Pure and Applied Logic

Master's Thesis

A classification of the set-theoretic total recursive functions of KPl

Author: Anton Fernández Dejean

Supervisors: Joost J. Joosten (University of Barcelona)
Juan Pablo Aguilera (TU Wien)

Barcelona, July 26, 2024

Abstract

The set theory KPl, which stands for Kripke-Platek-limit, roughly stipulates that there are unboundedly many admissible sets. Admissible sets are models of the Kripke-Platek set-theory KP which is a very weak fragment of ZFC. In [1], J. Cook and M. Rathjen classify the provably total set functions in KP using a proof system based on an ordinal notation system for the Bachmann-Howard ordinal relativized to a fixed set. In this paper, we adapt this result to the KPl set theory. We consider set functions which are provably total in KPl and Σ-definable by the same formula in any admissible set. We prove that, if f is such a function then, for any set x in the universe, the value $f(x)$ always belongs to an initial segment of $L(x)$, the constructible hierarchy relativized to the transitive closure of x, at a level below the relativized Takeuti-Feferman-Buchholz ordinal (the TFB ordinal is the proof-theoretic ordinal of KPl).
To prove this result, we first construct an ordinal notation system based on [2] for KPl relativized to a fixed set that we will use in order to build a logic dependent on this fixed set where we will embed KPl. Thanks to this relativized system, we will be able to bound the value of the function at this fixed set.

Acknowledgments

I want to thank my advisors Joost and Juan Pablo for teaching me so much about this topic and about logic and mathematics in general. Particularly, I want to thank Joost for the weekly meetings that I enjoyed a lot and Juan Pablo for coming up with such an interesting topic and for all the help and the disposition he had throughout the semester. Moreover, they also helped me to learn how to write mathematics in a more clear and rigorous manner. I could not have asked for better supervisors.

On the non-academic side, I would also like to enormously thank my family and friends, both from the master's program and outside, for always encouraging me and cheering me up. Their unwavering support and presence have motivated me a lot these last months. I could not ask to be surrounded by nicer people.

I am also grateful to all the people who have contributed to this thesis, whether it was through a private conversation or an email exchange.

Contents

1 Introduction **4**

2 The axiomatic set theory KPl **8**

3 Relativized ordinal notation system **14**
3.1 Preliminaries 14
3.2 The ordinal notation system 24

4 The system $\mathsf{RS}_l(X)$ **32**
4.1 The terms and formulas of $\mathsf{RS}_l(X)$ 32
4.2 Operator-controlled derivations 40
4.3 Cut-elimination for $\mathsf{RS}_l(X)$ 46
4.4 The Collapsing Theorem 55

5 Embedding KPl into $\mathsf{RS}_l(X)$ **62**
5.1 The $\Vdash$ relation 62
5.2 The embedding Theorem 67

6 The provably total set-recursive functions of KPl **92**

1 Introduction

Set theory has several different axiomatizations. The most studied one is ZFC (Zermelo-Fraenkel with the axiom of Choice), but many others are interesting as well. In particular, if we exclude the Power Set axiom and restrict the Separation and Collection schemas to Δ_0-formulas (formulas that do not have any unbounded quantifier) from ZF, we end up with KP, which stands for Kripke-Platek Set Theory. The transitive set models of KP are known as *admissible* sets, while an ordinal is called *admissible* whenever L_α is admissible, where L_α is the α-th rank of Gödel's constructible hierarchy. Admissible sets and admissible ordinals have been widely studied and are related to numerous areas of computability theory. They appear in hyperarithmetical theory: the first admissible ordinal above ω is the Church-Kleene ordinal ω_1^{ck}, which is the first non-recursive ordinal. Admissible ordinals are crucial in α-recursion theory, which is a generalization of computability theory to subsets of admissible ordinals. Moreover, admissible sets are related to E-recursion theory, a generalization of the theory of computability from natural numbers to arbitrary sets, via Van de Wiele's theorem, that states that a function is E-recursive if and only if the function is uniformly Σ_1-definable on every admissible set (see [16]). It is also worth mentioning that KP, and by extension admissible sets, is also related to (subsystems of) second-order arithmetic.

In this thesis, we are interested in a set theory that states that the universe is a limit of admissible sets. This theory, called KPl for Kripke-Platek-limit, is not exactly an extension of KP. The language of most of the set-theoretic theories, for example the language of KP, is $\{\in\}$. As for the language of KPl, it expands the language of KP with the predicate Ad, that intends to mean that a set is admissible. As we will see in Section 2, KPl consists in the axioms of KP (minus Δ_0-Collection) written in the language $\{\in, Ad\}$ together with some axioms defining the predicate Ad and the Limit axiom. Roughly, the Limit axiom states that there are unboundedly many admissible sets. The main objective of this thesis is to classify the provably total and Σ-definable set-recursive functions of KPl.

Those kind of classification results for a given theory rely on the ordinal analysis of the studied theory. Research in mathematics at the beginning of the XXth century was deeply influenced by Hilbert's programme. The programme seeked to provide a grounding for mathematics by showing the consistency of formal systems by finitistic means. Gentzen started Hilbert's programme by providing such a finitistic proof for number theory. In [6], he proved the consistency of Peano Arithmetic, PA, from a

finitistically acceptable theory (e.g. Primitive Recursive Arithmetic) using transfinite induction up to ε_0. The ordinal ε_0 is the first fixpoint of the ordinal function $\beta \mapsto \omega^\beta$, also defined as the limit of $\{\omega, \omega^\omega, \omega^{\omega^\omega}, \dots\}$. Gentzen's proof relies on a recursive representation system for ordinals below ε_0. Moreover, Gentzen also showed that PA proves transfinite induction for arithmetic predicates up to any ordinal below ε_0, but not up to ε_0. This way, the "strength" of PA is measured by ε_0: we say that ε_0 is the proof-theoretic ordinal of PA. Gentzen's work marks the start of modern ordinal analysis.
Later, the ordinal analysis of many theories has been obtained. First, it was for subsytems of second order arithmetic. For example, Takeuti gave an ordinal analysis of $(\mathsf{\Pi}^1_1 - \mathsf{CA})$ in 1967 in [18]. Takeuti's work was followed by himself and by many other mathematicians, such as Buchholz and Pohlers, that refined his techniques and simplified the treatment of ordinal analysis. Next, manly due to the work of Jäger and Pohlers ([7]), the ordinal analysis of set theories was achieved, starting with KP. Ordinal analysis is still nowadays a widely studied branch of proof theory.

Thanks to ordinal analysis, classification results of provably total recursive functions of many theories have been obtained. The starting point is the classification of the provably total recursive functions of PA, that has been proved in different ways (e.g. by Kreisel in 1951-1952 in [9] and [10], by Buchholz and Wainer in 1987 in [3] or by A. Weiermann in 1996 in [19]). To state and prove such a classification result, we establish a representation system for the ordinals below ε_0 (the proof-theoretic ordinal of PA). Then, we define a hierarchy of functions, called the *fast growing hierarchy*, based on the ordinals below ε_0. It can be shown that any provably total recursive function of PA has to be majorised[1] by some function of the fast growing hierarchy at level less than ε_0.

In 2016, Jacob Cook and Michael Rathjen in [5] gave a classification of the provably total set-recursive functions of KP using tools from ordinal analysis and proof theory. The result we show in this thesis is similar to Theorem 6.2 in [5] (and, in fact, the proof uses the same methods) but for KPl total set-recursive functions. Nonetheless, there is a constraint: the existence of many non-recursive ordinals (e.g. ω_1^{ck}) is provable in KPl. The image of a set x by a function could be contained by an admissible set not containing x. To avoid this obstacle, we have to restrict our study to functions that have the property that the image of any set x belongs to the same admissible sets as x does. The statement of our main theorem will be the

[1] We say that a function f is majorised by the function g whenever there is some natural number n such that $\forall m > n(f(m) < g(m))$

following. Given any set-recursive function f such that KPI proves that f is total and Σ-definable with the same formula in any admissible set, for any set x, the value $f(x)$ belongs to an initial segment of the constructible hierarchy relativized to x at level below the relativized Takeuti-Feferman-Buchholz ordinal, where the TFB ordinal is the proof-theoretic ordinal of KPI. We will call this bounding set $\hat{G}_n(x)$, where n is some natural number that can be read of the proof of totality so n is related to the adequate level of the relativized constructible hierarchy. The formal statement of our main theorem will be the following.

Theorem 6.3. *Let f be a set-recursive function such that* KPI *proves that f is total and uniformly Σ-definable in any admissible set. Then, there is some $n < \omega$ such that*

$$V \models \forall x(f(x) \in \hat{G}_n(x)).$$

In the proof of our main theorem, given such a function f, we will fix a set X and show that, indeed, the value $f(X)$ is bounded as stated. To do that, we will construct a logic depending on X in which we will embed KPI and through which we will be able to show that $f(X)$ belongs to $\hat{G}_n(X)$. This system, that we will call $\mathsf{RS}_l(X)$ for Ramified Set Theory relativized to X (the l stands for limit), will be based on a relativized ordinal notation system for KPI. By means of this relativized notation system we will give names to relevant ordinals. In particular, we will write the first $(\omega+1)$-many regular uncountable cardinals above the set-theoretic rank of X as Ω_n for each $n \leq \omega$. In $\mathsf{RS}_l(X)$, the admissible ordinals will be exactly the Ω_n's and we will use the ordinals from the recursive set of notations in many ways (to define the terms of the system, the length of the derivations, the complexity of the formulas, etc.).

As we said earlier, the skeleton of this thesis will be similar to the skeleton of Cook and Rathjen's paper [5]. In fact, the main references used in thesis are Cook and Rathjen's paper [5] together with Pohlers' article [12].
In Section 2 we will give the formal definition of KP and KPI. In Section 3 we will introduce and study our relativized notation system, based on [2] and [12]. We will employ this notation system to define the system $\mathsf{RS}_l(X)$ in Section 4, where we will show some proof-theoretic properties of this system such that, under certain conditions, cuts can be eliminated. In this section we also prove an important result that will help to eliminate cuts, the Collapsing Theorem, that allows to collapse the length of a proof and the complexity of the cuts in a proof below a certain Ω_n, for

$n < \omega$. In Section 5 we embed KPI into $\mathsf{RS}_l(X)$. Finally, we will put all the tools and results together to prove the main theorem in Section 6.

2 The axiomatic set theory KPl

In this section, we define the axiomatic set theory KPl. This theory is strongly related to KP, Kripke-Platek: the theory KPl states that the universe is a limit of models of KP. This is why we will deal, at the beginning, with two languages, the language of KP, which is $\mathcal{L} := \{\in, \notin\}$ (we will use the connective $\neg$ as a defined notion) and the language of KPl, where we add a predicate that intends to mean that a given set is a model of KP together with the negation of this predicate. So we define $\mathcal{L}' := \{\in, \notin, Ad, \neg Ad\}$.
Moreover, since KP and KPl are first-order axiomatic theories, we use the usual first-order language to construct formulas. We have the connectives $\{\vee, \wedge\}$, the quantifiers $\{\forall, \exists\}$, the auxiliary symbols $\{(,)\}$ and an infinite set of variables $\mathcal{V}$. We will often use the letters x, y and z, as well as a and b, to denote both variables and sets (it should be easy to discern if a letter is used as a variable or as a set by context).
We define the following notation to say that some axiom or formula is satisfied in a given set.

Definition 2.1. *Let A be a formula in any of the languages $\mathcal{L}$ or $\mathcal{L}'$. Let x be a set. The formula obtained by restricting all the unbounded quantifiers of A to x will be denoted by A^x.*

Moreover, we will use the equality symbol as a defined notion and we will define the connectives $\neg$ and $\rightarrow$. The following definition works for both languages $\mathcal{L}$ and $\mathcal{L}'$.

Definition 2.2. *The formula $a \subseteq b$ will stand for $\forall x \in a(x \in b)$.*
The formula $a = b$ will stand for $a \subseteq b \wedge b \subseteq a$.

The formula $\neg A$ is obtained by replacing in A the symbol $\in$ by $\notin$ and vice-versa, $\wedge$ by $\vee$ and vice-versa, $\forall$ by $\exists$ and vice-versa, and $Ad(\cdot)$ by $\neg Ad(\cdot)$ and vice-versa.
We define $a \rightarrow b \equiv \neg a \vee b$.
The formula $a \neq b$ will stand for $\neg a = b$.

We will start by stating all of the axioms and schemas that we will use and, after that, we will introduce both KP and KPl. We will work with finite sets of formulas. In the following definition, we consider Γ to be any finite set of formulas in the language $\mathcal{L}$ or $\mathcal{L}'$.

Definition 2.3.

Logical axioms:	$\Gamma, A, \neg A$ *for any formula* A,
Leibniz Principle:	$\Gamma, (a = b \wedge B(a)) \rightarrow B(b)$ *for any formula* $B(a)$,
Pair:	$\Gamma, \exists z(a \in z \wedge b \in z)$,
Union:	$\Gamma, \exists z \forall y \in a \forall x \in y(x \in z)$,
Δ_0-*Separation:*	$\Gamma, \exists y[\forall x \in y(x \in a \wedge B(x)) \wedge \forall x \in a(B(x) \rightarrow x \in y)]$ *for any* Δ_0-*formula* $B(a)$,
Class Induction:	$\Gamma, \forall x[\forall y \in x B(y) \rightarrow B(x)] \rightarrow \forall x B(x)$ *for any formula* $B(a)$,
Infinity:	$\Gamma, \exists x[\exists z \in x(z \in x) \wedge \forall y \in x \exists z \in x(y \in z)]$,
Δ_0-*Collection:*	$\Gamma, \forall x \in a \exists y B(x, y) \rightarrow \exists z \forall x \in a \exists y \in z B(x, y)$ *for any* Δ_0-*formula* $B(a, b)$.

Now, the axioms defining the Ad predicate are the following (the set ω is the first ordinal that contains all the natural numbers; the predicate $Tran(x)$ means that x is transitive, i.e. for any $y \in x$ all the elements of y are elements of x).

Ad1:	$\Gamma, \forall x[Ad(x) \rightarrow \omega \in x \wedge \ Tran(x)]$,
Ad2:	$\Gamma, \forall x \forall y[Ad(x) \wedge Ad(y) \rightarrow x \in y \vee x = y \vee y \in x]$,
Ad3:	$\Gamma, \forall x[Ad(x) \rightarrow (Pair)^x \wedge (Union)^x \wedge (\Delta_0 - Sep)^x \wedge (\Delta_0) - Coll)^x]$,
Lim:	$\Gamma, \forall x \exists y[Ad(y) \wedge x \in y]$.

The KP theory is a weak subtheory of ZF, where we get rid of the Power Set axiom and we restrict the Separation and Collection axiom schemas to Δ_0-formulas. One of the most interesting properties about KP is the fact that this theory is sufficient to construct Gödel's constructible hierarchy. In fact, the constructible universe L is a model of both ZFC and KP.

Definition 2.4. *The axioms of* KP *are the Logical axioms and the following axioms and schemas in the language* $\mathcal{L} = \{\in, \notin\}$.

1. *Leibniz Principle,*
2. *Pair,*
3. *Union,*
4. Δ_0-*Separation,*
5. *Class Induction,*
6. *Infinity,*

7. Δ_0-*Collection.*

Transitive set models of KP are called *admissible* sets or classes. In KPl, the (*Lim*) axiom states that for any set x we can find an admissible set that contains x. Therefore, starting with some set (for instance the empty set), we can build a chain of any length of admissible sets, each containing the previous one. This gives the existence of unboundedly many admissible sets in our universe.

We remind that the language of KPl is $\mathcal{L}' = \{\in, \notin, Ad, \neg Ad\}$, where, in general, the Ad predicate intends to mean that a given set is admissible. Actually, the predicate Ad in this thesis does not just mean admissible. Admissible sets are not linearly ordered by the $\in$-relation, and so axiom ($Ad2$) should in general not hold. Nonetheless, we will restrict our study of admissible sets to those belonging to a particular hierarchy of sets, which will be linearly ordered, and so the predicate Ad will mean "is admissible and belongs to this hierarchy". That is reason why we include ($Ad2$) as an axiom of KPl.
We now introduce the axioms of KPl. We observe that we don't have Δ_0-Collection here.

Definition 2.5. *The axioms of* KPl *are the Logical axioms and the following axioms and schemas in the language $\mathcal{L}'$.*

1. *Leibniz Principle,*
2. *Pair,*
3. *Union,*
4. Δ_0-*Separation,*
5. *Class Induction,*
6. *Infinity,*
7. *Ad1,*
8. *Ad2,*
9. *Ad3,*
10. *Lim.*

We will now introduce the rules of inference that we will use for KPI. These rules are written in Tait's sequent style. Both in the premises and in the conclusion of each rule, we have finite sets of $\mathcal{L}'$-formulas. That means that, below, the symbol Γ represents any finite set of $\mathcal{L}'$-formulas, just as in the formulation of the axioms. Moreover, A and B are any $\mathcal{L}'$-formulas. The formulas derived in the conclusion can intuitively be thought as a disjunction. That is, when we derive A, B, it means that either A or B is true.

Definition 2.6. *The rules of inference of* KPI *are the following.*

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

$$(\vee)\ \frac{\Gamma, A}{\Gamma, A \vee B} \qquad\qquad (\vee)\ \frac{\Gamma, B}{\Gamma, A \vee B}$$

$$(b\exists)\ \frac{\Gamma, a \in b \wedge B(a)}{\Gamma, \exists x \in b\ B(x)}$$

$$(\exists)\ \frac{\Gamma, B(a)}{\Gamma, \exists x\ B(x)}$$

$$(b\forall)\ \frac{\Gamma, a \in b \rightarrow B(a)}{\Gamma, \forall x \in b\ B(x)}$$

$$(\forall)\ \frac{\Gamma, B(a)}{\Gamma, \forall x\ B(x)}$$

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

In the rules $(b\forall)$ *and* $(\forall)$*, it must be the case that* a *does not occur in the conclusion.*

We write $\mathsf{KPI} \vdash \Gamma$ *whenever* Γ *is an axiom or there is one (or two) set(s) of formulas* Δ *(and* Δ'*) such that* $\mathsf{KPI} \vdash \Delta$ *(and* $\mathsf{KPI} \vdash \Delta'$*) and we can obtain* Γ *from* Δ *(and* Δ'*) by the application of a rule.*

The rules of KP are the same but in the language $\mathcal{L}$.
We will make explicit an important difference between KP and KPl. We will show that KP proves Σ-Reflection. A formula A is Σ if there is no unbounded universal quantifier occurring in A. The proof of Σ-Reflection uses the Δ_0-Collection axiom and thus cannot be performed in KPl. Since we only want to highlight a major difference between KP and KPl, we will omit some details. We will be following the proof of Barwise, in [1].

Theorem 2.7 (Σ-Reflection principle)**.** *For any Σ-formula A in the language $\mathcal{L} = \{\in, \notin\}$, we have*

$$\mathsf{KP} \vdash A \to \exists x \ A^x.$$

Proof. The central idea of the proof is to use the following claim, that states that if a Σ-formula holds in some set x, then the formula holds in any superset of x.

Claim 2.7.1. *Let B be a Σ-formula in the language $\mathcal{L}$. Then we have*

$$\mathsf{KP} \vdash B^x \wedge x \subseteq y \to B^y.$$

We prove Claim 2.7.1 by induction on the construction of B. We fix a model of KP and we fix two sets in this model (the interpretations of x and y in the model, that we will call just x and y).

Base Case. If B is a Δ_0-formula, then $B^x \equiv B^y$ and so the claim is reduced to showing $\mathsf{KP} \vdash B \to B$, which always holds.

We suppose $B \equiv B_0 \wedge B_1$. We want to show

$$\mathsf{KP} \vdash (B_0 \wedge B_1)^x \wedge x \subseteq y \to (B_0 \wedge B_1)^y.$$

So we assume that $(B_0 \wedge B_1)^x$ holds in our model. Hence, the formulas B_0^x and B_1^x also hold. The induction hypothesis gives

$$\mathsf{KP} \vdash B_i^x \wedge x \subseteq y \to B_i^x \tag{1}$$

for $i \in \{0, 1\}$. Therefore, since B_i^x and $x \subseteq y$ are true in the model, B_i^y is also true, for $i \in \{0, 1\}$. Hence, $(B_0 \wedge B_1)^y$ holds in the model. This means that, in any model of KP, $(B_0 \wedge B_1)^y$ is true whenever $(B_0 \wedge B_1)^x \wedge y \subseteq x$ is true. Hence,

$$\mathsf{KP} \vdash (B_0 \wedge B_1)^x \wedge x \subseteq y \to (B_0 \wedge B_1)^y.$$

The other cases are similar. Since we do not want to insist in this proof, we consider that Claim 2.7.1 is shown.
We prove Theorem 2.7 by induction on the construction of A.
We treat the most interesting case for us, where Δ_0-Collection shows up. We suppose that A is $\forall a \in b\ B(a)$. By the induction hypothesis we have

$$\mathsf{KP} \vdash B(a) \to \exists x\ B(a)^x.$$

To verify that KP proves $\forall a \in b\ B(a) \to \exists x \forall a \in b\ B(a)^x$, we fix a set b in a fixed model of KP and we assume that $\forall a \in b\ B(a)$ holds in the model. Our aim is to show that $\exists x \forall a \in b\ B(a)^x$ holds in this model. In the model, for every a in b there is a set y_a such that $B(a)^{y_a}$ holds by the induction hypothesis applied to $B(a)$. This means that $\forall a \in b\, \exists y_a B(a)^{y_a}$ holds. Now we use Δ_0-Collection to get $\exists y \forall a \in b\, \exists y_a \in y\ B(a)^{y_a}$. We take a witness y and we define $Y := \bigcup y$. Then, for each $a \in b$, we have some y_a such that $B(a)^{y_a} \wedge y_a \subseteq Y$ holds. Therefore, by Claim 2.7.1, we obtain $\forall a \in b\ B(a)^Y$. Hence, taking Y as our witness, we obtain that $\exists x \forall a \in b\ B(a)^x$ holds in the model.

$\square$

We cannot perform the proof of Σ-Reflection within KPl due to the absence of Δ_0-Collection. This difference between KP and KPl has further implications: the principle of Σ-Reflection is used to prove more results, like Σ-Collection or Δ-Replacement, that are used to prove Σ-Recursion (see [1]). So in KP, mainly due to Δ_0-Collection, we are able to define an $n+1$-ary Σ function from an $n+2$-ary Σ function by recursion. This definition cannot be done within KPl, where only a weaker result can be proved (see [12]).

3 Relativized ordinal notation system

We recall that our main goal in this thesis is to show that, given a set-recursive Σ-definable function f such that KPI proves that f is total and definable with the same Σ-formula in any admissible set, the image of any set under f is bounded. In the proof of this main theorem, our way to show that $f(x)$ is bounded for every x will consist on fixing a set X and seeing that $f(X)$ is indeed bounded by a bound that depends on X. Once our set X is fixed, we will build a proof system dependent on this X where we will embed KPI in some way, in order to reason in this new system. This system, that we will call $\mathsf{RS}_l(X)$, for Ramified Set Theory relativized to X, will be defined based on ordinals: since KPI states that there are unboundedly many admissible sets, we will consider the sequence of the first ω-many different admissible sets from the constructible hierarchy relativized to X, defined as follows. The set $TC(\{X\})$ is the transitive closure of $\{X\}$, the smallest transitive set that includes $\{X\}$.

Definition 3.1. *Let X be any set. We define for every ordinal α the set $L_\alpha(X)$ as:*

$L_0(X) = TC(\{X\})$,
$L_{\alpha+1}(X) = \{Y \subseteq L_\alpha(X) : Y \text{ is definable over } \langle L_\alpha(X), \in\rangle \text{ with parameters in } L_\alpha(X)\}$,
$L_\gamma(X) = \bigcup_{\alpha<\gamma} L_\alpha(X)$ *if γ is a limit.*

Moreover, having a fixed set X we let θ be the set-theoretic rank of X (where $\text{rank}(y) = sup\{\text{rank}(z)+1 : z \in y\}$ for any set y). Let $\lambda\beta.\Omega_\beta$ enumerate the sequence of consecutive cardinals κ such that $\kappa \geq \theta$. In particular, for every ordinal β we have that $L_{\Omega_\beta}(X)$ is admissible. Actually, we will only be interested in those cardinals up to Ω_ω, which contains ω many X-admissible ordinals (each Ω_n for $n < \omega$).
We will be using ordinals to a great extent in $\mathsf{RS}_l(X)$. Since this system will depend on the fixed set X, we need the ordinals that we will employ to depend on X too. That is why we firstly have to thoroughly define the set of ordinals that we will be using in a recursive way from X. This is exactly what this section is about: we are going to define for a fixed set X with set-theoretic rank θ a recursive set $T(\theta)$ of strings representing ordinals.

3.1 Preliminaries

In this subsection we develop the tools that we need to define $T(\theta)$. We will define the Veblen functions φ_α for each ordinal α. The general idea is that φ_α enumerates the fixpoints of the previous φ_β for $\beta < \alpha$. To understand the construction of the Veblen functions, we first need some ordinal theory. We refer the interested reader not familiar with ordinals and cardinals to Chapter I sections 6-10 of [11] for an overview, or Chapters 2 and 3 of [8] for a more complete study.

Definition 3.2. *A set x is transitive iff $\forall\, y{\in}x\, \forall\, z{\in}y\, (z \in x)$.*
A set x is well-ordered by a relation R iff every non-empty subset y of x has a least element in the R-ordering.
An ordinal is a set α such that α is transitive and well-ordered by the $\in$-relation. We write On to denote the class of all the ordinals.

We also define an ordering on On.

Definition 3.3. *Let $\alpha, \beta \in On$. We define $\alpha < \beta$ iff $\alpha \in \beta$.*

Ordinals fall into three categories, defined as follows.

Definition 3.4. *Let α be any ordinal. Then either*

1. *$\alpha = 0 := \emptyset$, or*
2. *$\alpha = \beta + 1$ for some β, in which case α is called a successor ordinal, or*
3. *For every $\beta < \alpha$ we have $\beta + 1 < \alpha$, in which case α is called a limit ordinal.*

To define the Veblen functions, we will need some basic ordinal arithmetic.

Definition 3.5. *Let α be an ordinal.*
We define $\alpha + \beta$ for every ordinal β by recursion on β.

1. *$\alpha + 0 = \alpha$,*
2. *$\alpha + (\beta + 1) = (\alpha + \beta) + 1$,*
3. *$\alpha + \beta = sup(\alpha + \delta : \delta < \beta)$ if β is a limit ordinal.*

We define $\alpha \cdot \beta$ for every ordinal β by recursion on β.

1. *$\alpha \cdot 0 = 0$,*
2. *$\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$,*
3. *$\alpha \cdot \beta = sup(\alpha \cdot \delta : \delta < \beta)$ if β is a limit ordinal.*

We define α^β for every ordinal β by recursion on β.

1. $\alpha^0 = 1$,
2. $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$,
3. $\alpha^\beta = sup(\alpha^\delta : \delta < \beta)$ *if* β *is a limit ordinal.*

We refer to page 23 of [8] to see some properties of the operations defined above. We need the notion of cardinal that we define as follows.

Definition 3.6. *Let x be any set. Then, we write $|x|$ to denote the cardinality of x, i.e. the least ordinal κ such that there is a bijection between x and κ.*
An ordinal κ is called a cardinal iff $|\kappa| = \kappa$.

We are specially interested in regular cardinals, that we define now.

Definition 3.7. *A cardinal κ is called regular iff for every ordinal β, if there is a sequence of ordinals $\langle \lambda_\alpha : \alpha < \beta \rangle$ such that $\lambda_\alpha < \kappa$ for every $\alpha < \beta$ and $\kappa = \bigcup_{\alpha<\beta} \lambda_\alpha$, then $\beta \geq \kappa$.*

Usually, we will use greek letters α, β to denote ordinals. We reserve κ to denote cardinals.

Now, before presenting the Veblen functions, we need two important concepts related to a regular cardinal κ. First, we define the notion of *closed* and *unbounded* in κ class of ordinals. Next, we define the notion of *normal* function in κ.

Definition 3.8. *Let M be a class of ordinals and let κ be a regular cardinal. We say that M is closed in κ iff for any subset $U \subseteq M \cap \kappa$ with $|U| < \kappa$ we have that $sup(U) \in M$.*
We say that M is unbounded in κ iff for any $\alpha < \kappa$ there is some $\beta \in M$ such that $\alpha < \beta < \kappa$.
If M is closed and unbounded in κ, we say that M is κ-club.

Moreover, if a class M is κ-club for any regular cardinal κ, then we will say that M is *club*.

Definition 3.9. *Let $f : On \to On$ be an ordinal function. We say that f is order-preserving iff for any ordinals α and β we have*

$$\alpha \leq \beta \to f(\alpha) \leq f(\beta).$$

We say that f is κ-normal iff f is order-preserving and

- $\kappa \subseteq dom(f)$,
- *for any subset* $U \subseteq \kappa$ *with* $|U| < \kappa$ *we have that*

$$sup(f[U]) = f(sup(U)).$$

Moreover, if an ordinal function f is κ-normal for any regular cardinal κ we will say that f is *normal.*

Lemma 3.10. *Given an order-preserving function* f*, for any ordinal* β *we have that* $\beta \leq f(\beta)$*. In particular, this result applies for* κ*-normal functions for any regular cardinal* κ*.*

Proof. We assume that this does not hold. Then, consider

$$\beta = min(\{\delta : \delta > f(\delta)\}).$$

By definition, we have that $f(\beta) < \beta$. Since f preserves the order, $f(f(\beta)) < f(\beta)$, contradicting the minimality of β. □

The next lemma relates κ-club classes to κ-normal functions. In particular, we will use *enumerating functions* of classes of ordinals. The enumerating function of M enumerates M by associating to each ordinal of the order-type of M an element of M in a strictly increasing manner.

Lemma 3.11. *Let* M *be a class of ordinals and let* κ *be a regular cardinal. Then,*

$$M \text{ is } \kappa\text{-club iff the enumerating function } en_M \text{ of } M \text{ is } \kappa\text{-normal.}$$

Proof. Let us start with the left to right implication.
We notice that $otype(M \cap \kappa) = \kappa$. Otherwise, it would be $otype(M \cap \kappa) = \alpha < \kappa$ and so $\kappa = \bigcup_{\beta<\alpha} en_{M\cap\kappa}(\beta)$, against the regularity of κ. This shows that $\kappa = dom(en_{M\cap\kappa}) \subseteq dom(en_M)$.
Now, let $U \subseteq M \cap \kappa$ with $|U| < \kappa$. Then $|en_M[U]| < \kappa$ and so $sup(en_M[U]) \in \kappa \cap M$ since κ is regular and M is closed. Therefore, there is $\alpha < \kappa$ such that $sup(en_M[U]) = en_M(\alpha)$. Let us see that $\alpha = sup(U)$. First, if $\beta \in U$ then $en_M(\beta) < en_M(\alpha)$ and so $\beta < \alpha$. This means that $sup(U) \leq \alpha$. Now, if the inequality was strict, we would have that $en_M(sup(U)) < en_M(\alpha) = sup(en_M[U])$ and so there would be some $\beta \in U$ such that $en_M(sup(U)) < en_M(\beta)$, against $\beta \leq sup(U)$. Hence, the desired result is obtained.

Now, suppose that en_M is κ-normal. Let $\alpha < \kappa$. Then, by Lemma 3.10, we have

$$\alpha < \alpha + 1 \leq en_M(\alpha + 1).$$

Also, $en_M(\alpha + 1) < \kappa$ and so M is unbounded in κ.
Let $U \subseteq M \cap \kappa$ with $|U| < \kappa$. Then, we have

$$sup(U) = sup(en_M[otype(U)]) = en_M(sup(otype[U])) \in M.$$

Hence, M is closed in κ. □

For an ordinal function f, we say that $x \in dom(f)$ is a fixpoint of f if it belongs to the class $Fix(f) := \{x \in dom(f) : f(x) = x\}$. We will now show that the fixpoints of a κ-normal function form a κ-club class. This means that the enumerating function of $Fix(f)$ is κ-normal and we can form a chain of κ-club classes and κ-normal functions by iterating the processus of taking fixpoints of the enumerating function of the previous class. This is how, in particular, Veblen functions work.

Lemma 3.12. *Let κ be a regular cardinal. Let f be a κ-normal function. Then, the class $Fix(f) = \{x \in dom(f) : f(x) = x\}$ of the fixpoints of f is κ-club.*

Proof. Let $\alpha < \kappa$. We define $\beta_0 = \alpha + 1$ and $\beta_{n+1} = f(\beta_n)$ for any $n < \omega$. Then, we define $\beta := sup\{\beta_n : n < \omega\} < \kappa$. By κ-normality, we get

$$f(\beta) = f(sup(\beta_n : n < \omega)) = sup(f(\beta_n : n < \omega)) = sup(\beta_{n+1} : n < \omega) = \beta.$$

Therefore, we obtain $\alpha < \beta \in Fix(f) \cap \kappa$. Hence, the class $Fix(f)$ is unbounded in κ.
Now, let $U \subseteq Fix(f) \cap \kappa$ with $|U| < \kappa$. We get

$$f(sup(U)) = sup(f[U]) = sup(U).$$

Thus, we obtain $sup(U) \in Fix(f)$. Hence, the class $Fix(f)$ is closed in κ. □

We are now able to define the Veblen functions.

Definition 3.13. *We define simultaneously $Cr(\alpha)$ and φ_α for every ordinal α by induction on α.*
$Cr(0) := \{\alpha : \forall \beta, \delta{<}\alpha \, (\beta + \delta < \alpha)\}$ and φ_0 enumerates $Cr(0)$.
$Cr(\alpha) := \{\beta : \forall \delta{<}\alpha \, (\varphi_\delta(\beta) = \beta)\}$ and φ_α enumerates $Cr(\alpha)$.

Sometimes, we will write $\varphi\alpha\beta$ instead of $\varphi_\alpha(\beta)$. Since each φ_α enumerates a class of ordinals, every φ_α is an ordinal order-preserving function.
The following results show that each $Cr(\alpha)$ is club and each φ_α is normal. We first see that this holds for $\alpha = 0$.

Lemma 3.14. *$Cr(0)$ is κ-club for any regular κ.*

Proof. To see that $Cr(0)$ is uncountable, let $\alpha < \kappa$. Let $\beta_0 = \alpha+1$ and $\beta_{n+1} = \beta_n+\beta_n$. It follows that $\beta := sup(\{\beta_n : n < \omega\}) < \kappa$. Therefore, we have the inequalities

$$\alpha < \beta_0 \leq \beta < \kappa.$$

Moreover β is additive principal: let $\delta, \gamma < \beta$. In particular, we have $\delta, \gamma < \beta_n$ for some n. Hence, $\delta + \gamma < \beta_n + \beta_n = \beta_{n+1} < \beta$.
Let $U \subseteq Cr(0) \cap \kappa$ with $U < \kappa$. Let $\alpha, \beta < sup(U)$. In particular, $\alpha, \beta < \gamma$ for some $\gamma \in U$. Then, since $\gamma \in Cr(0)$, by definition we obtain $\alpha + \beta < \gamma \leq sup(U)$. □

Corollary 3.15. *φ_0 is κ-normal for any regular κ.*

Proof. By Lemma 3.17, the class $Cr(0)$ is κ-club for any regular κ and so by Lemma 3.11 its enumerating function φ_α is κ-normal for any regular κ. □

Now, we prove this result for any ordinal α.

Lemma 3.16. *For any ordinal α and any uncountable regular cardinal $\kappa > \alpha$,*

$$Cr(\alpha) \text{ is } \kappa\text{-club and } \varphi_\alpha \text{ is } \kappa\text{-normal.}$$

Proof. We proceed by induction on α.

- The case $\alpha = 0$ is already done in Lemma 3.17 and Corollary 3.15.
- We suppose that $Cr(\alpha)$ is κ-club and φ_α is κ-normal. Then $Cr(\alpha+1)$ consists on the fixpoints of φ_α and so $Cr(\alpha+1)$ is κ-club by Lemma 3.12. Since $\varphi_{\alpha+1}$ enumerates $Cr(\alpha+1)$, we have that $\varphi_{\alpha+1}$ is κ-normal by Lemma 3.11.
- We suppose that α is a limit. Then $Cr(\alpha) = \bigcap_{\beta<\alpha} Cr(\beta)$. Since the intersection of less than κ many κ-club sets is κ-club (cf. Theorem 8.3 of [8]), the class $Cr(\alpha)$ is again κ-club. It follows that φ_α is κ-normal by Lemma 3.11.

□

The next Lemma gives an explicit description of the elements of $Cr(0)$, which are the additive principal ordinals.

Lemma 3.17. *Let α be an ordinal. Then $\alpha \in Cr(0)$ iff $\alpha = \omega^\beta$ for some ordinal β .*

Proof. First, we show that all ordinals of the form ω^β are additive principal by induction on β.

- $\omega^0 = 1$, which is additive principal.
- We suppose that ω^β is additive principal. Then, we have

$$\omega^{\beta+1} = \omega^\beta \cdot \omega = sup(\{\omega^\beta \cdot n : n < \omega\}.$$

  Let $\delta, \gamma < \omega^{\beta+1}$. That is, there is $n < \omega$ such that $\delta, \gamma < \omega^\beta \cdot n$. Therefore, we get

$$\delta + \gamma < \omega^\beta \cdot n + \omega^\beta \cdot n = \omega^\beta \cdot (2n) < \omega^{\beta+1}.$$

- We suppose that β is a limit ordinal. Then, we have

$$\omega^\beta = sup(\{\omega^\xi : \xi < \beta\}).$$

  Let $\delta, \gamma < \omega^\beta$. That is, there is $\xi < \beta$ such that $\delta, \gamma < \omega^\xi$. By the induction hypothesis, we obtain $\delta + \gamma < \omega^\xi \leq \omega^\beta$.

Now, we will see that any additive principal ordinal is of the form ω^β. We suppose not. The, let $\alpha \in Cr(0)$ such that $\alpha \neq \omega^\beta$ for any β. Since the class $\{\omega^\beta : \beta \text{ is an ordinal}\}$ is unbounded, there is some β such that $\alpha < \omega^\beta$. In fact, the least such ordinal, that we call $\beta = min(\{\delta : \alpha < \omega^\delta\})$, has to be a successor ordinal. Otherwise, $\alpha > \omega^\delta$ for all $\delta < \beta$ and so $\alpha \geq sup(\{\omega^\delta : \delta < \beta\}) = \omega^\beta$, a contradiction.
We now know that there is β such that

$$\omega^\beta < \alpha < \omega^{\beta+1} = \omega^\beta \cdot \omega = sup(\{\omega^\beta \cdot n : n < \omega\}).$$

Therefore, there is some $n < \omega$ such that

$$\omega^\beta \cdot n < \alpha < \omega^\beta \cdot (n+1) = \omega^\beta \cdot n + \omega^\beta.$$

But since α is additive principal, we have

$$\omega^\beta \cdot n + \omega^\beta \cdot n < \alpha < \omega^\beta \cdot n + \omega^\beta < \omega^\beta \cdot n + \omega^\beta \cdot n,$$

which is a contradiction. Hence $\alpha \notin Cr(0)$. □

The first element of $Cr(1)$ is ε_0, which is the limit of $\{\omega, \omega^\omega, \omega^{\omega^\omega}, \dots\}$.
By means of Lemma 3.17, we can write each ordinal as the sum of a finite number of additive principal ordinals of the form ω^β. This is how we define the Cantor normal form of an ordinal.

Definition 3.18. *Let α be an ordinal. We define the Cantor normal form of α as follows.*

$$\alpha =_{NF} \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \text{ if } \alpha \geq \alpha_1 \geq \cdots \geq \alpha_n \text{ and } \alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}.$$

If $\alpha = 0$, then taking $n = 0$ we get an empty sum, which by convention equals 0.

We show that Veblen functions are monotone on the first component.

Lemma 3.19. *Let α and β be ordinals. If $\alpha < \beta$, then $Cr(\beta) \subsetneq Cr(\alpha)$ and $\varphi_\alpha(\gamma) \leq \varphi_\beta(\gamma)$ for any ordinal γ.*

Proof. We suppose that $\alpha < \beta$. If $\delta \in Cr(\beta)$, then $\varphi_\xi(\delta) = \delta$ for all $\xi < \beta$ and so, in particular, for all $\xi < \alpha$, showing that $\delta \in Cr(\alpha)$.
Let us see that the inclusion is proper. Let $\delta = \varphi_\alpha(0) \in Cr(\alpha)$. Then $\delta > 0$ (since in particular $0 \notin Cr(0)$). Therefore, we have $\varphi_\alpha(\delta) > \varphi_\alpha(0) = \delta$ and so $\delta \notin Cr(\beta)$. □

The next lemma shows that the outputs of each φ_β are fixpoints of the previous Veblen functions.

Lemma 3.20. *Let α and β be ordinals. If $\alpha < \beta$ then $\varphi_\alpha(\varphi_\beta(\gamma)) = \varphi_\beta(\gamma)$ for any ordinal γ.*

Proof. We observe that $\varphi_\beta(\gamma) \in Cr(\beta)$. Since $\alpha < \beta$, it must be a fixpoint of φ_α, and so $\varphi_\alpha(\varphi_\beta(\gamma)) = \varphi_\beta(\gamma)$. □

Since we want to represent ordinals with Veblen functions, we are interested in knowing how some ordinal can be the output of different Veblen functions.

Lemma 3.21. *Let $\alpha_1, \alpha_2, \beta_1, \beta_2$ be ordinals. Then*

$$\varphi\alpha_1\beta_1 = \varphi\alpha_2\beta_2 \Leftrightarrow \begin{cases} \alpha_1 = \alpha_2 \wedge \beta_1 = \beta_2, \text{ or} \\ \alpha_1 < \alpha_2 \wedge \beta_1 = \varphi\alpha_2\beta_2, \text{ or} \\ \alpha_2 < \alpha_1 \wedge \beta_2 = \varphi\alpha_1\beta_1. \end{cases}$$

Proof. If $\alpha_1 = \alpha_2$, then ($\varphi\alpha_1\beta_1 = \varphi\alpha_1\beta_2$ iff $\beta_1 = \beta_2$), since $\varphi\alpha_1$ is an enumerating function.
If $\alpha_1 < \alpha_2$, then $\varphi\alpha_1(\varphi\alpha_2\beta_2) = \varphi\alpha_2\beta_2$ by Lemma 3.21. Therefore, we obtain the equivalence

$$(\varphi\alpha_1\beta_1 = \varphi\alpha_2\beta_2 \text{ iff } \varphi\alpha_1\beta_1 = \varphi\alpha_1(\varphi\alpha_2\beta_2) \text{ iff } \beta_1 = \varphi\alpha_2\beta_2).$$

The case $\alpha_2 < \alpha_1$ is analogous. ☐

A similar proof shows the following Lemma.

Lemma 3.22. *Let $\alpha_1, \alpha_2, \beta_1, \beta_2$ be ordinals. Then*

$$\varphi\alpha_1\beta_1 < \varphi\alpha_2\beta_2 \Leftrightarrow \begin{cases} \alpha_1 = \alpha_2 \wedge \beta_1 < \beta_2, \text{ or} \\ \alpha_1 < \alpha_2 \wedge \beta_1 < \varphi\alpha_2\beta_2, \text{ or} \\ \alpha_2 < \alpha_1 \wedge \beta_2 < \varphi\alpha_1\beta_1. \end{cases}$$

Lemma 3.23. *For any ordinal α, all fixpoints of φ_α are limit ordinals.*

Proof. First, we observe that $0 < \varphi\alpha 0$ for any α and so 0 is not a fixpoint for any α. Let us see that successor ordinals are not fixpoints.
We suppose that $\beta + 1 = \varphi_\alpha(\beta + 1)$. Then, given $\delta \leq \alpha$, we have

$$\varphi_\delta(\beta + 1) = \varphi_\delta(\varphi_\alpha(\beta + 1)) = \varphi_\alpha(\beta + 1) = \beta + 1.$$

In particular, taking $\delta = 0$ yields

$$\omega^{\beta+1} = \varphi_0(\beta + 1) = \beta + 1,$$

a contradiction since ω^ξ is always a limit. ☐

Lemma 3.24. *Let α and $\beta > 0$ be ordinals. Then $\alpha \leq \varphi\alpha 0 < \varphi\alpha\beta$.*

Proof. We proceed by induction on α.

- $0 < 1 = \varphi_0(0)$.
- We suppose that $\alpha \leq \varphi_\alpha(0)$. But $\varphi_\alpha(0) < \varphi_{\alpha+1}(0)$ and so $\alpha < \varphi_{\alpha+1}(0)$. It follows that $\alpha + 1 \leq \varphi_{\alpha+1}(0)$. In fact, by the previous Lemma, we get that $\alpha + 1 < \varphi_{\alpha+1}(0)$.
- We suppose that α is a limit and $\beta \leq \varphi_\beta(0)$ for all $\beta < \alpha$. Fix $\beta < \alpha$. Then $\varphi_\beta(\varphi_\alpha(0)) = \varphi_\alpha(0) > \varphi_\beta(0) \geq \beta$. Therefore, for any $\beta < \alpha$ we have $\beta < \varphi_\alpha(0)$. It follows that $\alpha \leq \varphi\alpha 0$.

☐

We show that any additive principal ordinal α can be written as the output of some Veblen function with second argument below α.

Lemma 3.25. *Let $\alpha \in Cr(0)$. There are uniquely determined ordinals δ, β such that $\alpha = \varphi\delta\beta$ and $\beta < \alpha$.*

Proof. We start by proving the existence of δ and β.
Let $\gamma = min(\{\delta : \alpha < \varphi\delta\alpha\})$. This minimum exists because $\alpha \leq \varphi\alpha 0 < \varphi_\alpha(\alpha)$ by Lemma 3.24 (in the case $\alpha = 0$ we have $0 < \varphi_0(0)$). If $\gamma = 0$, then $\alpha < \varphi 0\alpha$. But $\alpha = \varphi 0\beta$ for some β and $\varphi 0\beta = \alpha < \varphi 0\alpha$. This shows that $\beta < \alpha$.
We suppose that $\gamma \neq 0$. Then, we have $\alpha \geq \varphi\delta\gamma$ for all $\delta < \gamma$. Since $\alpha \leq \varphi\delta\alpha$, it must be $\alpha = \varphi\delta\alpha$ for all $\delta < \gamma$, and so $\alpha \in Cr(\delta)$ for all $\delta < \gamma$. Therefore, $\alpha \in Cr(\gamma)$ and so there is some β with $\alpha = \varphi\gamma\beta$. But by the definition of γ, $\varphi\gamma\beta = \alpha < \varphi\gamma\alpha$ and so $\beta < \alpha$.

We prove uniqueness. We suppose that $\alpha = \varphi\delta\beta = \varphi\gamma\xi$ with $\beta, \xi < \alpha$.
If $\delta < \gamma$, then $\alpha = \varphi\delta\beta < \varphi\delta\alpha = \varphi\delta(\varphi\gamma\xi) = \varphi\gamma\xi = \alpha$, a contradiction. Hence, we have $\delta = \gamma$. It follows that $\beta = \xi$. □

We define the class of *strongly critical* ordinals. An ordinal α is strongly critical if α is a fixpoint of every φ_β, for $\beta < \alpha$.

Definition 3.26. $SC := \{\alpha : \alpha \in Cr(\alpha)\}$.
The function $\lambda\beta.\Gamma_\beta$ enumerates SC.

Lemma 3.27. *Let α be an ordinal. Then*

$$\alpha \in Cr(\alpha) \textit{ iff } \varphi\beta\delta < \alpha \textit{ for any } \beta, \delta < \alpha.$$

Proof. First, we notice that ($\alpha \in Cr(\alpha)$ iff $\alpha = \varphi\alpha 0$) since $\alpha \leq \varphi\alpha 0 < \varphi\alpha\beta$ for any $\beta > 0$.
We suppose that $\alpha \in Cr(\alpha)$ and so $\alpha = \varphi\alpha 0$. Let $\beta, \delta < \alpha$. Then

$$\varphi\beta\alpha = \varphi_\beta(\varphi\alpha 0) = \varphi\alpha 0 = \alpha,$$

and so $\alpha = \varphi\beta\alpha$. Therefore, we obtain $\varphi\beta\delta < \varphi\beta\alpha = \alpha$.
Now, suppose that $\varphi\beta\delta < \alpha$ for any $\beta, \delta < \alpha$. Fix $\beta < \alpha$. Then, we have $\varphi\beta\delta < \alpha$ for any $\delta < \alpha$. It follows that

$$\varphi\beta\alpha = sup(\{\varphi\beta\delta : \delta < \alpha\}) \leq \alpha.$$

But $\alpha \leq \varphi\beta\alpha$. Hence, we get $\alpha = \varphi\beta\alpha$, and so $\alpha \in Cr(\beta)$ for any $\beta < \alpha$. This means that $\alpha \in Cr(\alpha)$. □

We notice that the proof shows in particular that

$$\begin{aligned} \alpha \in SC & \text{ iff } \alpha \in Cr(\alpha) \\ & \text{ iff } \alpha = \varphi\alpha 0 \\ & \text{ iff } \varphi\beta\delta < \alpha \text{ for any } \beta, \delta < \alpha \\ & \text{ iff } \alpha = \varphi\beta\alpha \text{ for any } \beta < \alpha. \end{aligned}$$

A deeper study of Veblen functions can be found in [13]or in [17].

3.2 The ordinal notation system

Now, we will construct the ordinal notation system $T(\theta)$ that we will use to build the $\mathsf{RS}_l(X)$-logic. Recall that $\mathsf{RS}_l(X)$ will depend on some fixed set X.

Reading Convention 3.28. *The set X is a fixed set. The set-theoretic rank of X is θ. The sequence $\langle \Omega_n : n \leq \omega \rangle$ enumerates the first $\omega + 1$ uncountable regular cardinals κ such that $\kappa > \theta$. We have that $L_{\Omega_n}(X) \vDash \mathsf{KP}$ for every $n \leq \omega$.*

We first have to study the ordinal machinery needed to define $T(\theta)$. One of the features of $\mathsf{RS}_l(X)$ for which we will use ordinals will be to associate an ordinal level to each term of the system. This means that, since the starting point of the relativized constructible hierarchy is the transitive closure of X, the sets of $TC(\{X\})$ will be seen as basic elements, and the level of those basic terms will be some initial ordinals. Thus, we need one initial ordinal per element of $TC(\{X\})$, that is, we need "θ-many" ordinals. Those ordinals will be chosen to be the first "θ-many" strongly critical ordinals. That is why we will include in $T(\theta)$ a representation of every Γ_β for $\beta \leq \theta$ (see Definition 3.26).
Moreover, we want the set $T(\theta)$ to contain a representation of each ordinal Ω_n for $n \leq \omega$ because the $\mathsf{RS}_l(X)$-system will be mainly based on the admissible sets of the relativized constructible hierarchy, that is, on each $\mathbb{L}_{\Omega_n}(X)$ for $n < \omega$. Finally, the set of ordinals represented by a string in $T(\theta)$ will be closed under addition, the Veblen function φ and some Collapsing functions ψ_n for $n < \omega$ that we will use in order to collapse ordinals under every Ω_n for $n < \omega$. We define the ψ_n functions as follows, given Reading Convention 3.28.

Definition 3.29. *We define simultaneously the sets $B_n(\alpha)$ of ordinals and the ordinal function $\psi_n(\alpha)$. For every $n < \omega$, we define $B_n(\alpha) = \bigcup_{k<\omega} B_n^k(\alpha)$, where $B_n^k(\alpha)$ is defined by double induction on n and k as follows.*

- $B_0^0(\alpha) = \{0\} \cup \{\Gamma_\beta : \beta \leq \theta\} \cup \{\Omega_m : m \leq \omega\}$,
  $B_0^{k+1}(\alpha) = B_0^k \cup \{\delta + \delta', \varphi\delta\delta' : \delta, \delta' \in B_0^k(\alpha)\} \cup \{\psi_m(\delta) : \delta \in B_0^k(\alpha) \wedge \delta < \alpha \wedge m < \omega\}$.

- $B_{n+1}^0(\alpha) = \Omega_n \cup \{\Omega_m : m \leq \omega\}$,
  $B_{n+1}^{k+1}(\alpha) = B_{n+1}^k \cup \{\delta + \delta', \varphi\delta\delta' : \delta, \delta' \in B_{n+1}^k(\alpha)\} \cup \{\psi_m(\delta) : \delta \in B_{n+1}^k(\alpha) \wedge \delta < \alpha \wedge m < \omega\}$.

The ordinal collapsing function ψ_n is defined as $\psi_n(\alpha) = min\{\beta : \beta \notin B_n(\alpha)\}$.

We will sometimes write $\psi_n\alpha$ instead of $\psi_n(\alpha)$. Now, we will be studying the sets $B_n(\alpha)$ and the collapsing functions ψ_n. The next lemma shows that the function ψ_0 takes values on the strongly critical ordinals between $\Gamma_{\theta+1}$ and Ω_0, and that ψ_{n+1} takes values on the strongly critical ordinals between Ω_n and Ω_{n+1} for every $n < \omega$. In particular, for any natural number n the function $\psi_n(\alpha)$ collapses any ordinal under Ω_n: this is an important property, since, in our system $\mathsf{RS}_l(X)$, we will want to collapse some ordinal bounds under some Ω_n, as we will see later.

Lemma 3.30. *For every ordinal α and every natural number n, we have:*

1. *$\psi_n(\alpha)$ is a strongly critical ordinal,*
2. *$|B_0(\alpha)| = \max(\aleph_0, |\theta|)$ and $|B_{n+1}(\alpha)| = \Omega_n$,*
3. *$\Gamma_{\theta+1} \leq \psi_0(\alpha) < \Omega_0$ and $\Omega_n < \psi_{n+1}(\alpha) < \Omega_{n+1}$.*

Proof.

1. We will show that $\psi_n(\alpha) = \varphi_{\psi_n(\alpha)}(0)$. First, we write $\psi_n(\alpha)$ in Cantor normal form to show, in the first place, that $\psi_n(\alpha)$ is additive principal:

$$\psi_n(\alpha) = \omega^{\alpha_1} + \cdots + \omega^{\alpha_m},$$

with $\alpha_1 \geq \cdots \geq \alpha_m$. We need to see that $m = 1$. We suppose not. Then, $\alpha_1, \ldots, \alpha_m < \psi_n(\alpha)$ and so $\alpha_1, \ldots, \alpha_m \in B_n(\alpha)$. But then $\psi_n(\alpha) = \varphi_0(\alpha_1) + \cdots + \varphi_0(\alpha_m) \in B_n(\alpha)$, a contradiction with the definition of the ψ_n function. Therefore, $\psi_n(\alpha) = \varphi_0(\alpha_1)$ with $\alpha_1 \leq \psi_n(\alpha)$ and so $\psi_n(\alpha)$ is additive principal. Now, by Lemma 3.25 we can find some ordinals $\delta \leq \psi_n(\alpha)$ and $\gamma < \psi_n(\alpha)$ such that $\psi_n(\alpha) = \varphi\delta\gamma$. We will show that $\gamma = 0$, which yields $\delta = \psi_n(\alpha)$. So, suppose that $\gamma \neq 0$. It follows that $\delta \leq \varphi\delta 0 < \varphi\delta\gamma = \psi_n(\alpha)$. That is, both

γ and δ are strictly below $\psi_n(\alpha)$, which means that $\gamma, \delta \in B_n(\alpha)$. But then $\psi_n(\alpha) = \varphi\delta\gamma \in B_n(\alpha)$, in contradiction with the definition of the ψ_n function. Hence,

$$\psi_n(\alpha) = \varphi_{\psi_n(\alpha)}(0),$$

which means that $\psi_n(\alpha)$ is strongly critical.

2. First, we study the case $n = 0$. We suppose that $|\theta|$ is infinite. We have

$$\begin{aligned} |B_0^0(\alpha)| &= |\{0\} \cup \{\Gamma_\beta : \beta \leq \theta\} \cup \{\Omega_m : m \leq \omega\}| \\ &= |\{\Gamma_\beta : \beta < \theta\}| \\ &= |\theta| \end{aligned}$$

and, if $|B_0^k(\alpha)| = |\theta|$ then also $|B_0^{k+1}(\alpha)| = |\theta|$. This shows that $|B_0^k(\alpha)| = |\theta|$ for all $k < \omega$. Therefore, $B_0(\alpha)$ is a countable union of sets of cardinality $|\theta|$. Thus, the cardinality of $B_0(\alpha)$ is $|\theta|$.
If $|\theta|$ is finite, then $|B_0^0(\alpha)| = |\{\Omega_n : n < \omega\}| = \aleph_0$ and by induction we can prove that $B_0(\alpha)$ is a countable union of sets of cardinality $\aleph_0$, and so $|B_0(\alpha)| = \aleph_0$.
Now, we assume that $n > 0$. Then $|B_n^0(\alpha)| = \Omega_n$ and, whenever $|B_n^m(\alpha)| = \Omega_n$, also $|B_n^{m+1}(\alpha)| = \Omega_n$ since $B_n^{m+1}(\alpha)$ is a countable union of sets of cardinality Ω_n. Thus, $B_n(\alpha)$ is a countable union of sets of cardinality Ω_n. Hence, the cardinality of $B_n(\alpha)$ is Ω_n.

3. We first study the case $n = 0$. By the first item of this lemma, we have that $\psi_0(\alpha)$ is strongly critical. Moreover, by definition $\{\Gamma_\beta : \beta \leq \theta\} \subseteq B_0(\alpha)$. This means that the first strongly critical ordinals up to and including Γ_θ must be below $\psi_0(\alpha)$. Therefore, $\Gamma_{\theta+1} \leq \psi_0(\alpha)$. Furthermore, if $\Omega_0 \leq \psi_0(\alpha)$ we would have that $\Omega_0 \subseteq B_0(\alpha)$, contradicting Item 2.
Now, we assume that $n > 0$. By definition, for any $n < \omega$ we have that $\Omega_n \subseteq B_{n+1}(\alpha)$ and so $\Omega_n < \psi_{n+1}(\alpha)$. If $\psi_{n+1}(\alpha) \geq \Omega_{n+1}$ then $\Omega_{n+1} \subseteq B_{n+1}(\alpha)$, in contradiction with Item 2.

□

The next lemma states more results about the sets $B_n(\alpha)$ and the functions ψ_n. In particular, Items 6., 7. and 8. mean that when $\alpha \in B_n(\alpha)$, the function ψ_n on input $\alpha + 1$ grows to the next strongly critical ordinal and, when $\alpha \notin B_n(\alpha)$, the function ψ_n on input $\alpha + 1$ stays at the same value as on input α.

Lemma 3.31. *For any ordinals α and β and for any natural number n, we have:*

1. *If* $\beta < \alpha$ *then* $B_n(\beta) \subseteq B_n(\alpha)$ *and* $\psi_n(\beta) \leq \psi_n(\alpha)$,
2. *If* $\beta \in B_n(\alpha) \cap \alpha$ *then* $\psi_n(\beta) < \psi_n(\alpha)$,
3. *If* $\beta \leq \alpha$ *and* $[\beta, \alpha) \cap B_n(\alpha) = \emptyset$ *then* $B_n(\beta) = B_n(\alpha)$,
4. $B_n(\alpha) \cap \Omega_n = \psi_n(\alpha)$,
5. *If* α *is a limit then* $B_n(\alpha) = \bigcup_{\beta<\alpha} B_n(\beta)$ *and* $\psi_n(\alpha) = sup(\psi_n(\beta) : \beta < \alpha)$,
6. $\psi_n(\alpha+1) \in \{\psi_n(\alpha), (\psi_n(\alpha))^\Gamma\}$, *where* δ^Γ *is defined as the first strongly critical ordinal above* δ,
7. *If* $\alpha \in B_n(\alpha)$ *then* $\psi_n(\alpha+1) = (\psi_n(\alpha))^\Gamma$,
8. *If* $\alpha \notin B_n(\alpha)$ *then* $\psi_n(\alpha+1) = \psi_n(\alpha)$.

Proof. We prove Item 1. Fix some natural number n.
Let $\beta < \alpha$. First, we note that $B_n^0(\beta) \subseteq B_n^0(\alpha)$. But from $\beta < \alpha$ we get

$$\psi_n \restriction \beta = (\psi_n \restriction \alpha) \restriction \beta. \qquad (\star)$$

Now, suppose that $B_n^k(\beta) \subseteq B_n^k(\alpha)$. We have that $B_n^{k+1}(\alpha)$ is closed under every function under which $B_0^{k+1}(\beta)$ is closed by $\star$. This means that $B_n^{k+1}(\beta) \subseteq B_n^{k+1}(\alpha)$. Hence, we obtain $B_n(\beta) \subseteq B_n(\alpha)$. Moreover, we get by definition $\psi_n(\beta) \leq \psi_n(\alpha)$.

The proof of the rest of the items is analogous to the proof of Lemma 2.6 in [5] (page 5). □

Thanks to those properties, we are able to define the *normal form* of an ordinal α. This normal form is an extension of Cantor normal form that gives representations to some strongly critical ordinals.

Definition 3.32. *Let* α *be an ordinal. We define the normal form of* α *as follows.*

1. $\alpha =_{NF} \alpha_1 + \cdots + \alpha_n$ *iff* $\alpha = \alpha_1 + \cdots + \alpha_n$, $n > 1$, *where the ordinals* $\alpha_1, \ldots, \alpha_n$ *are written in normal form and are additive principal and* $\alpha_1 \geq \cdots \geq \alpha_n$,
2. $\alpha =_{NF} \varphi\alpha_1\alpha_2$ *iff* $\alpha = \varphi\alpha_1\alpha_2$ *with* $\alpha_1, \alpha_2 < \alpha$ *and* α_1, α_2 *are written in normal form,*
3. $\alpha =_{NF} \psi_n(\alpha_1)$ *iff* $\alpha = \psi_n(\alpha_1)$ *with* $\alpha_1 \in B_n(\alpha_1)$ *and* α_1 *is written in normal form.*

The idea is that, starting with some basic ordinals (the ordinal 0, the first strongly critical ordinals up to Γ_θ and the Ω_n ordinals), we will construct ordinals using these normal forms, so that each ordinal constructed this way has a unique representation. This is the reason why we add the condition $\alpha_1 \in B_n(\alpha_1)$ in Item 3.: it may be the case that $\alpha = \psi_n(\beta) = \psi_n(\beta+1) = \cdots = \psi_n(\alpha_1)$, but the condition $\alpha_1 \in B(\alpha_1)$ forces $\psi_n(\alpha_1+1)$ to take another value (for instance, the first strongly critical ordinal above α). Therefore, α_1 is the greatest ordinal with image α, and we choose $\psi_n(\alpha_1)$ to represent α.
This also motivates the next lemma that states that an ordinal is in some $B_n(\alpha)$ if and only if its "normal form components" are.

Lemma 3.33. *Let α, γ be any ordinals and let m be any natural number.*

1. *If $\alpha =_{NF} \alpha_1 + \cdots + \alpha_n$ then $[\alpha \in B_m(\gamma)$ iff $\alpha_1, \ldots, \alpha_n \in B_m(\gamma)]$,*
2. *If $\alpha =_{NF} \varphi\alpha_1\alpha_2$ then $[\alpha \in B_m(\gamma)$ iff $\alpha_1, \alpha_2 \in B_m(\alpha)]$,*
3. *If $\alpha =_{NF} \psi_m(\alpha_1)$ then $[\alpha \in B_m(\gamma)$ iff $\alpha_1 \in B_m(\gamma) \cap \gamma]$.*

Proof. 1. Let $\alpha =_{NF} \alpha_1 + \cdots \alpha_n$. If $\alpha_1, \ldots, \alpha_n \in B_m(\gamma)$, then, since $B_m(\gamma)$ is closed under addition, we have $\alpha \in B_m(\gamma)$ too.
We suppose now that $\alpha \in B_m(\gamma)$. We define for any ordinal β in normal form the set $AP(\beta)$ of additive predecessors of β as follows

$$AP(\beta) = \begin{cases} \emptyset & \text{if } \beta = 0 \\ \{\beta\} & \text{if } \beta \text{ is additive principal} \\ \{\beta_1, \ldots, \beta_k\} & \text{if } \beta =_{NF} \beta_1 + \cdots + \beta_k \end{cases}$$

We also define the set

$$Y = \{\delta \in B_m(\gamma) : AP(\delta) \subseteq B_m(\gamma)\}.$$

The inclusion $Y \subseteq B_m(\gamma)$ is obvious by definition. Our objective is to show that $B_m(\gamma) \subseteq Y$. This way, we will have $B_m(\gamma) = Y$ and so, since $\alpha \in B_m(\gamma) = Y$ we will conclude that $\alpha_1, \ldots, \alpha_n \in Y = B_m(\gamma)$.
Now, we have $AP(0) = \emptyset \subseteq B_m(\gamma)$. Also, we have $AP(\delta) = \{\delta\} \subseteq B_m(\gamma)$ for any $\delta \in \{\Omega_k : k \leq \omega\} \cup \{\Gamma_\beta : \beta \leq \theta\}$. Therefore, we get $\{0\} \cup \{\Omega_k : k \leq \omega\} \cup \{\Gamma_\beta : \beta \leq \theta\} \subseteq Y$. Also, in the case that $m \neq 0$, if $\beta < \Omega_m$, then $AP(\beta) \subseteq \Omega_m \subseteq B_m(\gamma)$ and so $AP(\beta) \subseteq Y$. Now, the set Y is closed under addition, any Veblen function and the ψ_m function restricted to γ. Indeed, let $\delta, \xi \in Y$. Then

$AP(\delta+\xi) \subseteq AP(\delta) \cup AP(\xi) \subseteq B_m(\gamma)$. Since $\varphi\delta\xi$ is always additive principal, we have $AP(\varphi\delta\xi) = \{\varphi\delta\xi\} \subseteq B_m(\gamma)$. At last, if $\delta < \gamma$ then $AP(\psi_m(\delta)) = \{\psi_m(\delta)\} \subseteq B_m(\gamma)$ since $\psi_m(\delta)$ is always additive principal.

2. Let $\alpha =_{NF} \varphi\alpha_1\alpha_2$. If $\alpha_1, \alpha_2 \in B_m(\gamma)$, then, since $B_m(\gamma)$ is closed under φ, we have $\alpha \in B_m(\gamma)$ too.
We suppose now that $\alpha \in B_m(\gamma)$. We define for any ordinal β in normal form the set $PP(\beta)$ of predicative predecessors of β as follows

$$PP(\beta) = \begin{cases} \emptyset & \text{if } \beta = 0 \\ \{\beta\} & \text{if } \beta \text{ is strongly critical} \\ \{\beta_1, \beta_2\} & \text{if } \beta =_{NF} \varphi\beta_1\beta_2 \end{cases}$$

We also define the set

$$Y = \{\delta \in B_m(\gamma) : PP(\delta) \subseteq B_m(\gamma)\}.$$

By a reasoning analogous to the one in the first item of this lemma we obtain $Y = B_m(\gamma)$, and so $\alpha_1, \alpha_2 \in B_m(\gamma)$.

3. Let $\alpha = \psi_m(\alpha_1)$. If $\alpha_1 \in B_m(\gamma)$ then $\alpha \in B_m(\gamma)$ because $B_m(\gamma)$ is closed under $\psi_m \restriction \gamma$.
We suppose now that $\alpha \in B_m(\gamma)$. Then we get $\psi_m(\alpha_1) < \psi_m(\gamma)$ which means that $\alpha_1 < \gamma$. But, by Definition 3.32, we have $\alpha_1 \in B_m(\alpha_1)$. Therefore, $B_m(\alpha_1) \subseteq B_m(\gamma)$, which yields $\alpha_1 \in B_m(\gamma) \cup \gamma$. □

At this point, we can do a first step in the definition of $T(\theta)$: we construct $R(\theta)$, the set of ordinals that the strings of $T(\theta)$ will intend to denote.

Definition 3.34. *We inductively define the set $R(\theta)$ together with the complexity $C\alpha \in \omega$ of its elements.*

1. $0 \in R(\theta)$ *and* $C0 = 0$.
2. *For every* $n < \omega$, $\Omega_n \in R(\theta)$ *and* $C\Omega_n = 0$.
3. *For every* $\beta \leq \theta$, $\Gamma_\theta \in R(\theta)$ *and* $C\Gamma_\theta = 0$.
4. *If* $\alpha_1, \ldots, \alpha_n \in R(\theta)$ *and* $\alpha =_{NF} \alpha_1 + \cdots + \alpha_n$ *then* $\alpha \in R(\theta)$ *and* $C\alpha = \max(C\alpha_1, \ldots, C\alpha_n) + 1$.

5. *If* $\alpha_1, \alpha_2 \in R(\theta)$ *and* $\alpha =_{NF} \varphi\alpha_1\alpha_2$ *then* $\alpha \in R(\theta)$ *and* $C\alpha = \max(C\alpha_1, C\alpha_2) + 1$.

6. *If* $\alpha_1 \in R(\theta)$ *and* $\alpha =_{NF} \psi_n\alpha_1$ *then* $\alpha \in R(\theta)$ *and* $C\alpha = C\alpha_1 + 1$.

Let $\alpha \in R(\theta)$. If α is either 0, Γ_β for some $\beta \leq \theta$ or Ω_n for some $n < \omega$, then α has no normal forms and so $C\alpha$ is uniquely determined to be 0. If α is not an $R(\theta)$-basic ordinal, then by Lemma 3.33 it is included in $R(\theta)$ due to exactly one of the rules of this definition, and so $C\alpha$ is uniquely determined.
Recall that we want to transform $R(\theta)$ into a recursive representation system. So the first problem is that we have to computably deal with the condition $\alpha_1 \in B_n(\alpha_1)$ in Definition 3.32(3.). To do this, we define for each $n < \omega$ and each $\alpha \in R(\theta)$ the set of ordinals $\mathrm{Arg}_n(\alpha)$, that consists in all the ordinals that occur in the normal form of α as an argument of the ψ_n function.

Definition 3.35. *Let* $n < \omega$. *We define for each* $\alpha \in R(\theta)$ *the set of ordinals* $\mathrm{Arg}_n(\alpha)$ *by induction on* $C\alpha$ *as follows*

1. $\mathrm{Arg}_n(0) = \mathrm{Arg}_n(\Gamma_\beta) = \mathrm{Arg}_n(\Omega_m) = \emptyset$ *for all* $\beta \leq \theta$ *and all* $m < \omega$,

2. *If* $\alpha =_{NF} \alpha_1 + \cdots + \alpha_m$ *then* $\mathrm{Arg}_n(\alpha) = \mathrm{Arg}_n(\alpha_1) \cup \cdots \cup \mathrm{Arg}_n(\alpha_m)$,

3. *If* $\alpha =_{NF} \varphi\alpha_1\alpha_2$ *then* $\mathrm{Arg}_n(\alpha) = \mathrm{Arg}_n(\alpha_1) \cup \mathrm{Arg}_n(\alpha_2)$,

4. *If* $\alpha =_{NF} \psi_m(\alpha_1)$ *with* $m \neq n$ *then* $\mathrm{Arg}_n(\alpha) = \mathrm{Arg}_n(\alpha_1)$,

5. *If* $\alpha =_{NF} \psi_n(\alpha_1)$ *then* $\mathrm{Arg}_n(\alpha) = \{\alpha_1\} \cup \mathrm{Arg}_n(\alpha_1)$.

An easy induction on $C\alpha$ shows the next lemma, that gives a recursive equivalence to the condition $\alpha_1 \in B_n(\alpha_1)$ in Definition 3.32.

Lemma 3.36. *Let* $\alpha, \beta \in R(\theta)$. *Let* $n < \omega$. *Then,*

$$\alpha \in B_n(\beta) \text{ iff } \forall\delta \in \mathrm{Arg}_n(\alpha)(\delta < \beta).$$

We define $T(\theta)$ as the set of unique representations of ordinals in $R(\theta)$.

Definition 3.37. *We define* $T(\theta)$ *as the set of strings in the language* $\{0, +, \varphi\} \cup \{\Gamma_\beta : \beta < \theta\} \cup \{\Omega_n : n \leq \omega\} \cup \{\psi_n : n < \omega\}$ *corresponding to ordinals in* $R(\theta)$ *written in normal form, as in Definition* 3.32.

Strings in $T(\theta)$ *are ordered by the order induced from the ordering of ordinals in* $R(\theta)$. *Let* $\prec$ *denote this order.*

Theorem 3.38. *The set $T(\theta)$ and the relation $\prec$ on $T(\theta)$ are primitive recursive in θ.*

Again, this theorem can be proved by induction on $C\alpha$. We refer to [5] (Theorem 2.13 in page 9) for the proof.

4 The system $\mathsf{RS}_l(X)$

Now that we have the needed ordinal machinery, we can develop our system $\mathsf{RS}_l(X)$ for each set X. We will first define the terms of the system, which will correspond to elements of the constructible hierarchy relativized to X (see Definition 3.1). Then, we will define the formulas of the system together with some syntactical features related to the rules of derivations, that we will define after that.
Since the objective is to embed KPI into $\mathsf{RS}_l(X)$, we need to have a (Cut) rule. Nonetheless, it is complicated to trace back a proof from a derivation that has used the (Cut) rule since the active formulas in the premises do not appear in the conclusion. We recall that, in our main theorem, we consider a provably in KPI total and uniformly Σ-definable set-recursive function f. We want to show that for any set x the value $f(x)$ belongs to an initial segment of the constructible universe relativized to x at level below the relativized Takeuti-Feferman-Buchholz ordinal. In the proof of our main theorem, we will eventually show that there is a derivation in $\mathsf{RS}_l(X)$ of $\exists y(A_f(X,y))^{\mathbb{L}_\alpha(X)}$ for some α, where $\mathbb{L}_\alpha(X)$ is an $\mathsf{RS}_l(X)$-term representing the α-th stage of the constructible hierarchy and A_f is the formula that uniformly defines f. From this result, we will be able to show by induction on the length of the $\mathsf{RS}_l(X)$-proof that, indeed, there is a set y in the set $L_\alpha(X)$ such that $A_f(X,y)$ is satisfied in $L_\alpha(X)$. But to perform this induction we need, in fact, a derivation of $\exists y(A_f(X,y))^{\mathbb{L}_\alpha(X)}$ in $\mathsf{RS}_l(X)$ without cuts. To obtain such a derivation, we will need to eliminate cuts.
So, after defining the system $\mathsf{RS}_l(X)$, the main goal of this section is to show that, for certain derivations (that are conveniently the ones we need) cuts can be eliminated.

4.1 The terms and formulas of $\mathsf{RS}_l(X)$

We define and study the system $\mathsf{RS}_l(X)$ for a fixed set X, following Reading Convention 3.28. This means that θ and the Ω_n for $n \leq \omega$ ordinals are also fixed. The set of strings $T(\theta)$ defined in the previous section is fixed too, as it depends on X. From now on, we assume that all the ordinals we use are represented by a string in the set $T(\theta)$. Actually, we won't make a difference between ordinals and representations and we will simply talk about ordinals. We are going to define $\mathcal{T}$, our set of $\mathsf{RS}_l(X)$-terms. Each term t will have a level $|t|$. Below, by $\mathrm{rank}(u)$ we will mean the set-theoretic rank of u.

Definition 4.1. *The set $\mathcal{T}$ of $\mathsf{RS}_l(X)$-terms is defined as follows.*

- $\overline{u} \in \mathcal{T}$ *for every* $u \in TC(\{X\})$ *and* $|\overline{u}| = \Gamma_{\mathrm{rank}(u)}$. *Those are called basic terms.*

- $\mathbb{L}_\alpha(X) \in \mathcal{T}$ *for every* $\alpha \leq \Omega_\omega$ *and* $|\mathbb{L}_\alpha(X)| = \Gamma_{\theta+1} + \alpha$.
- $[x \in \mathbb{L}_\alpha(X) : B(x, s_1, \dots, s_n)^{\mathbb{L}_\alpha(X)}] \in \mathcal{T}$ *for every* $\alpha < \Omega_\omega$, *for every* KPI-*formula* $B(x, y_1, \dots, y_n)$ *and every* $s_1, \dots, s_n \in \mathcal{T}$ *with* $|s_1|, \dots, |s_n| < \Gamma_{\theta+1} + \alpha$. *Moreover,* $|[x \in \mathbb{L}_\alpha(X) : B(x, s_1, \dots, s_n)^{\mathbb{L}_\alpha(X)}]| = \Gamma_{\theta+1} + \alpha$.

Usually, we will just write $[x \in \mathbb{L}_\alpha(X) : B(x)]$ for terms of the third kind. We notice that the level of $\mathbb{L}_{\Omega_n}(X)$ is Ω_n for every $n < \omega$.
Now, we define the $\mathsf{RS}_l(X)$-formulas, together with their *type*. The type of a formula is strongly related to the rules of inference that we will define later: a formula will have $\bigvee$-type whenever we can derive it from a single premise, and a formula will have $\bigwedge$-type whenever we need all the formulas from a given set of premises to derive the formula.

Definition 4.2. *The* $\mathsf{RS}_l(X)$*-formulas are exactly the* KPI*-formulas replacing free variables by* $\mathsf{RS}_l(X)$*-terms and restricting all unbounded quantifiers to* $\mathsf{RS}_l(X)$*-terms. The* $\mathsf{RS}_l(X)$*-formulas of the form* $\overline{u} \in \overline{v}$ *or* $\overline{u} \notin \overline{v}$ *are called basic.*
Moreover, each non-basic $\mathsf{RS}_l(X)$*-formula of the form* $s \in t$, $A \vee B$, $Ad(t)$ *and* $\exists x \in t\ G(x)$ *has* $\bigvee$*-type and each non-basic* $\mathsf{RS}_l(X)$*-formula of the form* $s \notin t$, $A \wedge B$, $\neg Ad(t)$, $\forall x \in t\ G(x)$ *has* $\bigwedge$*-type.*

We observe that, by definition, there are no free variables in the $\mathsf{RS}_l(X)$-formulas. From now on, we will call $\mathsf{RS}_l(X)$-terms and $\mathsf{RS}_l(X)$-formulas simply terms and formulas.

We will say that a formula $A(s_1, \dots, s_n)^{\mathbb{L}_{\Omega_n}(X)}$ is Σ^{Ω_n} iff $A(x_1, \dots, x_n)$ is a KPI Σ-formula and $|s_1|, \dots, |s_n| < \Omega_n$.

We want to keep track of the terms that appear in the formulas. This motivates the following definition.

Definition 4.3. *For a formula* A*, we define* $k(A) = \{|t| : t \text{ occurs in } A \text{ including subterms}\}$. *For a finite set of formulas* Γ*, we define* $k(\Gamma) = \bigcup_{A \in \Gamma} k(A)$.

For example, we have $k([x \in \mathbb{L}_{\Omega_n}(X) : x \in \mathbb{L}_{\Omega_m}(X)] \in \mathbb{L}_{\Omega_m}(X)) = \{\Omega_n, \Omega_m\}$.
We will use the following abbreviations. Whereas Items 2., 3. and 4. from below are the standard standard abbreviations already introduced in Section 2, we will use the symbol $\dot{\in}$ defined in Item 1. to ease and simplify definitions and proofs. Some examples will be provided when we use this abbreviation.

Definition 4.4.

1. *Let s and t be terms such that $|s| < |t|$. For $\circ \in \{\wedge, \rightarrow\}$, we define*

$$s \mathrel{\dot{\in}} t \circ A(s,t) = \begin{cases} \overline{u} \in \overline{v} \circ A(\overline{u}, \overline{v}) & \textit{if } s \in t \equiv \overline{u} \in \overline{v}, \\ A(s,t) & \textit{if } t = \mathbb{L}_\alpha(X), \\ B(s) \circ A(s,t) & \textit{if } t = [x \in \mathbb{L}_\alpha(X) : B(x)]. \end{cases}$$

2. *$s = t$ will stand for $\forall x \in s(x \in t) \wedge \forall x \in t(x \in s)$.*

3. *$\neg A$ is obtained from A by replacing $\in$ by $\notin$ and vice-versa, $\vee$ by $\wedge$ and vice-versa, $\forall$ by $\exists$ and vice-versa and $Ad(\cdot)$ by $\neg Ad(\cdot)$ and vice-versa.*

4. *$A \rightarrow B$ will stand for $\neg A \vee B$.*

We now define two objects that will characterize formulas to build the logic $\mathsf{RS}_l(X)$. First, we define for each non-basic formula A the set $\mathcal{C}(A)$ of its characteristic subsentences. It contains all the premises that allow the derivation of A.

Definition 4.5. *We define $\mathcal{C}(A)$ for a non-basic formula A of $\bigvee$-type.*

$\mathcal{C}(r \in t) = \{s \mathrel{\dot{\in}} t \wedge r = s : |s| < |t|\}$;
$\mathcal{C}(A \vee B) = \{A, B\}$;
$\mathcal{C}(Ad(t)) = \{t = \mathbb{L}_{\Omega_n}(X) : n < \omega \wedge \Omega_n \leq |t|\}$;
$\mathcal{C}(\exists x \in t\ A(x)) = \{s \mathrel{\dot{\in}} t \wedge A(s) : |s| < |t|\}$.

Now, given a non-basic formula A of $\bigwedge$-type, we define $\mathcal{C}(A) = \{\neg B : B \in \mathcal{C}(\neg A)\}$.

In some of the above definitions the symbol $\dot{\in}$ introduced in Definition 4.4 appears. Actually, we have three different definitions of $\mathcal{C}(A)$ in those cases, depending on the form of some term that appears in A.

Example. The set $\mathcal{C}(\exists x \in t\ A(x))$ is different depending on the form of t.

- If $t \equiv \overline{u}$ then $\mathcal{C}(\exists x \in \overline{u}\ A(x)) = \{s \in \overline{u} \wedge A(s) : |s| < \Gamma_{\mathrm{rank}(u)}\}$,
- If $t \equiv \mathbb{L}_\alpha(X)$ for some $\alpha \leq \Omega_\omega$ then $\mathcal{C}(\exists x \in \mathbb{L}_\alpha(X)\ A(s)) = \{A(s) : |s| < \Gamma_{\theta+1} + \alpha\}$,
- If $t \equiv [x \in \mathbb{L}_\alpha(X) : B(x)]$ for some $\alpha < \Omega_\omega$ then $\mathcal{C}(\exists x \in t\ A(s)) = \{B(s) \wedge A(s) : |s| < \Gamma_{\theta+1} + \alpha\}$.

This makes sense: we give an intuitive explanation. We focus on the second item of the example. We suppose that a term s intends to mean some set Y and s has level less than the level of the term $\mathbb{L}_\alpha(X)$. Then, the set Y belongs to the set $L_\alpha(X)$, because it is below in the constructible hierarchy. Thus, to derive that there is some x in $\mathbb{L}_\alpha(X)$ that satisfies some formula we only need to know that some term s of level less than $\Gamma_{\theta+1}+\alpha$ satisfies that formula (because we already know that, granted that $|s| < \Gamma_{\theta+1} + \alpha$, we have that s "belongs" to $\mathbb{L}_\alpha(X)$).
A similar comment can be made for the third case: if $|s| < \Gamma_{\theta+1} + \alpha$ then s already "belongs" to $\mathbb{L}_\alpha(X)$, and so in order to derive $\exists x \in [x \in \mathbb{L}_\alpha(X) : B(x)]\ A(x)$ it suffices that some s with $|s| < \Gamma_{\theta+1}+\alpha$ satisfies B (and so s "belongs" to $[x \in \mathbb{L}_\alpha(X) : B(x)]$) and A.

Now, we define the rank of a term or formula by recursion. This notion will measure the complexity of the formulas that are active in an application of the (Cut) rule.

Definition 4.6. *We define the rank of a term or formula by recursion.*

- $\mathrm{rk}(\overline{u}) = \Gamma_{\mathrm{rank}(u)}$,
- $\mathrm{rk}(\mathbb{L}_\alpha(X)) = \Gamma_{\theta+1} + \omega \cdot \alpha$,
- $\mathrm{rk}([x \in \mathbb{L}_\alpha(X) : B(x)]) = \max(\Gamma_{\theta+1} + \omega \cdot \alpha + 1, \mathrm{rk}(B(\overline{\emptyset})) + 2)$,

- $\mathrm{rk}(s \in t) = \mathrm{rk}(s \notin t) = \max(\mathrm{rk}(s) + 6, \mathrm{rk}(t) + 1)$,
- $\mathrm{rk}(Ad(t)) = \mathrm{rk}(\neg Ad(t)) = \mathrm{rk}(t) + 5$,
- $\mathrm{rk}(A \vee B) = \mathrm{rk}(A \wedge B) = \max(\mathrm{rk}(A), \mathrm{rk}(B)) + 1$,
- $\mathrm{rk}(\exists x \in t\ A(x)) = \mathrm{rk}(\forall x \in t\ A(x)) = \max(\mathrm{rk}(t), \mathrm{rk}(A(\overline{\emptyset})) + 2)$.

Remark. Since $s = t$ is an abbreviation for

$$\forall x \in s(x \in t) \wedge \forall x \in t(x \in s),$$

we have that

$$\begin{aligned}\mathrm{rk}(s = t) &= \max(\mathrm{rk}(\forall x \in s(x \in t)), \mathrm{rk}(\forall x \in t(x \in s))) + 1\\ &= \max(\max(\mathrm{rk}(s), \mathrm{rk}(t) + 3), \max(\mathrm{rk}(t), \mathrm{rk}(s) + 3)) + 1\\ &= \max(\mathrm{rk}(s), \mathrm{rk}(t)) + 4.\end{aligned}$$

Since we have defined many complexity measures with similar names, we are going to summarize all the measures we are using in the following table.

Measure	**Introduced in**	**Explanation**
$\mathrm{rank}(u)$	Page 14	The set-theoretic rank of the set u.
$\lvert t\rvert$	Definition 4.1	The level of the $\mathsf{RS}_l(X)$-term t. For basic terms, the definition of the level of the term $\overline{u}$ uses the rank of the set u.
$k(A)$	Definition 4.3	The set that contains all the levels of the $\mathsf{RS}_l(X)$-terms appearing in the $\mathsf{RS}_l(X)$-formula A.
$k(\Gamma)$	Definition 4.3	The set that contains all the levels of the $\mathsf{RS}_l(X)$-terms appearing in the $\mathsf{RS}_l(X)$-formulas of Γ.
$\mathrm{rk}(t)$	Definition 4.6	The rank of the $\mathsf{RS}_l(X)$-term t. It measures the complexity of the term t and can be written using the level $\lvert t\rvert$ by Lemma 4.7.
$\mathrm{rk}(A)$	Definition 4.6	The rank of the $\mathsf{RS}_l(X)$-formula A. It measures the complexity of the $\mathsf{RS}_l(X)$-formula A. We will use $\mathrm{rk}(A)$ to measure the complexity of the cuts performed in the derivations in $\mathsf{RS}_l(X)$. By Lemma 4.7 we can write $\mathrm{rk}(A)$ in terms of $k(A)$.

We now state and show some properties about the rank of a formula. We will prove in Lemma 4.8 that the complexity of the characteristic formulas (the possible/needed premises) of a formula is always below the complexity of that formula. First, we show some technical results.

Lemma 4.7.

1. *Let t be any term. Then there is $n < \omega$ such that* $\mathrm{rk}(t) = \omega \cdot \lvert t\rvert + n$.
2. *Let A be any formula. Then there is $n < \omega$ such that* $\mathrm{rk}(A) = \omega \cdot \max(k(A)) + n$.
3. *Let A be any formula and s be any term. If* $\lvert s\rvert < \max(k(A(s)))$ *then* $\mathrm{rk}(A(s)) = \mathrm{rk}(A(\overline{\emptyset}))$.

Proof. 1. We consider cases based on the form of t.

Case 1. We suppose $t \equiv \overline{u}$. Then $\mathrm{rk}(\overline{u}) = \Gamma_{\mathrm{rank}(u)} = \lvert\overline{u}\rvert$, and $\Gamma_{\mathrm{rank}(u)} = \omega \cdot \Gamma_{\mathrm{rank}(u)}$, so taking $n = 0$ we obtain the result.

Case 2. We suppose $t \equiv \mathbb{L}_\alpha(X)$. Then

$$\mathrm{rk}(\mathbb{L}_\alpha(X)) = \Gamma_{\theta+1} + \omega \cdot \alpha = \omega \cdot (\Gamma_{\theta+1} + \alpha) = \omega \cdot |\mathbb{L}_\alpha(X)|.$$

Again, we take $n = 0$ to complete this case.

Case 3. We suppose $t \equiv [x \in \mathbb{L}_\alpha(X) : B(x)]$.
We assume first that $\Gamma_{\theta+1} + \omega \cdot \alpha + 1 \geq \mathrm{rk}(B(\overline{\emptyset})) + 2$. This means that

$$\mathrm{rk}(t) = \Gamma_{\theta+1} + \omega \cdot \alpha + 1 = \omega \cdot |t| + 1.$$

We assume now that $\Gamma_{\theta+1} + \omega \cdot \alpha + 1 < \mathrm{rk}(B(\overline{\emptyset})) + 2$. It follows that $\mathrm{rk}(t) = \mathrm{rk}(B(\overline{\emptyset})) + 2$. But observe that, following Definition 4.1, the formula $B(\overline{\emptyset})$ has all its terms of level less than or equal to $\Gamma_{\theta+1} + \alpha$. But $B(\overline{\emptyset})$ must have at least a term of level $\Gamma_{\theta+1} + \alpha$, so that $\Gamma_{\theta+1} + \omega \cdot \alpha + 1 < \mathrm{rk}(B(\overline{\emptyset})) + 2$. This means that $\mathrm{rk}(B(\overline{\emptyset})) = \Gamma_{\theta+1} + \omega \cdot \alpha + m$ for some $m < \omega$. This m comes from the form of B, following Definition 4.6. Hence,

$$\mathrm{rk}(t) = \mathrm{rk}(B(\overline{\emptyset})) = \omega \cdot (\Gamma_{\theta+1} + \alpha) + m + 2 = \omega \cdot |t| + m + 2$$

for some $m < \omega$.

2. We consider cases based on the form of A.

Case 1. We suppose $A \equiv s \in t$ or $A \equiv s \notin t$. Then

$$\mathrm{rk}(A) = \max(\mathrm{rk}(s) + 6, \mathrm{rk}(t) + 1) = \max(\omega \cdot |s| + n_s + 6, \omega \cdot |t| + n_t + 1)$$

for some $n_s, n_t < \omega$.
If $|s| < |t|$ then $\mathrm{rk}(A) = \omega \cdot |t| + n_t + 1 = \omega \cdot \max(k(A)) + n$ with $n = n_t + 1$.
If $|t| < |s|$ then $\mathrm{rk}(A) = \omega \cdot |s| + n_s + 6 = \omega \cdot \max(k(A)) + n$ with $n = n_s + 1$.
If $|s| = |t|$ then $r(A) = \omega \cdot \max(k(A)) + n$ with $n = \max(n_s + 6, n_t + 1)$.

Case 2. We suppose $A \equiv Ad(t)$ or $A \equiv \neg Ad(t)$. Then $\mathrm{rk}(A) = \mathrm{rk}(t) + 5 = \omega \cdot |t| + 5 = \omega \cdot \max(k(A)) + 5$, as $k(A) = \{|t|\}$.

Case 3. We suppose $A \equiv A_0 \vee A_1$ or $A \equiv A_0 \wedge A_1$. Then by the induction hypothesis

$\mathrm{rk}(A_0) = \omega \cdot \max(k(A_0)) + n_0$ for some $n_0 < \omega$ and $\mathrm{rk}(A_1) = \omega \cdot \max(k(A_1)) + n_1$ for some n_1. Therefore,

$$\mathrm{rk}(A) = \max(\omega \cdot \max(k(A_0)) + n_0, \omega \cdot \max(k(A_1)) + n_1) + 1 = \omega \cdot \max(k(A)) + n$$

where $n = n_0 + 1$ if $\max(k(A_0)) > \max(k(A_1))$, or $n = n_1 + 1$ if $\max(k(A_0)) < \max(k(A_1))$, or $n = \max(n_0, n_1) + 1$ if $\max(k(A_0)) = \max(k(A_1))$.

Case 4. We suppose $A \equiv \exists x \in t\ B(x)$ or $A \equiv \forall x \in t\ B(x)$. By the induction hypothesis, we have $\mathrm{rk}(B(\overline{\emptyset})) = \omega \cdot \max(k(B)) + n_0$ for some $n_0 < \omega$ and $\mathrm{rk}(t) = \omega \cdot |t| + n_1$ for some $n_1 < \omega$.
If $\mathrm{rk}(t) > \mathrm{rk}(B(\overline{\emptyset})) + 2$, then $|t| > \max(k(B))$ and so

$$\mathrm{rk}(A) = \mathrm{rk}(t) = \omega \cdot |t| + n_1 = \omega \cdot \max(k(A)) + n_1$$

If $\mathrm{rk}(t) < \mathrm{rk}(B(\overline{\emptyset}))+2$, then either $|t| = \max(k(B)) = \max(k(A))$ or $|t| < \max(k(B)) = \max(k(A))$. In both cases, we get

$$\mathrm{rk}(A) = \omega \cdot \max(k(B)) + n = \omega \cdot \max(k(A)) + n$$

for some $n < \omega$.

3. This proof goes by induction using Items 1. and 2. of this lemma. We refer the interested reader to [5] (Lemma 3.12. in page 17). □

Lemma 4.8. *Let A be any non-basic formula and let $B \in \mathcal{C}(A)$. Then, we have*

$$\mathrm{rk}(B) < \mathrm{rk}(A).$$

Proof. We prove this result by induction on the construction of A, supposing that A has $\bigvee$ type.

Case 1. We suppose $A \equiv s \in t$. We split the cases depending on the form of t.
Subcase 1.1. We assume $t \equiv \overline{u}$. Then s is not basic since A is not basic by assumption. Therefore, $\mathrm{rk}(A) = \mathrm{rk}(s)+6$ and if $B \in \mathcal{C}(A)$ then B is of the form $r \in \overline{u} \wedge r = s$ for some r with $|r| < |\overline{u}|$. It follows from this last condition that r is basic, say $r \equiv \overline{v}$, and then

$$\begin{aligned}\mathrm{rk}(B) &= \max(\mathrm{rk}(\overline{v} \in \overline{u}), \mathrm{rk}(\overline{v} = s) + 1\\ &= \mathrm{rk}(\overline{v} = s) + 1\\ &= \mathrm{rk}(s) + 5\\ &< \mathrm{rk}(s) + 6 = \mathrm{rk}(A).\end{aligned}$$

Subcase 1.2. We assume $t \equiv \mathbb{L}_\alpha(X)$. Then $\mathrm{rk}(A) = \max(\mathrm{rk}(s) + 6, \mathrm{rk}(t) + 1)$. If $B \in \mathcal{C}(A)$, then B is of the form $r = s$ for some r with $|r| < |t|$. Therefore,

$$\mathrm{rk}(B) = \max(\mathrm{rk}(r), \mathrm{rk}(s)) + 4 < \max(\mathrm{rk}(t) + 1, \mathrm{rk}(s) + 6) = \mathrm{rk}(A).$$

Subcase 1.3. We assume $t \equiv [x \in \mathbb{L}_\alpha(X) : F(x)]$. Then $\mathrm{rk}(A) = \max(\mathrm{rk}(s) + 6, \mathrm{rk}(t) + 1)$. If $B \in \mathcal{C}(A)$, then B is of the form $F(r) \wedge s = r$ for some r with $|r| < |t|$. Therefore,

$$\begin{aligned}\mathrm{rk}(B) &= \max(\mathrm{rk}(F(r)), \mathrm{rk}(s = r)) + 1 \\ &= \max(\mathrm{rk}(F(r)), \mathrm{rk}(s) + 4, \mathrm{rk}(r) + 4) + 1.\end{aligned}$$

But $\mathrm{rk}(s) + 5, \mathrm{rk}(r) + 5 < \max(\mathrm{rk}(s) + 6, \mathrm{rk}(t) + 1) = \mathrm{rk}(A)$. Also, $\mathrm{rk}(F(r)) < \mathrm{rk}(t)$. Indeed, if $\max(k(F(r))) \leq |s|$, then $\mathrm{rk}(F(r)) + 1 < \omega \cdot |s| + \omega \leq \mathrm{rk}(t)$ by Lemma 4.7(2.); if $|s| < \max(k(F(r)))$, then by Lemma 4.7(3.) we have

$$\mathrm{rk}(F(r)) + 1 = \mathrm{rk}(F(\overline{\emptyset})) + 1 < \max(\Gamma_{\theta+1} + \omega \cdot \alpha + 1, \mathrm{rk}(F(\overline{\emptyset})) + 2) = \mathrm{rk}(t).$$

Hence, from $\mathrm{rk}(F(r)) < \mathrm{rk}(t)$ we get $\mathrm{rk}(F(r)) + 1 < \mathrm{rk}(A)$. Gathering everything, we obtain

$$\mathrm{rk}(B) = \max(\mathrm{rk}(F(r)), \mathrm{rk}(s) + 4, \mathrm{rk}(r) + 4) + 1 < \mathrm{rk}(A).$$

Case 2. We suppose $A \equiv B \vee C$. Then $\mathrm{rk}(A) = \max(\mathrm{rk}(B), \mathrm{rk}(C)) + 1$ and so $\mathrm{rk}(B), \mathrm{rk}(C) < \mathrm{rk}(A)$.

Case 3. We suppose $A \equiv \exists x \in t\ F(x)$. We split subcases based on the form of t.

Subcase 3.1. We assume $t \equiv \overline{u}$. Then $\mathrm{rk}(A) = \max(\mathrm{rk}(\overline{u}) + 3, \mathrm{rk}(F(\overline{\emptyset})) + 2)$. We have $B \equiv \overline{v} \in \overline{u} \wedge F(\overline{v})$ for some basic term with $|\overline{v}| < |\overline{v}|$. This means that

$$\mathrm{rk}(B) = \max(\mathrm{rk}(\overline{u}) + 2, \mathrm{rk}(F(\overline{v})) + 1).$$

If $\mathrm{rk}(\overline{u}) + 2 \geq \mathrm{rk}(F(\overline{v})) + 1$ then

$$\begin{aligned}\mathrm{rk}(A) &= \max(\mathrm{rk}(\overline{u}) + 3, \mathrm{rk}(F(\overline{v})) + 2) \\ &= \mathrm{rk}(\overline{u}) + 3 \\ &< \mathrm{rk}(\overline{u}) + 2 \\ &= \mathrm{rk}(B).\end{aligned}$$

We assume now that $\mathrm{rk}(\overline{u}) + 2 < \mathrm{rk}(F(\overline{v})) + 1$. If $|\overline{v}| \geq \max(k(F(\overline{v})))$, then, using Lemma 4.7 (1.), we get $\mathrm{rk}(B) = \mathrm{rk}(F(\overline{v})) + 1 < \mathrm{rk}(\overline{u}) < \mathrm{rk}(A)$. If $|\overline{v}| < \max(k(F(\overline{v}),$ then, using Lemma 4.7(3.), we get $\mathrm{rk}(B) = \mathrm{rk}(F(\overline{v})) + 1 = \mathrm{rk}(F(\overline{\emptyset})) + 1 < \mathrm{rk}(F(\overline{\emptyset})) + 2 = \mathrm{rk}(A)$.

Subcase 3.2. We suppose $t \equiv \mathbb{L}_\alpha(X)$ for some ordinal α. This means that $\mathrm{rk}(A) = \max(\mathrm{rk}(t), \mathrm{rk}(F(\overline{\emptyset})) + 2)$. Moreover, we have $B = F(s)$ for some term s with $|s| < |t|$, and so $\mathrm{rk}(B) = \mathrm{rk}(F(s))$. If $|s| \geq \max(k(F(s)))$, then $\mathrm{rk}(B) = \mathrm{rk}(F(s)) < \mathrm{rk}(t) \leq \mathrm{rk}(A)$ by Lemma 4.7(1. and 2.). If $|s| < \max(k(F(s)))$, then $\mathrm{rk}(B) = \mathrm{rk}(F(s)) = \mathrm{rk}(F(\overline{\emptyset})) < \mathrm{rk}(A)$ by Lemma 4.7(3.).

Subcase 3.3. We suppose $t \equiv [y \in \mathbb{L}_\alpha(X) : C(y)]$. We have $\mathrm{rk}(A) = \max(\mathrm{rk}(t), \mathrm{rk}(F(\overline{\emptyset})) + 2)$. Moreover, we have $B = C(s) \wedge F(s)$ for some term s with $|s| < |t|$, and so

$$\mathrm{rk}(B) = \max(\mathrm{rk}(C(s)), \mathrm{rk}(F(s))) + 1.$$

Subcase 3.3.1. We first assume that $|s| < \max(k(F(s)))$. In this case $\mathrm{rk}(F(s)) + 1 = \mathrm{rk}(F(\overline{\emptyset})) + 2 < \mathrm{rk}(A)$. Moreover, we also have $\mathrm{rk}(C(s)) < \mathrm{rk}(A)$: if $\max(k(C(s))) < |t|$, then $\mathrm{rk}(B) < \mathrm{rk}(C(s)) + 1 < \mathrm{rk}(t) \leq \mathrm{rk}(A)$ by Lemma 4.7(1. and 2.); if $\max(k(C(s))) \geq |t|$, then by Lemma 4.7(3.), we get $\mathrm{rk}(C(s)) + 1 = \mathrm{rk}(C(\overline{\emptyset})) + 1 < \mathrm{rk}(A)$.
This means that $\mathrm{rk}(B) < \mathrm{rk}(A)$.

Subcase 3.3.2. Now, we assume $|s| \geq \max(k(F(s)))$. In this case, we have $\mathrm{rk}(F(s)) < \mathrm{rk}(t)$ by Lemma 4.7(1. and 2.) and so $\mathrm{rk}(F(s)) + 1 < \mathrm{rk}(A)$. Moreover, by the same reasoning than in Subcase 3.3.1, we get $\mathrm{rk}(C(s)) + 1 < \mathrm{rk}(A)$.
Hence, $\mathrm{rk}(B) < \mathrm{rk}(A)$.

An analogous proof shows that the result holds if A has $\bigwedge$-type. □

4.2 Operator-controlled derivations

Now, we define the derivations of the $\mathsf{RS}_l(X)$-proof system. Derivations will be controlled by operators, that are some kind of functions between sets of ordinals. We need operators to control the depth of the proofs as well as the formulas that are being derived at each step, as we will explain later. Moreover, the key use of operators is that, when we eliminate cuts, we somehow transfer the complexity that was patent in the cut formulas (that is, their rank) to the operator, under the form of adding more ordinals to the set controlling the derivations. We now define the notion of operator. This notion depends on the fixed set X. Recall that, following Reading Convention 3.28, we have a set X fixed together with the cardinals Ω_n for $n \leq \omega$.

Definition 4.9. *Consider the class $\mathcal{P}(ON) = \{Y : Y \text{ is a set of ordinals}\}$. An operator is a function $\mathcal{H} : \mathcal{P}(ON) \to \mathcal{P}(ON)$ such that for every $Y, Y' \in \mathcal{P}(ON)$ the following conditions are satisfied.*

1. $\{0\} \cup \{\Gamma_\beta : \beta \leq \theta + 1\} \cup \{\Omega_i : i \leq \omega\} \subseteq \mathcal{H}(Y)$.
2. *Let* $\alpha =_{NF} \alpha_1 + \dots + \alpha_n$*. Then,* $\alpha \in \mathcal{H}(Y)$ *iff* $\alpha_1, \dots, \alpha_n \in \mathcal{H}(Y)$.
3. *Let* $\alpha =_{NF} \varphi\alpha_1\alpha_2$*. Then,* $\alpha \in \mathcal{H}(Y)$ *iff* $\alpha_1, \alpha_2 \in \mathcal{H}(Y)$.
4. $Y \subseteq \mathcal{H}(Y)$.
5. *If* $Y \subseteq \mathcal{H}(Y')$ *then* $\mathcal{H}(Y) \subseteq \mathcal{H}(Y')$.

Moreover, we will use the following abbreviations.

- $\mathcal{H}$ *will often denote* $\mathcal{H}(\emptyset)$.
- *For a term* t*,* $\mathcal{H}[t](Y)$ *will mean* $\mathcal{H}(Y \cup \{|t|\})$.
- *For a formula* A*,* $\mathcal{H}[A](Y)$ *will mean* $\mathcal{H}(Y \cup k(A))$.
- *For a finite set of formulas* Γ*,* $\mathcal{H}[\Gamma](Y)$ *will mean* $\mathcal{H}(Y \cup k(\Gamma))$.

We notice that we add extra conditions in comparison with the common definition of operator because, in this thesis, we only care about some specific operators (e.g. we need operators to contain the Ω_n ordinals - this also means that the definition of operator is different for each set X). Operators are functions but we can treat them as sets of ordinals: if we write $\mathcal{H}(Y)$ for some operator $\mathcal{H}$ and some set of ordinals Y, we are considering the image of Y by $\mathcal{H}$, which is a set of ordinals. As well, when we write $\mathcal{H}$, we are considering the set of ordinals $\mathcal{H}(\emptyset)$.
As we said earlier, operators will control the depth of the proof and the formulas derived. This will be done in the following sense. There will be an ordinal associated to each derivation. Each application of a rule will increase this ordinal. To make a derivation controlled by an operator $\mathcal{H}$ of a set of formulas with some ordinal α associated to the derivation, the ordinal α must be in $\mathcal{H}$. As well, the operator $\mathcal{H}$ must control the formulas that are derived by containing the levels of the terms appearing in the derived formulas.

Definition 4.10. *Let $\mathcal{H}$ be an operator and let Γ be a set of formulas. We have that Γ is derived by an $\mathcal{H}$-controlled derivation with ordinal α whenever $\{\alpha\} \cup k(\Gamma) \subseteq \mathcal{H}$ and one of the following axioms or rules can be applied.*

Axioms:

$$\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \overline{u} \in \overline{v} \text{ for any } u, v \in TC(\{X\}) \text{ such that } u \in v,$$
$$\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \overline{u} \notin \overline{v} \text{ for any } u, v \in TC(\{X\}) \text{ such that } u \notin v.$$

Rules:

$$(\wedge)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, A \wedge B, A \qquad \mathcal{H} \overset{\alpha_1}{\vdash} \Gamma, A \wedge B, B}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, A \wedge B} \qquad \alpha_0, \alpha_1 < \alpha$$

$$(\vee)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, A \vee B, A}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, A \vee B} \qquad \alpha_0 < \alpha$$

$$(\vee)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, A \vee B, B}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, A \vee B} \qquad \alpha_0 < \alpha$$

$$(\in)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, r \in t, s \mathrel{\dot{\in}} t \wedge r = s}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, r \in t} \qquad \begin{array}{l} \alpha_0 < \alpha, \\ |s| < |t|, \\ |s| < \Gamma_{\theta+1} + \alpha, \\ r \in t \textit{ not basic.} \end{array}$$

$$(\notin)\ \frac{\mathcal{H}[s] \overset{\alpha_s}{\vdash} \Gamma, r \notin t, s \mathrel{\dot{\in}} t \to r \neq s \textit{ for all } |s| < |t|}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, r \notin t} \qquad \begin{array}{l} \alpha_s < \alpha, \\ r \in t \textit{ not basic.} \end{array}$$

$$(b\exists)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, \exists x \in t\ B(x), s \dot{\in} t \wedge B(s)}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \exists x \in t\ B(x)} \qquad \begin{array}{l} \alpha_0 < \alpha, \\ |s| < |t|, \\ |s| < \Gamma_{\theta+1} + \alpha. \end{array}$$

$$(b\forall)\ \frac{\mathcal{H}[s] \overset{\alpha_s}{\vdash} \Gamma, \forall x \in t\ B(x), s \dot{\in} t \to B(s) \textit{ for all } |s| < |t|}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \forall x \in t\ B(x)} \qquad \alpha_s < \alpha$$

$$(Ad)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, Ad(t), t = \mathbb{L}_{\Omega_n}(X)}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, Ad(t)} \qquad \begin{array}{l} \alpha_0 < \alpha, \\ n \leq \omega, \\ \Omega_n < |t|. \end{array}$$

$$(\neg Ad)\ \frac{\mathcal{H} \overset{\alpha_n}{\vdash} \Gamma, \neg Ad(t), t \neq \mathbb{L}_{\Omega_n}(X) \textit{ for all } n \leq \omega}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \neg Ad(t)} \qquad \alpha_n < \alpha$$

$$(Cut)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, A \qquad \mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, \neg A}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma} \qquad \alpha_0 < \alpha$$

$$(\mathsf{Ref}_n)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, \exists z \in \mathbb{L}_{\Omega_n}(X)\ A^z, A^{\mathbb{L}_{\Omega_n}(X)}}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \exists z \in \mathbb{L}_{\Omega_n}(X)\ A^z} \qquad \begin{array}{l} \alpha_0, \Omega_n < \alpha, \\ A \textit{ is a } \Sigma \textit{ formula.} \end{array}$$

Besides (Cut), each rule supplies a new formula in the conclusion. This formula is called the principal formula of the inference. Likewise, each rule withholds a (some) formula(s) of the premise(s). Those kind of formulas are called the active formulas of the derivation. Any other formula is called passive.

We observe that the principal formula of a given rule already appears in the premise(s). Usually, rules are used to derive a formula by introducing the formula in the conclusion from the active formulas in the premises. However, we take the convention of writing the principal formula in the premises of the rule to have access to the Weakening principle, that we will show in Lemma 4.15. Roughly, weakening states that whenever we derive a finite set of formulas Γ we can also derive Γ, Γ' for another finite set of formulas Γ' (we can add formulas to the conclusion of the derivation). We cannot prove this principle without taking the convention of including the principal formula in the premises of each rule for technical proof-theoretic details.

Remind that the symbol $\dot{\in}$ is an abbreviation introduced in Definition 4.4. This means that whenever this symbol appears, we have three different rules that depend on the form of a term occuring in the principal formula.
We now give an example to illustrate this remark. We will make explicit the three different rules condensed in the $(b\exists)$ rule of Definition 4.10 thanks to the $\dot{\in}$ symbol.

Example. The rule $(b\exists)$ has three different forms.

$$(b\exists)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, \exists x \in \overline{u}\ B(x), s\ \in\ \overline{u} \wedge B(s)}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \exists x \in \overline{u}\ B(x)} \qquad \begin{array}{l} \alpha_0 < \alpha \\ |s| < \Gamma_{\mathrm{rank}(u)}, \\ |s| < \Gamma_{\theta+1} + \alpha. \end{array}$$

$$(b\exists)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, \exists x \in \mathbb{L}_\gamma(X)\ B(x), B(s)}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \exists x \in \mathbb{L}_\gamma(X)\ B(x)} \qquad \begin{array}{l} \alpha_0 < \alpha, \\ |s| < \Gamma_{\theta+1} + \gamma, \\ |s| < \Gamma_{\theta+1} + \alpha. \end{array}$$

$$(b\exists)\ \frac{\mathcal{H} \overset{\alpha_0}{\vdash} \Gamma, \exists x \in t\ B(x), F(s) \wedge B(s)}{\mathcal{H} \overset{\alpha}{\vdash} \Gamma, \exists x \in t\ B(x)} \qquad \begin{array}{l} t \equiv [x \in \mathbb{L}_\gamma(X) : F(x)], \\ \alpha_0 < \alpha, \\ |s| < |t|, \\ |s| < \Gamma_{\theta+1} + \alpha. \end{array}$$

Instead of writing those three rules, we condense them in one rule using the defined abbreviation.
Moreover, in the rules $(\in)$ and $(b\exists)$ we impose the condition $|s| < \Gamma_{\theta+1} + \alpha$ to restrict the search of witnesses to coherent terms. For example, in the above example, we made explicit the $(b\exists)$ rule when the term that bounds the quantification is $\overline{u}$. To derive $\exists x \in \overline{u}\ B(x)$, we are searching for a term s that "belongs" to $\overline{u}$ and satisfies B. Therefore, we are only interested in the terms of level below the level of $\overline{u}$, which is $\Gamma_{\mathrm{rank}(u)}$.

To show how derivations work, we exhibit a proof in $\mathsf{RS}_l(X)$. The objective of the next lemma is only to display a proof in $\mathsf{RS}_l(X)$ and will not have any further importance. To simplify the proof we omit to write the principal formula of each inference in the premises (actually, to be rigous we should write all the formulas that appear in the derivation at the leaves of the proof-tree, and the application of a rule would remove the premise). Since we want to exhibit a proof, we write it without repeating formulas to make the inferences clearer. In fact, once we have proved weakening, we will adopt Reading Convention 4.16 to omit the repeated formulas in the premises along this thesis.

Lemma 4.11. *Let $\mathcal{H}$ be any operator and $X \neq \emptyset$. Then,*

1. $\mathcal{H} \overset{\alpha}{\vdash} \overline{\emptyset} \in \mathbb{L}_0(X)$, *and*
2. $\mathcal{H} \overset{\alpha}{\vdash} \exists x \in \overline{\{\emptyset\}}(x \in \mathbb{L}_0(X))$.

Proof. 1. We observe that, since Γ_1 is the minimum of the levels of the $\mathsf{RS}_l(X)$-terms, the first line of each side of the following derivation hold vacuously.

$$(\wedge)\ \dfrac{(b\forall)\ \dfrac{\mathcal{H}[s] \overset{0}{\vdash} s \in \overline{\emptyset} \to s \in \overline{\emptyset} \text{ for all } |s| < \Gamma_1}{\mathcal{H} \overset{1}{\vdash} \forall x \in \overline{\emptyset}(x \in \overline{\emptyset})} \qquad (b\forall)\ \dfrac{\mathcal{H} \overset{0}{\vdash} s \in \overline{\emptyset} \to s \in \overline{\emptyset} \text{ for all } |s| < \Gamma_1}{\mathcal{H} \overset{1}{\vdash} \forall x \in \overline{\emptyset}(x \in \overline{\emptyset})}}{(\in)\ \dfrac{\mathcal{H} \overset{2}{\vdash} \overline{\emptyset} = \overline{\emptyset}}{\mathcal{H} \overset{3}{\vdash} \overline{\emptyset} \in \mathbb{L}_0(X)}}$$

2. From Item 1., we get the first line of the right branch of the following derivation.

$$(b\exists)\ \dfrac{(\wedge)\ \dfrac{\dfrac{\mathit{Axiom}}{\mathcal{H} \overset{3}{\vdash} \overline{\emptyset} \in \overline{\{\emptyset\}}} \qquad \mathcal{H} \overset{3}{\vdash} \overline{\emptyset} \in \mathbb{L}_0(X)}{\mathcal{H} \overset{4}{\vdash} \overline{\emptyset} \in \overline{\{\emptyset\}} \wedge \overline{\emptyset} \in \mathbb{L}_0(X)}}{\mathcal{H} \overset{5}{\vdash} \exists x \in \overline{\{\emptyset\}}(x \in \mathbb{L}_0(X))}$$

□

We will see more involved examples of $\mathsf{RS}_l(X)$ derivations in Section 5.

Now, focusing on the definition of the rules of inference, we notice that the ordinals that appear in each derivation are related to the depth of the proof in the sense that the application of a rule increases this ordinal. In the following lemmas, we will be proving results of the form "if $\mathcal{H} \overset{\alpha}{\vdash} \Gamma$ then P" for some conclusion P, and to prove some result like this we will say that we argue "by induction α". This means that we will be reasoning by induction on the length of the derivation.

We observe that if (R) is a rule with principal formula A, the active formula in the premise(s) of (R) is always some $B \in \mathcal{C}(A)$, as in Definition 4.5. Moreover, if A is not a disjunction or a conjunction, each $B \in \mathcal{C}(A)$ is of the form $F(t)$ for a so-called characteristic term t.

Definition 4.12. *If A is a formula with $\bigwedge$-type different from a conjunction, we define $t_A(B) := t$ for $B \in \mathcal{C}(A)$, where t is the characteristic term of the premise $B \equiv F(t)$ in the derivation of A.*
If A has $\bigvee$-type or is a conjunction, then we define $t_A(B) := \mathbb{L}_0(X)$ for any $B \in \mathcal{C}(A)$.

For example, given a formula $F := \forall x \in t\ A(x)$ and a term s such that $|s| < |t|$, we have $t_F(s \dot{\in} t \to A(s)) = s$. Again, the explicit version without the abbreviating symbol $\dot{\in}$ of this last set will depend on the form of t.
Moreover, we conveniently define $t_A(B)$ this way for conjunctions and $\bigvee$-type formulas because it allows us to uniformize the controlling operator in the premise of a derivation.
Indeed, we observe that in the rules deriving a formula with $\bigwedge$-type, the operator that controls the premise with active formula $B \in \mathcal{C}(A)$ is exactly $\mathcal{H}[t_A(B)]$, no matter the form of A. That is the reason why we took the convention to define $t_A(B) = \mathbb{L}_0(X)$ for any non $\bigwedge$-type formula (or any conjunction) A. In fact, for a conjunction or a $\bigvee$-type formula, the operator $\mathcal{H}[t_A(B)]$ is exactly $\mathcal{H}$ for any $B \in \mathcal{C}(A)$ and so, while talking about the premise(s) of A we can freely use the controlling operator $\mathcal{H}[t_A(B)]$ for formulas of both $\bigwedge$-type and $\bigvee$-type.

4.3 Cut-elimination for $\mathsf{RS}_l(X)$

We recall that the rank of a formula was introduced in Definition 4.6. We will write the complexity of the formulas that have been removed by an application of the (Cut) rule in some derivation as a subscript, as follows.

Definition 4.13. *We will write $\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma$ whenever $\mathcal{H} \overset{\alpha}{\vdash} \Gamma$ and all the active formulas of the premise of an inference using (Cut) occurring in the derivation have rank strictly less than ρ. In this case, we will say that the cut complexity of the derivation is bounded by ρ.*

Recall that the objective is to eliminate cuts from derivations. First, we show some immediate results about operator controlled derivations that will be useful to prove the cut-elimination lemmas.
Henceforth, we will use y as the set of indexes of the possibly many premises used to derive some conclusion. The next lemma states that an operator appearing in the conclusion of an inference is always contained in the operator(s) appearing in the premise(s).

Lemma 4.14. *Let $\mathcal{H}$ be any operator. We suppose that $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$ can be obtained by the application of an inference rule with premises $\mathcal{H}_i \left|\frac{\alpha_i}{\rho_i}\right. \Gamma_i$ with $\alpha_i < \alpha$ and $\mathcal{H}_i$ operators, for $i \in y$. Then, $\mathcal{H} \subseteq \mathcal{H}_i$.*

Proof. First, we suppose that the principal formula is $A \in \Gamma$. If A has $\bigvee$-type, then there is only one premise, with control operator $\mathcal{H}_i = \mathcal{H}$. If A has $\bigwedge$-type, then $\mathcal{H}_i = \mathcal{H}[t_A(B)] \supseteq \mathcal{H}$ for some $B \in \mathcal{C}(A)$.
Now, if $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$ has been obtained by (Cut) then $\mathcal{H}_i = \mathcal{H}$ in both premises. If the derivation has been obtain by applying ($\mathrm{Ref}_n(X)$) for some $n < \omega$ then $\mathcal{H}_i = \mathcal{H}$ in the unique premise. □

Now, we prove a very useful lemma that we will often use throughout this thesis. It states that we can weaken derivations as long as the operator controls the added formulas, we can increase the ordinal of the derivation as long as the new ordinal belongs to the operator, and we can increase the upper bound of the complexity of the cuts.

Lemma 4.15. *Let $\mathcal{H}$ be an operator. Let α, α', ρ and ρ' be ordinals. Let Δ and Γ be finite sets of formulas. If $\alpha \leq \alpha' \in \mathcal{H}$, $\rho \leq \rho'$, $k(\Delta) \subseteq \mathcal{H}$ and $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$ then $\mathcal{H} \left|\frac{\alpha'}{\rho'}\right. \Gamma, \Delta$.*

Proof. We proceed by induction on α. If Γ is an axiom, then Γ, Δ is an axiom too and, since $\{\alpha'\} \cup k(\Gamma \cup \Delta) \subseteq \mathcal{H}$ and an axiom has no cuts (which means that the complexity of the cut formulas is bounded by any ordinal), we have that $\mathcal{H} \left|\frac{\alpha'}{\rho'}\right. \Gamma, \Delta$.
We suppose that Γ has been derived by the application of a rule (R) and consider cases based on (R).

Case 1. We assume that (R) is (Cut). Then we have $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$ and the premises are

$$\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A, \text{ and} \tag{1}$$

$$\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, \neg A \tag{2}$$

with $\alpha_0 < \alpha < \alpha'$ for some formula A with $\text{rank}(A) < \rho$.
First, we assume that A or $\neg A$ are in Δ. Then, we use the induction hypothesis on (1) or (2) respectively to obtain

$$\mathcal{H} \left|\frac{\alpha'}{\rho'}\right. \Gamma, \Delta.$$

We assume now that A and $\neg A$ are not in Δ. Then, since $\alpha_0 < \alpha$, by the induction hypothesis, we have $\mathcal{H} \left|\frac{\alpha}{\rho'}\right. \Gamma, \Delta, A$ and $\mathcal{H} \left|\frac{\alpha}{\rho'}\right. \Gamma, \Delta, \neg A$. We apply (Cut), observing that the cut complexity of the derivation does not increase since $\text{rank}(A) < \rho < \rho'$, and we obtain $\mathcal{H} \left|\frac{\alpha'}{\rho'}\right. \Gamma, \Delta, \neg A$.

Case 2. Suppose that $A \in \Gamma$ is the principal formula of the last derivation obtained by an application of the rule (R) and A has $\bigvee$-type. Then, we have $\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A, B$ for some $B \in \mathcal{C}(A)$ (we write Γ, A, which is the same as Γ, to make explicit the principal formula A). If $B \in \Delta$, by the induction hypothesis, we get

$$\mathcal{H} \left|\frac{\alpha'}{\rho'}\right. \Gamma, \Delta, A.$$

We assume that $B \notin \Delta$. Then, the induction hypothesis gives

$$\mathcal{H} \left|\frac{\alpha}{\rho'}\right. \Gamma, \Delta, A, B.$$

We use the rule (R) to obtain

$$\mathcal{H} \left|\frac{\alpha'}{\rho'}\right. \Gamma, \Delta, A.$$

The cases where the principal formula of the last derivation has $\bigwedge$-type and the case where the last rule is some $(\mathsf{Ref_n})$ are analogous. □

Now that we have proved that the Weakening rule is acceptable in $\mathsf{RS}_l(X)$, we will drop the repetition of the principal formula of a rule in the premises to simplify proofs following the next reading convention.

Reading Convention 4.16. *We suppose that we have the following inference with principal formula A by applying the rule (R).*

$$(R)\ \frac{\mathcal{H} \left|\frac{\alpha_i}{}\right. \Gamma_i, A}{\mathcal{H} \left|\frac{\alpha}{}\right. \Gamma, A}$$

We will often omit the repeated formula A in the premises and, instead, write

$$(R)\ \frac{\mathcal{H} \left|\frac{\alpha_i}{}\right. \Gamma_i}{\mathcal{H} \left|\frac{\alpha}{}\right. \Gamma, A}$$

We can prove some kind of Inversion for $\bigwedge$-type formulas.

Lemma 4.17 (Inversion)**.** *Let $\mathcal{H}$ be any operator. Let A be a $\bigwedge$-type formula and let Γ be a finite set of formulas. Let α and ρ be ordinals. If $\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, A$ then $\mathcal{H}[t_A(B)] \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, B$ for every $B \in \mathcal{C}(A)$.*

Proof. We proceed by induction on α. If Γ, A is an axiom, then Γ is an axiom since A is non-basic. Therefore, Γ, B is an axiom for each $B \in \mathcal{C}(A)$ and so $\mathcal{H}[t_B(A)] \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, B$. We suppose that Γ, A has been obtained by the application of a rule with A not principal. Then, we can apply the induction hypothesis to the premises of this inference and use the rule again to obtain the result.
We suppose that Γ, A has been obtain by the application of a rule (R) with principal formula A. This means that, for each $B \in \mathcal{C}(A)$, we have

$$\mathcal{H}[t_A(B)] \overset{\alpha_B}{\underset{\rho}{\vdash}} \Gamma, A, B$$

with $\alpha_B < \alpha$. By the induction hypothesis, for every $B \in \mathcal{C}(A)$ we have

$$\mathcal{H}[t_A(B)][t_A(B)] \overset{\alpha_B}{\underset{\rho}{\vdash}} \Gamma, B, B. \tag{1}$$

But $\mathcal{H}[t_A(B)][t_A(B)] = \mathcal{H}[t_A(B)]$ since $t_A(B) \in \mathcal{H}[t_A(B)]$, and so (1) is exactly

$$\mathcal{H}[t_A(B)] \overset{\alpha_B}{\underset{\rho}{\vdash}} \Gamma, B.$$

Again by Lemma 4.15, we obtain

$$\mathcal{H}[t_A(B)] \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, B$$

for every $B \in \mathcal{C}(A)$. □

The next lemma will only be used at the end of the thesis, but we include it here since it is a technical result that is proved the same way as many others in this section. This result follows the idea that conclusions of derivation can be thought to be disjunctions.

Lemma 4.18. *Let $\mathcal{H}$ be any operator. Let $\Gamma \cup \{A, B\}$ be a finite set of formulas. If $\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, A \vee B$ then $\mathcal{H} \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, A, B$.*

Proof. We proceed by induction on α. If $\Gamma, A \vee B$ is an axiom, then Γ is an axiom ($A \vee B$ cannot be an axiom because it is not a basic formula). This means that Γ, A, B is also an axiom, and the result holds.

We suppose that $\Gamma, A \vee B$ has been obtained by the application of a rule. If the principal formula is not $A \vee B$, then we use the induction hypothesis on the premise(s) and use again the rule to obtain the result.
So we assume that the principal formula is $A \vee B$. We have $\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A \vee B, C$ where C is either A or B. By the inudction hypothesis, we get $\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A, B, C$, which is exactly $\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A, B$. By means of Lemma 4.15, we obtain $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, A, B$. □

We are now ready to prove the Predicative Cut Elimination Theorem. To prepare for the proof, we first show the Reduction Lemma. This result will only be used to simplify the proof the upcoming Predicative Cut Elimination Theorem and will not appear further in the thesis.

Lemma 4.19 (Reduction)**.** *Let $\mathcal{H}$ be any operator. Let α be an ordinal. Let Γ and Δ be finite sets of formulas. Let A be a formula with $\mathrm{rk}(A) = \rho \notin \{\Omega_n : n < \omega\}$. If A is $\overline{u} \in \overline{v}$ for some $u, v \in TC(\{X\})$ or A has $\bigvee$-type, and both $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \neg A$ and $\mathcal{H} \left|\frac{\beta}{\rho}\right. \Delta, A$ hold, then $\mathcal{H} \left|\frac{\alpha+\beta}{\rho}\right. \Gamma, \Delta$ also holds.*

Proof. We will consider cases based on the form of A. Before considering the case where A is a basic formula, we show the following claim:

Claim 4.19.1.

1. *If $u \in v$ is true in $TC(\{X\})$ and $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \neg\overline{u} \in \overline{v}$ then $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$.*
2. *If $u \notin v$ is true in $TC(\{X\})$ and $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \overline{u} \in \overline{v}$, then $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$.*

We prove Item 1. of Claim 4.19.1 by induction on α. If $\Gamma, \neg\overline{u} \in \overline{v}$ is an axiom, then Γ is an axiom since $u \notin v$ is false in $TC(\{X\})$, and so $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$ holds.
We suppose that $\Gamma, \neg\overline{u} \in \overline{v}$ was derived by an application of a rule (R). Then $\neg\overline{u} \in \overline{v}$ is a passive formula since it is basic. Therefore, we have (possibly many) premises of the form $\mathcal{H}_i \left|\frac{\alpha_i}{\rho}\right. \Gamma_i, \neg\overline{u} \in \overline{v}$ with $\alpha_i < \alpha$ for every $i \in y$. By the induction hypothesis, we obtain $\mathcal{H}_i \left|\frac{\alpha_i}{\rho}\right. \Gamma_i$, with $\alpha_i < \alpha$, for every $i \in y$. Finally, we apply the rule (R) to get $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$.
The proof of Item 2. is analogous, and so Claim 4.19.1 is shown.

We now start the proof of the Reduction Lemma.
We suppose that $A \equiv \overline{u} \in \overline{v}$. Then, by Claim 4.19.1, we have either $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$ or $\mathcal{H} \left|\frac{\beta}{\rho}\right. \Delta$ depending on whether $u \in v$ holds in $TC(\{X\})$ or not. Therefore, by Lemma 4.15 we obtain $\mathcal{H} \left|\frac{\alpha+\beta}{\rho}\right. \Gamma, \Delta$.

We suppose now that A has $\bigvee$-type. We have

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \neg A, \tag{1}$$

$$\mathcal{H} \left|\frac{\beta}{\rho}\right. \Delta, A. \tag{2}$$

We proceed by induction on β. If Δ, A is an axiom, then Δ is an axiom since A is non-basic and so Γ, Δ is also an axiom, showing that $\mathcal{H} \left|\frac{\alpha+\beta}{\rho}\right. \Gamma, \Delta$ holds.

Now, we assume that Δ, A has been obtained by the application of a rule (R). We suppose that A is not the principal formula in this last derivation. Then, we can apply the induction hypothesis to the premises and apply again the rule (R) to obtain the result.
We suppose now that A is the principal formula of the last derivation. We notice that (R) cannot be any (Ref_n) rule since, if it were, we would have $\mathrm{rk}(A) = \Omega_n$, against the hypothesis. Therefore, the rule (R) is either $(\vee)$, $(\in)$, $(b\exists)$ or (Ad), and we have the only premise

$$\mathcal{H} \left|\frac{\beta_0}{\rho}\right. \Delta, A, B, \tag{3}$$

with $\beta_0 < \beta$. We apply the induction hypothesis to (1) and (3) to obtain

$$\mathcal{H} \left|\frac{\alpha+\beta_0}{\rho}\right. \Gamma, \Delta, B. \tag{4}$$

On the other hand, from (1), we use Lemma 4.17 (Inversion) and we get

$$\mathcal{H}[t_A(B)] \left|\frac{\alpha}{\rho}\right. \Gamma, \neg B. \tag{5}$$

But $t_A(B) \in \mathcal{H}$ by (4), and so $\mathcal{H}[t_A(B)] = \mathcal{H}$. Thus, Lemma 4.15 on (5) gives

$$\mathcal{H} \left|\frac{\alpha+\beta_0}{\rho}\right. \Gamma, \Delta, \neg B. \tag{6}$$

Finally, we apply (Cut) to (4) and (6) to obtain $\mathcal{H} \left|\frac{\alpha+\beta}{\rho}\right. \Gamma, \Delta$. We note that, by Lemma 4.8, we have $\mathrm{rk}(B) < \mathrm{rk}(A) = \rho$ and so the complexity of the cuts done in the derivation is still bounded by ρ. □

At last, we state and prove the Predicative Cut Elimination Theorem. If we have a derivation $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$ with $\rho < \Omega_k$ where $k = min(n < \omega : \rho < \Omega_n)$, this results allows us to lower the bound of the complexity of the cuts up to $\Omega_{k-1} + 1$, or to 0 if $k = 0$.

Theorem 4.20 (Predicative Cut Elimination). *Let $\mathcal{H}$ be any operator closed under φ. Let $\alpha \in \mathcal{H}$ and ρ be ordinals such that $\Omega_n \notin [\rho, \rho + \omega^\alpha)$ for any $n < \omega$. We have that if $\mathcal{H} \frac{\ \beta\ }{\rho+\omega^\alpha} \Gamma$ then $\mathcal{H} \frac{\ \varphi\alpha\beta\ }{\rho} \Gamma$.*

Proof. First of all, we observe that $\varphi\alpha\beta \in \mathcal{H}$ since $\alpha, \beta \in \mathcal{H}$ and so the ordinal bound of the derivation appearing in the conclusion of the theorem is coherent. We proceed by induction on β. If Γ is an axiom then trivially $\mathcal{H} \frac{\ \varphi\alpha\beta\ }{\rho} \Gamma$.
We suppose that Γ has been obtained by the application of a rule (R). We distinguish cases according to whether (R) is the rule (Cut) or (R) is any other rule.

Case 1. We suppose that (R) is not (Cut). Then we have the premise(s) $\mathcal{H}_i \frac{\ \beta_i\ }{\rho+\omega^\alpha} \Gamma_i$ with $\beta_i < \beta$ for each $i \in y$. By the induction hypothesis, we get $\mathcal{H}_i \frac{\ \varphi\alpha\beta_i\ }{\rho} \Gamma_i$ and, since $\varphi\alpha\beta_i < \varphi\alpha\beta$ for all $i \in y$ we obtain $\mathcal{H} \frac{\ \varphi\alpha\beta\ }{\rho} \Gamma$ by an application of (R).

Case 2. We suppose now that (R) is (Cut). This means that the premises are

$$\mathcal{H} \frac{\ \beta_0\ }{\rho+\omega^\alpha} \Gamma, B \text{ with } \beta_0 < \beta \tag{1}$$

$$\mathcal{H} \frac{\ \beta_0\ }{\rho+\omega^\alpha} \Gamma, \neg B \text{ with } \beta_0 < \beta \tag{2}$$

for some formula B. By the induction hypothesis applied to (1) and (2), we have

$$\mathcal{H} \frac{\ \varphi\alpha\beta_0\ }{\rho} \Gamma, B \tag{3}$$

$$\mathcal{H} \frac{\ \varphi\alpha\beta_0\ }{\rho} \Gamma, \neg B. \tag{4}$$

We observe that $\text{rk}(B) < \rho + \omega^\alpha$ (if the rank of B was greater than $\rho + \omega^\alpha$ we could not have derived Γ with cuts of complexity bounded by $\rho + \omega^\alpha$). If $\text{rk}(B) < \rho$, then we can apply (Cut) to (3) and (4) and get $\mathcal{H} \frac{\ \varphi\alpha\beta\ }{\rho} \Gamma$ since $\varphi\alpha\beta_0 < \varphi\alpha\beta$.
If $\text{rk}(B) \in [\rho, \rho + \omega^\alpha)$ we cannot apply (Cut) efficiently because it would increase the complexity of the cuts. In this case, we write $\text{rk}(B) = \rho + \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ for some $\alpha > \alpha_1 \geq \cdots \geq \alpha_n$. By Lemma 4.15 on (3) and (4) we are able to have the cut complexity bound equal to the rank of B:

$$\mathcal{H} \frac{\ \varphi\alpha\beta_0\ }{\rho+\omega^{\alpha_1}+\cdots+\omega^{\alpha_n}} \Gamma, B \tag{5}$$

$$\mathcal{H} \frac{\ \varphi\alpha\beta_0\ }{\rho+\omega^{\alpha_1}+\cdots+\omega^{\alpha_n}} \Gamma, \neg B. \tag{6}$$

We observe that either B or $\neg B$ has $\bigvee$-type, and so we can apply Lemma 4.19 (Reduction) to (5) and (6) and obtain

$$\mathcal{H} \vdash^{\varphi\alpha\beta_0+\varphi\alpha\beta_0}_{\rho+\omega^{\alpha_1}+\cdots+\omega^{\alpha_n}} \Gamma.$$

By Lemma 4.15, since $\varphi\alpha\beta_0 + \varphi\alpha\beta_0 < \varphi\alpha\beta$ and $\varphi\alpha\beta \in \mathcal{H}$, we get

$$\mathcal{H} \vdash^{\varphi\alpha\beta}_{\rho+\omega^{\alpha_1}+\cdots+\omega^{\alpha_n}} \Gamma. \tag{7}$$

Now, since $\rho+\omega^{\alpha_1}+\cdots+\omega^{\alpha_n} = \mathrm{rk}(B) \in \mathcal{H}$ (this comes from the fact that $k(B) \subseteq \mathcal{H}$ by (3) and $\mathrm{rk}(B) = \omega \cdot \max(k(B)) + n'$ for some $n' < \omega$), we have that $\alpha_1, \ldots, \alpha_n \in \mathcal{H}$. Therefore, as $\alpha_n < \alpha$, we can apply the induction hypothesis to (7) and obtain

$$\mathcal{H} \vdash^{\varphi_{\alpha_n}(\varphi\alpha\beta)}_{\rho+\omega^{\alpha_1}+\cdots+\omega^{\alpha_{n-1}}} \Gamma. \tag{8}$$

But $\alpha_n < \alpha$ shows that $\varphi_{\alpha_n}(\varphi\alpha\beta) = \varphi\alpha\beta$ by Lemma 3.20, and so (8) is exactly

$$\mathcal{H} \vdash^{\varphi\alpha\beta}_{\rho+\omega^{\alpha_1}+\cdots+\omega^{\alpha_{n-1}}} \Gamma. \tag{9}$$

Repeating this $n-1$ times, we finally obtain

$$\mathcal{H} \vdash^{\varphi\alpha\beta}_{\rho} \Gamma. \tag{10}$$

□

We will use Theorem 4.20 in two different ways. The optimal use of this theorem will come when the bound of the complexity of the cuts is $\delta < \Omega_0$. In this case, it will be that $\delta = 0 + \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ for some $\alpha_1, \ldots, \alpha_n < \Omega_0$, and so n consecutive applications of the Predicative Cut Elimination Theorem will provide a derivation of the same formulas but with no cuts. The second case appears when the bound of the complexity of the cuts is $\Omega_n < \delta < \Omega_{n+1}$ for some $n < \omega$. Here, the Predicative Cut Elimination Theorem will eliminate cuts up to $\Omega_n + 1$. This is why we need another result that will allow us to collapse bounds below the Ω_n ordinals. This will be provided by Theorem 4.24 (Collapsing Theorem), that we will state and prove in the next subsection. Before that, we show another very useful predicative result.

Lemma 4.21 (Boundedness)**.** *Let $\mathcal{H}$ be any operator. Let n be any natural number. Let ρ be an ordinal. Let $A^{\mathbb{L}_{\Omega_n}(X)}$ be a Σ^{Ω_n} formula and let α, β be ordinals such that $\beta \in \mathcal{H}$ and $\alpha \leq \beta < \Omega_n$. If $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, A^{\mathbb{L}_{\Omega_n}(X)}$ then $\mathcal{H} \vdash^{\alpha}_{\rho} \Gamma, A^{\mathbb{L}_{\beta}(X)}$.*

Proof. We proceed by induction on α. If $\Gamma, A^{\mathbb{L}_{\Omega_n}(X)}$ is an axiom then $\Gamma, A^{\mathbb{L}_\beta(X)}$ is also an axiom (if Γ is an axiom this is clear; if $A^{\mathbb{L}_{\Omega_n}(X)}$ is an axiom then A has no quantifiers and therefore $A^{\mathbb{L}_{\Omega_n}(X)} \equiv A \equiv A^{\mathbb{L}_\beta(X)}$).
We suppose that the derivation has been obtained by the application of a rule (R). If $A^{\mathbb{L}_{\Omega_n}(X)}$ is not the principal formula, then we can apply the induction hypothesis to the premise(s) and use the rule (R) again to obtain the result. So assume that $A^{\mathbb{L}_{\Omega_n}(X)}$ is the principal formula. We distinguish 3 cases: the rule (R) is not (Ref_n) and the formula $A^{\mathbb{L}_{\Omega_n}(X)}$ has $\bigwedge$-type, the rule (R) is not (Ref_n) and the formula $A^{\mathbb{L}_{\Omega_n}(X)}$ has $\bigvee$-type, or the rule (R) is (Ref_n). We observe that the last rule can not be (Ref_m) with $m > n$ since all the terms that appear in $A^{\mathbb{L}_{\Omega_n}(X)}$ have level less than Ω_n. The case where (R) is (Ref_m) with $m < n$ falls within the Case 2, where $A^{\mathbb{L}_{\Omega_n}(X)}$ has $\bigvee$-type.

Case 1. We suppose that $A^{\mathbb{L}_{\Omega_n}(X)}$ has $\bigwedge$-type. We notice that, since $A^{\mathbb{L}_{\Omega_n}(X)}$ is $\Sigma^{\mathbb{L}_{\Omega_n}(X)}$, every formula in $\mathcal{C}(A^{\mathbb{L}_{\Omega_n}(X)})$ is also Σ^{Ω_n}. This means that formulas in $\mathcal{C}(A^{\mathbb{L}_{\Omega_n}(X)})$ are of the form $B^{\mathbb{L}_{\Omega_n}(X)}$ and $B^{\mathbb{L}_\beta(X)} \in \mathcal{C}(A^{\mathbb{L}_\beta(X)})$ for every $B^{\mathbb{L}_{\Omega_n}(X)} \in \mathcal{C}(A^{\mathbb{L}_{\Omega_n}(X)})$. So, we have

$$\mathcal{H}[t_{A^{\mathbb{L}_{\Omega_n}(X)}}(B^{\mathbb{L}_{\Omega_n}(X)})] \left|\frac{\alpha_B}{\rho}\right. \Gamma, A^{\mathbb{L}_{\Omega_n}(X)}, B^{\mathbb{L}_{\Omega_n}(X)}.$$

with $\alpha_B < \alpha$, for every $B^{\mathbb{L}_{\Omega_n}(X)} \in \mathcal{C}(A^{\mathbb{L}_{\Omega_n}(X)})$. Since $\alpha_B < \alpha$, we can use the induction hypothesis, that gives

$$\mathcal{H}[t_{A^{\mathbb{L}_{\Omega_n}(X)}}(B^{\mathbb{L}_{\Omega_n}(X)})] \left|\frac{\alpha_B}{\rho}\right. \Gamma, A^{\mathbb{L}_\beta(X)}, B^{\mathbb{L}_\beta(X)}.$$

Applying the rule (R), we obtain as desired

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, A^{\mathbb{L}_\beta(X)}.$$

Case 2. We suppose that $A^{\mathbb{L}_{\Omega_n}(X)}$ has $\bigvee$-type and is not of the form $\exists x \in \mathbb{L}_{\Omega_n}(X)\ C(x)^{\mathbb{L}_{\Omega_n}(X)}$. Then, again each formula in $\mathcal{C}(A^{\mathbb{L}_{\Omega_n}(X)})$ is $\Sigma^{\mathbb{L}_{\Omega_n}(X)}$. So, there is $B^{\mathbb{L}_{\Omega_n}(X)} \in \mathcal{C}(A^{\mathbb{L}_{\Omega_n}(X)})$ such that

$$\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A^{\mathbb{L}_{\Omega_n}(X)}, B^{\mathbb{L}_{\Omega_n}(X)},$$

with $\alpha_0 < \alpha$. By the induction hypothesis, we have

$$\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A^{\mathbb{L}_\beta(X)}, B^{\mathbb{L}_\beta(X)}.$$

We apply the rule (R) and we obtain

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, A^{\mathbb{L}_\beta(X)}.$$

We notice that we can do this even in the case that $A^{\mathbb{L}_{\Omega_n}(X)} \equiv \exists x \in \mathbb{L}_{\Omega_n}(X)\, C(x)^{\mathbb{L}_{\Omega_n}(X)}$. If this is the case, we would have

$$\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A^{\mathbb{L}_{\Omega_n}(X)}, C(t)^{\mathbb{L}_{\Omega_n}(X)},$$

for some term t with $|t| < \mathbb{L}_{\Omega_n}(X)$ and $|t| < \Gamma_{\theta+1} + \alpha$. By the previous reasoning, we get

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, A^{\mathbb{L}_\beta(X)}, C(t)^{\mathbb{L}_\beta(X)}.$$

The thing is that $\alpha < \beta$ shows that $\Gamma_{\theta+1} + \alpha < \Gamma_{\theta+1} + \beta$, and so $|t| < \Gamma_{\theta+1} + \beta$, meaning that we can apply $(b\exists)$ to get

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \exists x \in \mathbb{L}_\beta(X)\ C(x)^{\mathbb{L}_\beta(X)},$$

as desired.

Case 3. We suppose $A^{\mathbb{L}_{\Omega_n}(X)} \equiv \exists x \in \mathbb{L}_{\Omega_n}(X)\ C^x$ for some formula C and (R) is (Ref_n). This means that we have

$$\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A^{\mathbb{L}_{\Omega_n}(X)}, C^{\mathbb{L}_{\Omega_n}(X)},$$

with $\alpha_0 < \alpha$. We will use the induction hypothesis on $A^{\mathbb{L}_{\Omega_n}(X)}$ and on $C^{\mathbb{L}_{\Omega_n}(X)}$ separately: on $A^{\mathbb{L}_{\Omega_n}(X)}$ to get $A^{\mathbb{L}_\beta(X)}$ and on C to get $C^{\mathbb{L}_{\alpha_0}(X)}$. We obtain the following.

$$\mathcal{H} \left|\frac{\alpha_0}{\rho}\right. \Gamma, A^{\mathbb{L}_\beta(X)}, C^{\mathbb{L}_{\alpha_0}(X)}.$$

We obtain by an application of $(b\exists)$ the desired result:

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \exists x \in \mathbb{L}_\beta(X)\ C^x.$$

□

4.4 The Collapsing Theorem

Unfortunately, we are unable to collapse ordinal bounds of derivations below some Ω_n for every operator. We are going to define some specific operators that will allow us to collapse cuts below each Ω_n with our ψ_n functions.

Definition 4.22. *We define $\mathcal{H}_\beta$ as follows. For any set of ordinal Y we let*

$$\mathcal{H}_\beta(Y) = \bigcap \{B_n(\alpha) : Y \subseteq B_n(\alpha) \text{ with } \beta < \alpha \text{ and } n < \omega\}.$$

We notice that $\mathcal{H}_\beta = \bigcap\{B_n(\alpha) : \beta < \alpha \wedge n < \omega\} = B_0(\beta+1)$ since if $\beta < \alpha$ then $\beta + 1 \leq \alpha$ and so $B_n(\beta+1) \subseteq B_n(\alpha)$ for any $n < \omega$ and $B_0(\beta+1) \subseteq B_n(\beta+1)$ for any $n < \omega$.
It is straightforward to show that $\mathcal{H}_\beta$ is an operator closed under Veblen functions for every ordinal β.
Next, we show a result that holds for arbitrary operators but that we will use only in the proof of the Collapsing Theorem.

Lemma 4.23. *Let $\mathcal{H}$ be any operator. Let α, β, γ and ρ be ordinals. Let $\Gamma \cup \{A\}$ be a finite set of formulas. If $\beta > \gamma \in \mathcal{H}$ and $\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \forall x \in \mathbb{L}_\beta(X) A(x)$ then*

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \forall x \in \mathbb{L}_\gamma(X) A(x).$$

Proof. We proceed by induction on α. If $\Gamma, \forall x \in \mathbb{L}_\beta(X) A(x)$ is an axiom, then Γ is an axiom. Therefore, $\Gamma, \forall x \in \mathbb{L}_\gamma(X) A(x)$ is an axiom too.
We suppose that $\Gamma, \forall x \in \mathbb{L}_\beta(X) A(x)$ has been obtained by the application of a rule. If the principal formula is not $\forall x \in \mathbb{L}_\beta(X) A(x)$, then we just use the induction hypothesis on the premise(s) and apply the rule again. So assume that $\forall x \in \mathbb{L}_\beta(X) A(x)$ is the principal formula of the last derivation, and the rule applied is $(b\forall)$. The premises are $\mathcal{H}[s] \left|\frac{\alpha_s}{\rho}\right. \Gamma, \forall x \in \mathbb{L}_\gamma(X) A(x), A(s)$ with $\alpha_s < \alpha$ for all terms the terms s with $|s| < \Gamma_{\theta+1} + \beta$. By the induction hypothesis, we get

$$\mathcal{H}[s] \left|\frac{\alpha_s}{\rho}\right. \Gamma, A(s), \forall x \in \mathbb{L}_\gamma(X) A(x) \text{ for all } |s| < \Gamma_{\theta+1} + \beta. \tag{1}$$

But $A(s) \equiv s \dot{\in} \mathbb{L}_\gamma(X) \rightarrow A(s)$ for all $|s| < \Gamma_{\theta+1} + \gamma$, and so focusing on those $|s| < \Gamma_{\theta+1} + \gamma$ in (1) we have in particular

$$\mathcal{H}[s] \left|\frac{\alpha_s}{\rho}\right. \Gamma, s \dot{\in} \mathbb{L}_\gamma(X) \rightarrow A(s), \forall x \in \mathbb{L}_\gamma(X) A(x) \text{ for all } |s| < \Gamma_{\theta+1} + \gamma.$$

We apply $(b\forall)$ and get

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \forall x \in \mathbb{L}_\gamma(X)\ A(x), \forall x \in \mathbb{L}_\gamma(X)\ A(x),$$

which is the same as

$$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma, \forall x \in \mathbb{L}_\gamma(X)\ A(x).$$

□

We recall that, given some $m \leq \omega$, a Σ^{Ω_m} formula is an $\mathsf{RS}_l(X)$-formula that has been obtained from a KPl Σ-formula by restricting all the unbounded quantifiers to $\mathbb{L}_{\Omega_m}(X)$ and by replacing free variables by terms of level strictly less than Ω_m.

Theorem 4.24 (Collapsing Theorem)**.** *Let $n \leq \omega$ and let $m < \omega$. Let Γ be a set of Σ^{Ω_m}-formulas and let α and β be ordinals with $\beta \in \mathcal{H}_\beta$.*
If $\mathcal{H}_\beta \overset{\alpha}{\underset{\Omega_n+1}{\vdash}} \Gamma$ then $\mathcal{H}_{\beta+\omega^{\Omega_n+1+\alpha}} \overset{\psi_m(\beta+\omega^{\Omega_n+1+\alpha})}{\underset{\psi_m(\beta+\omega^{\Omega_n+1+\alpha})}{\vdash}} \Gamma$.

Proof. To simplify notation we define for every ordinal α

$$\hat{\alpha} = \beta + \omega^{\Omega_n+1+\alpha}.$$

We are going to prove a more general claim to deal with cases where some terms might be added to the operator:

Claim 4.24.1. *Let Γ, α, β, n and m as in the assumption of the theorem. Let Δ be any finite set of formulas such that $k(\Delta) \subseteq B_m(\beta+1)$.*

$$\textit{If } \mathcal{H}_\beta[\Delta] \overset{\alpha}{\underset{\Omega_n+1}{\vdash}} \Gamma \textit{ then } \mathcal{H}_{\hat{\alpha}}[\Delta] \overset{\psi_m\hat{\alpha}}{\underset{\psi_m\hat{\alpha}}{\vdash}} \Gamma.$$

With Claim 4.24.1, taking $\Delta = \emptyset$ we obtain the theorem.
First, we observe that, from $0, 1, \alpha, \Omega_n, \beta \in \mathcal{H}_\beta[\Delta] = \bigcap\{B_k(\gamma) : k(\Delta) \subseteq B_k(\gamma) \wedge k < \omega\}$, we get that

$$\hat{\alpha} = \beta + \varphi_0(\Omega_n + 1 + \alpha) \in \mathcal{H}_\beta[\Delta]$$

since each $B_k(\alpha)$ is closed under addition and Veblen functions. Therefore, also $\hat{\alpha} \in \mathcal{H}_{\hat{\alpha}}[\Delta]$, and so $\psi_m(\hat{\alpha}) \in \mathcal{H}_{\hat{\alpha}}[\Delta]$.
Now, we prove Claim 4.24.1 by induction on Ω_n with a subsidiary induction on α.
If Γ is an axiom, then the claim is trivial.

We suppose that Γ has been obtained by the application of a rule. We run through the cases based upon this last inference rule.

Case 1. We suppose that the last rule applied has principal formula A of $\bigvee$-type. Then $\mathcal{H}_\beta[\Delta] \overset{\alpha}{\underset{\Omega_n+1}{\vdash}} \Gamma', A$ by hypothesis and $\mathcal{H}_\beta[\Delta] \overset{\alpha_0}{\underset{\Omega_n+1}{\vdash}} \Gamma', B$ with $\alpha_0 < \alpha$ for some $B \in \mathcal{C}(A)$. By the induction hypothesis, $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \overset{\psi_m\hat{\alpha}_0}{\underset{\psi_m\hat{\alpha}_0}{\vdash}} \Gamma', B$. But $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \subseteq \mathcal{H}_{\hat{\alpha}}[\Delta]$ and so we get $\mathcal{H}_{\hat{\alpha}}[\Delta] \overset{\psi_m\hat{\alpha}_0}{\underset{\psi_m\hat{\alpha}}{\vdash}} \Gamma', B$ by means of Lemma 4.15. Since $\psi_m\hat{\alpha}_0 < \psi_m\hat{\alpha}$ and $\psi_m\hat{\alpha} \in \mathcal{H}_{\hat{\alpha}}[\Delta]$, we can apply the rule to obtain $\mathcal{H}_{\hat{\alpha}}[\Delta] \overset{\psi_m\hat{\alpha}}{\underset{\psi_m\hat{\alpha}}{\vdash}} \Gamma', A$.

Case 2. We suppose that the last rule applied has principal formula A of $\bigwedge$-type. Then $\mathcal{H}_\beta[\Delta] \overset{\alpha}{\underset{\Omega_n+1}{\vdash}} \Gamma', A$ by hypothesis and $\mathcal{H}_\beta[\Delta \cup t_A(B)] \overset{\alpha_B}{\underset{\Omega_n+1}{\vdash}} \Gamma', B$ with $\alpha_B < \alpha$ for each $B \in \mathcal{C}(A)$. We need to show that $k(\Delta \cup t_A(B)) \subseteq B_m(\beta+1)$ for any $B \in \mathcal{C}(A)$ to be able to use the induction hypothesis. So we prove the following claim.

Claim 4.24.2. *With the hypothesis of Case 2., we have that* $k(\Delta \cup t_A(B)) \subseteq B_m(\beta + 1)$.

We prove Claim 4.24.2 by considering cases based on the form of A.

Subcase 2.1. We assume $A \equiv \neg Ad(t)$ for some term t. Then, since A is Σ^m, we have $|t| < \Omega_m$. Any premise in $\mathcal{C}(A)$ is of the form $t \neq \mathbb{L}_{\Omega_l}(X)$ for some $l < \omega$, and so $k(t \neq \mathbb{L}_{\Omega_l}(X)) = \{|t|, \Omega_l\} \subseteq B_m(\beta + 1)$.

Subcase 2.2. We assume $A \equiv B_0 \wedge B_1$. Then $k(A) = k(B_0) \cup k(B_1)$. Since $k(A) \subseteq B_m(\beta + 1)$, we also have $k(B_0), k(B_1) \subseteq B_m(\beta + 1)$.

Subcase 2.3. We assume $A \equiv s \notin t$ or $A \equiv \forall x \in t\ C(x)$ for some term(s) t and s with $|t|, |s| < \Omega_m$ (because A is Σ_m). Then, the characteristic term $t_A(B)$ for any $B \in \mathcal{C}(A)$ has always level below $|t|$. Therefore, $\{|t_A(B)| : B \in \mathcal{C}(A)\} \subseteq \delta$, where $\delta := |t| < \Omega_m$. But then $\delta \in k(A) \cap \Omega_m \subseteq \mathcal{H}_\beta[\Delta] \cap \Omega_m \subseteq B_m(\beta + 1) \cap \Omega_m$. The last inclusion comes from the equality $\mathcal{H}[\Delta] = \bigcap\{B_k(\gamma) : \beta < \gamma \wedge k(\Delta) \subseteq B_k(\gamma) \wedge k < \omega\}$ and the fact that $k(\Delta) \subseteq B_m(\beta + 1)$.
Gathering everything together, we have that $k(t_A(B)) \subseteq B_m(\beta+1)$ for all $B \in \mathcal{C}(A)$.

Thus, Claim 4.24.2 is shown, and we can use the induction hypothesis, that yields $\mathcal{H}_{\hat{\alpha}_B}[\Delta \cup t_A(B)] \left|\frac{\psi_m \hat{\alpha}_B}{\psi_m \hat{\alpha}_B}\right. \Gamma', B$ for every $B \in \mathcal{C}(A)$. But, for each $B \in \mathcal{C}(A)$, we have $\psi_m \hat{\alpha}_B < \psi_m \hat{\alpha}$ and $\psi_m \hat{\alpha} \in \mathcal{H}_{\hat{\alpha}}[\Delta]$. Therefore, by an application of the rule, we obtain $\mathcal{H}_{\hat{\alpha}}[\Delta] \left|\frac{\psi_m \hat{\alpha}}{\psi_m \hat{\alpha}}\right. \Gamma', A$.

Case 3. We suppose that the last rule applied is (Ref_k) for some $k \leq m$. Then, we have $\mathcal{H}_\beta[\Delta] \left|\frac{\alpha}{\Omega_n+1}\right. \Gamma', \exists z \in \mathbb{L}_{\Omega_k}(X) F^z$, where F is a Σ-formula and $\mathcal{H}_\beta[\Delta] \left|\frac{\alpha_0}{\Omega_n+1}\right.$ $\Gamma', F^{\mathbb{L}_{\Omega_k}(X)}$. By the induction hypothesis, $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_m \hat{\alpha}_0}{\psi_m \hat{\alpha}_0}\right. \Gamma', F^{\mathbb{L}_{\Omega_k}(X)}$. We cannot use again the rule since maybe $\psi_m \hat{\alpha} \geq \Omega_k$. Instead, we use Lemma 4.21 (Boundedness) to obtain $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_m \hat{\alpha}_0}{\psi_m \hat{\alpha}_0}\right. \Gamma', F^{\mathbb{L}_{\psi_k \hat{\alpha}_0}(X)}$. Moreover, since $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \subseteq \mathcal{H}_{\hat{\alpha}}[\Delta]$, we apply Lemma 4.15 to change the controlling operator and increase the bound of the complexity of the cuts and get $\mathcal{H}_{\hat{\alpha}}[\Delta] \left|\frac{\psi_m \hat{\alpha}_0}{\psi_m \hat{\alpha}}\right. \Gamma', F^{\mathbb{L}_{\psi_k \hat{\alpha}_0}(X)}$. Now, an application of $(b\exists)$ yields $\mathcal{H}_{\hat{\alpha}}[\Delta] \left|\frac{\psi_m \hat{\alpha}}{\psi_m \hat{\alpha}}\right. \Gamma', \exists z \in \mathbb{L}_{\Omega_k}(X) F^z$.

Case 4. We suppose that the last rule applied is (Cut). Then, we have $\mathcal{H}_\beta[\Delta] \left|\frac{\alpha}{\Omega_n+1}\right. \Gamma$. We also have the premises $\mathcal{H}_\beta[\Delta] \left|\frac{\alpha_0}{\Omega_n+1}\right. \Gamma, A$ and $\mathcal{H}_\beta[\Delta] \left|\frac{\alpha_0}{\Omega_n+1}\right. \Gamma, \neg A$ with $\alpha_0 < \alpha$ and $\text{rk}(A) < \Omega_n + 1$. We will run through cases based upon the ordering relation between $\mathrm{rk}(A)$ and Ω_m.

Subcase 4.1. We assume $\mathrm{rk}(A) < \Omega_m$. First of all, we observe that

$$\mathcal{H}_\beta[\Delta] = \cap\{B_l(\delta) : k(\Delta) \subseteq B_l(\delta) \wedge \beta < \delta \wedge l < \omega\} \subseteq B_m(\beta+1)$$

since $k(\Delta) \subseteq B_m(\beta+1)$ by assumption. Thus, since $\mathrm{rk}(A) \in k(A) \subseteq \mathcal{H}_\beta[\Delta]$, we have

$$\mathrm{rk}(A) \in \mathcal{H}_\beta[\Delta] \cap \Omega_m \subseteq B_m(\beta+1) \cap \Omega_m = \psi_m(\beta+1) \leq \psi_m\hat{\alpha}.$$

By the induction hypothesis, we have $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_m\hat{\alpha}_0}{\psi_m\hat{\alpha}_0}\right. \Gamma, A$ and $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_m\hat{\alpha}_0}{\psi_m\hat{\alpha}_0}\right. \Gamma, \neg A$. Now, taking as the operator control $\mathcal{H}_{\hat{\alpha}}[\Delta]$ by means of Lemma 4.15, an application of (Cut) yields $\mathcal{H}_{\hat{\alpha}}[\Delta] \left|\frac{\psi_m\hat{\alpha}}{\psi_m\hat{\alpha}}\right. \Gamma$ as desired. Cut complexity is not increased since $\mathrm{rk}(A) < \psi_m\hat{\alpha}$.

Subcase 4.2. We assume $\Omega_m \leq \mathrm{rk}(A) < \Omega_n + 1$. We notice that we are not able to proceed as in Subcase 4.1 because the complexity of the cuts in the last derivation go beyond $\psi_m\hat{\alpha}$. We prove the following claim.

Claim 4.24.3. *Let $\beta \leq \eta < \hat{\alpha}$ such that $\eta \in \mathcal{H}_\eta$, let $k = min(l < \omega : \mathrm{rk}(A) < \Omega_l)$. If $\mathcal{H}_\eta[\Delta] \left|\frac{\delta}{\delta}\right. \Gamma, A$ and $\mathcal{H}_\eta[\Delta] \left|\frac{\delta}{\delta}\right. \Gamma, \neg A$ for some $\delta < \Omega_k$ then*

$$\mathcal{H}_{\hat{\alpha}}[\Delta] \left|\frac{\psi_m\hat{\alpha}}{\psi_m\hat{\alpha}}\right. \Gamma.$$

We show Claim 4.24.3.
We let $\mu = \max(\mathrm{rk}(A), \delta) + 1$. We notice that $\mu \leq \omega^\mu < \Omega_k$. Let $\rho = \Omega_{k-1} + 1$ (we know $k > 0$ since $\Omega_m \leq rk(A)$).
Then, we have $\Omega_{k-1} < \rho < \rho + \omega^\mu < \Omega_k \leq \Omega_n$.
From the hypothesis of Claim 4.24.3, we have

$$\mathcal{H}_\eta[\Delta] \left|\frac{\delta}{\delta}\right. \Gamma, A \tag{1}$$

and

$$\mathcal{H}_\eta[\Delta] \left|\frac{\delta}{\delta}\right. \Gamma, \neg A. \tag{2}$$

By an application of (Cut), we obtain

$$\mathcal{H}_\eta[\Delta] \left|\frac{\delta+1}{\rho+\omega^\mu}\right. \Gamma.$$

We observe that the complexity of the cuts is in fact bounded by μ, and so it is also bounded by $\rho + \omega^\mu$.

Obviously, there is no Ω_l in the interval $[\rho, \rho + \omega^\mu)$. Moreover, $\omega^\mu \in \mathcal{H}_\eta[\Delta]$ since $\delta, \mathrm{rk}(A) \in \mathcal{H}_\eta[\Delta]$. We can use Theorem 4.20 (Predicative Cut Elimination) and we get

$$\mathcal{H}_\eta[\Delta] \overset{\varphi_\mu(\delta+1)}{\underset{\rho}{\vdash}} \Gamma.$$

Now, since $\beta \leq \eta$, we have that $k(\Delta) \subseteq B_m(\beta+1) \subseteq B_m(\eta+1)$. Also, $\eta \in \mathcal{H}_\eta[\Delta]$ by assumption. Thus, the conditions of the Collapsing Theorem (in fact, the conditions of the general Claim 4.24.1 we are proving) are met and so, since $\rho < \Omega_n$, we use the main induction hypothesis to obtain

$$\mathcal{H}_{\eta+\omega^{\rho+\varphi_\mu(\delta+1)}}[\Delta] \overset{\psi_m(\eta+\omega^{\rho+\varphi_\mu(\delta+1)})}{\underset{\psi_m(\eta+\omega^{\rho+\varphi_\mu(\delta+1)})}{\vdash}} \Gamma. \tag{3}$$

Here, we have used the induction hypothesis from the main induction on Ω_n.
It remains to show that $\psi_m(\eta+\omega^{\rho+\varphi_\mu(\delta+1)}) \leq \psi_m(\beta+\omega^{\Omega_n+1+\alpha})$ in order to use Lemma 4.15 and obtain the conclusion of Claim 4.24.3.
From $\rho + \varphi_\mu(\delta+1) < \Omega_{k+1} \leq \Omega_n$ we get that

$$\omega^{\rho+\varphi_\mu(\delta+1)} < \omega^{\Omega_n}.$$

Now, we observe that, since $\beta \leq \eta < \beta + \omega^{\Omega_n+1+\alpha}$ we can write

$$\eta = \beta + \zeta \text{ for some } \zeta < \omega^{\Omega_n+1+\alpha}.$$

It follows that

$$\begin{aligned}
\eta + \omega^{\Omega_k+\varphi_\mu(\delta+1)} &= \beta + \zeta + \omega^{\Omega_k+\varphi_\mu(\delta+1)} \\
&< \beta + \zeta + \omega^{\Omega_n} \\
&\leq \beta + \omega^{\Omega_n+1+\alpha} \text{ since } \zeta, \omega^{\Omega_n} < \omega^{\Omega_n+1+\alpha} \text{ and } \omega^{\Omega_n+1+\alpha} \text{ is additive principal} \\
&= \hat{\alpha}.
\end{aligned}$$

From this inequality we derive that

$$\mathcal{H}_\eta + \omega^{\rho+\varphi_\mu(\delta+1)}[\Delta] \subseteq \mathcal{H}_{\hat{\alpha}} \text{ and } \psi_m(\eta + \omega^{\rho+\varphi_\mu(\delta+1)}) < \psi_m(\hat{\alpha}).$$

Lemma 4.15 on (3) yields

$$\mathcal{H}_{\hat{\alpha}}[\Delta] \overset{\psi_m(\hat{\alpha})}{\underset{\psi_m(\hat{\alpha})}{\vdash}} \Gamma.$$

Hence, Claim 4.24.3 is verified and we can continue analyzing Subcase 4.2 as follows.

Subsubcase 4.2.1. We assume $\mathrm{rk}(A) \neq \Omega_j$ for any $j < \omega$. Let $\Omega_k = min(\Omega_i :$

$\mathrm{rk}(A) < \Omega_i$). We notice that $\Gamma \cup \{A, \neg A\}$ is in particular a set of Σ^{Ω_k} formulas because $\Omega_m < \mathrm{rk}(A) < \Omega_k$. By the induction hypothesis, we have $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_k \hat{\alpha}_0}{\psi_k \hat{\alpha}_0}\right. \Gamma, A$ and $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_k \hat{\alpha}_0}{\psi_k \hat{\alpha}_0}\right. \Gamma, \neg A$. We use Claim 4.24.3 with $\eta = \hat{\alpha}_0$, $\delta = \psi_k \hat{\alpha}_0$ and $k = k$ to get the result.

Subsubcase 4.2.2. We assume $\Omega_m \leq \mathrm{rk}(A) = \Omega_k \leq \Omega_n$. In this case, A and $\neg A$ are of the form $\exists x \in \mathbb{L}_{\Omega_k}(X) B(x)$ and $\forall x \in \mathbb{L}_{\Omega_k}(X) \neg B(x)$ (respectively or alternatively), and we have

$$\mathcal{H}_{\beta}[\Delta] \left|\frac{\alpha_0}{\Omega_n + 1}\right. \Gamma, \exists x \in \mathbb{L}_{\Omega_k}(X) B(x) \tag{4}$$

and

$$\mathcal{H}_{\beta}[\Delta] \left|\frac{\alpha_0}{\Omega_n + 1}\right. \Gamma, \forall x \in \mathbb{L}_{\Omega_k}(X) \neg B(x). \tag{5}$$

By the induction hypothesis on (4), we have $\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_k \hat{\alpha}_0}{\psi_k \hat{\alpha}_0}\right. \Gamma, \exists x \in \mathbb{L}_{\Omega_k}(X) B(x)$. Let $\xi = \psi_k(\Omega_n + \omega^{\beta + \alpha_0})$. We observe that $\xi \in \mathcal{H}_{\hat{\alpha}_0}[\Delta] \cap \Omega_k$. By Lemma 4.21 (Boundedness), we get

$$\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\psi_k \hat{\alpha}_0}{\psi_k \hat{\alpha}_0}\right. \Gamma, \exists x \in \mathbb{L}_{\xi}(X) B(x). \tag{6}$$

On the other hand, we apply Lemma 4.23 to (5) to obtain

$$\mathcal{H}_{\beta}[\Delta] \left|\frac{\alpha_0}{\Omega_n + 1}\right. \Gamma, \forall x \in \mathbb{L}_{\xi}(X) \neg B(x).$$

Since $\beta < \hat{\alpha}_0$, by Lemma 4.15 we get

$$\mathcal{H}_{\hat{\alpha}_0}[\Delta] \left|\frac{\alpha_0}{\Omega_n + 1}\right. \Gamma, \forall x \in \mathbb{L}_{\xi}(X) \neg B(x).$$

By the induction hypothesis,

$$\mathcal{H}_{\hat{\alpha}_0 + \omega^{\Omega_n + 1 + \alpha_0}}[\Delta] \left|\frac{\psi_k(\hat{\alpha}_0 + \omega^{\Omega_n + 1 + \alpha_0})}{\psi_k(\hat{\alpha}_0 + \omega^{\Omega_n + 1 + \alpha_0})}\right. \Gamma, \forall x \in \mathbb{L}_{\xi}(X) \neg B(x). \tag{7}$$

Now, we apply Claim 4.24.3 to (6) and (7) with $\delta = \psi_k(\hat{\alpha}_0 + \omega^{\Omega_n + 1 + \alpha_0})$, $\eta = \hat{\alpha}_0 + \omega^{\Omega_n + 1 + \alpha_0}$ and $k = k$ and obtain the result. $\square$

5 Embedding KPl into $\mathsf{RS}_l(X)$

In this section, we will prove that we can derive in $\mathsf{RS}_l(X)$ any given finite set of formulas provable in KPl changing free variables by terms and bounding the unbounded quantifiers of the formulas by $\mathbb{L}_{\Omega_\omega}(X)$ with a bounded cut complexity and depth. In particular, in Theorem 5.20 we will show that if $\mathsf{KPl} \vdash \Gamma(a_1, \ldots, a_n)$ then there is $m < \omega$ such that $\mathcal{H}[s_1, \ldots, s_n] \overset{\Omega_\omega \cdot \omega^m}{\underset{\Omega_\omega + m}{\vdash}} \Gamma(s_1, \ldots, s_n)^{\mathbb{L}_{\Omega_\omega}(X)}$ for any operator $\mathcal{H}$ and any terms $s_1, \ldots, s_n$ of level below Ω_ω.

5.1 The $\Vdash$ relation

We start by introducing the $\Vdash$ relation. The relation $\Vdash \Gamma$ will mean that a set of formulas Γ is derivable with the control of any operator with a reasonable depth depending on the rank of those formulas. The ordinal bounds of the derivations will use the following operation.

Definition 5.1. *Let $\alpha_1, \ldots, \alpha_n$ be ordinals. Let $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a function such that $\alpha_{\pi(1)} \geq \cdots \geq \alpha_{\pi(n)}$. We define*

$$\alpha_1 \# \ldots \# \alpha_n = \alpha_{\pi(1)} + \cdots + \alpha_{\pi(n)}.$$

With the operation $\#$, we can now define $\Vdash \Gamma$.

Definition 5.2. *Given a set of formulas $\Gamma = \{A_1, \ldots, A_n\}$, we define*

$$\#\Gamma = \omega^{\mathrm{rk}(A_1)} \# \cdots \# \omega^{\mathrm{rk}(A_n)}.$$

Now, we define the relation $\Vdash$.
We write $\Vdash \Gamma$ whenever for any operator $\mathcal{H}$ we have $\mathcal{H}[\Gamma] \overset{\#\Gamma}{\underset{0}{\vdash}} \Gamma$.
We write $\overset{\alpha}{\underset{\rho}{\Vdash}} \Gamma$ whenever for any operator $\mathcal{H}$ we have $\mathcal{H}[\Gamma] \overset{\#\Gamma\#\alpha}{\underset{\rho}{\vdash}} \Gamma$.

The first part of the next lemma shows that we can treat the $\Vdash$ relation as a logic in the sense that given the premise(s) Γ_i, for $i \in y$, of the conclusion Γ of some instance of an $\mathsf{RS}_l(X)$-rule, if $\Vdash \Gamma_i$ for all $i \in y$ then $\Vdash \Gamma$. The second part of the lemma shows that whenever the formulas A, B are derivable with $\Vdash$ then we can derive $A \vee B$ instead. This will be useful to derive formulas of the form $A \to B$. For instance, a general reasoning that we will use to prove that $\Vdash A \to B$ will be to first derive $\Vdash \neg A, B$ and, by means of this Lemma, we will get $\Vdash \neg A \vee B$, which is equivalent to $\Vdash A \to B$.
Along this section, we follow Reading Convention 4.16 and omit the repetition of principal formulas in the premises.

Lemma 5.3. *Let $\Gamma \cup \{A, B\}$ be a finite set of formulas. Let α and ρ be ordinals.*

1. *If $B \in \mathcal{C}(A)$ then $\#(\Gamma, B)\#\alpha < \#(\Gamma, A)\#\alpha$.*
   Moreover, if Γ, A follows from the premise(s) Γ, B_i, for $i \in y$, by a rule other than (Cut) and (Ref_n), with $n < \omega$, with principal formula A and active formulas B_i, then $\Vdash^{\alpha}_{\rho} \Gamma, A$ whenever $\Vdash^{\alpha}_{\rho} \Gamma, B_i$ for all $i \in y$.

2. *If $\Vdash^{\alpha}_{\rho} \Gamma, A, B$ then $\Vdash^{\alpha}_{\rho} \Gamma, A \vee B$.*

Proof. 1. First, by Lemma 4.8 we have $\mathrm{rk}(B) < \mathrm{rk}(A)$ whenever $B \in \mathcal{C}(A)$. It follows that $\omega^{\mathrm{rk}(B)} < \omega^{\mathrm{rk}(A)}$ for any $B \in \mathcal{C}(A)$ and so $\#(\Gamma, B) < \#(\Gamma, A)$ for every $B \in \mathcal{C}(A)$.
We suppose now that $\Vdash^{\alpha}_{\rho} \Gamma, B_i$ for all $i \in y$ and fix an operator $\mathcal{H}$. Then,

$$\mathcal{H}[\Gamma, B_i] \vdash^{\#(\Gamma,B_i)\#\alpha}_{\rho} \Gamma, B_i$$

for all $i \in y$. By an application of the rule, and since $\#(\Gamma, B_i)\#\alpha < \#(\Gamma, A)\#\alpha$ for all $i \in y$, we get $\mathcal{H}[\Gamma, A] \vdash^{\#(\Gamma,A)\#\alpha}_{\rho} \Gamma, A$. Hence, $\Vdash^{\alpha}_{\rho} \Gamma, A$.

2. We suppose $\Vdash^{\alpha}_{\rho} \Gamma, A, B$ and fix an operator $\mathcal{H}$. Then,

$$\mathcal{H}[\Gamma, A, B] \vdash^{\#(\Gamma,A,B)\#\alpha}_{\rho} \Gamma, A, B.$$

We apply $(\vee)$ twice (one on A and the other on (B)) to obtain

$$\mathcal{H}[\Gamma, A, B] \vdash^{\#(\Gamma,A,B)\#\alpha+2}_{\rho} \Gamma, A \vee B, A \vee B,$$

which is exactly $\mathcal{H}[\Gamma, A, B] \vdash^{\#(\Gamma,A,B)\#\alpha+2}_{\rho} \Gamma, A \vee B$. By Item 1. of this lemma and since $\omega^{\mathrm{rk}(A\vee B)}$ is additive principal, we have

$$\#(\Gamma, A, B)\#\alpha+2 = \#\Gamma\#\omega^{\mathrm{rk}(A)}\#\omega^{\mathrm{rk}(B)}\#\alpha+2 < \#\Gamma\#\omega^{\mathrm{rk}(A\vee B)}\#\alpha = \#(\Gamma, A\vee B)\#\alpha.$$

Therefore, by Lemma 4.15 and since $\mathcal{H}[\Gamma, A, B] = \mathcal{H}[\Gamma, A \vee B]$ we obtain

$$\mathcal{H}[\Gamma, A \vee B] \vdash^{\#(\Gamma,A\vee B\#\alpha)}_{\rho} \Gamma, A \vee B.$$

Hence, $\Vdash^{\alpha}_{\rho} \Gamma, A \vee B$.

$\square$

We explain the remark preceding the statement of Lemma 5.3 more precisely. This lemma says that whenever we can derive a set of formulas Γ with the $\Vdash$ relation and a set of formulas Γ' follows from Γ by an $\mathsf{RS}_l(X)$ rule (with the control of any operator), then we can write the derivation with the $\Vdash$ relation. For example, if for any operator $\mathcal{H}$ we have $\mathcal{H}[A] \left|\frac{\omega^{\mathrm{rk}(A)}}{0}\right. A$ then we also have $\mathcal{H}[A \vee B] \left|\frac{\omega^{\mathrm{rk}(A\vee B)}}{0}\right. A \vee B$. In this case, we will write

$$(\vee)\ \frac{\Vdash A}{\Vdash A \vee B}$$

This is how we will write derivations with the $\Vdash$ relation, for example, in the proof of the next lemma.

Lemma 5.4. *For any formula A we have $\Vdash A, \neg A$.*

Proof. We proceed by induction on the rank of A. We note that given any non-basic formula A, since $\mathrm{rk}(B) < \mathrm{rk}(A)$ for any $B \in \mathcal{C}(A)$, when proving $\Vdash A, \neg A$ we can suppose $\Vdash B, \neg B$ by the induction hypothesis. We consider cases based on the form of A.

Case 1. We suppose that $A \equiv \overline{u} \in \overline{v}$. This means that $A, \neg A$ is an axiom (since either $u \in v$ or $u \notin v$ holds).

Case 2. We suppose that $A \equiv r \in t$ is not a basic formula. This means that either r or t is not a basic term, and so $|r| > \Gamma_{\theta+1}$ or $|t| > \Gamma_{\theta+1}$. By the induction hypothesis, we have $\Vdash s \dot{\in} t \wedge r = s, \neg(s \dot{\in} t \wedge r = s)$, which is exactly $\Vdash s \dot{\in} t \wedge r = s, s \dot{\in} t \to r \neq s$, for all terms s with $|s| < |t|$. Therefore, we have the following derivation for every term s with $|s| < |t|$, where the first inference is applied to the first formula and the second inference is applied to the second formula:

$$\begin{array}{l} (\in)\ \dfrac{\Vdash s \dot{\in} t \wedge r = s, s \dot{\in} t \to r \neq s}{\Vdash r \in t, s \dot{\in} t \to r \neq s} \\ (\notin)\ \dfrac{\Vdash r \in t, s \dot{\in} t \to r \neq s}{\Vdash r \in t, r \notin t} \end{array}$$

Hence, we obtain $\Vdash A, \neg A$.

Case 3. We suppose that $A \equiv \exists x \in t\ B(x)$. By the inductive hypothesis, we have $\Vdash s \dot{\in} t \wedge B(s), \neg(s \dot{\in} t \wedge B(s))$, which is exactly $\Vdash s \dot{\in} t \wedge B(s), s \dot{\in} t \to \neg B(s)$, for all terms s with $|s| < |t|$. Therefore, we have the following derivation for every term s with $|s| < |t|$, where the first inference is applied to the first formula and the second inference is applied to the second formula:

$$
\begin{array}{rl}
(b\exists) & \dfrac{\Vdash s\dot{\in}t \wedge B(s), s\dot{\in}t \rightarrow \neg B(s)}{\Vdash \exists x \in t\ B(x), s\dot{\in}t \rightarrow \neg B(s)} \\
(b\forall) & \dfrac{}{\Vdash \exists x \in t\ B(x), \forall x \in t\ \neg B(x)}
\end{array}
$$

Hence, we obtain $\Vdash A, \neg A$.
The other cases are analogous to those ones, but using the appropriate rules. □

At some points we will need to write derived formulas in some equivalent expression, e.g. write $A \rightarrow B$ instead of $\neg A \vee B$ as already mentioned before. We will simply use the symbol $\equiv$ as the label of the derivation when this happens.
The next lemma states some results that will be helpful to embed the KPl axioms and rules into the $\mathsf{RS}_l(X)$-system.

Lemma 5.5. *Let s be any term. Then, we have*

1. $\Vdash s \notin s$,
2. *Given any term t, if $|s| < |t|$ then* $\Vdash s\dot{\in}t \rightarrow s \in t$,
3. $\Vdash s \subseteq s$,
4. $\Vdash s = s$,
5. *Let α be an ordinal. If $|s| < \Gamma_{\theta+1} + \alpha$ then* $\Vdash s \in \mathbb{L}_\alpha(X)$.

Proof. 1. We proceed by induction on $\mathrm{rk}(s)$. We consider cases based on the form of s.

Case 1. We suppose $s \equiv \overline{u}$. Then $s \notin s$ is an axiom.

Case 2. We suppose $s \equiv \mathbb{L}_\alpha(X)$. By the induction hypothesis, we get $\Vdash r \notin r$ for all $|r| < |s|$, which is the same as $\Vdash r\dot{\in}s \wedge r \notin r$ by Definition 4.4. Therefore, we obtain the following derivation for all $|r| < |s|$:

$$
\begin{array}{rl}
(b\exists) & \dfrac{\Vdash r\dot{\in}s \wedge r \notin r}{\Vdash \exists x \in s(x \notin r)} \\
(\vee) & \dfrac{}{\Vdash \exists x \in s(x \notin r) \vee \exists x \in r(x \notin s)} \\
\equiv & \dfrac{}{\Vdash s \neq r} \\
\equiv & \dfrac{}{\Vdash r\dot{\in}s \rightarrow s \neq r} \\
(\notin) & \dfrac{}{\Vdash s \notin s}
\end{array}
$$

Case 3. We suppose $s \equiv [x \in \mathbb{L}_\alpha(X) : B(x)]$. By Lemma 5.4 we have $\Vdash B(r), \neg B(r)$ for any term r, and in particular for any term r with $|r| < |s|$. Moreover, by the induction hypothesis we have $\Vdash r \notin r$ for all $|r| < |s|$. We get the following derivation for all $|r| < |s|$:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Vdash B(r), \neg B(r) \qquad \Vdash r \notin r}{\Vdash B(r) \wedge r \notin r, \neg B(r)}(\wedge)}{\Vdash \exists x \in s(x \notin r), \neg B(r)}(b\exists)}{\Vdash \exists x \in s(x \notin r) \vee \exists x \in r(x \notin s), \neg B(r)}(\vee)}{\Vdash s \neq r, \neg B(r)}\equiv}{\Vdash \neg B(r) \vee \neg s = r}\text{Lemma 5.3}}{\Vdash B(r) \to s \neq r}\equiv}{\Vdash r \dot{\in} s \to s \neq r}\equiv}{\Vdash s \notin s}(\notin)
$$

We prove 2. and 3. simultaneously. Actually, we show $\Vdash \forall x \in s(x \in s)$ by induction on $\mathrm{rk}(s)$ and considering cases based on the form of s, and 2. will be shown along the way.

Case 1. We suppose $s \equiv \overline{u}$. Then, given the basic term $\overline{v}$, either $\overline{v} \in \overline{u}$ or $\overline{v} \notin \overline{u}$ is an axiom, and so $\overline{v} \in \overline{u}, \overline{v} \notin \overline{u}$ is an axiom. We have the following derivation for any basic term $\overline{v}$:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\Vdash \overline{v} \in \overline{u}, \overline{v} \notin \overline{u}}{\Vdash \neg \overline{v} \in \overline{u} \vee \overline{v} \in \overline{u}}\text{Lemma 5.3}}{\Vdash \overline{v} \in \overline{u} \to \overline{v} \in \overline{u}}\equiv}{\Vdash \overline{v} \dot{\in} \overline{u} \to \overline{v} \in \overline{u}}\equiv}{\Vdash \forall x \in s(x \in s)}(b\forall)
$$

We notice that the second to last line is exactly $\Vdash \overline{v} \dot{\in} \overline{u} \to \overline{v} \in \overline{u}$, as in Item 2.

Case 2. We suppose $s \equiv \mathbb{L}_\alpha(X)$. By the induction hypothesis $\Vdash \forall x \in r(x \in r)$

for any $|r| < |s|$. Therefore, we have the following derivation for all $|r| < |s|$:

$$(\wedge)\ \dfrac{\Vdash \forall x \in r(x \in r) \qquad \Vdash \forall x \in r(x \in r)}{\Vdash \forall x \in r(x \in r) \wedge \forall x \in r(x \in r)}$$
$$\equiv\ \dfrac{\Vdash \forall x \in r(x \in r) \wedge \forall x \in r(x \in r)}{\Vdash r = r}$$
$$\equiv\ \dfrac{\Vdash r = r}{\Vdash r \dot{\in} s \wedge r = r}$$
$$(\in)\ \dfrac{\Vdash r \dot{\in} s \wedge r = r}{\Vdash r \in s}$$
$$\equiv\ \dfrac{\Vdash r \in s}{\Vdash r \dot{\in} s \rightarrow r \in s}$$
$$(b\forall)\ \dfrac{\Vdash r \dot{\in} s \rightarrow r \in s}{\Vdash \forall x \in s(x \in s)}$$

Case 3. We suppose $s \equiv [x \in \mathbb{L}_\alpha(X) : B(x)]$. By the induction hypothesis, we get $\Vdash \forall x \in r(x \in r)$ for all $|r| < |s|$. So, by Lemma 4.15, we obtain $\Vdash \forall x \in r(x \in r), \neg B(r)$. Applying $(\wedge)$, we have $\Vdash r = r, \neg B(r)$. Moreover, we have $\Vdash B(r), \neg B(r)$ by Item 1. We get the following derivation for all $|r| < |s|$:

$$(\wedge)\ \dfrac{\Vdash r = r, \neg B(r) \qquad \Vdash B(r), \neg B(r)}{\Vdash B(r) \wedge r = r, \neg B(r)}$$
$$\equiv\ \dfrac{\Vdash B(r) \wedge r = r, \neg B(r)}{\Vdash r \dot{\in} s \wedge r = r, \neg B(r)}$$
$$(\in)\ \dfrac{\Vdash r \dot{\in} s \wedge r = r, \neg B(r)}{\Vdash r \in s, \neg B(r)}$$
$$\text{Lemma 5.3}\ \dfrac{\Vdash r \in s, \neg B(r)}{\Vdash \neg B(r) \vee r \in s}$$
$$\equiv\ \dfrac{\Vdash \neg B(r) \vee r \in s}{\Vdash B(r) \rightarrow r \in s}$$
$$\equiv\ \dfrac{\Vdash B(r) \rightarrow r \in s}{\Vdash r \dot{\in} s \rightarrow r \in s}$$
$$(b\forall)\ \dfrac{\Vdash r \dot{\in} s \rightarrow r \in s}{\Vdash \forall x \in s(x \in s)}$$

4. The result follows from Item 3. by an application of the $(\wedge)$ rule.

5. First, for all $|s| < \Gamma_{\theta+1} + \alpha$ we have $\Vdash s = s$ by Item 3. Using Definition 4.4, this is equivalent to $\Vdash s \dot{\in} \mathbb{L}_\alpha(X) \wedge s = s$. We apply $(\in)$ to obtain the result. $\square$

5.2 The embedding Theorem

In this subsection, we will show that all the axioms of KPI can be embedded in $\mathsf{RS}_l(X)$ and we will give ordinal length and cut-complexity bounds of the derivations of the axioms in $\mathsf{RS}_l(X)$. It seems that, in order to prove the desired Theorem 5.20 stated at the beginning of this section, it is sufficient to prove that we can find ordinal bounds α and β such that, for any KPI axiom Ax and any operator $\mathcal{H}$, we have

$$\mathcal{H} \left|\frac{\alpha}{\beta}\right. (\mathsf{Ax})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Nonetheless, we need a stronger result: the third Ad axiom states that any admissible set has to satisfy basic axioms. This means that we have to prove that, given an axiom Ax among Leibniz Principle, Pair, Union, Δ_0-Separation and Δ_0-Collection

$$\mathcal{H} \overset{\alpha}{\underset{\beta}{\vdash}} (\mathsf{Ax})^{\mathbb{L}_{\Omega_n}(X)}$$

for all $n \leq \omega$ and for any operator $\mathcal{H}$. Actually, for Δ_0-Collection we only need $n < \omega$ since this axiom is not an axiom of KPl.

To prove that the Leibniz Principle can be embedded into the $\mathsf{RS}_l(X)$-system, we will use the following preliminary Lemma.

Lemma 5.6. *Let s, t be terms such that $|s| < |t|$. Let Γ be any finite set of formulas and let A and B be formulas. If $\Vdash \Gamma, A, B$ then*

$$\Vdash \Gamma, s\dot{\in}t \to A, s\dot{\in}t \wedge B.$$

Proof. We argue by splitting cases based on the form of t.

Case 1. We suppose that $t \equiv \overline{u}$. Then, $s \equiv \overline{v}$ since $|s| < |t|$, Therefore, by hypothesis $\Vdash \Gamma, A, B$ and by Lemma 5.4 (together with Lemma 4.15) we have $\Vdash \Gamma, \overline{v} \in \overline{u}, \overline{v} \notin \overline{u}$. We get the following derivaitons:

$$\dfrac{(\wedge)\ \dfrac{\equiv\ \dfrac{(\vee)\ \dfrac{\Vdash \Gamma, \overline{v} \in \overline{u}, \overline{v} \notin \overline{u}}{\Vdash \Gamma, \neg\overline{v} \in \overline{u} \vee A, \overline{v} \in \overline{u}}}{\Vdash \Gamma, \overline{v} \in \overline{u} \to A, \overline{v} \in \overline{u}} \qquad \equiv\ \dfrac{(\vee)\ \dfrac{\Vdash \Gamma, A, B}{\Gamma, \neg\overline{v} \in \overline{u} \vee A, B}}{\Gamma, \overline{v} \in \overline{u} \to A, B}}{\equiv\ \Vdash \Gamma, \overline{v} \in \overline{u} \to A, \overline{v} \in \overline{u} \wedge B}}{\Gamma, \overline{v}\dot{\in}\overline{u} \to A, \overline{v}\dot{\in}\overline{u} \wedge B}$$

Case 2. We suppose that $t \equiv \mathbb{L}_\alpha(X)$. Then $s\dot{\in}t \to A \equiv A$ and $s\dot{\in}t \wedge B \equiv B$ by Definition 4.4. The desired result is, in this case, exactly $\Vdash \Gamma, A, B$, which is the hypothesis.

Case 3. We suppose that $t \equiv [x \in \mathbb{L}_\alpha(X) : C(x)]$. Then $s\dot{\in}t \to A \equiv C(s) \to A$ and $s\dot{\in}t \wedge B \equiv C(s) \wedge B$. Now, by Lemma 5.4 (together with Lemma 4.15) we have $\Vdash \Gamma, C(s), \neg C(s)$ and by hypothesis we have $\Vdash \Gamma, A, B$. We obtain the following derivations:

$$\dfrac{(\wedge)\ \dfrac{\equiv\ \dfrac{(\vee)\ \dfrac{\Vdash \Gamma, C(s), \neg C(s)}{\Vdash \Gamma, \neg C(s) \vee A, C(s)}}{\Vdash \Gamma, C(s) \to A, C(s)} \qquad \equiv\ \dfrac{(\vee)\ \dfrac{\Vdash \Gamma, A, B}{\Gamma, \neg C(s) \vee A, B}}{\Gamma, C(s) \to A, B}}{\equiv\ \Vdash \Gamma, C(s) \to A, C(s) \wedge B}}{\Gamma, s\dot{\in}t \to A, s\dot{\in}t \wedge B}$$

□

Lemma 5.7 (Leibniz Principle)**.** *Let s and t be terms. Let $n \leq \omega$. For any formula* KPI*-formula $A(x)$, we have*

$$\Vdash s \neq t, \neg A(s)^{\mathbb{L}_{\Omega_n}(X)}, A(t)^{\mathbb{L}_{\Omega_n}(X)}.$$

Proof. We prove a more general claim.

Claim 5.7.1. *Let $s_1, \dots, s_k, t_1, \dots, t_k$ be terms. Let $n \leq \omega$. For any* KPI*-formula $B(x_1, \dots, x_k)$,*

$$\Vdash s_1 \overline{\neq} t_1, \dots, s_k \overline{\neq} t_k, \neg B(s_1, \dots, s_k)^{\mathbb{L}_{\Omega_n}(X)}, B(t_1, \dots, t_k)^{\mathbb{L}_{\Omega_n}(X)}, \tag{1}$$

where $x \overline{\neq} y := \neg x \subseteq y, \neg y \subseteq x$.

With Claim 5.7.1, let m be the number of appearances of the free variable x in $A(x)$. Define the KPI-formula $C(x_1, \dots, x_m)$ with m free variables as the formula $A(x)$ but replacing the first appearance of x with x_1, the second appearance of x with $x_2, \dots$, the $m - th$ appearance of x with x_m. This means that $A(x) \equiv C(x, \dots, x)$. We apply the Claim 5.7.1 to $C(x_1, \dots x_m)$ to obtain

$$\Vdash s \overline{\neq} t, \dots, s \overline{\neq} t, \neg C(s, \dots, s)^{\mathbb{L}_{\Omega_n}(X)}, C(t, \dots, t)^{\mathbb{L}_{\Omega_n}(X)}.$$

which is the same as

$$\Vdash s \overline{\neq} t, \neg A(s)^{\mathbb{L}_{\Omega_n}(X)}, A(t)^{\mathbb{L}_{\Omega_n}(X)}.$$

Applying $(\vee)$ to $s \overline{\neq} t$ yields

$$\Vdash s \neq t, \neg A(s)^{\mathbb{L}_{\Omega_n}(X)}, A(t)^{\mathbb{L}_{\Omega_n}(X)}.$$

We prove Claim 5.7.1 by induction and by splitting cases based on the form of the KPI-formula $B(x_1, \dots, x_k)$. First, we show a useful property about the symbol $\overline{\neq}$: given any terms a, b and any finite set of formulas Γ

$$\text{if } \Vdash \Gamma, a \overline{\neq} b, \text{ then } \Vdash \Gamma, a \neq b. \tag{2}$$

Indeed, we have the following derivation:

$$\begin{array}{ll}
 & \Vdash \Gamma, a \overline{\neq} b \\
\equiv & \overline{\Vdash \Gamma, \neg(a \subseteq b), \neg(b \subseteq a)} \\
(\vee) & \overline{\Vdash \Gamma, \neg(a \subseteq b) \vee \neg(b \subseteq a), \neg(b \subseteq a)} \\
(\vee) & \overline{\Vdash \Gamma, \neg(a \subseteq b) \vee \neg(b \subseteq a), \neg(a \subseteq b) \vee \neg(b \subseteq a)} \\
\equiv & \overline{\Vdash \Gamma, a \neq b, a \neq b} \\
\equiv & \overline{\Vdash \Gamma, a \neq b}
\end{array}$$

Case 1. We suppose that $B(x_1, \dots, x_k) \equiv B(x_i, x_j) \equiv x_i \in x_j$ for some $i, j \in \{1, \dots, n\}$.
Subcase 1.1. We assume $i = j$. We have $B(x_i) \equiv x_i \in x_i$ and so $B(s_i)^{\mathbb{L}_{\Omega_n}(X)} \equiv s_i \in s_i$. Then, $\neg B(s_i)^{\mathbb{L}_{\Omega_n}(X)} \equiv s_1 \notin s_1$. But, by Lemma 5.5(1.) (together with Lemma 4.15), we already have $\Vdash s_i \overline{\neq} t_i, s_i \notin s_i, t_i \in t_i$, which is the desired result.

Subcase 1.2. We assume $i \neq j$. Then $B(x_i, x_j) \equiv x_i \in x_j$. This means that $B(s_i, s_j)^{\mathbb{L}_{\Omega_n}(X)} \equiv s_i \in s_j$. We notice that any premise $C \in \mathcal{C}(s_i \in s_j)$ is of the form $s_i = s$ for $|s| < |s_j|$. By the induction hypothesis, we have

$$\Vdash s_i \overline{\neq} t_i, s \overline{\neq} t, s_i \neq s, t_i = t$$

for all $|s| < |s_j|$ and $|t| < |t_j|$. We obtain the following derivation for all $|s| < |s_j|$ and all $|t| < |t_j|$:

$$\text{(2)}\ \dfrac{\Vdash s_i \overline{\neq} t_i, s \overline{\neq} t, s_i \neq s, t_i = t}{\Vdash s_i \overline{\neq} t_i, s \neq t, s_i \neq s, t_i = t}$$

$$\text{Lemma 5.6}\ \dfrac{\Vdash s_i \overline{\neq} t_i, s \neq t, s_i \neq s, t_i = t}{\Vdash s_i \overline{\neq} t_i, t \dot{\in} t_j \to s \neq t, s_i \neq s, t \dot{\in} t_j \wedge t_i = t}$$

$$(\in)\ \dfrac{\Vdash s_i \overline{\neq} t_i, t \dot{\in} t_j \to s \neq t, s_i \neq s, t \dot{\in} t_j \wedge t_i = t}{\Vdash s_i \overline{\neq} t_i, t \dot{\in} t_j \to s \neq t, s_i \neq s, t_i \in t_j}$$

$$(\notin)\ \dfrac{\Vdash s_i \overline{\neq} t_i, t \dot{\in} t_j \to s \neq t, s_i \neq s, t_i \in t_j}{\Vdash s_i \overline{\neq} t_i, s \notin t_j, s_i \neq s, t_i \in t_j}$$

$$\text{Lemma 5.6}\ \dfrac{\Vdash s_i \overline{\neq} t_i, s \notin t_j, s_i \neq s, t_i \in t_j}{\Vdash s_i \overline{\neq} t_i, s \dot{\in} s_j \wedge s \notin t_j, s \dot{\in} s_j \to s_i \neq s, t_i \in t_j}$$

$$(b\exists)\ \dfrac{\Vdash s_i \overline{\neq} t_i, s \dot{\in} s_j \wedge s \notin t_j, s \dot{\in} s_j \to s_i \neq s, t_i \in t_j}{\Vdash s_i \overline{\neq} t_i, \exists x \in s_j (x \notin t_j), s \dot{\in} s_j \to s_i \neq s, t_i \in t_j}$$

$$(\notin)\ \dfrac{\Vdash s_i \overline{\neq} t_i, \exists x \in s_j (x \notin t_j), s \dot{\in} s_j \to s_i \neq s, t_i \in t_j}{\Vdash s_i \overline{\neq} t_i, \exists x \in s_j (x \notin t_j), s_i \notin s_j, t_i \in t_j}$$

$$(\vee)\ \dfrac{\Vdash s_i \overline{\neq} t_i, \exists x \in s_j (x \notin t_j), s_i \notin s_j, t_i \in t_j}{\Vdash s_i \overline{\neq} t_i, \exists x \in s_j (x \notin t_j) \vee \exists x \in t_j (x \notin s_j), s_i \notin s_j, t_i \in t_j}$$

$$\equiv\ \dfrac{\Vdash s_i \overline{\neq} t_i, \exists x \in s_j (x \notin t_j) \vee \exists x \in t_j (x \notin s_j), s_i \notin s_j, t_i \in t_j}{\Vdash s_i \overline{\neq} t_i, s_j \neq t_j, s_i \notin s_j, t_i \in t_j}$$

$$\text{(2)}\ \dfrac{\Vdash s_i \overline{\neq} t_i, s_j \neq t_j, s_i \notin s_j, t_i \in t_j}{\Vdash s_i \neq t_i, s_j \neq t_j, s_i \notin s_j, t_i \in t_j}$$

Case 2. We suppose that $B(x_1, \dots, x_k) \equiv \exists y \in x_i\, F(y, x_1, \dots, x_k)$ for some $1 \leq i \leq k$. Then $B(s_1, \dots, s_k)^{\mathbb{L}_{\Omega_n}(X)} \equiv \exists y \in s_i\ F(y, s_1, \dots, s_k)$.
All the premises in $\mathcal{C}(B(s_1, \dots, s_n)^{\mathbb{L}_{\Omega_n}(X)})$ are of the form $F(r, s_1, \dots, s_k)$ for $|r| < |s_i|$. By the induction hypothesis, we have

$$\Vdash s_1 \overline{\neq} t_1, \dots, s_k \overline{\neq} t_k, \neg F(r, s_1, \dots, s_k), F(r, t_1, \dots, t_k)$$

for all $|r| < |s_i|$. By Lemma 5.6, we get

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, r \dot{\in} s_i \to \neg F(r, s_1, \ldots, s_k), r \dot{\in} s_i \wedge F(r, t_1, \ldots, t_k).$$

for all $|r| < |s_i|$. We apply ($b\forall$) and we obtain

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \forall y \in s_i \ F(y, s_1, \ldots, s_k), F(r, t_1, \ldots, t_k)$$

for $|r| < |s_i|$. An application of ($b\exists$) yields

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \forall y \in s_i \ F(y, s_1, \ldots, s_k), \exists y \in t_i F(y, t_1, \ldots, t_k).$$

Case 3. We suppose $B(x_1, \ldots, x_k) \equiv F_1(x_1, \ldots, x_k) \vee F_2(x_1, \ldots, x_k)$. This means that $B(s_1, \ldots, s_k)^{\mathbb{L}_{\Omega_n}(X)} \equiv F_1(s_1, \ldots, s_k)^{\mathbb{L}_{\Omega_n}(X)} \vee F_2(s_1, \ldots, s_k)^{\mathbb{L}_{\Omega_n}(X)}$. By the induction hypothesis, we get

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \neg F_1(s_1, \ldots, s_k), F_1(t_1, \ldots, t_k) \tag{3}$$

and

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \neg F_2(s_1, \ldots, s_k), F_2(t_1, \ldots, t_k). \tag{4}$$

Using ($\vee$) on (2) and (3) gives

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \neg F_1(s_1, \ldots, s_k), F_1(t_1, \ldots, t_k) \vee F_2(t_1, \ldots, t_k)$$

and

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \neg F_2(s_1, \ldots, s_k), F_1(t_1, \ldots, t_k) \vee F_2(t_1, \ldots, t_k).$$

Finally, we apply ($\wedge$) to get

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \neg F_1(s_1, \ldots, s_k) \wedge \neg F_2(s_1, \ldots, s_k), F_1(t_1, \ldots, t_k) \vee F_2(t_1, \ldots, t_k),$$

which is equivalent to

$$\Vdash s_1 \overline{\neq} t_1, \ldots, s_k \overline{\neq} t_k, \neg (F_1(s_1, \ldots, s_k) \vee F_2(s_1, \ldots, s_k)), F_1(t_1, \ldots, t_k) \vee F_2(t_1, \ldots, t_k).$$

All the other cases are dual to the ones already shown. □

We show now that the Set Induction axiom can be embedded into the $\mathsf{RS}_l(X)$-system.

Lemma 5.8 (Set Induction)**.** *Let $n \leq \omega$. For any KPl-formula A, we have*

$$\Vdash^{\omega^{\mathrm{rk}(B)}} B \to \forall x \in \mathbb{L}_{\Omega_n}(X) \ A(x),$$

where $B \equiv \forall x \in \mathbb{L}_{\Omega_n}(X)(\forall x \in y \ A(y)^{\mathbb{L}_{\Omega_n}(X)} \to A(x)^{\mathbb{L}_{\Omega_n}(X)})$.

Proof. We will use the following claim.

Claim 5.8.1. *For any term s with $|s| < \Omega_n$ and any operator $\mathcal{H}$:*

$$\mathcal{H}[B,s] \left|\frac{\omega^{\mathrm{rk}(B)} \# \omega^{|s|+1}}{0}\right. \neg B, A(s)^{\mathbb{L}_{\Omega_n}(X)}. \tag{1}$$

Having Claim 5.8.1, we apply $(b\forall)$ to get

$$\mathcal{H}[B] \left|\frac{\omega^{\mathrm{rk}(B)} \# \Omega_n}{0}\right. \neg B, \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)}$$

for any operator $\mathcal{H}$. We apply $(\vee)$ twice (one time to $\neg B$ and the other to $\forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)}$) to obtain

$$\mathcal{H}[B] \left|\frac{\omega^{\mathrm{rk}(B)} \# \Omega_n + 2}{0}\right. \neg B \vee \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)}, \neg B \vee \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)}.$$

Now, removing one repeated formula and writing the formula as an implication, this is the same as

$$\mathcal{H}[B] \left|\frac{\omega^{\mathrm{rk}(B)} \# \Omega_n + 2}{0}\right. B \to \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)}.$$

But $\mathrm{rk}(\mathbb{L}_{\Omega_n}(X)) = \Omega_n$, and so obviously

$$\mathrm{rk}(B \to \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)})) > \Omega_n + 1.$$

This yields

$$\#(B \to \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)})) \geq \omega^{\Omega_n + 1}.$$

Hence, by Lemma 4.15, we obtain

$$\mathcal{H}[B] \left|\frac{\#(B \to \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)})) \# \omega^{\mathrm{rk}(B)}}{0}\right. B \to \forall x \in \mathbb{L}_{\Omega_n}(X) A(x)^{\mathbb{L}_{\Omega_n}(X)}.$$

Since this holds for any operator $\mathcal{H}$, we can conclude

$$\left\|\frac{\omega^{\mathrm{rk}(B)}}{}\right. B \to \forall x \in \mathbb{L}_{\Omega_n}(X)\ A(x).$$

We just need to prove Claim 5.8.1. So let $\mathcal{H}$ be any operator. Recall that we need to prove that

$$\mathcal{H}[B,s] \left|\frac{\omega^{\mathrm{rk}(B)} \# \omega^{|s|+1}}{0}\right. \neg B, A(s)^{\mathbb{L}_{\Omega_n}(X)}.$$

for any term s with $|s| < \Omega_\omega$. We proceed by induction on $|s|$. By the induction hypothesis, we get

$$\mathcal{H}[B,r] \left|\frac{\omega^{\mathrm{rk}(B)} \# \omega^{|r|+1}}{0}\right. \neg B, A(r)^{\mathbb{L}_{\Omega_n}(X)}. \tag{2}$$

From this we can obtain

$$\mathcal{H}[B,r,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|r|+1}+1}{\underset{0}{\vdash}} \neg B, r\dot{\in}s \to A(r)^{\mathbb{L}_{\Omega_n}(X)}, \tag{3}$$

no matter the form of s. Indeed,

- If $s \equiv \overline{u}$, we apply ($\vee$) to (2) to get

$$\mathcal{H}[B,r,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|r|+1}+1}{\underset{0}{\vdash}} \neg B, \neg r \in \overline{u} \vee A(r)^{\mathbb{L}_{\Omega_n}(X)},$$

  which, by Definition 4.4, is exactly

$$\mathcal{H}[B,r] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|r|+1}}{\underset{0}{\vdash}} \neg B, r\dot{\in}\overline{u} \to A(r)^{\mathbb{L}_{\Omega_n}(X)}.$$

- If $s \equiv \mathbb{L}_\alpha(X)$, then $r\dot{\in}\mathbb{L}_\alpha(X) \to A(r)^{\mathbb{L}_{\Omega_n}(X)} \equiv A(r)^{\mathbb{L}_{\Omega_\omega}(X)}$, and so the derivation (2) is exactly

$$\mathcal{H}[B,r] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|r|+1}}{\underset{0}{\vdash}} \neg B, r\dot{\in}\mathbb{L}_\alpha(X) \to A(r)^{\mathbb{L}_{\Omega_n}(X)}.$$

  By Lemma 4.15, we can add 1 to the ordinal bound to obtain the desired result.

- If $r \equiv [x \in \mathbb{L}_\alpha(X) : C(x)]$, then we apply ($\vee$) to the derivation (2) to get

$$\mathcal{H}[B,r] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|r|+1}+1}{\underset{0}{\vdash}} \neg B, \neg C(r) \vee A(r)^{\mathbb{L}_{\Omega_n}(X)},$$

  which, by Definition 4.4, is exactly the same as

$$\mathcal{H}[B,r] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|r|+1}+1}{\underset{0}{\vdash}} \neg B, r\dot{\in}s \to A(r)^{\mathbb{L}_{\Omega_n}(X)}.$$

  On the other hand, by Lemma 5.4 we have

$$\mathcal{H}[A,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|s|}+2}{\underset{0}{\vdash}} \neg A(s)^{\mathbb{L}_{\Omega_n}(X)}. \tag{4}$$

So, starting with (3) and (4), we have the following derivations for all $|r| < |s|$:

$$(b\forall)\ \dfrac{\mathcal{H}[B,r,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|r|+1}+1}{\underset{0}{\vdash}} \neg B, r\dot{\in}s \to A(r)^{\mathbb{L}_{\Omega_n}(X)}}{\mathcal{H}[B,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|s|}+2}{\underset{0}{\vdash}} \neg B, \forall y \in s\ A(y)^{\mathbb{L}_{\Omega_n}(X)}}$$

$$(\wedge)\ \dfrac{\mathcal{H}[B,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|s|}+2}{\underset{0}{\vdash}} \neg B, \forall y \in s\ A(y)^{\mathbb{L}_{\Omega_n}(X)} \qquad \mathcal{H}[A,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|s|}+2}{\underset{0}{\vdash}} \neg A(s)^{\mathbb{L}_{\Omega_n}(X)}}{\mathcal{H}[A,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|s|}+3}{\underset{0}{\vdash}} \neg B, \forall y \in s\ A(y)^{\mathbb{L}_{\Omega_n}(X)} \wedge \neg A(s)^{\mathbb{L}_{\Omega_n}(X)}, A(s)^{\mathbb{L}_{\Omega_n}(X)}}$$

$$(b\exists)\ \dfrac{\mathcal{H}[A,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|s|}+3}{\underset{0}{\vdash}} \neg B, \forall y \in s\ A(y)^{\mathbb{L}_{\Omega_n}(X)} \wedge \neg A(s)^{\mathbb{L}_{\Omega_n}(X)}, A(s)^{\mathbb{L}_{\Omega_n}(X)}}{\mathcal{H}[A,s] \overset{\omega^{\mathrm{rk}(B)}\#\omega^{|s|}+4}{\underset{0}{\vdash}} \neg B, \exists x \in \mathbb{L}_{\Omega_n}(X)[\forall y \in x\ A(x)^{\mathbb{L}_{\Omega_n}(X)} \wedge \neg A(x)^{\mathbb{L}_{\Omega_n}(X)}], A(s)^{\mathbb{L}_{\Omega_n}(X)}}$$

We notice that the second formula displayed in the last line is exactly $\neg B$. Moreover, we observe that $\omega^{\mathrm{rk}(B)}\#\omega^{|s|} + 4 < \omega^{\mathrm{rk}(B)}\#\omega^{|s|+1}$. Hence, we can remove one of the repeated formulas and also use Lemma 4.15 to obtain

$$\mathcal{H}[A, s] \left|\frac{\omega^{\mathrm{rk}(B)}\#\omega^{|s|+1}}{0}\right. \neg B, A(s)^{\mathbb{L}_{\Omega_n}(X)}.$$

and Claim 5.8.1 is verified. □

We continue with the Δ_0-Separation axiom schema.

Lemma 5.9 (Δ_0-Separation). *Let $n \leq \omega$. Let $A(x, y_1, \ldots, y_k)$ be a* KPI *Δ_0-formula. Let $s, t_1, \ldots, t_k$ be terms such that $|s|, |t_1|, \ldots, |t_k| < \Omega_n$. We will use the abbreviation $\vec{t} = t_0, \ldots, t_k$. Then,*

$$\Vdash \exists y \in \mathbb{L}_{\Omega_n}(X)[\forall x \in y(x \in s \wedge A(x, \vec{t})^{\mathbb{L}_{\Omega_n}(X)}) \wedge \forall x \in s(A(x, \vec{t})^{\mathbb{L}_{\Omega_n}(X)} \rightarrow x \in y)].$$

Proof. Before all else, we observe that A is a Δ_0-formula, which means that A does not have any unrestricted quantifier. This means that, since $|t_1|, \ldots, |t_n| < \Omega_n$, we have $A(\cdot, \vec{t})^{\mathbb{L}_{\Omega_\omega}(X)} \equiv A(\cdot, \vec{t})$.
Let $\alpha = \max\{|s|, |t_1|, \ldots, |t_n|\} + 1 < \Omega_n$. We define β. We define the term t as follows

$$t = [z \in \mathbb{L}_\beta(X) : z \in s \wedge A(z, \vec{t})].$$

First, we show that

$$\Vdash \forall x \in t(x \in s \wedge A(x, \vec{t})). \tag{1}$$

The following derivation gives this result (the first line follows Lemma 5.4):

$$\begin{array}{rl}
\text{Lemma 5.3(2.)} & \dfrac{\Vdash \neg(r \in s \wedge A(r, \vec{t})), r \in s \wedge A(r, \vec{t}) \text{ for all } |r| < \Omega_n}{\Vdash \neg(r \in s \wedge A(r, \vec{t})) \vee (r \in s \wedge A(r, \vec{t}))} \\
\equiv & \dfrac{}{\Vdash r \in s \wedge A(r, \vec{t}) \rightarrow (r \in s \wedge A(r, \vec{t}))} \\
\text{Definition 4.4} & \dfrac{}{\Vdash r \dot{\in} t \rightarrow (r \in s \wedge A(r, \vec{t}))} \\
(b\forall) & \dfrac{}{\Vdash \forall x \in t(x \in s \wedge A(x, \vec{t}))}
\end{array}$$

On the other hand, we will show that, for any $|r| < |s|$,

$$\Vdash \forall x \in s(A(x, \vec{t})^{\mathbb{L}_{\Omega_\omega}(X)} \rightarrow x \in t) \tag{2}$$

by splitting cases based on the form of s.

Case 1. We suppose $s \equiv \overline{u}$. Then we have the following derivation for all $|r| < |\overline{u}|$

$$
\cfrac{
  (\wedge)\,\cfrac{
    \cfrac{\text{Lemma 5.4}}{\Vdash \neg r \in \overline{u}, r \in \overline{u}}
    \qquad
    \cfrac{\text{Lemma 5.4}}{\Vdash \neg A(r,\vec{t}), A(r,\vec{t})}
  }{\Vdash \neg r \in \overline{u}, \neg A(r,\vec{t}), r \in \overline{u} \wedge A(r,\vec{t})}
  \qquad
  \cfrac{\text{Lemma 5.5(4.)}}{\Vdash r = r}
}{\Vdash \neg r \in \overline{u}, \neg A(r,\vec{t}), (r \in \overline{u} \wedge A(r,\vec{t})) \wedge r = r}\;(\wedge)
$$

$$
\begin{array}{rl}
 & \Vdash \neg r \in \overline{u}, \neg A(r,\vec{t}), (r \in \overline{u} \wedge A(r,\vec{t})) \wedge r = r \\ \hline
\text{Definition 4.4} & \Vdash \neg r \in \overline{u}, \neg A(r,\vec{t}), r \dot{\in} t \wedge r = r \\ \hline
(\in) & \Vdash \neg r \in \overline{u}, \neg A(r,\vec{t}), r \in t \\ \hline
\text{Lemma 5.3(2.)} & \Vdash \neg r \in \overline{u}, \neg A(r,\vec{t}) \vee r \in t \\ \hline
\equiv & \Vdash \neg r \in \overline{u}, A(r,\vec{t}) \to r \in t \\ \hline
\text{Lemma 5.3(2.)} & \Vdash \neg r \in \overline{u} \vee (A(r,\vec{t}) \to r \in t) \\ \hline
\equiv & \Vdash r \dot{\in} \overline{u} \to (A(r,\vec{t}) \to r \in t) \\ \hline
(b\forall) & \Vdash \forall x \in s (A(x,\vec{t}) \to x \in t)
\end{array}
$$

Case 2. We suppose that $s \equiv \mathbb{L}_\gamma(X)$. We have the following derivation for all $|r| < |s|$:

$$
\cfrac{
  (\wedge)\,\cfrac{
    \cfrac{\text{Lemma 5.5(5.)}}{\Vdash r \in \overline{u}}
    \qquad
    \cfrac{\text{Lemma 5.4}}{\Vdash \neg A(r,\vec{t}), A(r,\vec{t})}
  }{\Vdash \neg A(r,\vec{t}), r \in s \wedge A(r,\vec{t})}
  \qquad
  \cfrac{\text{Lemma 5.5(4.)}}{\Vdash r = r}
}{\Vdash\Vdash \neg A(r,\vec{t}), (r \in s \wedge A(r,\vec{t})) \wedge r = r}\;(\wedge)
$$

$$
\begin{array}{rl}
 & \Vdash\Vdash \neg A(r,\vec{t}), (r \in s \wedge A(r,\vec{t})) \wedge r = r \\ \hline
\text{Definition 4.4} & \Vdash \neg A(r,\vec{t}), r \dot{\in} t \wedge r = r \\ \hline
(\in) & \Vdash \neg A(r,\vec{t}), r \in t \\ \hline
\text{Lemma 5.3(2.)} & \Vdash \neg A(r,\vec{t}) \vee r \in t \\ \hline
\equiv & \Vdash A(r,\vec{t}) \to r \in t \\ \hline
\text{Definition 4.4} & \Vdash r \dot{\in} s \to (A(r,\vec{t}) \to r \in t) \\ \hline
(b\forall) & \Vdash \forall x \in s (A(x,\vec{t}) \to x \in t)
\end{array}
$$

Case 3. We suppose that $s \equiv [y \in \mathbb{L}_\gamma(X) : B(x)]$. We have the following derivation for all $|r| < |s|$.

$$
(\wedge)\ \dfrac{\dfrac{\text{Lemma 5.4}}{\Vdash B(r), \neg B(r)} \qquad \dfrac{\text{Lemma 5.5(4.)}}{\Vdash r = r}}{\Vdash \neg B(r), B(r) \wedge r = r}
$$

$$
(\wedge)\ \dfrac{(\in)\ \dfrac{\Vdash \neg B(r), B(r) \wedge r = r}{\Vdash \neg B(r), r \in s} \qquad \dfrac{\text{Lemma 5.4}}{\Vdash \neg A(r, \vec{t}), A(r, \vec{t})}}{\Vdash \neg B(r), \neg A(r, \vec{t}), r \in s \wedge A(r, \vec{t})}
$$

$$
(\wedge)\ \dfrac{\Vdash \neg B(r), \neg A(r, \vec{t}), r \in s \wedge A(r, \vec{t}) \qquad \dfrac{\text{Lemma 5.5(4.)}}{\Vdash r = r}}{\Vdash \neg B(r), \neg A(r, \vec{t}), (r \in s \wedge A(r, \vec{t})) \wedge r = r}
$$

$$
\begin{array}{rc}
\text{Definition 4.4} & \dfrac{\Vdash \neg B(r), \neg A(r, \vec{t}), (r \in s \wedge A(r, \vec{t})) \wedge r = r}{\Vdash \neg B(r), \neg A(r, \vec{t}), r \dot{\in} t \wedge r = r} \\
(\in) & \dfrac{}{\Vdash \neg B(r), \neg A(r, \vec{t}), r \in t} \\
\text{Lemma 5.3(2.)} & \dfrac{}{\Vdash \neg B(r), \neg A(r, \vec{t}) \vee r \in t} \\
\equiv & \dfrac{}{\Vdash \neg B(r), A(r, \vec{t}) \to r \in t} \\
\text{Lemma 5.3(2.)} & \dfrac{}{\Vdash \neg B(r) \vee (A(r, \vec{t}) \to r \in t)} \\
\equiv & \dfrac{}{\Vdash B(r) \to (A(r, \vec{t}) \to r \in t)} \\
\text{Definition 4.4} & \dfrac{}{\Vdash r \dot{\in} s \to (A(r, \vec{t}) \to r \in t)} \\
(b\forall) & \dfrac{}{\Vdash \forall x \in s(A(x, \vec{t}) \to x \in t)}
\end{array}
$$

Now, using (1) and (2) as a starting point, we have the following derivation.

$$
\begin{array}{rc}
(\wedge) & \dfrac{\Vdash \forall x \in t(x \in s \wedge A(x, \vec{t})) \qquad \Vdash \forall x \in s(A(x, \vec{t}) \to x \in t)}{\Vdash \forall x \in t(x \in s \wedge A(x, \vec{t})) \wedge \forall x \in s(A(x, \vec{t}) \to x \in t)} \\
\text{Definition 4.4} & \dfrac{}{\Vdash t \dot{\in} \mathbb{L}_{\Omega_n}(X) \wedge (\forall x \in t(x \in s \wedge A(x, \vec{t})) \wedge \forall x \in s(A(x, \vec{t}) \to x \in t))} \\
(b\exists) & \dfrac{}{\Vdash \exists y \in \mathbb{L}_{\Omega_n}(X)[\forall x \in y(x \in s \wedge A(x, \vec{t})) \wedge \forall x \in s(A(x, \vec{t}) \to x \in y)]}
\end{array}
$$

□

The proof of the next lemma can be found in [5]. To help the reader interested in those proofs, we give the correspondence between the results used in the proofs in Cook-Rathjen ([5]) and their version stated in this thesis.

Cook-Rathjen	**This thesis**
Abbreviations 3.5 ii)	Definition 4.4
Lemma 4.2 ii)	Lemma 5.3(2.)
Lemma 4.3 i)	Lemma 5.4
Lemma 4.3 iii)	Lemma 5.5(3.)
Lemma 4.3 vii)	Lemma 5.5(5.)

Lemma 5.10. *Here are the Infinity, Pairing and Union axioms embedded into the* $\mathsf{RS}_l(X)$*-system. Let* $n \leq \omega$*.*

1. $\Vdash \exists x \in \mathbb{L}_{\Omega_n}(X)[\exists z \in x(z \in x) \wedge \forall y \in x \exists z \in x(y \in z)]$.
2. *Let* s *and* t *be terms such that* $|s|, |t| < \Omega_n$*. Then*
$$\Vdash \exists z \in \mathbb{L}_{\Omega_n}(X)(s \in z \wedge t \in z).$$
3. *Let* s *be a term such that* $|s| < \Omega_n$*. Then*
$$\Vdash \exists z \in \mathbb{L}_{\Omega_n}(X)\forall y \in s \forall x \in y(x \in z).$$

Proof.
1. The proof is analogous to the proof of Lemma 4.6 of [5] (page 32).
2. The proof is analogous to the proof of Lemma 4.8 i) of [5] (page 34).
3. The proof is analogous to the proof of Lemma 4.8 ii) of [5] (page 34). □

Now, we focus on the axioms ruling the Ad predicate. To embed the first axiom, we need to know what $(Ad1)^{\mathbb{L}_{\Omega_\omega}(X)}$ will look like. In particular, we need to see what term ω translates to and what predicate $Tran$ translates to. We start with the latter.

Definition 5.11. *Let* t *be any term. We define the* $\mathsf{RS}_l(X)$*-predicate*

$$Tran(t) \equiv \forall x \in t \forall y \in x(y \in t).$$

We note that the KPl Δ_0-formula $Tran(x)$ with free variable x is translated to $Tran(t)$ in the $\mathsf{RS}_l(X)$-language, for some chosen term t. This is why we write both predicates the same way. Now, we notice that, in KPl, the ordinal ω is the first limit ordinal. This means that we can define ω as the unique ordinal containing infinitely many successor ordinals.

$$y = \omega \text{ iff } On(y) \wedge \forall z \in y[\exists z' \in y(z \in z') \wedge \exists u \in z \forall v \in z(u \neq z \rightarrow v \in u)].$$

Since the formula defining ω is Δ_0, the translation of the formula to the $\mathsf{RS}_l(X)$-system will be the same, and so the term

$$\underline{\omega} = [y \in \mathbb{L}_\omega(X) : \exists y \in x(On(y) \wedge \forall z \in y[(z = 0 \vee Succ(z)) \wedge \exists u \in y(z \in u)] \wedge Tran(x))]]$$

fully captures the set ω. This means that $(Ad1)^{\mathbb{L}_{\Omega_\omega}(X)}$ can be written as

$$\forall x \in \mathbb{L}_{\Omega_\omega}(X)[Ad(x) \to \underline{\omega} \in x \wedge Tran(x)].$$

We show that $Tran(\mathbb{L}_{\Omega_n}(X))$ holds for every natural number n.

Lemma 5.12. *Let $n < \omega$. Then,*

$$\Vdash Tran(\mathbb{L}_{\Omega_n}(X)).$$

Proof. Let t be any term such that $|t| < \Omega_n$. We show that for every term s with $|s| < |t|$

$$\Vdash s\dot{\in}t \to s \in \mathbb{L}_{\Omega_n}(X) \tag{1}$$

by fixing such an s and splitting cases based on the form of t. First, we observe that $|s| < \Omega_n$ and so, by Lemma 5.5(5.), we have

$$\Vdash s \in \mathbb{L}_{\Omega_n}(X). \tag{2}$$

Case 1. We suppose $t \equiv \overline{u}$. Then $s \equiv \overline{v}$. Therefore, using Lemma 4.15 on (2), we get

$$\Vdash \neg\overline{v} \in \overline{u}, \overline{v} \in \mathbb{L}_{\Omega_n}(X).$$

By Lemma 5.3(2.), we obtain

$$\Vdash \neg\overline{v} \in \overline{u} \vee \overline{v} \in \mathbb{L}_{\Omega_n}(X),$$

which is equivalent to

$$\Vdash \overline{v} \in \overline{u} \to \overline{v} \in \mathbb{L}_{\Omega_n}(X).$$

By Definition 4.4, this is

$$\Vdash \overline{v}\dot{\in}\overline{u} \to \overline{v} \in \mathbb{L}_{\Omega_n}(X).$$

Case 2. We suppose $t \equiv \mathbb{L}_\alpha(X)$. Then, by Definition 4.4 the derivation (2) is exactly

$$\Vdash s\dot{\in}t \to s \in \mathbb{L}_{\Omega_n}(X).$$

Case 3. We suppose $t \equiv [x \in \mathbb{L}_\alpha(X) : B(x)]$. Then, using Lemma 4.15 on (2) we get

$$\Vdash \neg B(s), s \in \mathbb{L}_{\Omega_n}(X).$$

By Lemma 5.3(2.), we obtain

$$\Vdash \neg B(s) \vee s \in \mathbb{L}_{\Omega_n}(X)$$

which is exactly

$$\Vdash B(s) \rightarrow s \in \mathbb{L}_{\Omega_n}(X).$$

This is equivalent, by Definition 4.4, to

$$\Vdash s \dot{\in} t \rightarrow s \in \mathbb{L}_{\Omega_n}(X).$$

We have shown (1) for any $|s| < |t|$. By an application of $(b\forall)$, we obtain

$$\Vdash \forall y \in t(y \in \mathbb{L}_{\Omega_n}(X)).$$

Again by Definition 4.4, this is the same as

$$\Vdash t \dot{\in} \mathbb{L}_{\Omega_n}(X) \rightarrow \forall y \in t(y \in \mathbb{L}_{\Omega_n}(X)).$$

Since this holds for any $|t| < \Omega_n$, another application of $(b\forall)$ yields

$$\Vdash \forall x \in \mathbb{L}_{\Omega_n}(X) \forall y \in x(y \in \mathbb{L}_{\Omega_n}(X)).$$

□

Similarly, we show that $\underline{\omega} \in \mathbb{L}_{\Omega_n}(X)$ holds for every $n < \omega$.

Lemma 5.13. *Let $n < \omega$. Then,*

$$\Vdash \underline{\omega} \in \mathbb{L}_{\Omega_n(X)}(X).$$

Proof. We observe that $|\underline{\omega}| = \Gamma_{\theta+1} + \omega$. Let $n < \omega$. Then, since $\Gamma_{\theta+1} + \omega < \Omega_n$, by Lemma 5.5(5.) we have

$$\Vdash \underline{\omega} \in L_{\Omega_n}(X).$$

□

We can now proceed to the proof of the embedding of $(Ad1)$ into $\mathsf{RS}_l(X)$.

Lemma 5.14 (Ad1)**.** *Let $\mathcal{H}$ be any operator. We have*

$$\mathcal{H} \left|\frac{\Omega_\omega+1}{\Omega_\omega}\right. \forall x \in \mathbb{L}_{\Omega_\omega}(X)(Ad(x) \rightarrow \underline{\omega} \in x \wedge Tran(x))$$

Proof. We need to derive

$$Ad(t) \rightarrow \underline{\omega} \in t \wedge Tran(t)$$

for every term t with $|t| < \Omega_\omega$ to be able to apply $(b\forall)$. So, fix a term t such that $|t| < \Omega_\omega$. In the following derivations, we fix a natural number n. By Lemmas 5.12 and 5.13, we have by an application of $(\wedge)$

$$\Vdash \underline{\omega} \in \mathbb{L}_{\Omega_n}(X) \wedge Tran(\mathbb{L}_{\Omega_n}(X)).$$

By means of Lemma 4.15, we get

$$\Vdash t \neq \mathbb{L}_{\Omega_n}(X), \underline{\omega} \in \mathbb{L}_{\Omega_n}(X) \wedge Tran(\mathbb{L}_{\Omega_n}(X)), \underline{\omega} \in t \wedge Tran(t). \tag{1}$$

On the other hand, by Lemma 5.7 we have

$$\Vdash t \neq \mathbb{L}_{\Omega_n}(X), \neg(\underline{\omega} \in \mathbb{L}_{\Omega_n}(X) \wedge Tran(\mathbb{L}_{\Omega_n}(X)), \underline{\omega} \in t \wedge Tran(t). \tag{2}$$

Fix an operator $\mathcal{H}$. An application of (Cut) to (1) and (2) gives

$$\mathcal{H}[t] \vdash^{\alpha}_{\Omega_{n+1}} t \neq \mathbb{L}_{\Omega_n}(X), \underline{\omega} \in t \wedge Tran(t),$$

where $\alpha = \#(t \neq \mathbb{L}_{\Omega_n}(X), \underline{\omega} \in t \wedge Tran(t)) < \Omega_\omega$. Since this derivation holds for every $n < \omega$, we apply $(\neg Ad)$ to get

$$\mathcal{H}[t] \vdash^{\Omega_\omega}_{\Omega_\omega} \neg Ad(t), \underline{\omega} \in t \wedge Tran(t).$$

By Lemma 5.3(2.), we obtain

$$\mathcal{H}[t] \vdash^{\Omega_\omega}_{\Omega_\omega} \neg Ad(t) \vee \underline{\omega} \in t \wedge Tran(t),$$

which is the same as

$$\mathcal{H}[t] \vdash^{\Omega_\omega}_{\Omega_\omega} Ad(t) \rightarrow \underline{\omega} \in t \wedge Tran(t).$$

This holds for any $|t| < \Omega_\omega$. An application of $(b\forall)$ yields

$$\mathcal{H} \vdash^{\Omega_\omega + 1}_{\Omega_\omega} \forall x \in \mathbb{L}_{\Omega_\omega}(X)(Ad(x) \rightarrow \underline{\omega} \in x \wedge Tran(x)).$$

□

We continue with the second axiom about the Ad predicate.

Lemma 5.15 (Ad2)**.** *Let $\mathcal{H}$ be any operator. We have*

$$\mathcal{H}\left|\frac{\Omega_\omega+3}{\Omega_\omega}\right. \forall x \in \mathbb{L}_{\Omega_\omega}(X)\forall y \in \mathbb{L}_{\Omega_\omega}(X)(Ad(x) \land Ad(y) \to x \in y \lor x = y \lor y \in x)$$

Proof. We need to derive

$$Ad(s) \land Ad(t) \to s \in t \lor s = t \lor t \in s$$

for all the terms s, t with $|s|, |t| < \Omega_\omega$ in order to apply $(b\forall)$ twice. In the following derivations, we fix two natural numbers n and m.
If $n = m$, then by Lemma 5.5(4.) we have

$$\Vdash \mathbb{L}_{\Omega_n}(X) = \mathbb{L}_{\Omega_m}(X).$$

If $n < m$, by Lemma 5.5(5.) we have

$$\Vdash \mathbb{L}_{\Omega_n}(X) \in \mathbb{L}_{\Omega_m}(X).$$

If $m < n$, by Lemma 5.5(5.) we have

$$\Vdash \mathbb{L}_{\Omega_m}(X) \in \mathbb{L}_{\Omega_n}(X).$$

In any case, two applications of $(\lor)$ give

$$\Vdash \mathbb{L}_{\Omega_n}(X) \in \mathbb{L}_{\Omega_m}(X) \lor \mathbb{L}_{\Omega_n}(X) = \mathbb{L}_{\Omega_m}(X) \lor \mathbb{L}_{\Omega_m}(X) \in \mathbb{L}_{\Omega_n}(X).$$

We use Lemma 4.15 to get

$$\begin{aligned}&\Vdash s \neq \mathbb{L}_{\Omega_n}(X), t \neq \mathbb{L}_{\Omega_m}(X), \mathbb{L}_{\Omega_n}(X) \in \mathbb{L}_{\Omega_m}(X) \lor \mathbb{L}_{\Omega_n}(X) = \mathbb{L}_{\Omega_m}(X) \lor \mathbb{L}_{\Omega_m}(X) \in \mathbb{L}_{\Omega_n}(X),\\ &\quad s \in t \lor s = t \lor t \in s.\end{aligned} \tag{1}$$

On the other hand, by Lemma 5.7 we have

$$\begin{aligned}&\Vdash s \neq \mathbb{L}_{\Omega_n}(X), t \neq \mathbb{L}_{\Omega_m}(X), \neg(\mathbb{L}_{\Omega_n}(X) \in \mathbb{L}_{\Omega_m}(X) \lor \mathbb{L}_{\Omega_n}(X) = \mathbb{L}_{\Omega_m}(X) \lor \mathbb{L}_{\Omega_m}(X) \in \mathbb{L}_{\Omega_n}(X)),\\ &\quad s \in t \lor s = t \lor t \in s.\end{aligned} \tag{2}$$

We fix an operator $\mathcal{H}$. An application of (Cut) on (1) and (2) yields

$$\mathcal{H}[s,t]\left|\frac{\alpha}{\Omega_{n+m}}\right. s \neq \mathbb{L}_{\Omega_n}(X), t \neq \mathbb{L}_{\Omega_m}(X), s \in t \lor s = t \lor t \in s,$$

where $\alpha = \#(s \neq \mathbb{L}_{\Omega_n}(X), t \neq \mathbb{L}_{\Omega_m}(X), s \in t \vee s = t \vee t \in s)$. Since this derivation holds for any $n, m < \omega$, we use $(\neg Ad)$ twice and we obtain

$$\mathcal{H}[s,t] \left|\frac{\Omega_\omega+1}{\Omega_\omega}\right. \neg Ad(s), \neg Ad(t), s \in t \vee s = t \vee t \in s.$$

We use Lemma 5.3(2.) to have

$$\mathcal{H}[s,t] \left|\frac{\Omega_\omega+1}{\Omega_\omega}\right. \neg Ad(s) \vee \neg Ad(t), s \in t \vee s = t \vee t \in s,$$

which is the same as

$$\mathcal{H}[s,t] \left|\frac{\Omega_\omega+1}{\Omega_\omega}\right. \neg(Ad(s) \wedge Ad(t)), s \in t \vee s = t \vee t \in s$$

Again by Lemma 5.3(2.), we get

$$\mathcal{H}[s,t] \left|\frac{\Omega_\omega+1}{\Omega_\omega}\right. \neg(Ad(s) \wedge Ad(t)) \vee s \in t \vee s = t \vee t \in s,$$

which is equivalent to

$$\mathcal{H}[s,t] \left|\frac{\Omega_\omega+1}{\Omega_\omega}\right. Ad(s) \wedge Ad(t) \to s \in t \vee s = t \vee t \in s$$

Therefore, since this last derivation holds for any t, s with $|t|, |s| < \Omega_\omega$, we can apply $(b\forall)$ twice to obtain

$$\mathcal{H} \left|\frac{\Omega_\omega+3}{\Omega_\omega}\right. \forall x \in \mathbb{L}_{\Omega_\omega}(X) \forall y \in \mathbb{L}_{\Omega_\omega}(X)(Ad(x) \wedge Ad(y) \to x \in y \vee x = y \vee y \in x).$$

□

We keep up with the third axiom about the Ad predicate. We already know that $\mathsf{RS}_l(X)$ shows that any X-admissible ordinal satisfies all the axioms appearing in $(Ad3)$ but Δ_0-Collection. So we first focus on Δ_0-Collection.

Lemma 5.16. *Let $A(x,y)$ be any Δ_0-formula of* KPI *with free variables x and y. Let $n < \omega$. Let s be a term such that $|s| < \Omega_n$. Then,*

$$\Vdash \forall x \in s \exists y \in \mathbb{L}_{\Omega_n}(X)\ A(x,y) \to \exists z \in \mathbb{L}_{\Omega_n}(X) \forall x \in s \exists y \in z\ A(x,y).$$

Proof. We fix an operator $\mathcal{H}$. To simplify formulas, we write $B \equiv \forall x \in s \exists y \in (X)\ A(x,y)$. This means that we will show

$$\Vdash B^{\mathbb{L}_{\Omega_n}(X)} \to \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z.$$

First, by Lemma 5.4 we have

$$\mathcal{H}[B^{\mathbb{L}_{\Omega_n}(X)}] \left|\frac{\omega^{\mathrm{rk}(B^{\mathbb{L}_{\Omega_n}(X)})} \# \omega^{\mathrm{rk}(\neg B^{\mathbb{L}_{\Omega_n}(X)})}}{0}\right. B^{\mathbb{L}_{\Omega_n}(X)}, \neg B^{\mathbb{L}_{\Omega_n}(X)}.$$

We apply the (Ref_n) rule to get

$$\mathcal{H}[B^{\mathbb{L}_{\Omega_n}(X)}] \left|\frac{\alpha}{0}\right. \exists z \in \mathbb{L}_{\Omega_n}(X) B^z, \neg B^{\mathbb{L}_{\Omega_n}(X)}$$

with $\alpha = \omega^{\mathrm{rk}(B^{\mathbb{L}_{\Omega_n}(X)})} \# \omega^{\mathrm{rk}(\neg B^{\mathbb{L}_{\Omega_n}(X)})} + 1$. Since $\mathrm{rk}(\mathbb{L}_{\Omega_n}(X)) = \Omega_n$, we have $\alpha > \Omega_n$ and we are able to use the rule with this ordinal bound.
We use ($\vee$) twice, the first time applying the rule to the first formula and the second time applying the rule to the second formula in order to obtain the same formula twice, as displayed here after removing one repeated formula:

$$\mathcal{H}[B^{\mathbb{L}_{\Omega_n}(X)}] \left|\frac{\alpha+2}{0}\right. \neg B^{\mathbb{L}_{\Omega_n}(X)} \vee \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z.$$

But since ω^β is additive principal for any ordinal β, we have

$$\alpha+2 = \omega^{\mathrm{rk}(B^{\mathbb{L}_{\Omega_n}(X)}) \# \omega^{\mathrm{rk}(\neg B^{\mathbb{L}_{\Omega_n}(X)})}} + 3 < \omega^{\mathrm{rk}(B^{\mathbb{L}_{\Omega_n}(X)})+1} = \#(B^{\mathbb{L}_{\Omega_n}(X)} \to \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z).$$

Therefore, by Lemma 4.15

$$\mathcal{H}[\neg B^{\mathbb{L}_{\Omega_n}(X)} \vee \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z] \left|\frac{\#(B^{\mathbb{L}_{\Omega_n}(X)} \to \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z)}{0}\right. \neg B^{\mathbb{L}_{\Omega_n}(X)} \vee \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z.$$

which is exactly the same as

$$\mathcal{H}[B^{\mathbb{L}_{\Omega_n}(X)} \to \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z] \left|\frac{\#(B^{\mathbb{L}_{\Omega_n}(X)} \to \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z)}{0}\right. B^{\mathbb{L}_{\Omega_n}(X)} \to \exists z \in \mathbb{L}_{\Omega_n}(X)\ B^z.$$

□

We have everything we needed to embed ($Ad3$).

Lemma 5.17 (Ad3). $\mathcal{H} \left|\frac{\Omega_\omega+1}{\Omega_\omega}\right. \forall x \in \mathbb{L}_{\Omega_\omega}(X)(Ad(x) \to (Pair)^x \wedge (Union)^x \wedge (\Delta_0 - Separation)^x \wedge (\Delta_0 - Collection)^x)$.

Proof. We need to derive

$$Ad(t) \to (Pair)^t \wedge (Union)^t \wedge (\Delta_0 - Separation)^t \wedge (\Delta_0 - Collection)^t)$$

for every t with $|t| < \Omega_\omega$ in order to apply $(b\forall)$ and obtain the desired derivation. So let t be any term with $|t| < \Omega_\omega$. Fix a natural number n. By Lemmas 5.10, 5.16 and 5.9, we have

$$\Vdash (\mathsf{Ax})^{\mathbb{L}_{\Omega_n}(X)}$$

for every axiom Ax among Pair, Union, Δ_0-Collection and Δ_0-Separation. We will use the following abbreviation: given any term (or variable) s, we will write $\bigwedge(\mathsf{Axioms})^s$ to denote

$$(Pair)^s \wedge (Union)^s \wedge (\Delta_0 - Separation)^s \wedge (\Delta_0 - Collection)^s$$

By three application of $(\wedge)$ starting on (1), we get

$$\Vdash \bigwedge(\mathsf{Axioms})^{\mathbb{L}_{\Omega_n}(X)}.$$

Lemma 4.15 gives

$$\Vdash t \neq \mathbb{L}_{\Omega_n}(X), \bigwedge(\mathsf{Axioms})^t, \bigwedge(\mathsf{Axioms})^{\mathbb{L}_{\Omega_n}(X)}. \tag{1}$$

On the other hand, by Lemma 5.7 we have

$$\Vdash t \neq \mathbb{L}_{\Omega_n}(X), \bigwedge(\mathsf{Axioms})^t, \neg\bigwedge(\mathsf{Axioms})^{\mathbb{L}_{\Omega_n}(X)}. \tag{2}$$

Therefore, for any operator $\mathcal{H}$, an application of (Cut) on (2) and (3) yields

$$\mathcal{H}[t] \,\Big|\frac{\alpha}{\Omega_{n+1}}\; t \neq \mathbb{L}_{\Omega_n}(X), \bigwedge(\mathsf{Axioms})^t,$$

where $\alpha = \#(t \neq \mathbb{L}_{\Omega_n}(X), \bigwedge(\mathsf{Axioms})^t)$. Since this derivation holds for any $n < \omega$, an application of $(\neg Ad)$ yields

$$\mathcal{H}[t] \,\Big|\frac{\Omega_\omega}{\Omega_\omega}\; \neg Ad(t), \bigwedge(\mathsf{Axioms})^t.$$

By Lemma 5.3(2.), we get for any operator $\mathcal{H}$

$$\mathcal{H}[t] \,\Big|\frac{\Omega_\omega}{\Omega_\omega}\; \neg Ad(t) \vee \bigwedge(\mathsf{Axioms})^t,$$

which is equivalent to

$$\mathcal{H}[t] \,\Big|\frac{\Omega_\omega}{\Omega_\omega}\; Ad(t) \rightarrow \bigwedge(\mathsf{Axioms})^t.$$

Therefore, for any operator $\mathcal{H}$, we obtain by an application of $(b\forall)$:

$$\mathcal{H} \,\Big|\frac{\Omega_\omega+1}{\Omega_\omega}\; \forall x \in \mathbb{L}_{\Omega_\omega}(X)(Ad(x) \rightarrow (Pair)^x \wedge (Union)^x \wedge (\Delta_0 - Separation)^x \wedge (\Delta_0 - Collection)^x).$$

$\square$

Finally, we embed the limit axiom. We state a preliminary Lemma that we will use only twice in this thesis. We will use it to embed (Lim) into $\mathsf{RS}_l(X)$ and in the proof of the main Theorem 6.3. This result is very natural, as it states that all the $\mathbb{L}_{\Omega_n}(X)$ for $n < \omega$ are admissible.

Lemma 5.18. *For any natural number n we have*

$$\Vdash Ad(\mathbb{L}_{\Omega_n}(X)).$$

Proof. Let $n < \omega$. By Lemma 5.5(4.), we have $\Vdash \mathbb{L}_{\Omega_n}(X) = \mathbb{L}_{\Omega_n}(X)$. We apply (Ad) to obtain the result. □

Lemma 5.19 (Lim)**.** *Let $\mathcal{H}$ be any operator. Then*

$$\mathcal{H} \left|\frac{\Omega_\omega \cdot \omega^2}{0}\right. \forall x \in \mathbb{L}_{\Omega_\omega}(X) \exists y \in \mathbb{L}_{\Omega_\omega}(X)(Ad(y) \wedge x \in y).$$

Proof. Let s be a term such that $|s| < \Omega_\omega$. Then there is $n < \omega$ such that $|s| < \Omega_n$. So let $n := min(m < \omega : |s| < \Omega_n)$. It follows that $\Vdash s \in \mathbb{L}_{\Omega_n}(X)$ by Lemma 5.5(5.). On the other hand, by Lemma 5.18 we have

$$\Vdash Ad(\mathbb{L}_{\Omega_n}(X)). \tag{1}$$

We have the following derivation for any $|s| < \Omega_\omega$:

$$\dfrac{\dfrac{\dfrac{\Vdash Ad(\mathbb{L}_{\Omega_n}(X)) \qquad \Vdash s \in \mathbb{L}_{\Omega_n}(X)}{\Vdash Ad(\mathbb{L}_{\Omega_n}(X)) \wedge s \in \mathbb{L}_{\Omega_n}(X)}(\wedge)}{\Vdash \exists y \in \mathbb{L}_{\Omega_\omega}(X)(Ad(y) \wedge s \in y)}(b\exists)}{\Vdash \forall x \in \mathbb{L}_{\Omega_\omega}(X) \exists y \in \mathbb{L}_{\Omega_\omega}(X)(Ad(y) \wedge x \in y)}(b\forall)$$

Since $\#((Lim)^{\mathbb{L}_{\Omega_\omega}(X)}) = \omega^{\mathrm{rk}(\mathbb{L}_{\Omega_\omega}(X))} = \omega^{\Omega_\omega + 2} = \omega^{\Omega_\omega} \cdot \omega^2 = \Omega_\omega \cdot \omega^2$, we finally obtain that for any operator $\mathcal{H}$

$$\mathcal{H} \left|\frac{\Omega_\omega \cdot \omega^2}{0}\right. (Lim)^{\mathbb{L}_{\Omega_\omega}(X)}.$$

□

We have successfully embedded all of the axioms of KPl into the $\mathsf{RS}_l(X)$-system. It is now time to state and show the full embedding theorem.

Theorem 5.20. *Let $\Gamma(a_1, \ldots, a_n)$ be a finite set of formulas with all the free variables displayed such that $\mathsf{KPl} \vdash \Gamma(a_1, \ldots, a_n)$. Then, there is some $m < \omega$ such that for any operator $\mathcal{H}$ and any terms $s_1, \ldots, s_n$ we have*

$$\mathcal{H}[s_1,\dots,s_n] \mathrel{\Big|}\!\frac{\Omega_\omega\cdot\omega^m}{\Omega_\omega+m}\; \Gamma(s_1,\dots,s_n)^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Proof. We proceed by induction on the KPl proof.
If $\Gamma(a_1,\dots,a_n)$ is an axiom of KPl, then by one of the Lemmas 5.7, 5.8, 5.10, 5.14, 5.15, 5.17 and 5.19, we obtain the result.
Now, we assume that $\Gamma(a_1,\dots,a_n)$ is obtained by a KPl rule (R). So, if $\Delta(a_1,\dots,a_n)$ is the premise, or one of the premises, of this inference, then KPl proves $\Delta(a_1,\dots,a_n)$. Therefore, by the induction hypothesis, we have

$$\mathcal{H}[s_1,\dots,s_n] \mathrel{\Big|}\!\frac{\Omega_\omega\cdot\omega^m}{\Omega_\omega+m}\; \Delta(s_1,\dots,s_n)^{\mathbb{L}_{\Omega_\omega}(X)}.$$

for any operator $\mathcal{H}$ and any terms $s_1,\dots,s_n$ of level below Ω_ω. This is the way we will reason for each case (cases correspond to KPl rules).
So, we fix an arbitrary operator $\mathcal{H}$ and arbitrary terms $s_1,\dots,s_n$ of level less than Ω_ω. To simply notation, we will write $\vec{a}=a_1,\dots,a_n$ and $\vec{s}=s_1,\dots,s_n$ (even though $\vec{a}$ is a vector of variables of KPl and $\vec{s}$ is a vector of $\mathsf{RS}_l(X)$-terms). As a reminder, when we write $A(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}$ we are meaning the formula A replacing free variables by the terms in $\vec{s}$ and bounding all unrestricted quantifiers to $\mathbb{L}_{\Omega_\omega}(X)$.

Case 1. We suppose that the last KPl rule applied is $(\wedge)$. This means that

$$\Gamma(\vec{a}) = \Gamma'(\vec{a}), A(\vec{a}) \wedge B(\vec{a}),$$

for some KPl formulas A and B. Therefore, we have

$$\mathsf{KPl} \vdash \Gamma'(\vec{a}), A(\vec{a}) \tag{1}$$

and

$$\mathsf{KPl} \vdash \Gamma'(\vec{a}), B(\vec{a}) \tag{2}$$

We apply the induction hypothesis to (1) to find some $m_0<\omega$ such that

$$\mathcal{H}[\vec{s}] \mathrel{\Big|}\!\frac{\Omega_\omega\cdot\omega^{m_0}}{\Omega_\omega+m_0}\; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, A(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}. \tag{3}$$

We also apply the induction hypothesis to (2) to find some $m_1<\omega$ such that

$$\mathcal{H}[\vec{s}] \mathrel{\Big|}\!\frac{\Omega_\omega\cdot\omega^{m_1}}{\Omega_\omega+m_1}\; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, B(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}. \tag{4}$$

We apply the $\mathsf{RS}_l(X)$ rule $(\wedge)$ to (3) and (4) to obtain

$$\mathcal{H}[\vec{s}] \mathrel{\Big|}\!\frac{\Omega_\omega\cdot\omega^{\max(m_0+m_1)+1}}{\Omega_\omega+\max(m_0,m_1)+1}\; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, A(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)} \wedge B(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)},$$

which is exactly

$$\mathcal{H}[\vec{s}] \,\big|\!\frac{\Omega_\omega \cdot \omega^{\max(m_0+m_1)+1}}{\Omega_\omega+\max(m_0,m_1)+1}\, \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, (A \wedge B)(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Case 2. We suppose that the last KPl rule applied is $(\vee)$. This means that

$$\Gamma(\vec{a}) = \Gamma'(\vec{a}), A(\vec{a}) \vee B(\vec{a}),$$

for some KPl formulas A and B. Therefore, we have

$$\mathsf{KPl} \vdash \Gamma'(\vec{a}), C(\vec{a}) \tag{5}$$

where $C \in \{A, B\}$. By the induction hypothesis, we can find some $m < \omega$ such that

$$\mathcal{H}[\vec{s}] \,\big|\!\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega+m}\, \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, C(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

We apply the $\mathsf{RS}_l(X)$ rule $(\vee)$ to obtain

$$\mathcal{H}[\vec{s}] \,\big|\!\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega+m}\, \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, A(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)} \vee B(\vec{s})^{\mathbb{L}_{\Omega_\omega(X)}},$$

which is exactly

$$\mathcal{H}[\vec{s}] \,\big|\!\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega+m}\, \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, (A \vee B)(\vec{s})^{\mathbb{L}_{\Omega_\omega(X)}}.$$

Case 3. We suppose that the last KPl rule applied is $(b\forall)$. This means that

$$\Gamma(\vec{a}) = \Gamma'(\vec{a}), \forall x \in a_i \; A(x, \vec{a})$$

for some KPl-formula A and some free variable a_i with $i \in \{1, \ldots, n\}$. Therefore, we have

$$\mathsf{KPl} \vdash \Gamma'(\vec{a}), b \in a_i \to A(b, \vec{a})$$

where $b \neq a_j$ for all $j \in \{1, \ldots, n\}$. By the induction hypothesis we can find some $m < \omega$ such that

$$\mathcal{H}[r, \vec{s}] \,\big|\!\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega+m}\, \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, r \in s_i \to A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \tag{6}$$

for all term r with $|r| < |s_i|$. We would like to apply $(b\forall)$ to end this case. Nonetheless, we need to have $r \dot{\in} s \to A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}$ instead of $r \in s \to A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}$ in the premise to be able to do this. We split subcases based on the form of s_i.
Subcase 3.1. We assume that $s_i \equiv \overline{u}$. Then, we have

$$r \in \overline{u} \to A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)} \equiv r \dot{\in} \overline{u} \to A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

We apply the $\mathsf{RS}_l(X)$-rule $(b\forall)$ to (6) to obtain

$$\mathcal{H}[\vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m+1}}{\Omega_\omega + m+1}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \forall x \in s_i\ A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)},$$

which is the same as

$$\mathcal{H}[\vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m+1}}{\Omega_\omega + m+1}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, (\forall x \in s_i\ A(x, \vec{s}))^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Subcase 3.2. We assume that $s_i \equiv \mathbb{L}_\alpha(X)$. We apply Lemma 4.18 to (6) to get

$$\mathcal{H}[r, \vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m}}{\Omega_\omega + m}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, r \notin s_i, A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}. \tag{7}$$

On the other hand, we have $|r| < |s_i|$. Therefore, we can use Lemma 5.5(5.) to get

$$\mathcal{H}[r, \vec{s}] \left|\frac{\omega^{\mathrm{rk}(r \in s)}}{0}\right. r \in s_i$$

and so, by Lemma 4.15 we can add more formulas and increase the bounds of the derivation to have

$$\mathcal{H}[r, \vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m}}{\Omega_\omega + m}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, r \in s_i, A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}. \tag{8}$$

Now, using (Cut) on (7) and (8) gives

$$\mathcal{H}[r, \vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m}+1}{\Omega_\omega + m}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

By Definition 4.4, this can be written as

$$\mathcal{H}[r, \vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m}+1}{\Omega_\omega + m}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, r \dot{\in} s_i \to A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Finally, we apply $(b\forall)$ and artificially increase the cut complexity bound to obtain

$$\mathcal{H}[\vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m+1}}{\Omega_\omega + m+1}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \forall x \in s_i\ A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Subcase 3.3. We assume that $s_i \equiv [x \in \mathbb{L}_\alpha(X)]$. We apply Lemma 4.18 to (6) to get

$$\mathcal{H}[r, \vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m}}{\Omega_\omega + m}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, r \notin s_i, A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Using Lemma 4.15, we obtain

$$\mathcal{H}[r, \vec{s}] \left|\frac{\Omega_\omega \cdot \omega^{m}}{\Omega_\omega + m}\right. \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, r \notin s_i, A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg B(r). \tag{9}$$

On the other hand, by Lemma 5.4 we have

$$\Vdash \neg B(r), B(r). \tag{10}$$

and by Lemma 5.5(4.) we have

$$\Vdash r = r. \tag{11}$$

Since $|r| < |s_i|$, we have the following derivation

$$\dfrac{\Vdash \neg B(r), B(r) \qquad r = r}{\dfrac{\Vdash \neg B(r), B(r) \wedge r = r}{\Vdash \neg B(r), r \in s_i}\,(\in)}\,(\wedge)$$

Using Lemma 4.15, we get

$$\mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg B(r), r \in s_i, A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}. \tag{12}$$

Applying (Cut) to (9) and (12) yields

$$\mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^m + 1}{\Omega_\omega + m} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg B(r), A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

We have the following derivation

$$\begin{array}{ll}
 & \mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^m + 1}{\Omega_\omega + m} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg B(r), A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)} \\
(\vee) & \overline{\mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^m + 2}{\Omega_\omega + m} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg B(r) \vee A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}} \\
(\vee) & \overline{\mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^m + 3}{\Omega_\omega + m} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg B(r) \vee A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg B(r) \vee A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}} \\
\equiv & \overline{\mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^m + 4}{\Omega_\omega + m} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, B(r) \to A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}} \\
(b\forall) & \overline{\mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^m + 5}{\Omega_\omega + m} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \forall r \in s_i \; A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}}
\end{array}$$

We use Lemma 4.15 to obtain the final bounds:

$$\mathcal{H}[r, \vec{s}] \Big|\frac{\Omega_\omega \cdot \omega^{m+5}}{\Omega_\omega + m + 5} \; \Gamma'(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \forall r \in s_i \; A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Case 4. We suppose that the last KPl rule applied is $(\forall)$. This means that

$$\Gamma(\vec{a}) = \Gamma'(\vec{a}), \forall x A(x, \vec{a})$$

for some KPl formula A. Therefore,

$$\mathsf{KPl} \vdash \Gamma'(\vec{a}), A(c, \vec{a}),$$

with $c \neq a_i$ for all $i \in \{1, \ldots, n\}$. By the induction hypothesis, we can find some $m < \omega$ such that

$$\mathcal{H}[r, \vec{s}] \,\Big|\!\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\, \Gamma'(\vec{s}), A(r, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)} \text{ for all terms } r \text{ with } |r| < \Omega_\omega.$$

We apply the $\mathsf{RS}_l(X)$ rule $(b\forall)$ to get

$$\mathcal{H}[\vec{s}] \,\Big|\!\frac{\Omega_\omega \cdot \omega^m + 1}{\Omega_\omega + m}\, \Gamma'(\vec{s}), \forall x \in \mathbb{L}_{\Omega_\omega}(X)\ A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Using Lemma 4.15 gives the desired bounds, as we obtain

$$\mathcal{H}[\vec{s}] \,\Big|\!\frac{\Omega_\omega \cdot \omega^{m+1}}{\Omega_\omega + m + 1}\, \Gamma'(\vec{s}), \forall x \in \mathbb{L}_{\Omega_\omega}(X)\ A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Case 5. We suppose that the last KPl rule applied is $(b\exists)$. This means that

$$\Gamma(\vec{a}) = \Gamma'(\vec{a}), \exists x \in a_i A(x, \vec{a})$$

for some KPl formula A and some free variable a_i, with $i \in \{1, \ldots, n\}$. Therefore, we have

$$\mathsf{KPl} \vdash \Gamma'(\vec{a})^{\mathbb{L}_{\Omega_\omega}(X)}, c \in a_i \wedge A(c, \vec{a})$$

for some free variable c. This case is done by splitting two subcases based on whether c is a_j for some $j \in \{1, \ldots, n\}$ or not. We divide subcases based on the form of s_i. We refer the reader interested in the details to [5], Theorem 4.10. (case 3 of the proof, page 36).

Case 6. We suppose that the last KPl rule applied is $(\exists)$. This means that

$$\Gamma(\vec{a}) = \Gamma'(\vec{a}), \exists x A(x, \vec{a}).$$

for some KPl formula A. Therefore,

$$\mathsf{KPl} \vdash \Gamma'(\vec{a}), A(c, \vec{a}). \tag{13}$$

We distinguish two cases depending on whether $c = a_i$ for some $i \in \{1, \ldots, n\}$ or not.
We suppose that $c = a_i$ with $i \in \{1, \ldots, n\}$. Then, the derivation (13) is

$$\mathsf{KPl} \vdash \Gamma'(\vec{a}), A(a_i, \vec{a}).$$

By the induction hypothesis, we can find some $m < \omega$ such that

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m} \Gamma'(\vec{s}), A(s_i, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Since $|s_i| < \Omega_\omega$, we use the $\mathsf{RS}_l(X)$ rule $(b\exists)$ to get

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^m + 1}{\Omega_\omega + m} \Gamma'(\vec{s}), \exists x A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Finally, by means of Lemma 4.15, we are able to obtain the desired ordinal bounds, as follows

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^{m+1}}{\Omega_\omega + m + 1} \Gamma'(\vec{s}), \exists x A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

We suppose now that $c \neq a_i$ for any $i \in \{1, \ldots, n\}$. Since in the translation of KPl formulas to $\mathsf{RS}_l(X)$-formulas, free variables become terms in, we can assign to c any term we want with level below Ω_ω, and so we choose $\overline{\emptyset}$ while applying the induction hypothesis, that produces an $m < \omega$ such that

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m} \Gamma'(\vec{s}), A(\overline{\emptyset}, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Here, we apply the $\mathsf{RS}_l(X)$ rule $(\exists)$ to obtain

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^m + 1}{\Omega_\omega + m} \Gamma'(\vec{s}), \exists x \in \mathbb{L}_{\Omega_\omega}(X) A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Finally, by means of Lemma 4.15, we obtain the desired ordinal bounds, as follows

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^{m+1}}{\Omega_\omega + m + 1} \Gamma'(\vec{s}), \exists x \in \mathbb{L}_{\Omega_\omega}(X) A(x, \vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

Case 7. We suppose that the last KPl rule applied is (Cut). This means that there is some KPl formula $A(\vec{a}, b_1, \ldots, b_k)$, where $b_1, \ldots, b_k$ are all the free variables occuring in A different from the free variables in $\vec{a}$, such that

$$\mathsf{KPl} \vdash \Gamma(\vec{a}), A(\vec{a}, b_1, \ldots, b_k) \text{ and } \mathsf{KPl} \vdash \Gamma(\vec{a}), \neg A(\vec{a}, b_1, \ldots, b_k).$$

Since the level of $\overline{\emptyset}$ is below Ω_ω, we can choose $\overline{\emptyset}$ as the term replacing b_j in the $\mathsf{RS}_l(X)$-formula $A^{\mathbb{L}_{\Omega_\omega}(X)}$ for all $j \in \{1, \ldots, k\}$, and so by the induction hypothesis we can find $m_0, m_1 < \Omega_\omega$ such that

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^{m_0}}{\Omega_\omega + m_0} \Gamma(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, A(\vec{s}, \overline{\emptyset}, \ldots, \overline{\emptyset})^{\mathbb{L}_{\Omega_\omega}(X)}. \tag{14}$$

and

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^{m_1}}{\Omega_\omega + m_1} \Gamma(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}, \neg A(\vec{s}, \overline{\emptyset}, \ldots, \overline{\emptyset})^{\mathbb{L}_{\Omega_\omega}(X)}. \tag{15}$$

We apply the $\mathsf{RS}_l(X)$ rule (Cut) to (14) and 15) and obtain

$$\mathcal{H}[\vec{s}] \mid\frac{\Omega_\omega \cdot \omega^{\max(m_0, m_1)+1}}{\Omega_\omega + \max(m_0, m_1)} \Gamma(\vec{s})^{\mathbb{L}_{\Omega_\omega}(X)}.$$

□

6 The provably total set-recursive functions of KPl

We are now ready to prove the main theorem of this thesis. We will give a bound for $f(x)$, where f is a set-recursive function such that KPl proves that f is total and uniformly Σ-definable in every admissible set, and x is any set. This bound will be $\hat{G}_n(x)$, defined as follows. This definition depends on a fixed set X, following Reading Convention 3.28. This means that the cardinals Ω_n for $n \leq \omega$ and the functions ψ_n for $n < \omega$ are also fixed and depend on X.

Definition 6.1. *We define the ordinal e_n by recursion on n as follows:*

- $e_0 = \Omega_\omega + 1,$
- $e_{n+1} = \omega^{e_n}.$

Now, we define for each $n < \omega$ the set $\hat{G}_n(x) = L_{\psi_0(e_{n+3})}(X)$.

Lemma 6.2. *For every natural number n we have*

$$e_n \in B_0(e_{n+1}).$$

Proof. We proceed by induction on n.
For $n = 0$, we have $e_0 = \Omega_\omega + 1 \in B_0(e_1)$ since, by definition, we have $1, \Omega_\omega \in B_0(e_1)$.
We suppose that $e_n \in B_0(e_{n+1})$. We show that $e_{n+1} \in B_0(e_{n+2})$.
We have $e_{n+1} = \omega^{e_n} = \varphi 0 e_n$. By the induction hypothesis,

$$e_n \in B_0(e_{n+1}) \subseteq B_0(e_{n+2}).$$

It follows that $0, e_n \in B_0(e_{n+2})$, and therefore

$$e_{n+1} = \varphi 0 e_n \in B_0(e_{n+2}).$$

□

We finally state and show our main theorem.

Theorem 6.3 (Main Theorem)**.** *Let f be a set-recursive function such that* KPl *proves that f is total and uniformly Σ-definable in any admissible set. Then, there is some $n < \omega$ such that*

$$V \models \forall x(f(x) \in \hat{G}_n(x)).$$

Proof. Let $A_f(\cdot,\cdot)$ be the Σ formula that defines f in any admissible set. In particular, KPI proves that f is total and $A_f(x, f(x))$ is satisfied in $L_{\Omega_0}(x)$ for any x. Therefore, we have that $\mathsf{KPI} \vdash Ad(u) \to [\forall x \in u \exists! y \in u\ A_f(x,y)^u]$.
Now, fix a set X and let θ be the set-theoretic rank of X, as in Reading Convention 3.28. By Theorem 5.20, we have

$$\mathcal{H}_0 \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\ Ad(\mathbb{L}_{\Omega_0}(X)) \to \forall x \in \mathbb{L}_{\Omega_0}(X) \exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(x,y)^{\mathbb{L}_{\Omega_0}(X)},$$

which is exactly

$$\mathcal{H}_0 \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\ \neg Ad(\mathbb{L}_{\Omega_0}(X)) \vee \forall x \in \mathbb{L}_{\Omega_0}(X) \exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(x,y)^{\mathbb{L}_{\Omega_0}(X)}.$$

Therefore, by Lemma 4.18 we obtain

$$\mathcal{H}_0 \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\ \neg Ad(\mathbb{L}_{\Omega_0}(X)), \forall x \in \mathbb{L}_{\Omega_0}(X) \exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(x,y)^{\mathbb{L}_{\Omega_0}(X)}. \tag{1}$$

On the other hand, we have by Lemma 5.18 and Lemma 4.15

$$\mathcal{H}_0 \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\ Ad(\mathbb{L}_{\Omega_0}(X)), \forall x \in \mathbb{L}_{\Omega_0}(X) \exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(x,y)^{\mathbb{L}_{\Omega_0}(X)}. \tag{2}$$

We apply (Cut) to (1) and (2) to obtain

$$\mathcal{H}_0 \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\ \forall x \in \mathbb{L}_{\Omega_0}(X) \exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(x,y)^{\mathbb{L}_{\Omega_0}(X)}.$$

We notice that $\mathrm{rk}(Ad(\mathbb{L}_{\Omega_0}(X))) = \Omega_0 + 5$ and so the complexity of the cuts has not been increased. We now apply Lemma 4.17 (Inversion) two times. First, since

$$\exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(X,y)^{\mathbb{L}_{\Omega_0}(X)} \in \mathcal{C}(\forall x \in \mathbb{L}_{\Omega_0}(X) \exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(x,y)^{\mathbb{L}_{\Omega_0}(X)}),$$

we get

$$\mathcal{H}_0 \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\ \exists! y \in \mathbb{L}_{\Omega_0}(X) A_f(X,y)^{\mathbb{L}_{\Omega_0}(X)}.$$

Now, since the symbol "!" acts like a conjunction, we finally obtain

$$\mathcal{H}_0 \Big|\frac{\Omega_\omega \cdot \omega^m}{\Omega_\omega + m}\ \exists y \in \mathbb{L}_{\Omega_0}(X) A_f(X,y)^{\mathbb{L}_{\Omega_0}(X)}. \tag{3}$$

We are now going to eliminate cuts from the derivation (3). By Theorem 4.20 (Predicative Cut Elimination), we have

$$\mathcal{H}_0 \Big|\frac{e_{m+1}}{\Omega_\omega + 1}\ \exists y \in \mathbb{L}_{\Omega_0}(X) A_f(X,y)^{\mathbb{L}_{\Omega_0}(X)}.$$

By the Collapsing Theorem 4.24, we get

$$\mathcal{H}_{e_{m+2}} \left|\frac{\psi_0(e_{m+2})}{\psi_0(e_{m+2})}\right. \exists y \in \mathbb{L}_{\Omega_0}(X) A_f(X,y)^{\mathbb{L}_{\Omega_0}(X)}.$$

Again by Theorem 4.20 (Predicative Cut Elimination), we have

$$\mathcal{H}_{e_{m+2}} \left|\frac{\alpha}{0}\right. \exists y \in \mathbb{L}_{\Omega_0}(X) A_f(X,y)^{\mathbb{L}_{\Omega_0}(X)}.$$

where $\alpha = \varphi(\psi_0(e_{m+2}))(\psi_0(e_{m+2}))$. By the Boundedness Lemma we get

$$\mathcal{H}_{e_{m+2}} \left|\frac{\alpha}{0}\right. \exists y \in \mathbb{L}_{\alpha}(X) A_f(X,y)^{\mathbb{L}_{\alpha}(X)}.$$

It follows that

$$L_\alpha(X) \vDash \exists y A_f(X,y).$$

Indeed, we prove a more general result.

Claim 6.3.1. *Given a $\Sigma^{\mathbb{L}_\delta(X)}$-formula $A^{\mathbb{L}_\delta(X)}$, if $\mathcal{H}_\gamma \left|\frac{\beta}{0}\right. A^{\mathbb{L}_\delta(X)}$ then $L_\delta(X) \vDash A$.*

Fix δ. We prove Claim 6.3.1 by induction on β. If $A^{\mathbb{L}_\delta(X)}$ is an axiom, then $A^{\mathbb{L}_\delta(X)}$ is a basic formula. We suppose that this formula is $\overline{u} \in \overline{v}$. This means that $u, v \in TC(\{X\})$ satisfy $u \in v$. But $TC(\{X\}) \subseteq L_\delta(X)$. It follows that $L_\delta(X) \vDash A$. We assume that $A^{\mathbb{L}_\delta(X)}$ has been derived using a rule (R) different from (Cut) and (Ref_n). Then, we have $\mathcal{H}_\gamma[t_A(B)] \left|\frac{\beta_B}{0}\right. B^{\mathbb{L}_\delta(X)}$ for some/any premise $B^{\mathbb{L}_\delta(X)} \in \mathcal{C}(A^{\mathbb{L}_\delta(X)})$. But all those premises are also $\Sigma^{\mathbb{L}_\delta(X)}$-formulas. By the induction hypothesis, $L_\delta(X) \vDash B$ for some/any B. Finally, apply the same rule (R) but in KPl to obtain $L_\delta(X) \vDash A$.

Hence, using Claim 6.3.1, we get $L_\alpha(X) \vDash \exists y\ A_f(X,y)$. This means that $f(x) \in L_\alpha(X)$. Finally, $L_\alpha(X) \subseteq \hat{G}_{m+3}(X)$. By Lemma 6.2, we have

$$e_{m+2} \in B_0(e_{m+3}).$$

It follows that $\psi_0(e_{m+2}) < \psi_0(e_{m+3})$. Thus,

$$\alpha = \varphi(\psi_0(e_{m+2}))(\psi_0(e_{m+2})) < \psi_0(e_{m+3}).$$

Hence,

$$L_\alpha(X) \subseteq L_{\psi_0(e_{m+3})}(X) = \hat{G}_{m+3}(X).$$

$\square$

References

[1] J. Barwise. *Admissible sets and structures.* Springer, 1975.

[2] W. Buchholz. *A new system of proof-theoretic ordinal functions.* In Annals of Pure and Applied Logic volume 32, pages 195-207, 1986.

[3] W. Buchholz and S. Wainer. *Provably computable functions and the fast growing hierarchy.* In Logic and Combinatorics, Contemporary Mathematics 65, pages 179-198. American Mathematical Society, Providence, 1986.

[4] W. Buchholz. *A simplified version of local predicativity.* In Aczel P, Simmons H, Wainer SS (eds.), Proof Theory: A Selection of Papers from the Leeds Proof Theory Programme 1990. Cambridge University Press; 1993:115-148.

[5] J. Cook, M. Rathjen. *Classifying the provably total set functions of KP and KP(P).* University of Leeds, 2016.

[6] G. Gentzen. *Die Widerspruchsfreiheit der reinen Zahlentheorie.* In Mathematische Annalen 112. 1936.

[7] G. Jäger. *Zur Beweistheorie der Kripke-Platek-Mengenlehre über den natürlichen Zahlen.* In Arch math Logik 22, pages 121-139. 1980.

[8] T. Jech. *Set Theory. The Third Millennium Edition, revised and expanded.* Springer, 2003.

[9] G. Kreisel. *On the interpretation of non-finitist proofs I.* In Journal of Symbolic Logic 16, 1951.

[10] G. Kreisel. *On the interpretation of non-finitist proofs II.* In Journal of Symbolic Logic 17, 1952.

[11] K. Kunen. *Set Theory. An Introduction to Independence Proofs.*

[12] W. Pohlers. *Subsystems of Set Theory and Second Order Number Theory.* In Samuel R. Buss (ed.), Handbook of proof theory, pages 137-209. Elsevier, 1998.

[13] W. Pohlers. *Proof Theory: The First Step into Impredicativity* (Universitext). Springer, 2009.

[14] M. Rathjen. *The Realm of Ordinal Analysis.* In S.B. Cooper and J.K. Truss (eds.), Sets and Proofs, pages 219-279. Cambridge University Press, 1999.

[15] M. Rathjen. *The Art of Ordinal Analysis*. In M. Sanz-Solé, J. Soria, J.L. Varona and J. Verdera (eds.), Proceedings of the International Congress of Mathematicians Madrid, August 22–30, 2006, Volume II. European Mathematical Society, 2007.

[16] G.E. Sacks. *Higher Recursion Theory*. Springer, 1990.

[17] K. Schütte. *Proof Theory*. Springer, 1977.

[18] G. Takeuti. *Consistency proofs of subsystems of classical analysis*. In Annals of Mathematics, Vol 86 No. 2, pages 299-348. Princeton University, 1967.

[19] A. Weiermann. *How to characterize provably total functions by local predicativity*. In Journal of Symbolic Logic 61, 1996.

Index of Symbols

$B_n(\alpha)$, 24
$L_\alpha(X)$, 14
$T(\theta)$, 30
$\#$, 62
$\mathcal{C}(A)$, 34
$\mathcal{H}$, 41
$\mathcal{H} \left|\frac{\alpha}{\rho}\right. \Gamma$, 46
$\mathcal{H} \left|\frac{\alpha}{\quad}\right. \Gamma$, 42
$\mathcal{H}_\beta$, 55
Ω_ω, 14
Σ-formula, 12
Σ^{Ω_n}-formula, 33
$\Vdash$, 62
$\left\|\frac{\alpha}{\rho}\right.$, 62
$\bigvee$-type and $\bigwedge$-type, 33
$\dot{\in}$, 34
$\hat{G}_n(X)$, 92
$\mathrm{rk}(\cdot)$, 35
$\varphi.(\cdot)$, 18
$\psi_n(\alpha)$, 25
θ, 14
e_n, 92
$k(A)$, 33
$t_A(B)$, 46

General Index

$\mathsf{RS}_l(X)$-formulas, 33
$\mathsf{RS}_l(X)$-terms, 32

Cardinal, 16

Normal Form, 27

Operator, 41

Ordinal, 15

Rank, 35

Transitive set, 15

Veblen functions, 18