that we know the set of functions from families of elements in $[\![\mathrm{V}]\!]$ to families of elements in $[\![\mathrm{V}]\!]$ only, when the definition of $[\![\mathrm{V}]\!]$ is complete. The solution there was to add $\widehat{\mathrm{U}}_{f,g}$ without knowing that f and g are defined on all of $[\![\mathrm{V}]\!]$ – all we needed was that they are defined on $\mathrm{U}_{f,g}$. Similarly, we will add in the construction of the autonomous Mahlo universe codes $\mathrm{v}_{d,f,g}$ to V^α using only the fact that f, g are locally defined and that $d \in \mathrm{Deg}^{<\alpha}$. In fact we will only consider codes for subuniverses of the form $\mathrm{v}_{d,f,g}$.

All other occurring universes will be modelled in this universe by codes of the form $\mathrm{v}_{d,f,g}$. Especially we will add a reduction rule $\mathrm{u}_{d,u,a,f,g} \longrightarrow \mathrm{v}_{\mathrm{d}_-,f,g}$ where $\mathrm{d}_- = \mathrm{subdeg}(d, a)$. This means of course that that if in the model we have $u = \mathrm{v}_{d',f,g} \in \mathrm{V}^\alpha$, it is not necessarily the case that f, g are total on V: if u represents an element $\mathrm{u}_{d'',u,a,f,g}$ in the type theory, then we know only that f, g are total on $\mathrm{U}_{d'',u}$, not on V. However, the type theory guarantees that $d' = \mathrm{subdeg}(d'', a) : \mathrm{Deg}$, therefore in the model d' is an element of $[\![\mathrm{Deg}]\!]$ – roughly speaking d' is a "total degree".

One could have defined a more sophisticated model by defining at each stage α the least subuniverse of $\mathrm{V}^{<\alpha}$ closed under the constructions needed as an interpretation of $u := \mathrm{v}_{d,f,g}$, and, if this set is sufficiently closed, take this set as an interpretation of $\mathrm{U}_{d,u}$. This would be technically more complicated and we use the current definition, since it is simpler and gives a model.

$[\![\mathrm{Univ}_d]\!]_\rho$ will be essentially the set of $\mathrm{v}_{d',f,g}$ (where $d' = [\![d]\!]_\rho$) occurring in $[\![\mathrm{V}]\!]$. (We have to vary this definition slightly in order to accommodate for the rule stating that $d = d' : \mathrm{Deg}$ implies $\mathrm{Univ}_d = \mathrm{Univ}_{d'} : \mathrm{Set}$.)

5.4 Formal Definition of the Model

We will work in $(\mathrm{KP}\omega + (\mathrm{Aut} - \mathrm{Mahlo}))^+$ as defined above.

The set of terms in the model. As for all models of type theory in [53], we will omit in $[\![\mathrm{Term}]\!]$ the dependency of $\widehat{\mathrm{N}}$, $\widehat{\Sigma}$ on V, U, d, u, and have for instance $[\![\widehat{\mathrm{N}}_\mathrm{V}]\!]_\rho := [\![\widehat{\mathrm{N}}_{\mathrm{U},d,u}]\!]_\rho := \widehat{\mathrm{N}}$.

Univ_d, $\mathrm{U}_{d,u}$ will be interpreted as subsets of $[\![\,\mathrm{V}\,]\!]$, and in the model we treat inductive subuniverses as if they were recursive subuniverses and we could omit the recursive embedding altogether, i.e. as if they were subsets. Note that in the original type theory, the embedding $\widehat{T}$ of an inductive subuniverse (U, T_{U}) into its encompassing universe (V, T_{V}) is a constructor. For instance, let $\widehat{\mathrm{N}}_{\mathrm{U}}$ and $\widehat{\mathrm{N}}_{\mathrm{V}}$ be the codes for the natural numbers in U and V. Then in the type theory we don't have $\widehat{T}(\widehat{\mathrm{N}}_{\mathrm{U}}) = \widehat{\mathrm{N}}_{\mathrm{V}}$ (however, we cannot prove inside type theory that they are not equal). All we have is that the sets they denote are the same, i.e. $T_{\mathrm{V}}(\widehat{\mathrm{N}}_{\mathrm{V}}) = T_{\mathrm{V}}(\widehat{T}(\widehat{\mathrm{N}}_{\mathrm{U}}))$. In the model we have instead that $\widehat{\mathrm{N}}_{\mathrm{V}}$ and $\widehat{T}(\widehat{\mathrm{N}}_{\mathrm{U}})$ are equal, and we will in the model omit the embedding $\widehat{T}$ altogether. Therefore $\widehat{\mathrm{T}}_{\mathrm{U},d,u}$, $\widehat{\mathrm{T}}_{\mathrm{u},d,u,a,f,g}$ will be treated as if they were identity functions, and $\widehat{\mathrm{U}}_{\mathrm{U},d,u}$, $\widehat{\mathrm{U}}_{\mathrm{u},d,u,a,f,g}$ will be identified with u and $\mathrm{u}_{d,u,a,f,g}$, respectively. So we have the following reduction rules:

$$
\begin{array}{rcl}
\widehat{\mathrm{T}}_{\mathrm{U},d,u}(a) & \longrightarrow & a \ , \\
\widehat{\mathrm{T}}_{\mathrm{u},d,u,a,f,g}(x) & \longrightarrow & x \ , \\
\widehat{\mathrm{U}}_{\mathrm{U},d,u} & \longrightarrow & u \ , \\
\widehat{\mathrm{U}}_{\mathrm{u},d,u,a,f,g} & \longrightarrow & \mathrm{u}_{d,u,a,f,g} \ .
\end{array}
$$

The reduction rules for bdeg and subdeg are as given by their equality rules, i.e.

$$
\begin{array}{rcl}
\mathrm{bdeg}(\mathrm{deg}(r,s)) & \longrightarrow & r \ , \\
\mathrm{subdeg}(\mathrm{deg}(r,s),a) & \longrightarrow & s(a) \ .
\end{array}
$$

$\widehat{\mathrm{f}}$ and $\widehat{\mathrm{g}}$ will be interpreted by their underlying functions, i.e. we have

$$
\begin{array}{rcl}
\widehat{\mathrm{f}}_{\mathrm{v},d,f,g}(x,y) & \longrightarrow & f(x,y) \ , \\
\widehat{\mathrm{f}}_{\mathrm{u},d,u,a,f,g}(x,y) & \longrightarrow & f(x,y) \ , \\
\widehat{\mathrm{g}}_{\mathrm{v},d,f,g}(x,y,z) & \longrightarrow & g(x,y,z) \ , \\
\widehat{\mathrm{g}}_{\mathrm{u},d,u,a,f,g}(x,y,z) & \longrightarrow & g(x,y,z) \ .
\end{array}
$$

(We have to be careful that this causes no problems with the equality rules for $\mathrm{T}_{\mathrm{U},d,u}$, $\mathrm{T}_{\mathrm{u},d,a,f,g}$, but one sees immediately that there are indeed no problems.)

Furthermore $\mathrm{u}_{d,u,a,f,g}$ will be identified with $\mathrm{v}_{\mathrm{subdeg}(d,a),f,g}$, so we have

$$
\mathrm{u}_{d,u,a,f,g} \quad \longrightarrow \quad \mathrm{v}_{\mathrm{subdeg}(d,a),f,g} \ .
$$

Closure properties of V^{β}: The crucial closure property of V^{α} apart from closing $\mathrm{V}^{<\alpha}$ under one step basic set constructions is when to add $\mathrm{v}_{d,f,g}$ to V^{α}: We add it, if $d \in \mathrm{Deg}^{<\alpha}$ and we have at a previous stage $\beta < \alpha$ found that $\mathrm{V}^{<\beta}$ has the closure properties needed so that it can be used as an interpretation of $\mathrm{v}_{d,f,g}$. This means that

- $\mathrm{V}^{<\beta}$ is closed under f, g.
- $\mathrm{V}^{<\beta}$ contains $\mathrm{bdeg}(d)$.

- $V^{<\beta}$ is closed under subuniverses relative to the subdegrees formed from the branching degrees just mentioned. This means that for any operator $\widetilde{f}, \widetilde{g}$ on families of sets internal to $V^{<\beta}$ and any $a \in T^{<\beta}(\mathrm{bdeg}(d))$ we have with $\mathrm{d}_{-} := \mathrm{subdeg}(d, a)$ that $\mathrm{v}_{\mathrm{d}_{-}, \widetilde{f}, \widetilde{g}}$ is in $V^{<\beta}$.

This means that $V^{<\beta}$ is sufficiently closed to serve as an interpretation of $\mathrm{v}_{d,f,g}$, and we interpret $\mathrm{v}_{d,f,g}$ as $V^{<\beta}$ for the minimal such β.

Since we are working in a PER model, we have to replace in the definitions reference to terms as elements of the sets by pairs of terms as elements of those sets.

Definition of V^{α}, T^{α}, Deg^{α}. As for the super and Mahlo universe in [53], we define $(V^{\alpha}, T^{\alpha}) \in [\![\mathrm{Fam}]\!](\mathrm{Set})$ by closing it under the one step standard set constructions. We will then interpret $[\![\mathrm{V}]\!] = V^{<\mathrm{AutM}}$, $[\![\mathrm{T_V}]\!] = T^{<\mathrm{AutM}}$.

In order to define when to add $\mathrm{v}_{d,f,g}$ we first introduce what it means for a family of sets to be closed under d:

Definition 5.1 *Let* $(U, T) \in [\![\mathrm{Fam}]\!](\mathrm{Set})$, $\langle d, d' \rangle \in [\![\mathrm{Term}]\!]^2$. (U, T) *is downward closed under* d, d'*, written as* $\mathrm{degClosure}(U, T, d, d')$*, if the following holds:*

(1) U is closed under $\langle \mathrm{bdeg}(d), \mathrm{bdeg}(d') \rangle$*, i.e.*

$$\langle \mathrm{bdeg}(d), \mathrm{bdeg}(d') \rangle \in U \quad .$$

(2) U is closed under the formation of subuniverses for subdegrees of d, d':

Assume $a, a', \widetilde{f}, \widetilde{f}', \widetilde{g}, \widetilde{g}' \in [\![\mathrm{Term}]\!]$ *such that the following holds:*

- $\langle a, a' \rangle \in T(\mathrm{bdeg}(d))$,
- $\mathrm{d}_{-} := \mathrm{subdeg}(d, a)$, $\mathrm{d}_{-}' := \mathrm{subdeg}(d', a')$,
- *U is closed under* $\widetilde{f}, \widetilde{f}', \widetilde{g}, \widetilde{g}'$*, i.e.*

$$\langle \langle \widetilde{f}, \widetilde{g} \rangle, \langle \widetilde{f}', \widetilde{g}' \rangle \rangle \in [\![\mathrm{IFamOper}]\!](U, T) \quad .$$

Then

$$\langle \mathrm{v}_{\mathrm{d}_{-}, \widetilde{f}, \widetilde{g}}, \mathrm{v}_{\mathrm{d}_{-}', \widetilde{f}', \widetilde{g}'} \rangle \in U \quad .$$

Assume now an ordinal α and $d, d,' f, f', g, g' \in [\![\mathrm{Term}]\!]$. Assume β s.t. $\beta + 1 < \alpha$. Assume

- $(V^{<\beta}, T^{<\beta})$ is closed under the universe constructions.
- $(V^{<\beta}, T^{<\beta})$ is closed under f, f', g, g':

$$\langle \langle f, g \rangle, \langle f', g' \rangle \rangle \in [\![\mathrm{IFamOper}]\!](V^{<\beta}, T^{<\beta}) \quad .$$

- $\langle d, d' \rangle \in \mathrm{Deg}^{<\alpha}$.
- $\mathrm{degClosure}(V^{<\beta}, T^{<\beta}, d, d')$.

We call the above conditions the "conditions for possibly interpreting $[\![\mathrm{T}]\!](\mathrm{v}_{d,f,g})$ as $\mathrm{V}^{<\beta}$." Let β be minimal such that the above holds. Then

$$\langle \mathrm{v}_{d,f,g}, \mathrm{v}_{d',f',g'} \rangle \in \mathrm{V}^{\alpha} \ , \qquad \mathrm{T}^{\alpha}(\mathrm{v}_{d,f,g}) := \mathrm{V}^{<\beta} \ .$$

Furthermore Deg^{α}, $\mathrm{Deg}^{<\alpha}$ are defined as follows simultaneously with defining V^{α}, T^{α}:

$$\begin{array}{lll} \mathrm{Deg}^{<\alpha} & := & \bigcup_{\beta<\alpha} \mathrm{Deg}^{\beta} \ , \\ \mathrm{Deg}^{\alpha} & := & \mathrm{Clos}_{\rightarrow}(\{\langle \deg(r,s), \deg(r',s') \rangle \mid \langle r, r' \rangle \in \mathrm{V}^{<\alpha} \wedge \langle s, s' \rangle \in \mathrm{T}^{<\alpha}(r) \, [\![\rightarrow]\!] \, \mathrm{Deg}^{<\alpha}\}) \ . \end{array}$$

Interpretation of the other sets. We define now for an autonomous Mahlo ordinal AutM

$$\begin{array}{rcl} [\![\mathrm{V}]\!]_{\rho} & := & \mathrm{V}^{<\mathrm{AutM}} \ , \\ [\![\mathrm{T}_{\mathrm{V}}(a)]\!]_{\rho} & := & [\![\mathrm{T}_{\mathrm{V}}]\!]([\![a]\!]_{\rho}) := \mathrm{T}^{<\mathrm{AutM}}([\![a]\!]_{\rho}) \ , \\ [\![\mathrm{Deg}]\!]_{\rho} & := & \mathrm{Deg}^{<\mathrm{AutM}} \ , \end{array}$$

and then

$$\begin{array}{rcl} [\![\mathrm{Univ}_d]\!]_{\rho} & := & \mathrm{Clos}_{\rightarrow}(\{\langle \mathrm{v}_{d'',f,g}, \mathrm{v}_{d''',f',g'} \rangle \in [\![\mathrm{Term}]\!]^2 \mid \\ & & \qquad \langle \mathrm{v}_{d',f,g}, \mathrm{v}_{d'',f,g} \rangle \in [\![\mathrm{V}]\!] \wedge \langle \mathrm{v}_{d'',f,g}, \mathrm{v}_{d''',f',g'} \rangle \in [\![\mathrm{V}]\!]\}) \ , \\ & & \text{where } d' = [\![d]\!]_{\rho}. \\ [\![\mathrm{U}_{d,u}]\!]_{\rho} & := & [\![\mathrm{T}_{\mathrm{V}}]\!]([\![u]\!]_{\rho}) \ . \\ [\![\mathrm{T}_{\mathrm{U},d,u}(a)]\!]_{\rho} & := & [\![\mathrm{T}_{\mathrm{V}}]\!]([\![a]\!]_{\rho}) \ . \end{array}$$

As for all previous models we have $[\![\mathrm{V}]\!] \in \mathrm{L}_{\mathrm{AutM}+1}$ and $[\![\mathrm{T}_{\mathrm{V}}]\!](a) \in \mathrm{L}_{\mathrm{AutM}}$ for $a \in [\![\mathrm{V}]\!]$ and define therefore

$$\mathrm{o}(\mathrm{U}) := \mathrm{AutM} + 1 \ , \ \ \mathrm{o}(\mathrm{T}_{\mathrm{V}}(a)) := \mathrm{AutM} \ .$$

The interpretation of the basic set constructions will be defined as in the previously introduced models of type theories, and this will require the use of finitely many admissibles on top of AutM.

We can easily see see that Assumption 3.2 of [53] holds.

5.5 Correctness of the Model

Basic correctness. The crucial part of the correctness proof for the model model will be to show that $[\![\mathrm{Deg}]\!]$ is closed under the introduction rule for Deg – this is where we use the fact that AutM is an autonomous Mahlo ordinal.

A minor complication arises as well when showing that, if d, d' are equal elements of $[\![\mathrm{Deg}]\!]$, then $[\![\mathrm{Univ}_d]\!]$ and $[\![\mathrm{Univ}_{d'}]\!]$ are equal. Before verifying this, we show the correctness of the other rules:

- The correctness of all equality rules follow by the corresponding reduction rules.

- We have that $[\![\,\mathrm{bdeg}\,]\!] \in [\![\,\mathrm{Deg}\,]\!]\,[\![\,\to\,]\!]\,[\![\,\mathrm{V}\,]\!]$, and for $\langle d, d' \rangle \in [\![\,\mathrm{Deg}\,]\!]$ and $\langle a, a' \rangle \in [\![\,\mathrm{T}\,]\!](\mathrm{bdeg}(d))$ we have that $\langle \mathrm{subdeg}(d, a), \mathrm{subdeg}(d', a') \rangle \in [\![\,\mathrm{Deg}\,]\!]$:

  Assume $\langle d, d' \rangle \in [\![\,\mathrm{Deg}\,]\!]$. Then $\langle d, d' \rangle \in \mathrm{Deg}^{\alpha}$ for some α. But then $d \longrightarrow \mathrm{deg}(r, s)$, $d' \longrightarrow \mathrm{deg}(r', s')$ s.t. $\langle r, r' \rangle \in \mathrm{V}^{<\alpha}$ and $\langle s, s' \rangle \in \mathrm{T}^{<\alpha}(r)\,[\![\,\to\,]\!]\,\mathrm{Deg}^{<\alpha}$. $\mathrm{bdeg}(d) \longrightarrow r$ $\mathrm{bdeg}(d') \longrightarrow r'$, therefore $\langle \mathrm{bdeg}(d), \mathrm{bdeg}(d') \rangle \in \mathrm{V}^{<\alpha} \subseteq [\![\,\mathrm{V}\,]\!]$. Furthermore assume $\langle a, a \rangle \in [\![\,\mathrm{T}\,]\!](\mathrm{bdeg}(d)) = \mathrm{T}^{<\alpha}(\mathrm{bdeg}(d))$. We have $\mathrm{subdeg}(d, a) \longrightarrow s(a)$, $\mathrm{subdeg}(d', a') \longrightarrow s'(a')$, $\langle s(a), s'(a') \rangle \in \mathrm{Deg}^{<\alpha} \subseteq [\![\,\mathrm{Deg}\,]\!]$, and therefore as well $\langle \mathrm{subdeg}(d, a), \mathrm{subdeg}(d', a') \rangle \in [\![\,\mathrm{Deg}\,]\!]$.

- We show that $[\![\,\mathrm{Deg}\,]\!]$ is closed under the introduction rule. Assume $\langle r, r' \rangle \in [\![\,\mathrm{V}\,]\!]$, and $\langle s, s' \rangle \in [\![\,\mathrm{T}\,]\!](r)\,[\![\,\to\,]\!]\,[\![\,\mathrm{Deg}\,]\!]$. Then $\langle r, r' \rangle \in \mathrm{V}^{<\alpha}$ for some $\alpha < \mathrm{AutM}$. $[\![\,\mathrm{T}\,]\!](r) = \mathrm{T}^{<\alpha}(r) \in \mathrm{L}_{\mathrm{AutM}}$, $\forall \langle a, a' \rangle \in \mathrm{T}^{<\alpha}(r).\exists \beta < \mathrm{AutM}.\langle s(a), s'(a') \rangle \in \mathrm{Deg}^{\beta}$. By AutM admissible there exists $\gamma < \mathrm{AutM}$ s.t. $\alpha < \gamma$ and the β just given can always to be chosen $< \gamma$. But then it follows $\langle \mathrm{deg}(r, s), \mathrm{deg}(r', s') \rangle \in \mathrm{Deg}^{\gamma} \subseteq [\![\,\mathrm{Deg}\,]\!]$.

- $[\![\,\mathrm{V}\,]\!]$ is a universe: One easily verifies that Assumption 3.2 of [53] is fulfilled with U replaced by V. Therefore, by Theorem 3.3 of [53] and the fact that AutM is a recursively inaccessible we get that $([\![\,\mathrm{V}\,]\!], [\![\,\mathrm{T_V}\,]\!])$ is closed under the universe constructions.

- That $([\![\,\mathrm{U}_{d,u}\,]\!]_\rho, [\![\,\mathrm{T}_{\mathrm{U},d,u}\,]\!]_\rho)$ are closed under the universe constructions follows by the construction. $[\![\,\mathrm{U}_{d,u}\,]\!]_\rho \subseteq [\![\,\mathrm{V}\,]\!]$, $\widehat{\mathrm{T}}_{\mathrm{U},d,u}(x) \longrightarrow x$ and $\widehat{\mathrm{U}}_{d,y} \longrightarrow u \in [\![\,\mathrm{V}\,]\!]$, therefore $([\![\,\mathrm{U}_{d,u}\,]\!]_\rho, [\![\,\mathrm{T}_{\mathrm{U},d,u}\,]\!]_\rho)$ fulfils the rules for being an inductive subuniverse of V represented in V.

- We prove the correctness of the introduction rule introducing $\mathrm{v}_{d,f,g}$. More generally we prove the following:

  - Assume $\alpha \in \mathrm{Ord}$, $\langle d, d' \rangle \in \mathrm{Deg}^{\alpha}$. Assume $\kappa \in \mathrm{Ord}$, κ hyper$^{\alpha}$-Mahlo, $\langle \langle f, g \rangle, \langle f', g' \rangle \rangle \in [\![\,\mathrm{IFamOper}\,]\!](\mathrm{V}^{<\kappa}, \mathrm{T}^{<\kappa})$. Then $\langle \mathrm{v}_{d,f,g}, \mathrm{v}_{d',f',g'} \rangle \in \mathrm{V}^{\kappa+2}$.

  If we have proved that statement above then the assertion follows: If $\langle d, d' \rangle \in [\![\,\mathrm{Deg}\,]\!]$, then $\langle d, d' \rangle \in \mathrm{Deg}^{\alpha}$ some $\alpha < \mathrm{AutM}$. AutM is hyper$^{\alpha+1}$-Mahlo, so if $\langle \langle f, g \rangle, \langle f', g' \rangle \rangle \in [\![\,\mathrm{IFamOper}\,]\!]([\![\,\mathrm{V}\,]\!], [\![\,\mathrm{T}\,]\!])$, then $\langle \langle f, g \rangle, \langle f', g' \rangle \rangle \in [\![\,\mathrm{IFamOper}\,]\!](\mathrm{V}^{<\kappa}, \mathrm{T}^{<\kappa})$ for some κ which is hyper$^{\alpha}$-Mahlo. Here we use the fact that $\langle \langle f, g \rangle, \langle f', g' \rangle \rangle \in [\![\,\mathrm{IFamOper}\,]\!]([\![\,\mathrm{V}\,]\!], [\![\,\mathrm{T}\,]\!])$ can be expressed as a Π_2-formula in AutM (similarly as it was the case in case of the Mahlo universe). But then $\langle \mathrm{v}_{d,f,g}, \mathrm{v}_{d',f',g'} \rangle \in \mathrm{V}^{\kappa+2} \subseteq [\![\,\mathrm{V}\,]\!]$.

  We prove now the statement above by induction over α.

  Let $r := \mathrm{bdeg}(d)$, $r' := \mathrm{bdeg}(d')$, $s(x) := \mathrm{subdeg}(d, x)$, $s'(x) := \mathrm{subdeg}(s', x)$. We have $\langle r, r' \rangle \in \mathrm{U}^{<\alpha} \subseteq \mathrm{U}^{<\kappa}$ since $\alpha \leq \kappa$, and for $\langle x, x' \rangle \in \mathrm{T}^{<\alpha}(r)$ we have that $\langle s(x), s'(x') \rangle \in \mathrm{Deg}^{<\alpha}$. In order to show that $\langle \mathrm{v}_{d,f,g}, \mathrm{v}_{d',f',g'} \rangle \in \mathrm{V}^{\kappa+2}$ it suffices to show that the conditions for possibly interpreting

$[\![\, \mathrm{T} \,]\!](\mathrm{v}_{d,f,g})$ as $\mathrm{V}^{<\kappa}$ are fulfilled. Since κ is Mahlo, it is inaccessible and therefore $(\mathrm{V}^{<\kappa}, \mathrm{T}^{<\kappa})$ is closed under the universe constructions. It is by assumption closed under f, f', g, g'. By $\alpha \leq \kappa$ we get $\langle d, d' \rangle \in \mathrm{Deg}^{<\kappa}$. So we need to show $\mathrm{degClosure}(\mathrm{V}^{<\kappa}, \mathrm{T}^{<\kappa}, d, d')$. $\langle r, r' \rangle \in \mathrm{U}^{<\kappa}$. Assume $a, a', \widetilde{f}, \widetilde{f}', \widetilde{g}, \widetilde{g}'$ s.t. $\langle a, a' \rangle \in \mathrm{T}^{<\kappa}(r)$, $d_- := \mathrm{subdeg}(d, a)$, $d'_- := \mathrm{subdeg}(d', a')$, and $\langle\langle \widetilde{f}, \widetilde{g} \rangle, \langle \widetilde{f}', \widetilde{g}' \rangle\rangle \in [\![\, \mathrm{IFamOper} \,]\!](\mathrm{U}^{<\kappa}, \mathrm{T}^{<\kappa})$. The above can be expressed as a Π_2-formula. $\langle d_-, d'_- \rangle \in \mathrm{Deg}^{<\alpha}$. Let $\langle d_-, d'_- \rangle \in \mathrm{Deg}^{\beta}$ for $\beta < \alpha$. By κ being hyper$^{\alpha}$-Mahlo there exists a $\kappa' < \kappa$ which is hyper$^{\beta}$-Mahlo, s.t. $\langle\langle \widetilde{f}, \widetilde{g} \rangle, \langle \widetilde{f}', \widetilde{g}' \rangle\rangle \in [\![\, \mathrm{IFamOper} \,]\!](\mathrm{U}^{<\kappa'}, \mathrm{T}^{<\kappa'})$. By IH $\langle \mathrm{v}_{d_-,\widetilde{f},\widetilde{g}}, \mathrm{v}_{d'_-,\widetilde{f}',\widetilde{g}'} \rangle \in \mathrm{U}^{\kappa'+2} \subseteq \mathrm{U}^{<\kappa}$.

- The closure of $[\![\, \mathrm{U}_{d,\mathrm{v}_{d,f,g}} \,]\!]_\rho$ under $\widehat{\mathrm{f}}_{\mathrm{v},d,f,g}$, $\widehat{\mathrm{g}}_{\mathrm{v},d,f,g}$, follows by the construction and the reduction rules for $\widehat{\mathrm{f}}$, $\widehat{\mathrm{g}}$.
- Similarly follows, with $\mathrm{d}_- := \mathrm{subdeg}(d, a)$ the closure of $[\![\, \mathrm{U}_{\mathrm{d}_-,\mathrm{u}_{d,u,a,f,g}} \,]\!]_\rho$ under $\widehat{\mathrm{f}}_{\mathrm{u},d,u,a,f,g}$, $\widehat{\mathrm{g}}_{\mathrm{u},d,u,a,f,g}$.
- The correctness of the introduction rule introducing $\mathrm{u}_{d,u,a,f,g}$ follows by the fact that $u \in \mathrm{Dom}([\![\, \mathrm{Univ}_d \,]\!]_\rho)$ must reduce to $\mathrm{v}_{d'',\widetilde{f},\widetilde{g}}$ where $\langle \mathrm{v}_{d',\widetilde{f},\widetilde{g}}, \mathrm{v}_{d'',\widetilde{f},\widetilde{g}} \rangle \in [\![\, \mathrm{V} \,]\!]$ $(d' := [\![\, d \,]\!]_\rho)$, that therefore $[\![\, \mathrm{U}_{d,u} \,]\!]_\rho = [\![\, \mathrm{T}_\mathrm{V} \,]\!](\mathrm{v}_{d',\widetilde{f},\widetilde{g}}) = [\![\, \mathrm{T}_\mathrm{V} \,]\!](\mathrm{v}_{d'',\widetilde{f},\widetilde{g}})$ is closed under the formation of subuniverses (interpreting $\mathrm{v}_{d_-,f,g}$ for subdegrees d_- of d), and that $\mathrm{u}_{d,u,a,f,g} \longrightarrow \mathrm{v}_{d_-,f,g}$ for $d_- = \mathrm{subdeg}(d, a)$.
- That with $u_- := \mathrm{u}_{d,u,a,f,g}$ and $d_- := \mathrm{subdeg}(d, a)$ we have that (under the assumptions given by the rules) the correctness of the rules expressing that $(\mathrm{U}_{d_-,u_-}, \mathrm{T}_{d_-,u_-})$ is an inductive subuniverse of $(\mathrm{U}_{d,u}, \mathrm{T}_{d,u})$ holds follows since u must reduce to some $\mathrm{v}_{d,\widetilde{f},\widetilde{g}}$, and $\mathrm{U}_{d,u}$ was interpreted as some $\mathrm{V}^{<\alpha}$ which contained $[\![\, \mathrm{U}_{d_-,u_-} \,]\!]$ as a subset and u_- as an element (the argument is essentially the same as verifying the correctness of $\mathrm{U}_{d,u}$ being an inductive subuniverse of V above).

Correctness of the equality rule for Univ_d. We show the correctness of the equality rule for Univ_d, which expresses that if $d = d' : \mathrm{Deg}$ then $\mathrm{Univ}_d = \mathrm{Univ}_{d'} : \mathrm{Set}$. First we see easily by main induction on α and side-induction on β that the following small lemma holds:

$$\langle a, b \rangle \in \mathrm{V}^{\alpha} \to \forall \beta < \alpha (\langle a, a \rangle \in \mathrm{V}^{\beta} \vee \langle b, b \rangle \in \mathrm{V}^{\beta}) \to \langle a, b \rangle \in \mathrm{V}^{\beta} \quad .$$

Now we show the following statement:

$$\begin{array}{l} \forall \alpha < \mathrm{AutM}. \forall d, d', f, g, a \in [\![\, \mathrm{Term} \,]\!]. \forall \langle d, d' \rangle \in [\![\, \mathrm{Deg} \,]\!]. \\ \quad (\langle \mathrm{v}_{d,f,g}, a \rangle \in \mathrm{V}^{\alpha} \to \langle \mathrm{v}_{d',f,g}, a \rangle \in \mathrm{V}^{\alpha}) \wedge \\ \quad (\langle a, \mathrm{v}_{d,f,g} \rangle \in \mathrm{V}^{\alpha} \to \langle a, \mathrm{v}_{d',f,g} \rangle \in \mathrm{V}^{\alpha}) \quad . \end{array}$$

Then, since $[\![\, \mathrm{V} \,]\!]$ is symmetric and transitive, we obtain that $\langle d, d' \rangle \in [\![\, \mathrm{Deg} \,]\!]$ implies $[\![\, \mathrm{Univ}_d \,]\!] = [\![\, \mathrm{Univ}_{d'} \,]\!]$.

The statement is shown by induction on α. We need to show only the first half, the second half follows by the symmetry of V^α (the symmetry is part of Assumption 3.2 of [53]). Assume α, $\langle d, d'\rangle \in [\![\mathrm{Deg}]\!]$, $\langle \mathrm{v}_{d,f,g}, a\rangle \in \mathrm{V}^\alpha$, and that the assertion holds for $\beta < \alpha$. Then $a \longrightarrow \mathrm{v}_{d'',f',g'}$ for some d'', f', g'. Let $\mathrm{T}^\alpha(\mathrm{v}_{d,f,g}) = \mathrm{V}^{<\beta}$ with $\beta + 1 < \alpha$.

Therefore $\mathrm{V}^{<\beta}$ is closed under the universe constructions,

$$\langle\langle f, g\rangle, \langle f', g'\rangle\rangle \in [\![\mathrm{IFamOper}]\!](\mathrm{V}^{<\beta}, \mathrm{T}^{<\beta}) \ ,$$

and we have $\mathrm{degClosure}(\mathrm{V}^{<\beta}, \mathrm{T}^{<\beta}, d, d'')$. We have that $\langle d, d'\rangle \in [\![\mathrm{Deg}]\!]$ implies $\langle \mathrm{bdeg}(d), \mathrm{bdeg}(d')\rangle \in [\![\mathrm{V}]\!]$ and for $\langle a, a'\rangle \in [\![\mathrm{T_V}]\!](\mathrm{bdeg}(d))$ we have $\langle \mathrm{subdeg}(d, a), \mathrm{subdeg}(d', a')\rangle \in [\![\mathrm{Deg}]\!]$. But then we obtain using the IH for ordinals $< \beta$ that $\mathrm{degClosure}(\mathrm{V}^{<\beta}, \mathrm{T}^{<\beta}, d', d'')$ holds. Therefore $\langle \mathrm{v}_{d',f,g}, a\rangle \in \mathrm{V}^\alpha$.

Remaining construction of the model. The remaining steps are as for the other models of type theory with universes in [53]: We need finitely many admissibles above AutM in order to interpret the basic set constructions on top of $\mathrm{V}, \mathrm{T_V}$ (each application of the W-type on top of V requires one more admissible). Therefore the type theory can be interpreted in Kripke Platek set theory plus one Π_3-reflecting ordinal and finitely many admissibles (i.e. for Meta-each n we have n admissibles) above it. So we have given the essence of a proof of the following theorem:

Theorem 5.2

- *(a) We can model* $\mathrm{ML_W} + (\mathrm{Aut} - \mathrm{Mahlo})$ *in* $(\mathrm{KP}\omega + (\mathrm{Aut} - \mathrm{Mahlo}))^+$.
- *(b)* $|\mathrm{ML_W} + (\mathrm{Aut} - \mathrm{Mahlo})| \leq |(\mathrm{KP}\omega + (\mathrm{Aut} - \mathrm{Mahlo}))^+|$.
- *(c) The previous statements hold as well if we replace intensional by extensional equality.*

Remark 5.3 *The status of the above theorem is as follows: Above we have developed the model and proved its correctness, including the details involving the autonomous Mahlo universe which go beyond standard Martin-Löf type theory. However, we are relying on the basic constructions of models of type theory in Kripke-Platek set theory, which includes how to interpret the basic set constructions, and the properties of the terms and reductions of terms used. In the author's PhD thesis [42], these details have been worked out in full detail, but the setting there was not very generic to be adaptable to the current setting. The author would have liked to publish a modernised presentation of this model construction before publishing the current article, which would have enabled him to carry out the details in full. However, he felt that presenting the autonomous Mahlo and the* Π_3*-reflecting universes is more urgent, and therefore decided to present this universe with a sketch of the model, rather than waiting before the infrastructure article has gone through the publishing process. So what makes the current model constructions a sketch is the general setting common to all models of type theory (and one can refer to my PhD thesis for those details), the correctness of the model regarding the details of the autonomous Mahlo universe have been worked out in this article in detail.*

6 Conclusion

We have shown how to develop an autonomous Mahlo universe. We have developed a model for the autonomous Mahlo universe and therefore determined an upper bound for its proof theoretic strength.

Future research. Apart from introducing the Π_3-reflecting universe together with a proof theoretical analysis, the next steps will be to introduce stronger universes, for which we already have developed draft versions: the Π_N-reflecting universe and a Π^1_1-reflecting universe. We need to carry out as well well-ordering proofs for these extensions (at the moment there are only very rough sketches). Another line of research would be to explore, whether variants of these universe constructions can be used in programming, especially in generic programming.

References

[1] P. Aczel. The type theoretic interpretation of Constructive Set Ttheory: choice principles. In A. S. Troelstra and D. van Dalen, editors, *The L.E.J¿ Brouwer Centenary Symposium*, pages 1 – 40, Amsterdam, 1982. North Holland.

[2] P. Aczel and W. Richter. Inductive definitions and analogues of large cardinals. In W. Hodges, editor, *Conference in Mathematical Logic, London '70*, Lecture Notes in Mathematics, Vol. 255, pages 1 – 9, Berlin, 1972. Springer.

[3] T. Arai. Proof theory of theories of ordinals I: Reflecting ordinals. Draft, 1996.

[4] T. Arai. Systems of ordinal diagrams. Draft, 1996.

[5] T. Arai. Proof theory of theories of ordinals II: Σ_1 stability. Draft, 1997.

[6] T. Arai. Proof theory of theories of ordinals III: Π_1 collection. Draft, 1997.

[7] T. Arai. Ordinal diagrams for Π_3-reflection. *J. Symbolic Logic*, 65(3):1375 – 1394, 2000.

[8] T. Arai. Ordinal diagrams for recursively Mahlo universes. *Arch. Math. Logic*, 39(5):353 – 391, 2000.

[9] T. Arai. Proof theory for theories of ordinals. I. Recursively Mahlo ordinals. *Ann. Pure Appl. Logic*, 122(1 – 3):1 – 85, 2003.

[10] T. Arai. Proof theory for theories of ordinals. II. Π_3-reflection. *Ann. Pure Appl. Logic*, 129(1 – 3):39 – 92, 2004.

[11] J. Barwise. *Admissible Sets and Structures. An Approach to Definability Theory.* Omega-series. Springer, 1975.

[12] M. Benke, P. Dybjer, and P. Jansson. Universes for generic programs and proofs in dependent type theory. *Nordic J. of Computing*, 10(4):265–289, 2003.

[13] W. Buchholz. A simplified version of local predicativity. In P. Aczel, H. Simmons, and S. S. Wainer, editors, *Proof Theory. A selection of papers from the Leeds Proof Theory Programme 1990*, pages 115 – 147, Cambridge, 1992. Cambridge University Press.

[14] P. Dybjer. A general formulation of simultaneous inductive-recursive definitions in type theory. *Journal of Symbolic Logic*, 65(2):525–549, June 2000.

[15] P. Dybjer and A. Setzer. Finite axiomatizations of inductive and inductive-recursive definitions. In *Workshop on Generic Programming, Marstrand, Sweden, 18 June 1998.* http://www.cs.ruu.nl/people/johanj/programme_wgp98.html, 1998.

[16] P. Dybjer and A. Setzer. A finite axiomatization of inductive-recursive definitions. In J.-Y. Girard, editor, *Typed Lambda Calculi and Applications*, volume 1581 of *Lecture Notes in Computer Science*, pages 129–146. Springer, April 1999.

[17] P. Dybjer and A. Setzer. Indexed induction-recursion. In R. Kahle, P. Schroeder-Heister, and R. Stärk, editors, *Proof Theory in Computer Science*, volume 2183 of *Lecture Notes in Computer Science*, pages 93 – 113. Springer, 2001.

[18] P. Dybjer and A. Setzer. Induction-recursion and initial algebras. *Annals of Pure and Applied Logic*, 124:1 – 47, 2003.

[19] P. Dybjer and A. Setzer. Indexed induction-recursion. *Journal of Logic and Algebraic Programming*, 66:1 – 49, 2006.

[20] P. Hancock and P. Hyvernat. Programming interfaces and basic topology. *Ann. Pure Appl. Logic*, 137(1-3):189–239, 2006.

[21] P. Hinman. *Recursion-Theoretic Hierarchies.* Springer, Berlin, Heidelberg, New York, 1978.

[22] M. Hofmann. Syntax and semantics of dependent types. In A. M. Pitts and P. Dybjer, editors, *Semantics and logics of computation*, pages 79 – 130, Cambridge, 1997. Cambridge University Press.

[23] G. Jäger. *Theories for Admissible Sets: A Unifying Approach to Proof Theory.* Bibliopolis, Naples, 1986.

[24] G. Jäger and T. Strahm. Reflections on reflections in explicit mathematics. *Ann. Pure Appl. Logic*, 136(1-2):116 – 133, 2005.

[25] P. Jansson. *Functional Polytypic Programming: Use and Implementation.* PhD thesis, Computing Science, Chalmers University of Technology, Gothenburg, Sweden, 1997.

[26] P. Martin-Löf. Constructive mathematics and computer programming. In L. Cohen, J. Los, H. Pfeiffer, and K.-P. Podewski, editors, *Logic, Methodology and Philosophy of Science VI*, pages 153–175, Amsterdam, 1982. North-Holland.

[27] P. Martin-Löf. *Intuitionistic type theory.* Bibliopolis, Naples, 1984.

[28] B. Nordström, K. Petersson, and J. Smith. *Programming in Martin-Löf's type theory. An Introduction.* Oxford University-Press, Oxford, 1990.

[29] B. Nordström, K. Petersson, and J. M. Smith. Martin-löf's type theory. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of logic in computer science, Vol. 5*, pages 1 – 37. Oxford Univ. Press, 2000.

[30] E. Palmgren. On universes in type theory. In G. Sambin and J. Smith, editors, *Twenty five years of constructive type theory*, pages 191 – 204, Oxford, 1998. Oxford University Press.

[31] W. Pohlers. Subsystems of set theory and second order number theory. In S. R. Buss, editor, *Handbook of Proof Theory*, pages 209 – 335. Elsevier, 1998.

[32] A. Ranta. *Type-theoretical Grammar.* Clarendon Press, Oxford, 1995.

[33] M. Rathjen. Ordinal notations based on a weakly Mahlo cardinal. *Arch. math. Logic*, 29:249 – 263, 1990.

[34] M. Rathjen. Proof-theoretical analysis of KPM. *Arch. math. Logic*, 30:377 – 403, 1991.

[35] M. Rathjen. Eine Ordinalzahlanalyse der Π_3–Reflexion. Habilitationsschrift, Westfälische Wilhelms–Universität Münster, 1992.

[36] M. Rathjen. Collapsing functions based on recursively large cardinals: A well-ordering proof for KPM. *Arch. math. Logic*, 33:35–55, 1994.

[37] M. Rathjen. Proof theory of reflection. *Ann. Pure Appl. Logic*, 68:181 – 224, 1994.

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

[39] M. Rathjen. An ordinal analysis of parameter free Π^1_2-comprehension. *Arch. Math. Log.*, 44(3):263 – 362, April 2005.

[40] M. Rathjen. An ordinal analysis of stability. *Arch. Math. Logic*, 44(1):1 – 62, January 2005.

[41] W. Richter and P. Aczel. Inductive definitions and reflecting properties of admissible ordinals. In J. E. Fenstad and P. G. Hinman, editors, *Generalized recursion theory*, pages 301 – 381, Amsterdam, 1973. North-Holland.

[42] A. Setzer. *Proof theoretical strength of Martin-Löf Type Theory with W-type and one universe.* PhD thesis, Universität München, 1993. Available from http://www.cs.swan.ac.uk/∼csetzer.

[43] A. Setzer. A model for a type theory with Mahlo Universe. 10pp. Draft, available via
http://www.cs.swan.ac.uk/∼csetzer/articles/uppermahlo.pdf, 1996.

[44] A. Setzer. An upper bound for the proof theoretical strength of Martin-Löf Type Theory with W-type and one Universe. 33pp. Draft, available via http://www.cs.swan.ac.uk/∼sim/articles/1papdiss.dvi or .ps or .pdf, 1996.

[45] A. Setzer. An introduction to well-ordering proofs in Martin-Löf's type theory. In G. Sambin and J. Smith, editors, *Twenty-five years of constructive type theory*, pages 245 – 263, Oxford, 1998. Clarendon Press.

[46] A. Setzer. Well-ordering proofs for Martin-Löf type theory. *Annals of Pure and Applied Logic*, 92:113 – 159, 1998.

[47] A. Setzer. Ordinal systems. In C. B. and J. Truss, editors, *Sets and Proofs*, pages 301 – 331, Cambridge, 1999. Cambridge University Press.

[48] A. Setzer. Extending Martin-Löf type theory by one Mahlo-universe. *Arch. math. Logic*, 39:155 – 181, 2000.

[49] A. Setzer. Extending Martin-Löf type theory by one Mahlo-universe. *Arch. Math. Log.*, 39:155 – 181, 2000.

[50] A. Setzer. Ordinal systems part 2: One inaccessible. In S. Buss, P. Hajek, and P. Pudlak, editors, *Logic Colloquium '98*, ASL Lecture Notes in Logic 13, pages 426 – 448, Massachusetts, 2000. Peters Ltd.

[51] A. Setzer. Proof theory of Martin-Löf type theory – an overview. *Mathematiques et Sciences Humaines*, 42 année, n°165:59 – 99, 2004.

[52] A. Setzer. Ordinal systems part 3: Mahlo. In preparation, 2006.

[53] A. Setzer. Universes in type theory part I – Inaccessibles and Mahlo. In A. Andretta, K. Kearnes, and D. Zambella, editors, *Logic Colloquium '04*, pages 123 – 156. Association of Symbolic Logic, Lecture Notes in Logic 29, Cambridge University Press, 2008.

[54] A. Troelstra and D. v. Dalen. *Constructivism in Mathematics. An Introduction, Vol. II.* North-Holland, Amsterdam, 1988.