Thesis for the Degree of Doctor of Philosophy in Computer Science

Sheaf Semantics in Constructive Algebra and Type Theory

Bassel Mannaa

UNIVERSITY OF GOTHENBURG

Department of Computer Science and Engineering
University of Gothenburg
Göteborg, Sweden 2016

Sheaf Semantics in Constructive Algebra and Type Theory

Bassel Mannaa

Technical Report 135D
ISBN 978-91-628-9985-1 (Print), 978-91-628-9986-8 (PDF)
Department of Computer Science and Engineering
Programming Logic Research Group

University of Gothenburg
SE-405 30 Göteborg
Sweden
Telephone +46 (0)31 786 0000

Printed at Chalmers Reproservice
Göteborg, Sweden 2016

Abstract

In this thesis we present two applications of sheaf semantics. The first is to give constructive proof of Newton–Puiseux theorem. The second is to show the independence of Markov's principle from type theory.

In the first part we study Newton–Puiseux algorithm from a constructive point of view. This is the algorithm used for computing the Puiseux expansions of a plane algebraic curve defined by an affine equation over an algebraically closed field. The termination of this algorithm is usually justified by non-constructive means. By adding a separability condition we obtain a variant of the algorithm, the termination of which is justified constructively in characteristic 0. To eliminate the assumption of an algebraically closed base field we present a constructive interpretation of the existence of the separable algebraic closure of a field by building, in a constructive metatheory, a suitable sheaf model where there is such separable algebraic closure. Consequently, one can use this interpretation to extract computational content from proofs involving this assumption. The theorem of Newton–Puiseux is one example. We then can find Puiseux expansions of an algebraic curve defined over a non-algebraically closed field K of characteristic 0. The expansions are given as a fractional power series over a finite dimensional K-algebra.

In the second part we show that Markov's principle is independent from type theory. The underlying idea is that Markov's principle does not hold in the topos of sheaves over Cantor space. The presentation in this part is purely syntactical. We build an extension of type theory where the judgments are indexed by basic compact opens of Cantor space. We give an interpretation for this extension of type theory by way of computability predicate and relation. We can then show that Markov's principle is not derivable in this extension and consequently not derivable in type theory.

Keywords: Newton–Puiseux theorem, Algebraic curve, Sheaf model, Dynamic evaluation, Type theory, Markov's Principle, Forcing.

The present thesis is an extended version of the papers

(i) *Dynamic Newton–Puiseux Theorem* in "The Journal of Logic and Analysis" [Mannaa and Coquand, 2013] and the paper

(ii) *A Sheaf Model of the Algebraic Closure* in "The Fifth International Workshop on Classical Logic and Computation" [Mannaa and Coquand, 2014].

(iii) *The Independence of Markov's Principle in Type Theory* in "The 1st International Conference on Formal Structures For Computation and Deduction"[Coquand and Mannaa, 2016].

Acknowledgments

I'm very grateful to my advisor, Thierry Coquand, for his mentorship, continuing inspiration and guidance.

My gratitude to Henri Lombardi and Marie-Françoise Roy from whom I learned a lot.

I thank Andreas Abel for the collaboration, discussions and interesting exchange of ideas.

Thanks to Peter Dybjer, Bengt Nordström with whom my interactions, however brief, have doubtlessly made me a better researcher.

Thanks to Fabian Ruch for the fruitful collaboration.

Thanks Simon Huber, Andrea Vezossi, Anders Mörtberg, Guilhem Moulin and Cyrill Cohen for the many spontaneous and fruitful discussions around the coffee machine, in the sauna and at the pub.

Finally I'd like to thank my family and friends for their emotional support.

Contents

Introduction

The notion of a sheaf over a topological spaces was first explicitly defined by Jean Leray [Miller, 2000]. Sheaves became an essential tool in the study algebraic topology, e.g. sheaf cohomology. Intuitively a sheaf attaches data (i.e. a set) to each open of the topological space in such a way that the data attached to an open set U is in one-to-one correspondence to the compatible data attached to the opens of a cover $\bigcup_i U_i = U$. Thus a sheaf allows us to pass from the local to the global and vice-a-versa. The most common example is that of a continuous function $f : U \to \mathbb{R}$. The function f gives rise to a continuous function $U_i \to \mathbb{R}$ when restricted to points in U_i. On the other hand, given a family of continuous functions $f_i : U_i \to \mathbb{R}$ such that each pair, f_i and f_j coincide on the intersection $U_i \cap U_j$ we can glue or piece together these functions into a continuous function $f : U \to \mathbb{R}$ that coincide with f_i when restricted to points in U_i.

For his work in algebraic geometry, Grothendieck and his collaborators generalized the notion of sheaf over a topological space to that of a sheaf over an arbitrary site. A site is a category with a topology. Just as an open set is covered by subopens in a topological space. An object in a site is covered by a collection of maps into it. Grothendieck gave the name *topos* to the category of sheaves over a site and considered the study of toposes to be purpose of topology "il semble raisonnable et légitime aux auteurs du présent séminaire de considérer que l'objet de la Topologie est l'étude des topos (et non des seuls espaces topologiques)."[Artin et al., 1972]. One should remark that the logic of a topos is not necessarily boolean. That is to say, the algebra of subsheaves of an arbitrary sheaf is not necessarily boolean. This is similar to the situation in topology where a complement of an open set is not necessarily open, hence the negation is given by the interior of the complement and the law of excluded middle does not necessarily hold.

Around this time in the early 1960's Cohen introduced his method of forcing and proved the independence of the continuum hypothesis from ZF [Cohen, 1963]. Soon after Scott, Solovay, and Vopěnka introduced boolean valued models in order to simplify Cohen's proof [Solovay, 1965; Vopěnka, 1965; Bell, 2011]. In 1966 Lawvere observed that boolean valued models and the independence of the continuum hypothesis should be presented in terms of Grothendieck toposes [Interview:Lawvere 2] [McLarty, 1990]. It was later that Lawvere presented this result [Tierney, 1972]. A couple of years earlier, Lawvere and Tierney's developed the theory of elementary topos 1969-1971 [Lawvere, 1970]. This is an elementary axiomatization of the notion of a topos of which sheaf toposes are instances. We interject to remark that the definition of elementary topos is impredicative. Sheaf toposes on the other hand can be described predicatively and thus are more amenable to development in a constructive metatheory [Palmgren, 1997].

Soon after the introduction of elementary toposes the notion of internal language of a topos and the correspondence between type theories and toposes was discovered independently by Mitchell [Mitchell, 1972] and Bénabou [Coste, 1972] among others. Various equivalent notions of semantics accompanied. Perhaps the most intuitive of these is Joyal's generalization of Kripke semantics to what is now known as Kripke–Joyal semantics [Osius, 1975] with the purpose of unifying the various notions of forcing as instances of forcing in a sheaf topos [Bell, 2005]. This style of semantics is in fact conceptually similar to Beth's semantics of intuitionistic logic [Beth, 1956][1]. Indeed it has become customary to use the term Beth–Kripke–Joyal for this kind of semantics.

In this monograph we present two applications of sheaf semantics to constructive mathematics and dependent type theory. The monograph is thus divided into two self contained parts A and B. We sum the results briefly below and defer more detailed introductions to the relevant parts.

In Part A we develop a constructive proof of Newton–Puiseux theorem based on one by Abhyankar [1990]. Though Abhyankar's proof is algorithmic in nature, i.e. it describes an algorithm for computing the Puiseux expansions of an algebraic curve over an algebraically closed ground field, it is nonetheless non-constructive. Our contribution is eliminating two assumptions from the classical proof. The termination of Abhyankar's algorithm depends on the assumption of decidable equality on the ring of power series over an algebraically closed field. By eliminating this assumption we obtain a constructive proof of termination. We then turn to eliminating the assumption of an (sepa-

[1] The history of Beth semantics is quite complicated. Beth developed his semantics over a long period of time 1947-1956. [Troelstra and van Ulsen, 1999].

rable) algebraic closure. This is where sheaf semantics comes in play. We give an interpretation of the separable algebraic closure of a field by building, in a constructive metatheory, a suitable site model where there is such separable algebraic closure. The model gives us a direct description of the computational content residing in the assumption of separable algebraic closure. That is, it gives us a direct description of a constructive proof of a more general statement of Newton-Puiseux theorem not involving the assumption of the separable algebraic closure. To quote Joyal [1975] "This method is quite in the spirit of Hilbert when he suggested a deeper understanding of the introduction and elimination of ideal objects in mathematical reasoning".

In Part B we present a proof of the independence of Markov's principle from type theory. The underlying idea is that Markov's principle fails to hold in the topos of sheaves over Cantor space. The presentation in this part is, however, purely syntactical and without direct reference to toposes. We design a forcing extension of type theory in which we replace the usual type theoretic judgments by local ones. These are judgments valid locally at compact opens of the space. We add formally a *locality* inference rule allowing us to glue local judgments into global ones. We describe a semantics for this extension by way of computability predicates and relations. We force a term $\mathsf{f} : N \to N_2$ representing a generic point (sequence) in the space $2^{\mathbb{N}}$ and show that while it is provably false that this sequence is never 0, i.e. $\neg\neg\exists n.\mathsf{f}(n) = 0$, it cannot be shown that it has the value 0 at any time, i.e. it cannot be shown that $\exists n.\mathsf{f}(n) = 0$.

A more direct approach to show the independence of Markov's principle from type theory would be to give an interpretation of type theory in the topos of sheaves over Cantor space. However, this was unattainable due to the known difficulty with the sheaf interpretation of the type theoretic universe, see [Xu and Escardó, 2016; Hofmann and Streicher, 199?]. The Hoffmann-Streicher interpretation of the universe given by the presheaf $U(X) = \{A \mid A \in \mathsf{Presh}(\mathcal{C}/X), A \text{ small}\}$ does not extend well to sheaves. Mainly the presheaf $U(X) = \{A \mid A \in \mathsf{Sh}(\mathcal{C}/X), A \text{ small}\}$ is not necessarily a sheaf, actually not even necessarily a separated presheaf. Interpreting the universe as the sheafification $\widetilde{U}$ of the presheaf U is inadequate since an element in $\widetilde{U}(X)$ is then an equivalence class and it is not clear how to define the decoding $\mathsf{El}[a]$ where $[a] \in \widetilde{U}(X)$ is an equivalence class of sheaves.

At the end of this monograph we outline a proposed solution for this problem of interpretation of the universe in a sheaf model. The idea is to interpret the universe by the *stack* $UX = \mathsf{Sh}(\mathcal{C}/X)$ where $\mathsf{Sh}(\mathcal{C}/X)$ is the groupoid of small sheaves over an object X of $\mathcal{C}$. It can be shown that U is indeed a stack [Vistoli, 2004; Grothendieck and Dieudonné, 1960]. We will thus outline an interpretation of type theory in stacks where small types are interpreted by small sheaves (more accurately, stacks of small discrete groupoids). The model combines the familiar sheaf/presheaf interpretation of types theory (e.g. as presented in [Huber, 2015]) with the groupoid interpretation of type theory [Hofmann and Streicher, 1998].

I

Categorical Preliminaries

In this chapter we give a brief outline of some of the notions and results that will be used in Chapter III. We assume that the reader is familiar with basic notions from category theory used in general algebra.

1 Functors and presheaves

A (covariant) *functor* $\mathbf{F} : \mathcal{C} \to \mathcal{D}$ between two categories $\mathcal{C}$ and $\mathcal{D}$ assigns to each object C of $\mathcal{C}$ an object $\mathbf{F}(C)$ of $\mathcal{D}$ and to each arrow $f : C \to B$ of $\mathcal{C}$ an arrow $\mathbf{F}(f) : \mathbf{F}(C) \to \mathbf{F}(B)$ of $\mathcal{D}$ such that $\mathbf{F}(1_C) = 1_{\mathbf{F}(C)}$ and $\mathbf{F}(fg) = \mathbf{F}(f)\mathbf{F}(g)$. A *natural transformation* Θ between two functors $\mathbf{F} : \mathcal{C} \to \mathcal{D}$ and $\mathbf{G} : \mathcal{C} \to \mathcal{D}$ is collection of arrows, indexed by objects of $\mathcal{C}$, of the form $\Theta_C : \mathbf{F}(C) \to \mathbf{G}(C)$ such that for each arrow $f : C \to A$ of $\mathcal{C}$ the diagram

$$\begin{array}{ccc} \mathbf{F}(C) & \xrightarrow{\Theta_C} & \mathbf{G}(C) \\ \downarrow{\scriptstyle \mathbf{F}(f)} & & \downarrow{\scriptstyle \mathbf{G}(f)} \\ \mathbf{F}(A) & \xrightarrow{\Theta_A} & \mathbf{G}(A) \end{array}$$

commutes.

A contravariant functor $\mathbf{G}$ between $\mathcal{C}$ and $\mathcal{D}$ is a covariant functor $\mathbf{G} : \mathcal{C}^{op} \to \mathcal{D}$. Thus for $f : C \to B$ of $\mathcal{C}$ we have $\mathbf{G}(f) : \mathbf{G}(B) \to \mathbf{G}(C)$ in $\mathcal{D}$ and $\mathbf{G}(fg) = \mathbf{G}(g)\mathbf{G}(f)$. The collection of functors between two categories $\mathcal{C}$ and $\mathcal{D}$ and natural transformation between them form a category $\mathcal{D}^{\mathcal{C}}$.

A functor $\mathbf{F} \in \mathbf{Set}^{\mathcal{C}^{op}}$ is called a *presheaf* of sets over/on the category $\mathcal{C}$. For an arrow $f : A \to B$ of $\mathcal{C}$ the map $\mathbf{F}(f) : \mathbf{F}(B) \to \mathbf{F}(A)$ is called a *restriction* map between the sets $\mathbf{F}(B)$ and $\mathbf{F}(A)$. An element $x \in \mathbf{F}(B)$

has a restriction $xf = (\mathbf{F}(f))(x) \in \mathbf{F}(A)$ called the restriction of x *along* f.

A category is *small* if the collection of objects in the category form a set. A category is *locally small* if the collection of morphisms between any two objects in the category is a set. The presheaf $\mathbf{y}_C := \mathrm{Hom}(-, C)$ of $\mathbf{Set}^{\mathcal{C}^{op}}$ associates to each object A of $\mathcal{C}$ the set $\mathrm{Hom}(A, C)$ of arrows $A \to C$ of $\mathcal{C}$. Let $g \in y_C(B)$ and let $f : A \to B$ be a morphism of $\mathcal{C}$ then $gf \in \mathbf{y}_C(A)$ is the restriction of g along f. The presheaf $\mathbf{y}_C$ is called the *Yoneda embedding* of C.

Fact 1.1 (Yoneda Lemma). *Let $\mathcal{C}$ be a locally small category and $\mathbf{F} \in \mathbf{Set}^{\mathcal{C}^{op}}$. We have an isomorphism $\mathrm{Nat}(\mathbf{y}_C, \mathbf{F}) \cong \mathbf{F}(C)$. Where $\mathrm{Nat}(\mathbf{y}_C, \mathbf{F})$ is the set of natural transformations $\mathrm{Hom}_{\mathbf{Set}^{\mathcal{C}^{op}}}(\mathbf{y}_C, \mathbf{F})$ between the presheaves $\mathbf{y}_C$ and $\mathbf{F}$.*

A *sieve* S on an object C of a small category $\mathcal{C}$ is a set of morphisms with codomain C such that if $f : D \to C \in S$ then for any g with codomain D we have $fg \in S$. Given a set S of morphisms with codomain C we define the sieve generated by S to be $(S) = \{fg \mid f \in S, \mathrm{cod}(g) = \mathrm{dom}(f)\}$. Note that in $\mathbf{Set}^{\mathcal{C}^{op}}$ a sieve uniquely determines a subobject of $\mathbf{y}_C$. Given $f : D \to C$ and S a collection of arrows with codomain C then $f^*(S) = \{g \mid \mathrm{cod}(g) = D, fg \in S\}$. When S is a sieve $f^*(S) = Sf$ is a sieve on D, the restriction of S along f in $\mathbf{Set}^{\mathcal{C}^{op}}$. Dually, given $g : C \to D$ and M a collection of arrows with domain C then $g_*(M) = \{h \mid \mathrm{dom}(h) = D, hg \in M\}$. The presheaf $\mathbf{\Omega}$ is the presheaf assigning to each object C the set $\mathbf{\Omega}(C)$ of sieves on C with restriction maps f^* for each morphism $f : D \to C$ of $\mathcal{C}$.

2 Elementary topos

An *elementary topos* [Lawvere, 1970] is a category $\mathcal{C}$ such that

1. $\mathcal{C}$ has all finite limits and colimits.
2. $\mathcal{C}$ is Cartesian closed. In particular for any two objects C and D of $\mathcal{C}$ there is an object D^C such that there is a one-to-one correspondence between the arrows $A \to D^C$ and the arrows $A \times C \to D$ for any object A of $\mathcal{C}$. For a locally small category this is expressed as $\mathrm{Hom}(A \times C, D) \cong \mathrm{Hom}(A, D^C)$.
3. $\mathcal{C}$ has a subobject classifier. That is, there is an object Ω and a map $1 \xrightarrow{true} \Omega$ such that for any object C of $\mathcal{C}$ there is a one-to-one correspondence between the subobjects of C given by monomorphisms with codomain C and the maps from C to Ω (called clas-

sifying/characteristic maps). A subobject is uniquely determined by the pullback of the map $1 \xrightarrow{true} \Omega$ along the characteristic map.

An elementary topos can be considered as a generalization of the category **Set** of sets. The category $\mathbf{Set}^{\mathcal{C}^{op}}$ of presheaves on a small category $\mathcal{C}$ is an elementary topos. The lattice of subobjects of an object C in an elementary topos $\mathcal{E}$ (monomorphisms with codomain C) is a *Heyting algebra*.

3 Grothendieck topos

In this section we define the notions of site, coverage, and sheaf following [Johnstone, 2002*b*,*a*].

Definition 3.1 (Coverage)**.** By a coverage on a category $\mathcal{C}$ we mean a function $\mathbf{J}$ assigning to each object C of $\mathcal{C}$ a collection $\mathbf{J}(C)$ of families of morphisms of the form $\{f_i : C_i \to C \mid i \in I\}$ such that :
If $\{f_i : C_i \to C \mid i \in I\} \in \mathbf{J}(C)$ and $g : D \to C$ is a morphism, then there exist $\{h_j : D_j \to D \mid j \in J\} \in \mathbf{J}(D)$ such that for any $j \in J$ we have $gh_j = f_i k$ for some $i \in I$ and some $k : D_j \to C_i$.

A *site* $(\mathcal{C}, \mathbf{J})$ is a small category $\mathcal{C}$ equipped with a coverage $\mathbf{J}$. A family $\{f_i : C_i \to C \mid i \in I\} \in \mathbf{J}(C)$ is called elementary cover or elementary covering family of C.

Definition 3.2 (Compatible family)**.** Let $\mathcal{C}$ be a category and $\mathbf{F} : \mathcal{C}^{op} \to \mathbf{Set}$ a presheaf. Let $\{f_i : C_i \to C \mid i \in I\}$ be a family of morphisms in $\mathcal{C}$. A family $\{s_i \in \mathbf{F}(C_i) \mid i \in I\}$ is compatible if for all $\ell, j \in I$ whenever we have $g : D \to C_\ell$ and $h : D \to C_j$ satisfying $f_\ell g = f_j h$ we have $\mathbf{F}(g)(s_\ell) = \mathbf{F}(h)(s_j)$.

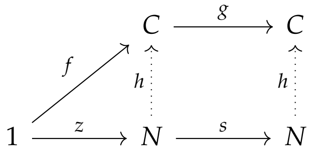
Definition 3.3 (The sheaf axiom)**.** Let $\mathcal{C}$ be a category. A presheaf $\mathbf{F} : \mathcal{C}^{op} \to \mathbf{Set}$ satisfies the sheaf axiom for a family of morphisms $\{f_i : C_i \to C \mid i \in I\}$ if whenever $\{s_i \in \mathbf{F}(C_i) \mid i \in I\}$ is a compatible family then there exist a unique $s \in \mathbf{F}(C)$ restricting to s_i along f_i for all $i \in I$. That is to say when there exist a unique s such that for all $i \in I$, $\mathbf{F}(f_i)(s) = s_i$. One usually refers to s as the *amalgamation* of $\{s_i\}_{i \in I}$.

Let $(\mathcal{C}, \mathbf{J})$ be a site. A presheaf $\mathbf{F} \in \mathbf{Set}^{\mathcal{C}^{op}}$ is a sheaf on $(\mathcal{C}, \mathbf{J})$ if it satisfies the sheaf axiom for each object C of $\mathcal{C}$ and each family of morphisms in $\mathbf{J}(C)$, i.e. if it satisfies the sheaf axiom for elementary covers.

The category of sheaves on a small site $\mathrm{Sh}(\mathcal{C}, \mathbf{J})$ is an elementary topos.

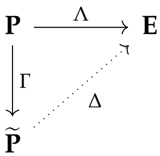
3.1 Natural numbers object and sheafification

A natural numbers object in a category with a terminal object is an object N along with two morphisms $z : 1 \to N$ and $s : N \to N$ such that for any diagram of the form $1 \xrightarrow{f} C \xrightarrow{g} C$ there is a unique morphism $h : N \to C$ making the diagram below commute.



Fact 3.4. *In* $\mathbf{Set}^{\mathcal{C}^{op}}$ *the constant presheaf* $\mathbf{N}$ *such that* $\mathbf{N}(C) = \mathbb{N}$ *and* $\mathbf{N}(f) = 1_{\mathbb{N}}$ *for every object* C *and morphism* f *of* $\mathcal{C}$ *is a natural numbers object.*

Let $(\mathcal{C}, \mathbf{J})$ be a site. The sheaf topos $\mathrm{Sh}(\mathcal{C}, \mathbf{J})$ is a full subcategory of the presheaf category $\mathbf{Set}^{\mathcal{C}^{op}}$. By the *sheafification* of a presheaf $\mathbf{P} \in \mathbf{Set}^{\mathcal{C}^{op}}$ we mean a sheaf $\widetilde{\mathbf{P}}$ of $\mathrm{Sh}(\mathcal{C}, \mathbf{J})$ along with a presheaf morphism $\Gamma : \mathbf{P} \to \widetilde{\mathbf{P}}$ such that for any sheaf $\mathbf{E}$ and any presheaf morphism $\Lambda : \mathbf{P} \to \mathbf{E}$ there is a unique sheaf morphism $\Delta : \mathbf{P} \to \mathbf{E}$ making the following diagram commute.



Fact 3.5. *Let* $(\mathcal{C}, \mathbf{J})$ *be a site. The sheaf topos* $\mathrm{Sh}(\mathcal{C}, \mathbf{J})$ *contains a natural numbers object* $\widetilde{\mathbf{N}}$ *where* $\widetilde{\mathbf{N}}$ *is the sheafification of the natural numbers presheaf* $\mathbf{N}$.

3.2 Kripke–Joyal sheaf semantics

We work with a typed language with equality $\mathcal{L}[V_1, ..., V_n]$ having the basic types $V_1, ..., V_n$ and type formers $- \times -, (-)^-, \mathcal{P}(-)$. The language $\mathcal{L}[V_1, ..., V_n]$ has typed constants and function symbols. For any type Y one has a stock of variables $y_1, y_2, ...$ of type Y. Terms and formulas of the language are defined as usual. We work within the proof theory of intuitionistic higher-order logic (IHOL). A detailed description of this deduction system is given in [Awodey, 1997].

The language $\mathcal{L}[V_1, ..., V_n]$ along with deduction system IHOL can be interpreted in an elementary topos in what is referred to as *topos semantics*. For a sheaf topos this interpretation takes a simpler form reminiscent of Beth semantics, usually referred to as *Kripke–Joyal sheaf semantics*. We describe this semantics here briefly following [Ščedrov, 1984]. Let $\mathcal{E} = \mathrm{Sh}(\mathcal{C}, \mathbf{J})$ be a sheaf topos. First we define a closure $\mathbf{J}^*$ of $\mathbf{J}$ as follows.

Definition 3.6 (Closure of a coverage)**.**

(i.) $\{C \xrightarrow{1_C} C\} \in \mathbf{J}^*(C)$ for all objects C in $\mathcal{C}$.

(ii.) If $\{C_i \xrightarrow{f_i} C\}_{i\in I} \in \mathbf{J}(C)$ then $\{C_i \xrightarrow{f_i} C\}_{i\in I} \in \mathbf{J}^*(C)$.

(iii.) If $\{C_i \xrightarrow{f_i} C\}_{i\in I} \in \mathbf{J}^*(C)$ and for each $i \in I$ we have $\{C_{ij} \xrightarrow{g_{ij}} C_i\}_{j\in J_i} \in \mathbf{J}^*(C_i)$ then $\{C_{ij} \xrightarrow{f_i g_{ij}} C\}_{i\in I, j\in J_i} \in \mathbf{J}^*(C)$.

An family $S \in \mathbf{J}^*(C)$ is called cover or covering family of C.

An interpretation of the language $\mathcal{L}[V_1, ..., V_n]$ in the topos $\mathcal{E}$ is given as follows: Associate to each basic type V_i of $\mathcal{L}[V_1, ..., V_n]$ an object $\mathbf{V}_i$ of $\mathcal{E}$. If Y and Z are types of $\mathcal{L}[V_1, ..., V_n]$ interpreted by objects $\mathbf{Y}$ and $\mathbf{Z}$, respectively, then the types $Y \times Z, Y^Z, \mathcal{P}(Z)$ are interpreted by $\mathbf{Y} \times \mathbf{Z}, \mathbf{Y}^{\mathbf{Z}}, \Omega^{\mathbf{Z}}$, respectively, where Ω is the subobject classifier of $\mathcal{E}$. A constant e of type E is interpreted by an arrow $\mathbf{1} \xrightarrow{\mathbf{e}} \mathbf{E}$ where $\mathbf{E}$ is the interpretation of E. For a term τ and an object $\mathbf{X}$ of $\mathcal{E}$, we write $\tau{:}\mathbf{X}$ to mean τ has a type X interpreted by the object $\mathbf{X}$.

Let $\phi(x_1, ..., x_n)$ be a formula with variables $x_1 : \mathbf{X}_1, ..., x_n : \mathbf{X}_n$. Let $c_1 \in \mathbf{X}_j(C), ..., c_n \in \mathbf{X}_n(C)$ for some object C of $\mathcal{C}$. We define the relation C *forces* $\phi(x_1, ..., x_n)[c_1, ..., c_n]$ written $C \Vdash \phi(x_1, ..., x_n)[c_1, ..., c_n]$ by induction on the structure of ϕ.

Definition 3.7 (Forcing)**.** First we replace the constants in ϕ by variables of the same type as follows: Let $e_1 : \mathbf{E}_1, ..., e_m : \mathbf{E}_m$ be the constants in $\phi(x_1, ..., x_n)$ then $C \Vdash \phi(x_1, ..., x_n)[c_1, ..., c_n]$ iff

$$C \Vdash \phi[y_1/e_1, ..., y_m/e_m](y_1, ..., y_m, x_1, ..., x_n)[\mathbf{e}_{1_C}(*), ..., \mathbf{e}_{m_C}(*), c_1, ..., c_n]$$

where $y_i{:}\mathbf{E}_i$ and $\mathbf{e}_i : \mathbf{1} \to \mathbf{E}_i$ is the interpretation of e_i.

Now it suffices to define the forcing relation for formulas free of constants by induction as follows:

$\boxed{\top}$ $C \Vdash \top$.

$\boxed{\bot}$ $C \Vdash \bot$ iff the empty family is a cover of C.

$\boxed{=}$ $C \Vdash (x_1 = x_2)[c_1, c_2]$ iff $c_1 = c_2$.

$\boxed{\wedge}$ $C \Vdash (\phi \wedge \psi)(x_1, ..., x_n)[c_1, ..., c_n]$ iff $C \Vdash \phi(x_1, ..., x_n)[c_1, ..., c_n]$ and $C \Vdash \psi(x_1, ..., x_n)[c_1, ..., c_n]$.

$\boxed{\vee}$ $C \Vdash (\phi \vee \psi)(x_1, ..., x_n)[c_1, ..., c_n]$ iff there exist a cover $\{C_i \xrightarrow{f_i} C\}_{i \in I} \in \mathbf{J}^*(C)$ such that for each $i \in I$ one has $C_i \Vdash \phi(x_1, ..., x_n)[c_1 f_i, ..., c_n f_i]$ or $C_i \Vdash \psi(x_1, ..., x_n)[c_1 f_i, ..., c_n f_i]$.

$\boxed{\Rightarrow}$ $C \Vdash (\phi \Rightarrow \psi)(x_1, ..., x_n)[c_1, ..., c_n]$ iff for all morphisms $f : D \to C$ whenever $D \Vdash \phi(x_1, ..., x_n)[c_1 f, ..., c_n f]$ then $D \Vdash \psi(x_1, ..., x_n)[c_1 f, ..., c_n f]$.

Let y be a variable of the type Y interpreted by the object $\mathbf{Y}$ of $\mathcal{E}$.

$\boxed{\exists}$ $C \Vdash (\exists y \phi(x_1, ..., x_n, y))[c_1, ..., c_n]$ iff there exist a cover $\{C_i \xrightarrow{f_i} C\}_{i \in I} \in \mathbf{J}^*(C)$ such that for each $i \in I$ one has $C_i \Vdash \phi(x_1, ..., x_n, y)[c_1 f_i, ..., c_n f_i, d]$ for some $d \in \mathbf{Y}(C_i)$.

$\boxed{\forall}$ $C \Vdash (\forall y \phi(x_1, ..., x_n, y))[c_1, ..., c_n]$ iff for all morphisms $f : D \to C$ and all $d \in \mathbf{Y}(D)$ one has $D \Vdash \phi(x_1, ..., x_n, y)[c_1 f, ..., c_n f, d]$.

We have the following derivable *local character* and *monotonicity* laws

$\boxed{\text{LC}}$ If $\{C_i \xrightarrow{f_i} C\}_{i \in I} \in \mathbf{J}^*(C)$ and $C_i \Vdash \phi(x_1, ..., x_n)[c_1 f_i, ..., c_n f_i]$ for all $i \in I$, then $C \Vdash \phi(x_1, ..., x_n)[c_1, ..., c_n]$.

$\boxed{\text{M}}$ If $C \Vdash \phi(x_1, ..., x_n)[c_1, ..., c_n]$ and $f : D \to C$ then $D \Vdash \phi(x_1, ..., x_n)[c_1 f, ..., c_n f]$.

Let T be a theory in the language $\mathcal{L}[V_1, ..., V_n]$ a model of a theory T in the topos $\mathcal{E}$ is given by an interpretation of $\mathcal{L}[V_1, ..., V_n]$ such that for all objects C of $\mathcal{C}$ one has $C \Vdash \phi$ for every sentence ϕ of T.

Fact 3.8. *The deduction system IHOL is sound with respect to topos semantics.* [Awodey, 1997]

Since Kripke–Joyal sheaf semantics is a special case of topos semantics [MacLane and Moerdijk, 1992, Ch. 6], this implies soundness of the deduction system with respect to Kripke–Joyal sheaf semantics.

Part A

Algebra: Newton-Puiseux Theorem

Introduction

Newton–Puiseux Theorem states that, for an algebraically closed field K of zero characteristic, given a polynomial $F \in K[[X]][Y]$ there exist a positive integer m and a factorization $F = \prod_{i=1}^{n}(Y - \eta_i)$ where each $\eta_i \in K[[X^{1/m}]][Y]$. These roots η_i are called the *Puiseux expansions* of F. The theorem was first proved by Newton [1736] with the use of Newton polygon. Later, Puiseux [1850] gave an analytic proof. It is worth mentioning that while the proof by Puiseux [1850] deals only with convergent power series over the field of complex numbers, the much earlier proof by Newton [1736] was algorithmic in nature and applies to both convergent and non-convergent power series [Abhyankar, 1976].

Newton–Puiseux Theorem is usually stated as: *The field of fractional power series (also known as the field of Puiseux series), i.e. the field* $K\langle\langle X \rangle\rangle = \bigcup_{m \in \mathbb{Z}^+} K((X^{1/m}))$*, is algebraically closed* [Walker, 1978].

Abhyankar [1990] presents another proof of this result, the "Shreedharacharya's Proof of Newton's Theorem". This proof is not constructive as it stands. Indeed it assumes decidable equality on the ring $K[[X]]$ of power series over a field, but given two arbitrary power series we cannot decide whether they are equal in *finite* number of steps. We explain in Chapter II how to modify his argument by adding a separability assumption to provide a constructive proof of the result: The field of fractional power series is *separably* closed. In particular, the termination of Newton–Puiseux algorithm is justified constructively in this case. This termination is justified by a non constructive reasoning in most references [Walker, 1978; Duval, 1989; Abhyankar, 1990], with the exception of [Edwards, 2005]. Following that, we show that the field of fractional power series algebraic over $K(X)$ is algebraically closed.

The remainder of this part is dedicated to analyzing in a constructive framework what happens if the field K is not supposed to be algebraically closed. This is achieved through the method of *dynamic evaluation* [Della Dora et al., 1985], which replaces factorization by gcd computations. The reference [Coste et al., 2001] provides a proof theoretic analysis of this method. In Chapter III, we build a sheaf theoretic model of dynamic evaluation. The site is given by the category of étale algebras over the base field with an appropriate Grothendieck topology. We prove constructively that the topos of sheaves on this site contains a separably closed extension of the base field. We also show that in characteristic 0 the *axiom of choice* fails to hold in this topos.

With this model we obtain, as presented in Chapter IV, a dynamic version of Newton–Puiseux theorem, where we compute the Puiseux expansions of a polynomial $F \in K[X, Y]$ where K is not necessarily algebraically closed. The Puiseux expansions in this case are fractional power series over an étale K-algebra. We then present a characterization of the minimal algebra extension of K required for factorization of F and we show that while there is more than one such minimal extension, any two of them are powers of a common K-algebra.

II

Constructive Newton–Puiseux Theorem

A polynomial over a ring is said to be *separable* if it is coprime with its derivative. A field K is *algebraically closed* if any polynomial over K has a root in K. A field K is *separably algebraically closed* if every separable polynomial over K has a root in K. The goal in this chapter is to prove using only constructive reasoning the statement:

Claim 0.1. *For an algebraically closed field K, the field $K\langle\langle X\rangle\rangle$ of franctional power series over K*

$$K\langle\langle X\rangle\rangle = \bigcup_{m\in\mathbb{Z}^+} K((X^{1/m}))$$

is separably algebraically closed.

The proof we present is based on a non-constructive proof by Abhyankar [1990].

1 Algebraic preliminaries

A (discrete) field is defined to be a non trivial ring in which any element is 0 or invertible. For a ring R, the formal power series ring $R[[X]]$ is the set of sequences $\alpha = \alpha(0) + \alpha(1)X + \alpha(2)X^2 + ...$, with $\alpha(i) \in R$ [Mines et al., 1988].

Definition 1.1 (Apartness)**.** A binary relation $R \subset S \times S$ on a set S is an *apartness* if for all $x, y, z \in S$

(i.) $\neg xRx$.

(ii.) $xRy \Rightarrow yRx$.

(iii.) $xRy \Rightarrow xRz \vee yRz$.

We write $x \# y$ to mean xRy where R is an apartness relation on the set of which x and y are elements. As is the case with equality, the set on which the apartness is defined it is usually clear from the context . An apartness is *tight* if it satisfies $\neg x \# y \Rightarrow x = y$.

Definition 1.2 (Ring with apartness)**.** A ring with apartness is a ring R equipped with an apartness relation # such that

(i.) $0 \# 1$.

(ii.) $x_1 + y_1 \# x_2 + y_2 \Rightarrow x_1 \# x_2 \vee y_1 \# y_2$.

(iii.) $x_1 y_1 \# x_2 y_2 \Rightarrow x_1 \# x_2 \vee y_1 \# y_2$.

See [Mines et al., 1988; Troelstra and van Dalen, 1988].

Next we define the apartness relation on power series as in [Troelstra and van Dalen, 1988, Ch 8].

Definition 1.3. Let R be a ring with apartness. For $\alpha, \beta \in R[[X]]$ we define $\alpha \# \beta$ if $\exists n\ \alpha(n) \# \beta(n)$.

The relation # on $R[[X]]$ as defined above is an apartness relation and makes $R[[X]]$ into a ring with apartness [Troelstra and van Dalen, 1988]. The relation # on $R[[X]]$ restricts to an apartness relation on the ring of polynomials $R[X] \subset R[[X]]$.

We note that, if K is a discrete field then for $\alpha \in K[[X]]$ we have $\alpha \# 0$ iff $\alpha(j)$ is invertible for some j. For $F = \alpha_0 Y^n + ... + \alpha_n \in K[[X]][Y]$, we have $F \# 0$ iff $\alpha_i(j)$ is invertible for some j and $0 \leq i \leq n$.

Let R be a commutative ring with apartness. Then R is an *integral domain* if it satisfies $x \# 0 \wedge y \# 0 \Rightarrow xy \# 0$ for all $x, y \in R$. A *Heyting* field is an integral domain satisfying $x \# 0 \Rightarrow \exists y\ xy = 1$. The Heyting field of fractions of R is the Heyting field obtained by inverting the elements $c \# 0$ in R and taking the quotient by the appropriate equivalence relation, see [Troelstra and van Dalen, 1988, Ch 8,Theorem 3.12]. For a and $b \# 0$ in R we have $a/b \# 0$ iff $a \# 0$.

For a discrete field K, an element $\alpha \# 0$ in $K[[X]]$ can be written as $X^m \sum_{i \in \mathbb{N}} a_i X^i$ with $m \in \mathbb{N}$ and $a_0 \neq 0$. It follows that the ring $K[[X]]$ is an integral domain. If $a_0 \neq 0$ we have that $\sum_{i \in \mathbb{N}} a_i X^i$ is invertible in

$K[[X]]$. We denote by $K((X))$, the Heyting field of fractions of $K[[X]]$, we also call it the Heyting field of Laurent series over K. Thus an element apart from 0 in $K((X))$ can be written as $X^n \sum_{i \in \mathbb{N}} a_i X^i$ with $a_0 \neq 0$ and $n \in \mathbb{Z}$, i.e. as a series where finitely many terms have negative exponents.

Unless otherwise qualified, in what follows, a field will always denote a discrete field.

Definition 1.4 (Separable polynomial)**.** Let R be a ring. A polynomial $p \in R[X]$ is separable if there exist $r, s \in R[X]$ such that $rp + sp' = 1$, where $p' \in R[X]$ is the derivative of p.

Lemma 1.5. *Let R be a ring and $p \in R[X]$ separable. If $p = fg$ then both f and g are separable.*

Proof. Let $rp + sp' = 1$ for $r, s \in R[X]$. Then $rfg + s(fg' + f'g) = (rf + sf')g + sfg' = 1$, thus g is separable. Similarly for f. □

Lemma 1.6. *Let R be a ring. If $p(X) \in R[X]$ is separable and $u \in R$ is a unit then $p(uY) \in R[Y]$ is separable.*

The following result is usually proved with the assumption that a polynomial over a field can be decomposed into irreducible factors. This assumption cannot be shown to hold constructively, see [Fröhlich and Shepherdson, 1956; Waerden, 1930]. We give a proof without this assumption. It works over a field of any characteristic.

Lemma 1.7. *Let f be a monic polynomial in $K[X]$ where K is a field. If f' is the derivative of f and g monic is the gcd of f and f' then writing $f = hg$ we have that h is separable.*

Proof. Let a be the gcd of h and h'. We have $h = l_1 a$. Let d be the gcd of a and a'. We have $a = l_2 d$ and $a' = m_2 d$, with l_2 and m_2 coprime.

The polynomial a divides $h' = l_1 a' + l_1' a$ and hence that $a = l_2 d$ divides $l_1 a' = l_1 m_2 d$. It follows that l_2 divides $l_1 m_2$ and since l_2 and m_2 are coprime, that l_2 divides l_1.

Also, if a^n divides p then $p = qa^n$ and $p' = q'a^n + nqa'a^{n-1}$. Hence da^{n-1} divides p'. Since l_2 divides l_1, this implies that $a^n = l_2 d a^{n-1}$ divides $l_1 p'$. So a^{n+1} divides $al_1 p' = hp'$.

Since a divides f and f', a divides g. We show that a^n divides g for all n by induction on n. If a^n divides g we have just seen that a^{n+1} divides $g'h$. Also a^{n+1} divides $h'g$ since a divides h'. So a^{n+1} divides $g'h + h'g = f'$. On the other hand, a^{n+1} divides $f = hg = l_1 ag$. So a^{n+1} divides g which is the gcd of f and f'.

This implies that a is a unit. □

The intuition is that the separable divisor h of a polynomial f is a separable polynomial that have a common root with f. However, this intuition is not entirely correct. Over a field with non-zero characteristic it could be the case that the derivative f' vanishes. In that case h is a unit, i.e. a constant polynomial.

Corollary 1.8. *Let K be a field of any characteristic and $f \in K[X]$ a non-constant monic polynomial. If the derivative $f' \neq 0$ then there is a non-constant separable divisor of f.*

Proof. By Lemma 1.7 we have $f = gh$ and $f' = gr$ where h is separable. Since f' is non-zero we have that g is a non-zero polynomial of degree less than or equal $\deg(f')$. But $\deg(f') < \deg(f)$ and thus $\deg(g) < \deg(f)$. We have then that h is non-constant □

Corollary 1.9. *Let K be a field of characteristic 0 and $f \in K[X]$ a non-constant monic polynomial. Then f has a non-constant separable divisor.*

Corollary 1.10. *Let K be a field of characteristic 0. If K is separably algebraically closed then K is algebraically closed*

If F is in $R[[X]][Y]$, by F_Y we mean the derivative of F with respect to Y.

Lemma 1.11. *Let K be a field and let $F = \sum_{i=0}^{n} \alpha_i Y^{n-i} \in K[[X]][Y]$ be separable over $K((X))$, then $\alpha_n \# 0 \vee \alpha_{n-1} \# 0$*

Proof. Since F is separable over $K((X))$ we have $PF + QF_Y = \gamma \# 0$ for $P, Q \in K[[X]][Y]$ and $\gamma \in K[[X]]$. From this we get that γ is equal to the constant term on the left hand side, i.e. $P(0)\alpha_n + Q(0)\alpha_{n-1} = \gamma \# 0$. Thus $\alpha_n \# 0 \vee \alpha_{n-1} \# 0$. □

2 Newton–Puiseux theorem

One key of Abhyankar's proof is Hensel's Lemma. Here we formulate a little more general version than the one in [Abhyankar, 1990] by dropping the assumption that the base ring is a field.

Lemma 2.1 (Hensel's Lemma)**.** *Let R be a ring and $F(X,Y) = Y^n + \sum_{i=1}^{n} a_i(X)\, Y^{n-i}$ be a monic polynomial in $R[[X]][Y]$ of degree $n > 1$. Given monic non-constant polynomials $G_0, H_0 \in R[Y]$ of degrees r and s respectively. Given $H^*, G^* \in R[Y]$ such that $F(0,Y) = G_0 H_0$, $r + s = n$ and $G_0 H^* + H_0 G^* = 1$. We can find $G(X,Y), H(X,Y) \in R[[X]][Y]$ of degrees r*

and s respectively, such that $F(X,Y) = G(X,Y)H(X,Y)$, $G(0,Y) = G_0$ and $H(0,Y) = H_0$.

Proof. The proof is almost the same as Abhyankar's [Abhyankar, 1990], we present it here for completeness.
Since $R[[X]][Y] \subsetneq R[Y][[X]]$, we can rewrite $F(X,Y)$ as a power series in X with coefficients in $R[Y]$. Let

$$F(X,Y) = F_0(Y) + F_1(Y)X + ... + F_q(Y)X^q + ...$$

with $F_i(Y) \in R[Y]$. Now we want to find $G(X,Y), H(X,Y) \in R[Y][[X]]$ such that $F = GH$. If we let $G = G_0 + \sum_{i=1}^{\infty} G_i(Y)X^i$ and $H = H_0 + \sum_{i=1}^{\infty} H_i(Y)X^i$, then for each q we need to find $G_i(Y), H_j(Y)$ for $i,j \leq q$ such that $F_q = \sum_{i+j=q} G_i H_j$. We also need $\deg G_k < r$ and $\deg G_\ell < s$ for $k, \ell > 0$.
We find such G_i, H_j by induction on q. We have that $F_0 = G_0 H_0$. Assume that for some $q > 0$ we have found all G_i, H_j with $\deg G_i < r$ and $\deg H_i < s$ for $1 \leq i < q$ and $1 \leq j < q$. Now we need to find H_q, G_q such that

$$F_q = G_0 H_q + H_0 G_q + \sum_{\substack{i+j=q \\ i<q, j<q}} G_i H_j$$

We let

$$U_q = F_q - \sum_{\substack{i+j=q \\ i<q, j<q}} G_i H_j$$

One can see that $\deg U_q < n$. We are given that $G_0 H^* + H_0 G^* = 1$. Multiplying by U_q we get $G_0 H^* U_q + H_0 G^* U_q = U_q$. By Euclidean division we can write $U_q H^* = E_q H_0 + H_q$ for some E_q, H_q with $\deg H_q < s$. Thus we write $U_q = G_0 H_q + H_0(E_q G_0 + G^* U_q)$. One can see that $\deg H_0(E_q G_0 + G^* U_q) < n$ since $\deg(U_q - G_0 H_q) < n$. Since H_0 is monic of degree s , $\deg(E_q G_0 + G^* U_q) < r$. We take $G_q = E_q G_0 + G^* U_q$. Now, we can write $G(X,Y)$ and $H(X,Y)$ as monic polynomials in Y with coefficients in $R[[X]]$, with degrees r and s respectively. □

It should be noted that the uniqueness of the factors G and H proven in [Abhyankar, 1990] may not necessarily hold when R is not an integral domain.

If $\alpha = \sum \alpha(i) X^i$ is an element of $R[[X]]$ we write $m \leqslant \operatorname{ord} \alpha$ to mean that $\alpha(i) = 0$ for $i < m$ and we write $m = \operatorname{ord} \alpha$ to mean furthermore that $\alpha(m)$ is invertible.

Lemma 2.2. *Let K be an algebraically closed field of characteristic zero. Let $F(X,Y) = Y^n + \sum_{i=1}^{n} \alpha_i(X)Y^{n-i} \in K[[X]][Y]$ be a monic non-constant polynomial of degree $n \geq 2$ separable over $K((X))$. Then there exist $m > 0$ and a proper factorization $F(T^m, Y) = G(T,Y)H(T,Y)$ with G and H in $K[[T]][Y]$.*

Proof. Assume w.l.o.g. that $\alpha_1(X) = 0$. This is Shreedharacharya's[1] trick [Abhyankar, 1990] (a simple change of variable $F(X, W - \alpha_1/n)$). The simple case is if we have ord $\alpha_i = 0$ for some $1 < i \leq n$. In this case $F(0,Y) = Y^n + d_2Y^{n-1} + ... + d_n \in K[Y]$ and $d_i \neq 0$. Thus $\forall a \in K\; F(0,Y) \neq (Y-a)^n$. For any root b of $F(0,b) = 0$ we have then a proper decomposition $F(0,Y) = (Y-b)^p H$ with $Y - b$ and H coprime, and we can use Hensel's Lemma 2.1 to conclude (In this case we can take $m = 1$).

In general, we know by Lemma 1.11 that for $k = n$ or $k = n-1$ we have $\alpha_k(X)$ is apart from 0. We then have $\alpha_k(\ell)$ invertible for some ℓ. We can then find p and m, $1 < m \leq n$, such that $\alpha_m(p)$ is invertible and $\alpha_i(j) = 0$ whenever $j/i < p/m$. We can then write

$$F(T^m, T^p Z) = T^{np}(Z^n + c_2(T)Z^{n-2} + \cdots + c_n(T))$$

with ord $c_m = 0$. As in the simple case, we have a proper decomposition

$$Z^n + c_2(T)Z^{n-2} + \cdots + c_n(T) = G_1(T,Z)H_1(T,Z)$$

with $G_1(T,Z)$ monic of degree l in Z and $H_1(T,Z)$ monic of degree q in Z, with $l + q = n$, $l < n$, $q < n$. We then take

$$G(T,Y) = T^{lp}G_1(T, Y/T^p)$$
$$H(T,Y) = T^{qp}H_1(T, Y/T^p)$$

□

Theorem 2.3. *Let K be an algebraically closed field of characteristic zero. Let $F(X,Y) = Y^n + \sum_{i=1}^{n} \alpha_i(X)Y^{n-i} \in K[[X]][Y]$ be a monic non-constant polynomial separable over $K((X))$. Then there exist a positive integer m and factorization*

$$F(T^m, Y) = \prod_{i=1}^{n} (Y - \eta_i) \qquad \eta_i \in K[[T]]$$

Proof. If $F(X,Y)$ is separable over $K((X))$ then $F(T^m, Y)$ for some positive integer m is separable over $K((T))$. The proof follows directly from Lemma 1.5 and Lemma 2.2 by induction. □

[1] Shreedharacharya's trick is also known as Tschirnhaus's trick [von Tschirnhaus and Green, 2003]. The technique of removing the second term of a polynomial equation was also known to Descartes [Descartes, 1637].

Corollary 2.4. *Let K be an algebraically closed field of characteristic zero. The Heyting field of fractional power series over K is separably algebraically closed.*

Proof. Let $F(X,Y) \in K((X))[Y]$ be a monic separable polynomial of degree $n > 1$. Let $\beta \# 0$ be the product of the denominators of the coefficients of F. Then we can write $F(X, \beta^{-1}Z) = \beta^{-n}G$ for $G \in K[[X]][Z]$. By Lemma 1.6 we get that F, hence G, is separable in Z over $K((X))$. By Theorem 2.3, $G(T^m, Z)$ factors linearly over $K[[T]]$ for some positive integer m. Consequently we get that $F(T^m, Y)$ factors linearly over $K((T))$. □

3 Related results

In the following we show that the elements in $K\langle\langle X\rangle\rangle$ algebraic over $K(X)$ form a discrete algebraically closed field.

Lemma 3.1. *Let K be a field and*

$$F(X,Y) = Y^n + b_1 Y^{n-1} + ... + b_n \in K(X)[Y]$$

be a non-constant monic polynomial such that $b_n \neq 0$. If $\gamma \in K((T))$ is a root of $F(T^q, Y)$, then ord $\gamma \leq d$ *for some positive integer d.*

Proof. We can find $h \in K[X]$ such that

$$G = hF = a_0(X)Y^n + a_1(X)Y^{n-1} + ... + a_n(X) \in K[X][Y]$$

with $a_n \neq 0$. Let $d = \text{ord}\ a_n(T^q)$. If ord $\gamma > d$ then so is ord $a_i\gamma^{n-i}$ for $0 \leq i < n$. But we know that in a_n there is a non-zero term with T-degree d. Thus $G(T^q, \gamma) \# 0$; Consequently $F(T^q, \gamma) \# 0$ □

Note that if $\alpha, \beta \in K\langle\langle X\rangle\rangle$ are algebraic over $K(X)$ then $\alpha + \beta$ and $\alpha\beta$ are algebraic over $K(X)$ [Mines et al., 1988, Ch 6, Corollary 1.4].

Lemma 3.2. *Let K be a field. The set of elements in $K\langle\langle X\rangle\rangle$ algebraic over $K(X)$ is a discrete set; More precisely # is decidable on this set.*

Proof. It suffices to show that for an element γ in this set $\gamma \# 0$ is decidable. Let $F = Y^n + a_1(X)Y^{n-1} + ... + a_n \in K(X)[Y]$ be a monic non-constant polynomial. Let $\gamma \in K((T))$ be a root of $F(T^q, Y)$. If $F = Y^n$ then $\neg\gamma \# 0$. Otherwise, F can be written as $Y^m(Y^{n-m} + ... + a_m)$ with $0 \leq m < n$ and $a_m \neq 0$. By Lemma 3.1 we can find d such that any element in $K((T))$ that is a root of $Y^{n-m} + ... + a_m$ has an order less than or equal to d. Thus $\gamma \# 0$ if an only if ord $\gamma \leq d$. □

If $\alpha \# 0 \in K\langle\langle X\rangle\rangle$ is algebraic over $K(X)$ then $1/\alpha$ is algebraic over $K(X)$. Thus the set of elements in $K\langle\langle X\rangle\rangle$ algebraic over $K(X)$ form a field $K\langle\langle X\rangle\rangle^{alg} \subset K\langle\langle X\rangle\rangle$. This field is in fact algebraically closed *in* $K\langle\langle X\rangle\rangle$ [Mines et al., 1988, Ch 6, Corollary 1.5].

Since for an algebraically closed field K we have shown $K\langle\langle X\rangle\rangle$ to be only *separably* algebraically closed, we need a stronger argument to show that $K\langle\langle X\rangle\rangle^{alg}$ is algebraically closed.

Lemma 3.3. *For an algebraically closed field K of characteristic zero, the field $K\langle\langle X\rangle\rangle^{alg}$ is algebraically closed.*

Proof. Let $F \in K\langle\langle X\rangle\rangle^{alg}[Y]$ be a monic non-constant polynomial of degree n. By Lemma 3.2 $K\langle\langle X\rangle\rangle^{alg}$ is a discrete field. By Lemma 1.7 we can decompose F as $F = HG$ with $H \in K\langle\langle X\rangle\rangle^{alg}[Y]$ a non-constant monic separable polynomial. By Corollary 2.4, H has a root η in $K\langle\langle X\rangle\rangle$. Since $K\langle\langle X\rangle\rangle^{alg}$ is algebraically closed in $K\langle\langle X\rangle\rangle$ we have that $\eta \in K\langle\langle X\rangle\rangle^{alg}$. □

We can draw similar conclusions in the case of real closed fields [2].

Lemma 3.4. *Let R be a real closed field. Then*

(i.) For any $\alpha \# 0 \in R\langle\langle X\rangle\rangle$ we can find $\beta \in R\langle\langle X\rangle\rangle$ such that $\beta^2 = \alpha$ or $-\beta^2 = \alpha$.

(ii.) A separable monic polynomial of odd degree in $R\langle\langle X\rangle\rangle[Y]$ has a root in $R\langle\langle X\rangle\rangle$.

Proof. Since R is real closed, the first statement follows from the fact an element $a_0 + a_1X + ... \in R[[X]]$ with $a_0 > 0$ has a square root in $R[[X]]$. Let $F(X,Y) = Y^n + \alpha_1 Y^{n-1} + ... + \alpha_n \in R[[X]][Y]$ be a monic polynomial of odd degree $n > 1$ separable over $R((X))$. We can assume w.l.o.g. that $\alpha_1 = 0$. Since F is separable, i.e. $PF + QF_Y = 1$ for some $P, Q \in R((X))[Y]$, then by a similar construction to that in Lemma 2.2 we can write $F(T^m, T^pZ) = T^{np}V$ for $V \in R[[T]][Z]$ such that $V(0,Z) \neq (Z+a)^n$ for all $a \in R$. Since R is real closed and $V(0,Z)$ has odd degree, $V(0,Z)$ has a root r in R. We can find proper decomposition into coprime factors $V(0,Z) = (Z-r)^\ell q$. By Hensel's Lemma 2.1, we lift those factors to factors of V in $R[[T]][Z]$ thus we can write $F = GH$ for monic non-constant $G, H \in R[[T]][Y]$. By Lemma 1.5 both G and H are separable. Either G or H has odd degree. Assuming G has odd degree greater than 1, we can further factor G into non-constant factors. The statement follows by induction. □

[2]We reiterate that by a field we mean a discrete field.

Let R be a real closed field. By Lemma 3.2 we see that $R\langle\langle X\rangle\rangle^{alg}$ is discrete. A non-zero element in $\alpha \in R\langle\langle X\rangle\rangle^{alg}$ can be written $\alpha = X^{m/n}(a_0 + a_1 X^{1/n} + ...)$ for $n > 0, m \in \mathbb{Z}$ with $a_0 \neq 0$. Then α is positive iff its initial coefficient a_0 is positive [Basu et al., 2006]. We can then see that this makes $R\langle\langle X\rangle\rangle^{alg}$ an ordered field.

Lemma 3.5. *For a real closed field R, the field $R\langle\langle X\rangle\rangle^{alg}$ is real closed.*

Proof. Let $\alpha \in R\langle\langle X\rangle\rangle^{alg}$. Since $R\langle\langle X\rangle\rangle^{alg}$ is discrete, by Lemma 3.4 we can find $\beta \in R\langle\langle X\rangle\rangle^{alg}$ such that $\beta^2 = \alpha$ or $-\beta^2 = \alpha$.
Let $F \in R\langle\langle X\rangle\rangle^{alg}[Y]$ be a monic polynomial of odd degree n. Applying Lemma 1.7 several times, by induction we have $F = H_1 H_2 .. H_m$ with $H_i \in R\langle\langle X\rangle\rangle^{alg}[Y]$ separable non-constant monic polynomial. For some i we have H_i of odd degree. By Lemma 3.4, H_i has a root in $R\langle\langle X\rangle\rangle^{alg}$. Thus F has a root in $R\langle\langle X\rangle\rangle^{alg}$. □

III

The Separable Algebraic Closure

In Section 1 we describe the category $\mathcal{A}_K$ of étale K-algebras. In Section 2 we specify a coverage $\mathbf{J}$ on the category $\mathcal{A}_K^{op}$. In Section 3 we demonstrate that the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ contains a separably algebraically closed extension of K. In Section 5 and Section 6 we look at the logical properties of the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ with respect to choice axioms and booleanness.

1 The category of Étale K-Algebras

We recall the definition of separable polynomial from Chapter II.

Definition 1.1 (Separable polynomial)**.** Let R be a ring. A polynomial $p \in R[X]$ is separable if there exist $r, s \in R[X]$ such that $rp + sp' = 1$, where $p' \in R[X]$ is the derivative of p.

Let K be a discrete field and A a K-algebra. An element $a \in A$ is *separable algebraic* if it is the root of a separable polynomial over K. The algebra A is *separable algebraic* if all elements of A are separable algebraic. An algebra over a field is said to be *finite* if it has finite dimension as a vector space over K. We note that if A is a finite K-algebra then we have a finite basis of A as a vector space over K.

Definition 1.2. An algebra A over a field K is *étale* if it is finite and separable algebraic.

It is worth mentioning that there is an elementary characterization of étale K-algebras given as follows: Let A be a finite K-algebra with basis $(a_1, ..., a_n)$. We associate to each element $a \in A$ the matrix representation $[m_a] \in \mathrm{M}(n, K)$ of the K-linear map $x \mapsto ax$. Let $\mathrm{Tr}_{A/K}(a)$ be the trace of $[m_a]$. Let $\mathrm{disc}_{A/K}(x_1, ..., x_n) = \det((\mathrm{Tr}_{A/K}(x_i x_j))_{1 \le i,j \le n})$. The algebra A is étale if $\mathrm{disc}_{A/K}(a_1, ..., a_n)$ is a unit. The equivalence between Definition 1.2 and this characterization is shown in [Lombardi and Quitté, 2011, Ch. 6, Theorem 1.7].

Definition 1.3 (Regular ring)**.** A commutative ring R is *(von Neumann)* regular if for every element $a \in R$ there exist $b \in R$ such that $aba = a$ and $bab = b$. This element b is called the quasi-inverse of a.

The quasi-inverse b of an element a is unique for a [Lombardi and Quitté, 2011, Ch 4]. We thus use the notation a^* to refer to the quasi-inverse of a. A ring is regular iff it is zero-dimensional, i.e. any prime ideal is maximal, and reduced, i.e. $a^n = 0 \Rightarrow a = 0$. To be von Neumann regular is equivalent to the fact that any principal ideal (and hence any finitely generated ideal) is generated by an idempotent. If a is an element in R then the element $e = aa^*$ is an idempotent such that $\langle e \rangle = \langle a \rangle$ and R is isomorphic to $R_0 \times R_1$ with $R_0 = R/\langle e \rangle$ and $R_1 = R/\langle 1 - e \rangle$. Furthermore a is 0 on the component R_0 and invertible on the component R_1.

Definition 1.4 (Fundamental system of orthogonal idempotents)**.** A family $(e_i)_{i \in I}$ of idempotents in a ring R is a fundamental system of orthogonal idempotents if $\sum_{i \in I} e_i = 1$ and $\forall i, j [i \neq j \Rightarrow e_i e_j = 0]$.

Lemma 1.5. *Given a fundamental system of orthogonal idempotents $(e_i)_{i \in I}$ in a ring A we have a decomposition $A \cong \prod_{i \in I} A/\langle 1 - e_i \rangle$.*

Proof. Follows directly by induction from the fact that $A \cong A/\langle e \rangle \times A/\langle 1 - e \rangle$ for an idempotent $e \in A$. □

Fact 1.6.

1. *An étale algebra over a field K is zero-dimensional and reduced, i.e. regular.*

2. *Let A be a finite K-algebra and $(e_i)_{i \in I}$ a fundamental system of orthogonal idempotents of A. Then A is étale if and only if $A/\langle 1 - e_i \rangle$ is étale for each $i \in I$.*

[Lombardi and Quitté, 2011, Ch 6, Fact 1.3].

Note that an étale K-algebra A is finitely presented, i.e. can be written as $K[X_1, ..., X_n]/\langle f_1, ..., f_m \rangle$.

We define strict Bézout rings as in [Lombardi and Quitté, 2011, Ch 4].

Definition 1.7. A ring R is a (strict) Bézout ring if for all $a, b \in R$ we can find $g, a_1, b_1, c, d \in R$ such that $a = a_1 g$, $b = b_1 g$ and $ca_1 + db_1 = 1$.

If R is a regular ring then $R[X]$ is a strict Bézout ring (and the converse is true [Lombardi and Quitté, 2011]). Intuitively we can compute the gcd as if R was a field, but we may need to split R when deciding if an element is invertible or 0. Using this, we see that given a, b in $R[X]$ we can find a decomposition $R_1, \ldots, R_n$ of R and for each i we have g, a_1, b_1, c, d in $R_i[X]$ such that $a = a_1 g$, $b = b_1 g$ and $ca_1 + db_1 = 1$ with g monic. The degree of g may depend on i.

Lemma 1.8. *If A is an étale K-algebra and p in $A[X]$ is a separable polynomial then $A[a] = A[X]/\langle p \rangle$ is an étale K-algebra.*

Proof. See [Lombardi and Quitté, 2011, Ch 6, Lemma 1.5]. □

By a *separable extension* of a ring R we mean a ring $R[a] = R[X]/\langle p \rangle$ where $p \in R[X]$ is non-constant, monic and separable.

In order to build the classifying topos of a coherent theory T it is customary in the literature to consider the category of all finitely presented T_0 algebras where T_0 is an equational subtheory of T. The axioms of T then give rise to a coverage on the dual category [Makkai and Reyes, 1977, Ch. 9]. For our purpose consider the category $\mathcal{C}$ of finitely presented K-algebras. Given an object R of $\mathcal{C}$, the axiom schema of separable algebraic closure and the field axiom give rise to families

(i.) $R \to R[X]/\langle p \rangle$ where $p \in R[X]$ is monic and separable.

(ii.) $R \nearrow R/\langle a \rangle$, $R \searrow R[\frac{1}{a}]$, for $a \in R$.

Dualized, these are covering families of R in $\mathcal{C}^{op}$. We observe however that we can limit our consideration only to étale K-algebras. In this case we can assume a is an idempotent.

We study the small category $\mathcal{A}_K$ of étale K-algebras over a fixed field K and K-homomorphisms. First we fix an infinite set of names S. An object of $\mathcal{A}_K$ is an étale algebra of the form $K[X_1, ..., X_n]/\langle f_1, ..., f_m \rangle$ where $X_i \in S$ for all $1 \leq i \leq n$. Note that for each object R, there is a unique morphism $K \to R$. If A and B are objects of $\mathcal{A}_K$ and $\varphi : A \to B$ is a morphism of $\mathcal{A}_K$, the diagram

$$\begin{array}{ccc} & K & \\ \swarrow & & \searrow \\ A & \xrightarrow{\varphi} & B \end{array}$$

commutes.The trivial ring 0 is the terminal object in the category $\mathcal{A}_K$ and K is its initial object.

2 A topology for $\mathcal{A}_K^{op}$

Next we specify a coverage $\mathbf{J}$ on the category $\mathcal{A}_K^{op}$ per Definition I.3.1. A coverage is specified by a collection $\mathbf{J}(A)$ of families of morphisms of $\mathcal{A}_K^{op}$ with codomain A for each object A. Rather than describing the collection $\mathbf{J}(A)$ directly, we define for each object A a collection $\mathbf{J}^{op}(A)$ of families of morphisms of $\mathcal{A}_K$ with domain A. Then we take $\mathbf{J}(A)$ to be the dual of $\mathbf{J}^{op}(A)$ in the sense that for any object A we have $\{\overline{\varphi_i} : A_i \to A\}_{i \in I} \in \mathbf{J}(A)$ *if and only if* $\{\varphi_i : A \to A_i\}_{i \in I} \in \mathbf{J}^{op}(A)$ where the morphism φ_i of $\mathcal{A}_K$ is the dual of the morphism $\overline{\varphi}_i$ of $\mathcal{A}_K^{op}$. We call $\mathbf{J}^{op}$ cocoverage and elements of $\mathbf{J}^{op}(A)$ elementary cocovers (elementary cocovering families) of A. Analogously we define the closure $\mathbf{J}^{*op}$ to be the dual of the closure $\mathbf{J}^*$(See Definition I.3.6). We call a family $T \in \mathbf{J}^{*op}(A)$ a cocover (cocovering family) of A.

Definition 2.1 (Topology for $\mathcal{A}_K^{op}$)**.** Let A be an object of $\mathcal{A}_K$.

(i.) If $(e_i)_{i \in I}$ is a fundamental system of orthogonal idempotents of A, then

$$\{A \xrightarrow{\varphi_i} A/\langle 1 - e_i \rangle\}_{i \in I} \in \mathbf{J}^{op}(A)$$

where for each $i \in I$, φ_i is the canonical homomorphism.

(ii.) Let $A[a]$ be a separable extension of A. We have

$$\{A \xrightarrow{\psi} A[a]\} \in \mathbf{J}^{op}(A)$$

where ψ is the canonical homomorphism.

Note that in particular 2.1.(i.) implies that the trivial algebra 0 is covered by the empty family of morphisms since an empty family of elements in this ring form a fundamental system of orthogonal idempotents (The empty sum equals $0 = 1$ and the empty product equals $1 = 0$). Also note that 2.1.(ii.) implies that $\{A \xrightarrow{1_A} A\} \in \mathbf{J}^{op}(A)$.

Lemma 2.2. *The collections* **J** *of Definition 2.1 is a coverage on* $\mathcal{A}_K^{op}$.

Proof. Let $\eta : R \to A$ be a morphism of $\mathcal{A}_K$ and let

$$S = \{\varphi_i : R \to R_i\}_{i \in I} \in \mathbf{J}^{op}(R)$$

We show that there exist a family $\{\psi_j : A \to A_j\}_{j \in J} \in \mathbf{J}^{op}(A)$ such that for each $j \in J$, $\psi_j \eta$ factors through φ_i for some $i \in I$. By duality, this implies **J** is a coverage on $\mathcal{A}_K^{op}$.

By case analysis on the clauses of Definition 2.1

(i.) If $S = \{\varphi_i : R \to R/\langle 1 - e_i \rangle\}_{i \in I}$, where $(e_i)_{i \in I}$ is a fundamental system of orthogonal idempotents of R. In A, the family $(\eta(e_i))_{i \in I}$ is fundamental system of orthogonal idempotents. We have an elementary cocover

$$\{\psi_i : A \to A/\langle 1 - \eta(e_i) \rangle\}_{i \in I} \in \mathbf{J}^{op}(A)$$

For each $i \in I$, the homomorphism η induces a K-homomorphism $\eta_{e_i} : R/\langle 1 - e_i \rangle \to A/\langle 1 - \eta(e_i) \rangle$ where $\eta_{e_i}(r + \langle 1 - e_i \rangle) = \eta(r) + \langle 1 - \eta(e_i) \rangle$. Since $\psi_i(\eta(r)) = \eta(r) + \langle 1 - \eta(e_i) \rangle$ we have that $\psi_i \eta$ factors through φ_i as illustrated in the commuting diagram below.

$$\begin{array}{ccc} R/\langle 1 - e_i \rangle & \xrightarrow{\eta_{e_i}} & A/\langle 1 - \eta(e_i) \rangle \\ \varphi_i \uparrow & & \uparrow \psi_i \\ R & \xrightarrow{\eta} & A \end{array}$$

(ii.) If $S = \{\varphi : R \to R[r]\}$ with $R[r]$ a separable extension, that is $R[r] = R[X]/\langle p \rangle$, with $p \in R[X]$ monic, non-constant, and separable. Let $sp + tp' = 1$. We have

$$\eta(s)\eta(p) + \eta(t)\eta(p') = \eta(s)\eta(p) + \eta(t)\eta(p)' = 1$$

Then $q = \eta(p) \in A[X]$ is separable. Let $A[a] = A[X]/\langle q \rangle$. We have an elementary cocover

$$\{\psi : A \to A[a]\} \in \mathbf{J}^{op}(A)$$

where ψ is the canonical embedding. Let $\zeta : R[r] \to A[a]$ be the K-homomorphism such that $\zeta|_R = \eta$ and $\zeta(r) = a$. For $b \in R$, we have $\psi(\eta(b)) = \zeta(\varphi(b))$, i.e. a commuting diagram

$$\begin{array}{ccc} R[r] & \xrightarrow{\zeta} & A[a] \\ \varphi \uparrow & & \uparrow \psi \\ R & \xrightarrow{\eta} & A \end{array}$$

□

Lemma 2.3. *Let* $\mathbf{P} : \mathcal{A}_K \to \mathbf{Set}$ *be a presheaf on* $\mathcal{A}_K^{op}$ *such that* $\mathbf{P}(0) = 1$. *Let* R *be an object of* $\mathcal{A}_K$ *and let* $(e_i)_{i \in I}$ *be a fundamental system of orthogonal idempotents of* R. *For each* $i \in I$, *let* $R_i = R/\langle 1 - e_i \rangle$ *and let* $\varphi_i : R \to R_i$ *be the canonical homomorphism. Any family* $\{s_i \in \mathbf{P}(R_i)\}$ *is a compatible family with respect to the family morphisms* $\{\varphi_i : R \to R_i\}_{i \in I}$.

Proof. For $i, j \in I$, let B be an object and let $\vartheta : R_i \to B$ and $\psi : R_j \to B$ be two morphisms such that $\vartheta\varphi_i = \psi\varphi_j$, i.e. we have a commuting diagram

$$\begin{array}{ccc} R/\langle 1 - e_i \rangle & \xrightarrow{\vartheta} & B \\ \varphi_i \uparrow & & \uparrow \psi \\ R & \xrightarrow{\varphi_j} & R/\langle 1 - e_j \rangle \end{array}$$

We will show that $\mathbf{P}(\vartheta)(s_i) = \mathbf{P}(\psi)(s_j)$.

(i.) If $i = j$, then since φ_i is surjective we have $\vartheta = \psi$ and $\mathbf{P}(\vartheta) = \mathbf{P}(\psi)$.

(ii.) If $i \neq j$, then since $e_i e_j = 0$, $\varphi_i(e_i) = 1$ and $\varphi_j(e_j) = 1$ we have $\varphi_j(e_i e_j) = \varphi_j(e_i) = 0$. But then

$$1 = \vartheta(1) = \vartheta(\varphi_i(e_i)) = \psi(\varphi_j(e_i)) = \psi(0) = 0$$

Hence B is the trivial algebra 0. By assumption $\mathbf{P}(0) = 1$, hence $\mathbf{P}(\vartheta)(s_i) = \mathbf{P}(\psi)(s_j) = *$.

□

Corollary 2.4. *Let* $\mathbf{F}$ *be a sheaf on* $(\mathcal{A}_K^{op}, \mathbf{J})$. *Let* R *be an object of* $\mathcal{A}_K$ *and* $(e_i)_{i \in I}$ *a fundamental system of orthogonal idempotents of* R. *Let* $R_i = R/\langle 1 - e_i \rangle$ *and* $\varphi_i : R \to R_i$ *be the canonical homomorphism. The map* $f : \mathbf{F}(R) \to \prod_{i \in I} \mathbf{F}(R_i)$ *such that* $f(s) = (\mathbf{F}(\varphi_i)s)_{i \in I}$ *is an isomorphism.*

Proof. Since $\mathbf{F}(0) = 1$, by Lemma 2.3 any family of elements of the form $\{s_i \in \mathbf{F}(R_i) \mid i \in I\}$ is compatible. Since $\mathbf{F}$ is a sheaf satisfying the sheaf axiom I.3.3, the family $\{s_i \in \mathbf{F}(R_i)\}_{i \in I}$ has a unique amalgamation $s \in \mathbf{F}(R)$ with restrictions $s\varphi_i = s_i$. The isomorphism is given by $fs = (s\varphi_i)_{i \in I}$. We can then use the tuple notation $(s_i)_{i \in I}$ to denote the element s in $\mathbf{F}(R)$. □

One say that a polynomial $f \in R[X]$ has a *formal degree* n if f can be written as $f = a_n X^n + ... + a_0$ which is to express that for any $m > n$

the coefficient of X^m is known to be 0. One, on the other hand, say that a polynomial f has a degree $n > 0$ if f has a formal degree n and the coefficient of X^n is not 0.

Lemma 2.5. *Let R be a regular ring and $p_1, p_2 \in R[X]$ be monic polynomials of degrees n_1 and n_2 respectively. Let $R[a,b] = R[X,Y]/\langle p_1(X), p_2(Y)\rangle$. Let $q_1, q_2 \in R[Z]$ be of formal degrees $m_1 < n_1$ and $m_2 < n_2$ respectively. If $q_1(a) = q_2(b)$ then $q_1 = q_2 = r \in R$.*

Proof. Let $q_1(a) = q_2(b)$, then in $R[X,Y]$

$$q_1(X) - q_2(Y) = f(X,Y)p_1(X) + g(X,Y)p_2(Y)$$

for some $f, g \in R[X,Y]$.

In $R[a][Y] = R[X,Y]/\langle p_1(X)\rangle$ we have $q_1(a) - q_2(Y) = g(a,Y)p_2(Y)$. But $p_2(Y)$ is monic of Y-degree n_2 while $q_2(Y) - q_1(a)$ has formal Y-degree $m_2 < n_2$, hence, the coefficients of $g(a,Y) \in R[a][Y]$ are all equal to 0 in $R[a]$. We have then that all coefficient of Y^ℓ with $\ell > 0$ in $q_2(Y)$ are equal 0. That is, $q_2 = r \in R$ and that $q_1(a)$ is equal to the constant coefficient r of $q_2(Y)$. Thus in $R[X]$ we have $q_1(X) - r = h(X)p_1(X)$ for some $h \in R[X]$. Similarly, since $(q_1(X) - r)$ has a formal X-degree m_1 and p_1 is monic of degree $n_1 > m_1$ we get that $q_1 = r \in R$. □

Corollary 2.6. *Let R be an object of $\mathcal{A}_K$ and $p \in R[X]$ separable and monic. Let $R[a] = R[X]/\langle p\rangle$ and $\varphi : R \to R[a]$ the canonical morphism. Let $R[b,c] = R[X,Y]/\langle p(X), p(Y)\rangle$. The commuting diagram*

$$\begin{array}{ccc} R[a] & \xrightarrow{\vartheta} & R[b,c] \\ \varphi\uparrow & & \uparrow\zeta \\ R & \xrightarrow{\varphi} & R[a] \end{array} \qquad \vartheta|_R(r) = \zeta|_R(r) = r,\ \vartheta(a) = b,\ \zeta(a) = c$$

is a pushout diagram of $\mathcal{A}_K$. Moreover, φ is the equalizer of ζ and ϑ.

Proof. Let $R[a] \underset{\psi}{\overset{\eta}{\rightrightarrows}} B$ be morphisms of $\mathcal{A}_K$ such that $\eta\varphi = \psi\varphi$. Then for all $r \in R$ we have $\eta(r) = \psi(r)$.

Let $\gamma : R[c,d] \to B$ be the homomorphism such that $\gamma(r) = \eta(r) = \psi(r)$ for all $r \in R$ while $\gamma(b) = \eta(a), \gamma(c) = \psi(a)$. Then γ is the unique map such that $\gamma\vartheta = \eta$ and $\gamma\zeta = \psi$ and we have proved that the above diagram is a pushout diagram.

Let A be an object of $\mathcal{A}_K$ and let $\varrho : A \to R[a]$ be a map such that $\zeta\varrho = \vartheta\varrho$. By Lemma 2.5 if for $f \in R[a]$ one has $\zeta(f) = \vartheta(f)$ then $f \in R$

(i.e. f is of degree 0 as a polynomial in a over R). Thus $\varrho(A) \subset R$ and we can factor ϱ uniquely (since φ is injective) as $\varrho = \varphi\eta$ for $\eta : A \to R$. $\square$

Now let $\{\varphi : R \to R[a]\}$ be a singleton elementary cocover. Since one can form the pushout of φ with itself, the compatibility condition on a singleton family $\{s \in \mathbf{P}(R[a])\}$ can be simplified as follows : Let $R \xrightarrow{\varphi} R[a] \underset{\vartheta}{\overset{\eta}{\rightrightarrows}} A$ be a pushout diagram. A family $\{s \in \mathbf{P}(R[a])\}$ is compatible if and only if $s\vartheta = s\eta$.

Corollary 2.7. *The coverage* $\mathbf{J}$ *is subcanonical. That is, all representable presheaves are sheaves on* $(\mathcal{A}_K^{op}, \mathbf{J})$.

Proof. Consider the presheaf $\mathbf{y}_A = \mathrm{Hom}_{\mathcal{A}_K}(A, -)$ for some object A of $\mathcal{A}_K^{op}$.

(i.) Given $(e_i)_{i \in I}$ a fundamental system of orthogonal idempotents, an elementary cocover $\{\varphi_i : R \to R/\langle 1 - e_i\rangle\}_{i \in I} \in \mathbf{J}^{op}(R)$ and a family $\{\eta_i : A \to R/\langle 1 - e_i\rangle\}_{i \in I}$. By the isomorphism $R \cong \prod_{i \in I} R/\langle 1 - e_i\rangle$ there is a unique $\eta : A \to R$ such that $\varphi_i\eta = \eta_i$.

(ii.) Let $R[a]$ be a separable extension of R. Consider the elementary cocover $\{R \xrightarrow{\varphi} R[a]\} \in \mathbf{J}^{op}(R)$ and let $\{A \xrightarrow{\psi} R[a]\}$ be a compatible family. By Corollary 2.6, one has a pushout diagram $R \xrightarrow{\varphi} R[a] \underset{\zeta}{\overset{\vartheta}{\rightrightarrows}} R[b, c]$. Compatibility implies that $\vartheta\psi = \zeta\psi$. But by Corollary 2.6 the canonical embedding φ is the equalizer of ϑ and ζ. Thus there exist a unique $A \xrightarrow{\eta} R \in \mathbf{y}_A(R)$ such that $\varphi\eta = \psi$.

$\square$

The terminal object in the category $\mathrm{Sh}(\mathcal{A}_K^{op}, J)$ is the sheaf sending each object to the set $\{*\} = 1$. This is the sheaf $\mathbf{y}_K$ since in $\mathcal{A}_K^{op}$ there is only one morphism between any object and the object K.

3 The separable algebraic closure

We define the presheaf $\mathbf{F} : \mathcal{A}_K \to \mathbf{Set}$ to be the forgetful functor. That is, for an object A of $\mathcal{A}_K$, $\mathbf{F}(A) = A$ and for a morphism $\varphi : A \to C$ of $\mathcal{A}_K$, $\mathbf{F}(\varphi) = \varphi$.

Lemma 3.1. **F** *is a sheaf of sets on the site* $(\mathcal{A}_K^{op}, \mathbf{J})$

Proof. We will show that the presheaf **F** satisfies the sheaf axiom (Definition I.3.3) for the elementary covers of any object of $\mathcal{A}_K^{op}$ by case analysis on the clauses of Definition 2.1. Again, we'll work directly with the category $\mathcal{A}_K$ rather than $\mathcal{A}_K^{op}$ with the definition of compatible family and the sheaf axiom translated accordingly.

(i.) Let R be an object of $\mathcal{A}_K$ and $(e_i)_{i\in I}$ a fundamental system of orthogonal idempotents of R. The presheaf **F** has the property $\mathbf{F}(0) = 1$. By Lemma 2.3 a family $\{a_i \in R/\langle 1-e_i\rangle\}_{i\in I}$ is a compatible family for the elementary cocover $\{\varphi_i : R \to R/\langle 1-e_i\rangle\}_{i\in I} \in \mathbf{J}^{op}(R)$. By the isomorphism $R \xrightarrow{(\varphi_i)_{i\in I}} \prod_{i\in I} R/\langle 1-e_i\rangle$ the element $a = (a_i)_{i\in I} \in R$ is the unique element such that $\varphi_i(a) = a_i$.

(ii.) Let R be an object of $\mathcal{A}_K$ and let $p \in R[X]$ be a monic, non-constant and separable polynomial. Let $R[a] = R[X]/\langle p\rangle$ and let $\{r \in R[a]\}$ be a compatible family for the elementary cocover $\{\varphi : R \to R[a]\} \in \mathbf{J}^{op}(R)$. Let $R \xrightarrow{\varphi} R[a] \overset{\vartheta}{\underset{\zeta}{\rightrightarrows}} R[b,c]$ be the pushout diagram of Corollary 2.6. Compatibility then implies $\vartheta(r) = \zeta(r)$ which by the same Corollary is true only if the element r is in R. We then have that r is the unique element restricting to itself along the embedding φ.

□

We fix a field K of any characteristic. Our goal is to show that the object $\mathbf{F} \in \mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ described above is a separably algebraically closed field containing the base field K, i.e. we shall show that **F** is a model, in Kripke–Joyal semantics, of an separably algebraically closed field containing K.

Let $\mathcal{L}[F,+,.]$ be a language with basic type F and function symbols $+ : F \times F \to F$ and $. : F \times F \to F$. We extend the language $\mathcal{L}[F,+,.]$ by adding to it a constant symbol $a : F$ for each element of $a \in K$, we then obtain an extended language $\mathcal{L}[F,+,.]_K$. Define $\mathrm{Diag}(K)$ as : if ϕ is an atomic $\mathcal{L}[F,+,.]_K$-formula or the negation of one such that $K \models \phi(a_1,...,a_n)$ then $\phi(a_1,...,a_n) \in \mathrm{Diag}(K)$. The theory T equips the type F with the geometric axioms of a separably algebraically closed field containing K.

Definition 3.2. The theory T has the following sentences (with all the variables having the type F).

1. Diag(K).
2. The axioms of commutative group.
   1. $\forall x\ [0 + x = x + 0 = x]$
   2. $\forall x \forall y \forall z\ [x + (y + z) = (x + y) + z]$
   3. $\forall x \exists y\ [x + y = 0]$
   4. $\forall x \forall y\ [x + y = y + x]$
3. The axioms of commutative ring.
   - 3.1. $\forall x\ [x1 = x]$
   - 3.2. $\forall x\ [x0 = 0]$
   - 3.3. $\forall x \forall y\ [xy = yx]$
   - 3.4. $\forall x \forall y \forall z\ [x(yz) = (xy)z]$
   - 3.5. $\forall x \forall y \forall z\ [x(y + z) = xy + xz]$
4. The field axioms.
   - 4.1. $\forall x\ [x = 0 \vee \exists y\ [xy = 1]]$
   - 4.2. $1 \neq 0$
5. The axiom schema of separable algebraic closure.
   - 5.1. $\forall a_1 \ldots \forall a_n [\mathrm{sep}_F(Z^n + \sum_{i=1}^{n} Z^{n-i} a_i) \Rightarrow \exists x\ [x^n + \sum_{i=1}^{n} x^{n-i} a_i = 0]]$, where $\mathrm{sep}_F(p)$ holds iff $p \in F[Z]$ is separable.
6. The axiom of separable algebraic extension.
   Let $K[Y]_{sep}$ be the set of separable polynomials in $K[Y]$.
   - 6.1. $\forall x [\bigvee_{p \in K[Y]_{sep}} p(x) = 0]$.

With these axioms the type F becomes the type of separable algebraic closure of K. We proceed to show that the object $\mathbf{F}$ is an interpretation of the type F, i.e. $\mathbf{F}$ is a model of the separable algebraic closure of K.

First note that since there is a unique map $K \to C$ for any object C of $\mathcal{A}_K$, an element $a \in K$ gives rise to a unique map $\mathbf{1} \xrightarrow{a} \mathbf{F}$, that is the map $* \mapsto a \in \mathbf{F}(K)$. Every constant $a \in K$ of the language is then interpreted by the corresponding unique arrow $\mathbf{1} \xrightarrow{a} \mathbf{F}$. (we used the same symbol for constants and their interpretation to avoid cumbersome notation). That $\mathbf{F}$ satisfy Diag(K) then follows directly.

Lemma 3.3. $\mathbf{F}$ *is a ring object.*

Proof. For an object C of $\mathcal{A}_K$ the object $\mathbf{F}(C)$ is a commutative ring. We can easily verify that C forces the axioms for commutative ring. □

Lemma 3.4. $\mathbf{F}$ *is a field.*

Proof. For any object R of $\mathcal{A}_K$ one has $R \Vdash 1 \neq 0$ since for any $R \xrightarrow{\varphi} C$ such that $C \Vdash 1 = 0$ one has that C is trivial and thus $C \Vdash \bot$.
We show that for x and y of type $\mathbf{F}$ and any object R of $\mathcal{A}_K^{op}$ we have

$$R \Vdash \forall x \; [x = 0 \vee \exists y \; [xy = 1]]$$

Let $\varphi : A \to R$ be a morphism of $\mathcal{A}_K^{op}$ and let $a \in \mathbf{F}(A) = A$. We need to show that $A \Vdash a = 0 \vee \exists y[ya = 1]$. The element $e = aa^*$ is an idempotent and we have a cover

$$\{\varphi_1 : A/\langle e\rangle \to A, \varphi_2 : A/\langle 1-e\rangle \to A\} \in \mathbf{J}(A)$$

We have

$$A/\langle e\rangle \Vdash a\varphi_1 = 0$$
$$A/\langle 1-e\rangle \Vdash (a\varphi_2)(a^*\varphi_2) = e\varphi_2 = 1$$

Hence by $\boxed{\exists}$ we have $A/\langle 1-e\rangle \Vdash \exists y(a\varphi_2)y = 1$. By $\boxed{\vee}$ we have that $A/\langle 1-e\rangle \Vdash a\varphi_2 = 0 \vee \exists y[(a\varphi_2)y = 1]$. Similarly, we have $A/\langle e\rangle \Vdash a\varphi_1 = 0 \vee \exists y[(a\varphi_1)y = 1]$. By $\boxed{\text{LC}}$ and $\boxed{\forall}$ we get $R \Vdash \forall x \; [x = 0 \vee \exists y \; [xy = 1]]$. □

Lemma 3.5. *The field object* $\mathbf{F} \in \mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ *is separably algebraically closed.*

Proof. We prove that for all n and all $(a_1, ..., a_n) \in \mathbf{F}^n(R) = R^n$, if $p = Z^n + \sum_{i=1}^n Z^{n-i}a_i$ is separable then one has

$$R \Vdash \exists x \; [x^n + \sum_{i=1}^{n} x^{n-i}a_i = 0].$$

Let $R[b] = R[Z]/\langle p\rangle$. We have a singleton cover $\{\varphi : R[b] \to R\}$ and $R[b] \Vdash b^n + \sum_{i=1}^{n} b^{n-i}(a_i\varphi) = 0$. By $\boxed{\exists}$ we conclude that $R \Vdash \exists x \; [x^n + \sum_{i=1}^{n} x^{n-i}a_i = 0]$ □

Lemma 3.6. $\mathbf{F}$ *is separable algebraic over* K.

Proof. Let R be an object of $\mathcal{A}_K$ and $r \in R$. Since R is étale then by definition r is separable algebraic over K, i.e. we have a separable $q \in K[X]$ with $q(r) = 0$. By $\boxed{\vee}$ we get $R \Vdash \bigvee_{p \in K[X]_{sep}} p(r) = 0$. □

Since $\mathbf{F}$ is a field we have that Lemma II.1.7 holds for polynomials over $\mathbf{F}$. This means that for all objects R of $\mathcal{A}_K^{op}$ we have $R \Vdash$ Lemma II.1.7. Thus we have the following Corollary of Lemma II.1.7.

Corollary 3.7. *Let R be an object of $\mathcal{A}_K$ and let f be a monic polynomial in $R[X]$. If f' is the derivative of f then there exist a cocover $\{\varphi_i : R \to R_i\}_{i \in I} \in \mathbf{J}^{*op}(R)$ and for each R_i we have $h, g, q, r, s \in R_i[X]$ such that $\varphi_i(f) = hg$, $\varphi_i(f') = qg$ and $rh + sq = 1$. Moreover, h is monic and separable.*

Lemma 3.8. *Let K be a field of characteristic 0. The sheaf $\mathbf{F} \in \mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ is algebraically closed.*

Proof. Let R be an object of $\mathcal{A}_K^{op}$ and $(a_1, ..., a_n) \in \mathbf{F}^n(R) = R^n$ and let $p = Z^n + \sum_{i=1}^n Z^{n-i} a_i$. By Corollary 3.7 we have a cover $\{\varphi_j : R_j \to R\}_{j \in J}$ with separable divisors $h_j \in R_j[Z]$ of p. That is, h_j is monic and separable dividing $Z^n + \sum_{i=1}^n Z^{n-i} a_i \varphi_j$. We note that since R_j has characteristic 0, whenever p is non-constant then so is h_j. By Lemma 3.5 we have that $R_j \Vdash \exists x h_j(x) = 0$. Consequently, $R_j \Vdash \exists x \; [x^n + \sum_{i=1}^{n} x^{n-i} a_i \varphi_j = 0]$. By $\boxed{\text{LC}}$ we get that $R \Vdash \exists x \; [x^n + \sum_{i=1}^{n} x^{n-i} a_i = 0]$ □

4 The power series object

To describe the object of power series over $\mathbf{F}$ we need to specify the natural numbers object in the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ first. One typically obtains this natural numbers object by sheafification of the constant presheaf of natural numbers. Here we describe this sheaf.

4.1 The constant sheaves of $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$

Let $\mathbf{P} : \mathcal{A}_K \to \mathbf{Set}$ be a constant presheaf associating to each object A of $\mathcal{A}_K$ a discrete set B. That is, $\mathbf{P}(A) = B$ and $\mathbf{P}(A \xrightarrow{\varphi} R) = 1_B$ for all objects A and all morphism φ of $\mathcal{A}_K$.

Let $\widetilde{\mathbf{P}} : \mathcal{A}_K \to \mathbf{Set}$ be the presheaf such that $\widetilde{\mathbf{P}}(A)$ is the set of elements of the form $\{(e_i, b_i) \mid i \in I\}$ where $(e_i)_{i \in I}$ is a fundamental system of orthogonal idempotents of A and for each i, $b_i \in B$. We express such an element as a formal sum $\sum_{i \in I} e_i b_i$.

Let $\varphi : A \to R$ be a morphism of $\mathcal{A}_K$, the restriction of $\sum_{i \in I} e_i b_i \in \widetilde{\mathbf{P}}(A)$ along φ is given by $(\sum_{i \in I} e_i b_i)\varphi = \sum_{i \in I} \varphi(e_i) b_i \in \widetilde{\mathbf{P}}(R)$.

Two elements $\sum_{i\in I} e_i b_i \in \widetilde{\mathbf{P}}(A)$ and $\sum_{j\in J} d_j c_j \in \widetilde{\mathbf{P}}(A)$ are equal if and only if $\forall i \in I, j \in J[b_i \neq c_j \Rightarrow e_i d_j = 0]$. This relation is indeed reflexive since $\forall i, \ell \in I[i \neq \ell \Rightarrow e_i e_\ell = 0]$. Symmetry is immediate. To show transitivity, assume we are given $\sum_{i\in I} e_i b_i$, $\sum_{j\in J} d_j c_j$ and $\sum_{\ell\in L} u_\ell a_\ell$ in $\widetilde{\mathbf{P}}(A)$ such that

$$\forall i \in I, j \in J\ [b_i \neq c_j \Rightarrow e_i d_j = 0]$$
$$\forall j \in J, \ell \in L\ [c_j \neq a_\ell \Rightarrow d_j u_\ell = 0]$$

Let $k \in I$ and $t \in L$ such that $a_t \neq b_k$. Since B is discrete, one can split the sum $\sum_{j\in J} d_j = 1$ into three sums of those d_j such that $c_j = b_k$ and those d_h such that $c_h = a_t$ and those d_m such that c_m is different from both a_t and b_k. Hence we have

$$e_k u_t = e_k u_t \sum_{j\in J} d_j = e_k u_t \Big(\sum_{j\in J}^{c_j = b_k} d_j + \sum_{h\in J}^{c_h = a_t} d_h + \sum_{m\in J}^{c_m \neq a_t, c_m \neq b_k} d_m \Big) = 0$$

Note in particular that for $\sum_{i\in I} e_i b_i \in \widetilde{\mathbf{P}}(A)$ and canonical morphisms $\varphi_i : A \to A/\langle 1 - e_i \rangle$, one has for any $j \in I$ that

$$\Big(\sum_{i\in I} e_i b_i\Big)\varphi_j = b_j \in \widetilde{\mathbf{P}}(A/\langle 1 - e_j \rangle).$$

To prove that $\widetilde{\mathbf{P}}$ is a sheaf we will need the following lemmas.

Lemma 4.1. *Let R be a regular ring and let $(e_i)_{i\in I}$ be a fundamental system of orthogonal idempotents of R. Let $R_i = R/\langle 1 - e_i \rangle$ and let $([d_j])_{j\in J_i}$ be a fundamental system of orthogonal idempotents of R_i, where $[d_j] = d_j + \langle 1 - e_i \rangle$. We have that $(e_i d_j)_{i\in I, j\in J_i}$ is a fundamental system of orthogonal idempotents of R.*

Proof. In R one has $\sum_{j\in J_i} e_i d_j = e_i \sum_{j\in J_i} d_j = e_i(1 + \langle 1 - e_i \rangle) = e_i$. Hence, $\sum_{i\in I, j\in J_i} e_i d_j = \sum_{i\in I} e_i = 1$. For some $i \in I$ and $t, k \in J_i$ we have $(e_i d_t)(e_i d_k) = e_i(0 + \langle 1 - e_i \rangle) = 0$ in R. Thus for $i, \ell \in I$, $j \in J_i$ and $s \in J_\ell$ one has $i \neq \ell \vee j \neq s \Rightarrow (e_i d_j)(e_\ell d_s) = 0$. □

Lemma 4.2. *Let R be a regular ring, $f \in R[Z]$ a polynomial of formal degree n and $p \in R[Z]$ a monic polynomial of degree $m > n$. If in $R[X, Y]$ one has*

$$f(Y)(1 - f(X)) = 0 \mod \langle p(X), p(Y) \rangle$$

then $f = e \in R$ with e an idempotent.

Proof. Let $f(Z) = \sum_{i=0}^{n} r_i Z^i$. By the assumption, for some $q, g \in R[X, Y]$

$$f(Y)(1 - f(X)) = \sum_{i=0}^{n} r_i (1 - \sum_{j=0}^{n} r_j X^j) Y^i = qp(X) + gp(Y)$$

One has $\sum_{i=0}^{n} r_i (1 - \sum_{j=0}^{n} r_j X^j) Y^i = g(X, Y)p(Y) \mod \langle p(X) \rangle$. Since $p(Y)$ is monic of Y-degree greater than n, one has that for all $0 \le i \le n$

$$r_i (1 - \sum_{j=0}^{n} r_j X^j) = 0 \mod \langle p(X) \rangle$$

But this means that $r_i r_n X^n + r_i r_{n-1} X^{n-1} + ... + r_i r_0 - r_i$ is divisible by $p(X)$ for all $0 \le i \le n$ which because $p(X)$ is monic of degree $m > n$ implies that all coefficients are equal to 0. In particular, for $1 \le i \le n$ one gets that $r_i^2 = 0$ and hence $r_i = 0$ since R is reduced. For $i = 0$ one gets that the constant coefficient $r_0 r_0 - r_0 = 0$ and thus r_0 is an idempotent of R. □

Lemma 4.3. *The presheaf $\widetilde{\mathbf{P}}$ described above is a sheaf on $(\mathcal{A}_K^{op}, \mathbf{J})$.*

Proof. We show that $\widetilde{\mathbf{P}}$ satisfy the sheaf axiom (Definition I.3.3) for the coverage $\mathbf{J}$ described in Definition 2.1.

(i.) Let $(e_i)_{i \in I}$ be a fundamental system of orthogonal idempotents of an object R of $\mathcal{A}_K$ with $R_i = R/\langle 1 - e_i \rangle$ and canonical morphisms $\varphi_i : R \to R_i$. Since $\widetilde{\mathbf{P}}(0) = 1$ by Lemma 2.3 any set of elements $\{s_i \in \widetilde{\mathbf{P}}(R_i)\}_{i \in I}$ is a compatible family on the elementary cocover $\{\varphi_i\}_{i \in I} \in \mathbf{J}^{op}(R)$. For each i, Let $s_i = \sum_{j \in J_i} [d_j] b_j$. By Lemma 4.1 we have an element $s = \sum_{i \in I, j \in J_i} (e_i d_j) b_j \in \widetilde{\mathbf{P}}(R)$ the restriction of which along φ_i is the element $\sum_{j \in J_i} [d_j] b_j \in \widetilde{\mathbf{P}}(R_i)$. It remains to show that this is the only such element.

Let there be an element $\sum_{\ell \in L} c_\ell a_\ell \in \widetilde{\mathbf{P}}(R)$ that restricts to $u_i = s_i$ along φ_i. We have $u_i = \sum_{\ell \in L} [c_\ell] a_\ell$. One has that for any $j \in J_i$ and $\ell \in L$, $b_j \neq a_\ell \Rightarrow [c_\ell d_j] = 0$ in R_i, hence, in R one has $b_j \neq a_\ell \Rightarrow c_\ell d_j = r(1 - e_i)$. Multiplying both sides of $c_\ell d_j = r(1 - e_i)$ by e_i we get $b_j \neq a_\ell \Rightarrow c_\ell (e_i d_j) = 0$. Thus proving $s = \sum_{\ell \in L} c_\ell a_\ell$.

(ii.) Let $p \in R[X]$ be a monic non-constant separable polynomial. One has an elementary cocover $\{\varphi : R \to R[a] = R[X]/\langle p \rangle\}$. Let the singleton $\{s \in \widetilde{\mathbf{P}}(R[a])\}$ be a compatible family on this cocover. Let $s = \sum_{i \in i} e_i b_i \in \widetilde{\mathbf{P}}(R[a])$. We can assume w.l.o.g. that $\forall i, j \in I \; [i \neq j \Rightarrow b_i \neq b_j]$ since if $b_k = b_\ell$ one has that

$$(e_k + e_\ell) b_l + \sum_{j \in I}^{j \neq \ell, j \neq k} e_j b_j = s$$

Note that an idempotent e_i of $R[a]$ is a polynomial $e_i(a)$ in a of formal degree less than $\deg p$. Let $R[c,d] = R[X,Y]/\langle p(X), p(Y) \rangle$, by Corollary 2.6, one has a pushout diagram

$$\begin{array}{ccccc} a & \longmapsto & c & & \\ R[a] & \xrightarrow{\vartheta} & R[c,d] & & d \\ \uparrow{\scriptstyle\varphi} & & \uparrow{\scriptstyle\zeta} & & \uparrow \\ R & \xrightarrow{\varphi} & R[a] & & a \end{array}$$

That the singleton $\{s\}$ is compatible then means

$$s\vartheta = \sum_{i \in I} e_i(c) b_i = s\zeta = \sum_{i \in I} e_i(d) b_i$$

i.e. $\forall i, j \in I \; [b_i \neq b_j \Rightarrow e_i(c) e_j(d) = 0]$. By the assumption that $b_i \neq b_j$ whenever $i \neq j$ this means that in $R[c,d]$ for any $i \neq j \in I$

$$e_j(d) e_i(c) = 0$$

Thus

$$e_j(d) \sum_{i \neq j} e_i(c) = e_j(d)(1 - e_j(c)) = 0$$

i.e. in $R[X,Y]$ one has $e_j(Y)(1 - e_j(X)) = 0 \mod \langle p(X), p(Y) \rangle$. By Lemma 4.2 we have that $e_j \in R$. Thus we proved that for the singleton family $\{s \in \widetilde{\mathbf{P}}(R[a])\}$ to be compatible, s is equal to $\sum_{j \in J} d_j b_j \in \widetilde{\mathbf{P}}(R[a])$ such that $d_j \in R$ for $j \in J$. That is $\sum_{j \in J} d_j b_j \in \widetilde{\mathbf{P}}(R)$. Thus we have found a unique (since $\widetilde{\mathbf{P}}(\varphi)$ is injective) element in $\widetilde{\mathbf{P}}(R)$ restricting to s along φ.

□

Lemma 4.4. *Let* $\mathbf{P}$ *and* $\widetilde{\mathbf{P}}$ *be as described above. Let* $\Gamma : \mathbf{P} \to \widetilde{\mathbf{P}}$ *be the presheaf morphism such that* $\Gamma_R(b) = b \in \widetilde{\mathbf{P}}(R)$ *for any object* R *and* $b \in B$*. If* $\mathbf{E}$ *is a sheaf and* $\Lambda : \mathbf{P} \to \mathbf{E}$ *is a morphism of presheaves, then there exist a unique* *sheave morphism* $\Delta : \widetilde{\mathbf{P}} \to \mathbf{E}$ *such that the following diagram (of* $\mathbf{Set}^{\mathcal{A}_K}$*) commutes.*

$$\begin{array}{ccc} \mathbf{P} & \xrightarrow{\Lambda} & \mathbf{E} \\ \Big\downarrow{\scriptstyle \Gamma} & \nearrow_{\Delta} & \\ \widetilde{\mathbf{P}} & & \end{array}$$

That is to say $\Gamma : \mathbf{P} \to \widetilde{\mathbf{P}}$ *is the sheafification of* $\mathbf{P}$.

Proof. Let $a = \sum_{i \in I} e_i b_i \in \widetilde{\mathbf{P}}(A)$ and let $A_i = A/\langle 1 - e_i \rangle$ with canonical morphisms $\varphi_i : A \to A_i$.

Let $\mathbf{E}$ and Λ be as in the statement of the lemma. If there exist a sheaf morphism $\Delta : \widetilde{\mathbf{P}} \to \mathbf{E}$, then Δ being a natural transformation forces us to have for all $i \in I$, $\mathbf{E}(\varphi_i)\Delta_A = \Delta_{A_i}\widetilde{\mathbf{P}}(\varphi_i)$. By Lemma 2.4, we know that the map $\mathbf{E}(A) \ni d \mapsto (\mathbf{E}(\varphi_i)d \in \mathbf{E}(A_i))_{i \in I}$ is an isomorphism. Thus it must be that $\Delta_A(a) = (\Delta_{A_i}\widetilde{\mathbf{P}}(\varphi_i)a)_{i \in I} = (\Delta_{A_i}(b_i))_{i \in I}$. But $\Delta_{A_i}(b_i) = \Delta_{A_i}\Gamma_{A_i}(b_i)$[1]. To have $\Delta\Gamma = \Lambda$ we must have $\Delta_{A_i}(b_i) = \Lambda_{A_i}(b_i)$. Hence, we are forced to have $\Delta_A(a) = (\Lambda_{A_i}(b_i))_{i \in I}$. Note that Δ is unique since its value $\Delta_A(a)$ at any A and a is forced by the commuting diagram above. □

The constant presheaf of natural numbers $\mathbf{N}$ is the natural numbers object in $\mathbf{Set}^{\mathcal{A}_K}$. We associate to $\mathbf{N}$ a sheaf $\widetilde{\mathbf{N}}$ as described above. As noted in Chapter I, this is the natural numbers object in $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$. Alternatively, from Lemma 4.4 one can easily show that $\widetilde{\mathbf{N}}$ satisfies the axioms of a natural numbers object.

Definition 4.5. Let $\mathbf{F}[[X]]$ be the presheaf mapping each object R of $\mathcal{A}_K$ to $\mathbf{F}[[X]](R) = R[[X]] = R^{\mathbb{N}}$ with the obvious restriction maps.

Lemma 4.6. $\mathbf{F}[[X]]$ *is a sheaf.*

Proof. The proof is immediate as a corollary of Lemma 3.1. □

Lemma 4.7. *The sheaf* $\mathbf{F}[[X]]$ *is naturally isomorphic to the sheaf* $\mathbf{F}^{\widetilde{\mathbf{N}}}$.

Proof. Let C be an object of $\mathcal{A}_K^{op}$. Since $\mathbf{F}^{\widetilde{\mathbf{N}}}(C) \cong \mathbf{y}_C \times \widetilde{\mathbf{N}} \to \mathbf{F}$, an element $\alpha_C \in \mathbf{F}^{\widetilde{\mathbf{N}}}(C)$ is a family (indexed by object of $\mathcal{A}_K^{op}$) of elements of the form $\alpha_{C,D} : \mathbf{y}_C(D) \times \widetilde{\mathbf{N}}(D) \to \mathbf{F}(D)$ where D is an object of $\mathcal{A}_K^{op}$.

[1] Note that the b_i in the expression $\Delta_{A_i}(b_i)$ is an element of $\widetilde{\mathbf{P}}(A_i)$ while the b_i in the expression $\Gamma_{A_i}(b_i)$ is an element of $\mathbf{P}(A_i) = B$.

Define $\Theta : \mathbf{F}^{\tilde{\mathbf{N}}} \to \mathbf{F}[[X]]$ as $(\Theta\alpha)_C(n) = \alpha_{C,C}(1_C, n)$. Define $\Lambda : \mathbf{F}[[X]] \to \mathbf{F}^{\tilde{\mathbf{N}}}$ as

$$(\Lambda\beta)_{C,D}(C \xrightarrow{\varphi} D, \sum_{i\in I} e_i n_i) = (\vartheta_i \varphi(\beta_C(n_i)))_{i\in I} \in \mathbf{F}(D)$$

where $D \xrightarrow{\vartheta_i} D/\langle 1 - e_i \rangle$ is the canonical morphism. Note that by Lemma 2.4 one indeed has that $(\vartheta_i\varphi(\beta_C(n_i)))_{i\in I} \in \prod_{i\in I} \mathbf{F}(D_i) \cong \mathbf{F}(D)$. One can easily verify that Θ and Λ are natural. To show the isomorphism we will show $\Lambda\Theta = 1_{\mathbf{F}^{\tilde{\mathbf{N}}}}$ and $\Theta\Lambda = 1_{\mathbf{F}[[X]]}$. We have

$$\begin{aligned}(\Lambda\Theta\alpha)_{C,D}(\varphi, \sum_{i\in I} e_i n_i) &= (\vartheta_i\varphi((\Theta\alpha)_C(n_i)))_{i\in I} \\ &= (\vartheta_i\varphi(\alpha_{C,C}(1_C, n_i)))_{i\in I} \\ &= ((\alpha_{C,D_i}(\vartheta_i\varphi, n_i)))_{i\in I} \\ &= \alpha_{C,D}(\varphi, \sum_{i\in I} e_i n_i)\end{aligned}$$

Thus showing $\Lambda\Theta = 1_{\mathbf{F}^{\tilde{\mathbf{N}}}}$. Next we show $\Theta\Lambda = 1_{\mathbf{F}[[X]]}$.

$$\begin{aligned}(\Theta\Lambda\beta)_C(n) &= (\Lambda\beta)_{C,C}(1_C, n) \\ &= 1_C 1_C(\beta_C(n)) = \beta_C(n)\end{aligned}$$

□

Lemma 4.8. *The power series object* $\mathbf{F}[[X]]$ *is a ring object.*

Proof. A Corollary to Lemma 3.3. □

5 Choice axioms

The axiom of choice fails to hold (even in a classical metatheory) in the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ whenever the field K is not separably algebraically closed. To show this we will show that there is an epimorphism in $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ with no section.

Fact 5.1. *Let* $\Theta : \mathbf{P} \to \mathbf{G}$ *be a morphism of sheaves on a site* $(\mathcal{C}, \mathbf{J})$*. Then* Θ *is an epimorphism if for each object* C *of* $\mathcal{C}$ *and each element* $c \in \mathbf{G}(C)$ *there is a cover* S *of* C *such that for all* $f : D \to C$ *in the cover* S *the element* cf *is in the image of* Θ_D. [MacLane and Moerdijk, 1992, Ch 3].

First we consider a simple case where the base filed K has characteristic 0 and there exist an element in K which has no square root in K. Consider the algebraically closed sheaf $\mathbf{F}$ and the natural transformation $\Theta : \mathbf{F} \to \mathbf{F}$ where for $c \in C$ we have $\Theta_C(c) = c^2$. Consider an element $y \in \mathbf{F}(C)$ and let $\{\varphi_i : C_i \to C\}_{i \in I}$ be a cover with $p_i \in C_i[X]$ a separable divisor of $X^2 - y$. Since K has characteristic 0, p_i is non-constant. Let $C_i[x_i] = C_i[X]/\langle p_i \rangle$. We have a cover $\{\vartheta_i : C_i[x_i] \to C\}_{i \in I}$ of C and $\Theta_{C_i[x_i]}(x_i) = {x_i}^2$. By construction, $p_i(x_i) = 0$ and since p_i divides $X^2 - y\vartheta_i$ we have ${x_i}^2 - y\vartheta_i = 0$, that is $\Theta_{C_i[x_i]}(x_i) = y\vartheta_i$. Thus Θ is an epimorphism of $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$.

Lemma 5.2. *The epimorphism Θ have no section.*

Proof. Suppose Θ have a section Δ. Then for an object C of $\mathcal{A}_K^{op}$, $\Delta_C : \mathbf{F}(C) \to \mathbf{F}(C)$ would need to map an element $y \in \mathbf{F}(C)$ to its square root in $\mathbf{F}(C)$ which is not in general possible since C doesn't necessarily contain the square root of each of its elements. In particular, by assumption there is an element $a \in K$ with no square root in K. For example, let the base field be $\mathbb{Q}$ and take $C = \mathbb{Q}$. We need Δ such that $\Theta_C \Delta_C(2) = (\Delta_C(2))^2 = 2$ but there no element $a \in \mathbb{Q}$ such that $a^2 = 2$. □

This construction can be easily generalized to show that the axiom of choice does not hold in $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ for any non-algebraically closed field K of characteristic 0.

Lemma 5.3. *Let K be a field of characteristic 0 not algebraically closed. There is an epimorphism in $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ with no section.*

Proof. Let $f = X^n + \sum_{i=1}^n r_i X^{n-i} \in K[X]$ be a non-constant polynomial for which no root in K exist. w.l.o.g. we assume f separable. One can construct $\Lambda : \mathbf{F} \to \mathbf{F}$ defined by $\Lambda_C(c) = c^n + r_1 c^{n-1} + \cdots + r_{n-1} c \in C$. Given $d \in \mathbf{F}(C)$, let $g = X^n + \sum_{i=1}^{n-1} r_i X^{n-i} - d$. By Corollary 3.7 there is a cover $\{C_\ell \xrightarrow{\varphi_\ell} C\}_{\ell \in L} \in \mathbf{J}^*(C)$ with $h_\ell \in C_\ell[X]$ a separable non-constant polynomial dividing g. Let $C_\ell[x_\ell] = C_\ell[X]/\langle h_\ell \rangle$ one has a singleton cover $\{C_\ell[x_\ell] \xrightarrow{\vartheta_\ell} C_\ell\}$ and thus a composite cover $\{C_\ell[x_\ell] \xrightarrow{\varphi_\ell \vartheta_\ell} C\}_{\ell \in L} \in \mathbf{J}^*(C)$. Since x_ℓ is a root of $h_\ell \mid g$ we have $\Lambda_{C_\ell[x_\ell]}(x_\ell) = x_\ell^n + \sum_{i=1}^{n-1} r_i x_\ell^{n-i} = d$ or more precisely $\Lambda_{C_\ell[x_\ell]}(x_\ell) = d\varphi_\ell \vartheta_\ell$. Thus, Λ is an epimorphism (by Fact 5.1) and it has no section, for if it had a section $\Psi : \mathbf{F} \to \mathbf{F}$ then one would have $\Psi_K(-r_n) = a \in K$ such that $a^n + \sum_{i=1}^n r_i a^{n-i} = 0$ which is not true by assumption. □

For a field of non-zero characteristic the situation is analogous, albeit a bit trickier. The proof of Lemma 5.3 depended on the fact that when given a non-constant monic polynomial $f \in C[X]$ and $c \in C$, where C is an étale K-algebra, then $f + c$ has a non-constant separable divisor. This is not necessarily the case when K has characteristic p. What we aim to show is that a similar statement holds when f is separable.

Lemma 5.4. *Let R be a regular ring and $f = uX^n + \sum_{i=1}^{n} c_i X^{n-i} \in R[X]$ where u is a unit. Let $g \in R[X]$ and $f = gh$ for some h. Then* $\deg(g) \leq \deg(f)$.

Proof. Let $g = \sum_{i=0}^{m} a_i X^i$ and $h = \sum_{i=0}^{t} b_i X^i$. Assume $m > n$ and let $e = a_m a_m^*$. Since $m > n$ we have that $a_m b_t = 0$.

But a_m is a unit in $R/\langle 1-e \rangle$, thus $b_t = 0$ in $R/\langle 1-e \rangle$ and h can be written $h = \sum_{i=0}^{t-1} b_i X^i \in R/\langle 1-e \rangle[X]$. But then again since $m > n$ we have $a_m b_{t-1} = 0$ in $R/\langle 1-e \rangle$...etc. By induction on the degree of h it follows that $b_i = 0$ in $R/\langle 1-e \rangle$ for all $0 \leq i \leq t$. But then $f = 0$, and hence $u = 0$ in $R/\langle 1-e \rangle$. Since u is a unit we have that $1 = 0$ in $R/\langle 1-e \rangle$ which means $a_m = 0$. ☐

We have then the following version of Corollary II.1.8 for étale K-algebra.

Corollary 5.5. *Let R be an étale K-algebra. Let f be a non-constant monic polynomial with a derivative $f' = uX^n + \sum_{i=1}^{n} c_i X^{n-i} \in R[X]$. If u is a unit then there is a cover $\{\varphi_i : R_i \to R\}_{i \in I}$ where f has a non-constant separable divisor in $R_i[X]$ for all i.*

Proof. By Corollary 3.7 we have a cover $\{R_i \to R\}_{i \in I}$ and a separable divisor $h_i \in R_i[X]$ of f where $f = h_i g_i$ and $f' = q_i g_i$. By Lemma 5.4 $\deg(g_i) \leq \deg(f_i)$. But then $\deg(g_i) < \deg(f)$ and thus h_i is non-constant. ☐

Corollary 5.6. *Let K be a field of any characteristic and $f \in K[X]$ a non-constant separable polynomial. Let R be an étale K-algebra. Let $p \in R[X]$ such that $f - p$ is a constant in R. There is a cover $\{\varphi_i : R_i \to R\}_{i \in I}$ where in each R_i there is a non-constant separable divisor $h_i \in R_i[X]$ of p.*

Proof. Since f is separable $g = \gcd(f, f')$ is a unit and thus $f' \neq 0$. It then follows that f' has the form $f' = uX^n + \sum_{i=1}^{n} c_i X^{n-i}$ where u is a unit. Since $f - p$ is constant we have $f' = p' = uX^n + \sum_{i=1}^{n} c_i X^{n-i}$. By Corollary 5.5 there is a cover $\{\varphi_i : R_i \to R\}_{i \in I}$ where in each R_i there is a non-constant separable divisor h_i of p. ☐

Lemma 5.7. *Let K be a field of any characteristic not separably algebraically closed. There is an epimorphism in* $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ *with no section.*

Proof. Let $f = X^n + \sum_{i=1}^{n} r_i X^{n-i} \in K[X]$ be a non-constant separable polynomial for which no root in K exist. One can construct $\Lambda : \mathbf{F} \to \mathbf{F}$ defined by $\Lambda_C(c) = c^n + r_1 c^{n-1} + \cdots + r_{n-1}c \in C$.
Given $d \in \mathbf{F}(C)$, let $g = X^n + \sum_{i=1}^{n-1} r_i X^{n-i} - d$. Since $f - g = d \in C$, by Corollary 5.6 there is a cover $\{C_\ell \xrightarrow{\varphi_\ell} C\}_{\ell \in L} \in \mathbf{J}^*(C)$ with $h_\ell \in C_\ell[X]$ a separable non-constant polynomial dividing g. Let $C_\ell[x_\ell] = C_\ell[X]/\langle h_\ell \rangle$ one has a singleton cover $\{C_\ell[x_\ell] \xrightarrow{\vartheta_\ell} C_\ell\}$ and thus a composite cover $\{C_\ell[x_\ell] \xrightarrow{\varphi_\ell \vartheta_\ell} C\}_{\ell \in L} \in \mathbf{J}^*(C)$. Since x_ℓ is a root of $h_\ell \mid g$ we have $\Lambda_{C_\ell[x_\ell]}(x_\ell) = x_\ell^n + \sum_{i=1}^{n-1} r_i x_\ell^{n-i} = d$ or more precisely $\Lambda_{C_\ell[x_\ell]}(x_\ell) = d\varphi_\ell \vartheta_\ell$. Thus, Λ is an epimorphism (by Fact 5.1) and it has no section, for if it had a section $\Psi : \mathbf{F} \to \mathbf{F}$ then one would have $\Psi_K(-r_n) = a \in K$ such that $a^n + \sum_{i=1}^{n} r_i a^{n-i} = 0$ which is not true by assumption. ☐

Theorem 5.8. *Let K be a field not separably algebraically closed. The axiom of choice fails to hold in the topos* $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$. ☐

We demonstrate further that when the base field is $\mathbb{Q}$ the weaker axiom of *dependent choice* does not hold (internally) in the topos $\mathrm{Sh}(\mathcal{A}_{\mathbb{Q}}^{op}, \mathbf{J})$. For a relation $R \subset Y \times Y$ the axiom of dependent choice is stated as

$$\forall x \exists y R(x,y) \Rightarrow \forall x \exists g \in Y^N [g(0) = x \wedge \forall n R(g(n), g(n+1))] \quad \text{(ADC)}$$

Theorem 5.9. $\mathrm{Sh}(\mathcal{A}_{\mathbb{Q}}^{op}, \mathbf{J}) \Vdash \neg\mathrm{ADC}$.

Proof. Consider the binary relation on the algebraically closed object $\mathbf{F}$ defined by the characteristic function $\phi(x,y) := y^2 - x = 0$. Assume $C \Vdash \mathrm{ADC}$ for some object C of $\mathcal{A}_K$. Since $C \Vdash \forall x \exists y [y^2 - x = 0]$ we have $C \Vdash \forall x \exists g \in \mathbf{F}^{\widetilde{\mathbf{N}}}[g(0) = x \wedge \forall n [g(n+1)^2 = g(n)]]$. That is for all morphisms $C \xrightarrow{\zeta} A$ of $\mathcal{A}_K$ and elements $a \in \mathbf{F}(A)$ one has $A \Vdash \exists g \in \mathbf{F}^{\widetilde{\mathbf{N}}}[g(0) = a \wedge \forall n [g(n+1)^2 = g(n)]]$. Taking $a = 2$ we have $A \Vdash \exists g \in \mathbf{F}^{\widetilde{\mathbf{N}}}[g(0) = 2 \wedge \forall n [g(n+1)^2 = g(n)]]$. Which by $\boxed{\exists}$ implies the existence of a cocover $\{\eta_i : A \to A_i\}_{i \in I}$ and power series $\alpha_i \in \mathbf{F}^{\widetilde{\mathbf{N}}}(A_i)$ such that $A_i \Vdash \alpha_i(0) = 2 \wedge \forall n [\alpha_i(n+1)^2 = \alpha_i(n)]]$. By Lemma 4.7 we have $\mathbf{F}^{\widetilde{\mathbf{N}}}(A_i) \cong A_i[[X]]$ and thus the above forcing implies the existence of a series $\alpha_i = 2 + 2^{1/2} + ... + 2^{1/2^j} + ... \in A_i[[X]]$. But this holds only if A_i contains a root of $X^{2^j} - 2$ for all j which implies A_i is trivial as will shortly show after the following remark.

Consider an algebra R over $\mathbb{Q}$. Assume R contains a root of $X^{2^n} - 2$ for some n. Then letting $\mathbb{Q}[x] = \mathbb{Q}[X]/\langle X^{2^n} - 2\rangle$, one will have a homomorphism $\xi : \mathbb{Q}[x] \to R$. By Eisenstein's criterion the polynomial $X^{2^n} - 2$ is irreducible over $\mathbb{Q}$, making $\mathbb{Q}[x]$ a field of dimension 2^n and ξ either an injection with a trivial kernel or $\xi = \mathbb{Q}[x] \to 0$.

Now we continue with the proof. Until now we have shown that for all $i \in I$, the algebra A_i contains a root of $X^{2^j} - 2$ for all j. For each $i \in I$, let A_i be of dimension m_i over $\mathbb{Q}$. We have that A_i contains a root of $X^{2^{m_i}} - 2$ and we have a homomorphism $\mathbb{Q}(\sqrt[2^{m_i}]{2}) \to A_i$ which since A_i has dimension $m_i < 2^{m_i}$ means that A_i is trivial for all $i \in I$. Hence, $A_i \Vdash \bot$ and consequently $C \Vdash \bot$. We have shown that for any object D of $\mathcal{A}_{\mathbb{Q}}^{op}$ if $D \Vdash \mathrm{ADC}$ then $D \Vdash \bot$. Hence $\mathrm{Sh}(\mathcal{A}_{\mathbb{Q}}^{op}, \mathbf{J}) \Vdash \neg\mathrm{ADC}$. □

We say that a topos satisfy the internal axiom of choice if the axiom of choice holds in the internal language of this topos, or equivalently if for each object B the endofunctor $(-)^B$ preserves epimorphisms [MacLane and Moerdijk, 1992].

Since ADC does not hold internally in $\mathrm{Sh}(\mathcal{A}_{\mathbb{Q}}^{op}, \mathbf{J})$ by Theorem 5.9, it follows that the internal axiom of choice does not hold in this topos either.

6 The logic of $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$

In this section we will demonstrate that in a classical metatheory one can show that the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ is boolean. In fact we will show that, in a classical metatheory, the boolean algebra structure of the subobject classifier is the one specified by the boolean algebra of idempotents of the algebras in $\mathcal{A}_K$. Except for Theorem 6.8 the reasoning in this section is classical.

Recall that the idempotents of a commutative ring form a boolean algebra. In terms of ring operations the logical operators are defined as follows.

1. $e_1 \wedge e_2 = e_1 e_2$
2. $e_1 \vee e_2 = e_1 + e_2 - e_1 e_2$
3. $\neg e = 1 - e$
4. $e_1 \leq e_2$ iff $e_1 \wedge e_2 = e_1$ and $e_1 \vee e_2 = e_2$
5. $\top = 1$

6. $\perp = 0$

A sieve S on an object C is said to cover a morphism $f : D \to C$ if $f^*(S)$ contains a cover of D. Dually, a cosieve M on C is said to cover a morphism $g : C \to D$ if the sieve dual to M covers the morphism dual to g. i.e. a morphism is covered by a sieve (cosieve) when there is a cover of its domain (cocover of its codomain) such that its composition with each element in the cover (cocover) lies in the sieve (cosieve).

Definition 6.1 (Closed cosieve)**.** A sieve M on an object C of $\mathcal{C}$ is closed if $\forall f : D \to C[M \text{ covers } f \Rightarrow f \in M]$. A closed cosieve on an object C of $\mathcal{C}^{op}$ is the dual of a closed sieve in $\mathcal{C}$.

Fact 6.2 (Subobject classifier)**.** *The subobject classifier in the category of sheaves on a site $(\mathcal{C}, \mathbf{J})$ is the presheaf $\mathbf{\Omega}$ where for an object C of $\mathcal{C}$ the set $\mathbf{\Omega}(C)$ is the set of closed sieves on C and for each $f : D \to C$ we have a restriction map $M \mapsto \{h \mid \text{cod}(h) = D, fh \in M\}$.*

Lemma 6.3. *Let R be an object of $\mathcal{A}_K$. If R is a field the closed cosieves on R are the maximal cosieve generated by the singleton $\{1_R : R \to R\}$ and the minimal cosieve $\{R \to 0\}$.*

Proof. Let S be a closed cosieve on R and let $\varphi : R \to A \in S$ and let I be a maximal ideal of A. If A is nontrivial we have a field morphism $R \to A/I$ in S where A/I is a finite field extension of R. Let $A/I = R[a_1, ..., a_n]$. But then the morphism $\vartheta : R \to R[a_1, ..., a_{n-1}]$ is covered by S. Thus $\vartheta \in S$ since S is closed. By induction on n we get that a field morphism $\eta : R \to R$ is in S but η in turn covers the identity $R \xrightarrow{1_R} R$. Thus $1_R \in S$. □

Corollary 6.4. *For an object R of $\mathcal{A}_K$. If R is a field $\mathbf{\Omega}(R)$ is a 2-valued boolean algebra.*

Proof. This is a direct Corollary of Lemma 6.3. The maximal cosieve (1_R) correspond to the idempotent 1 of R, that is the idempotent e such that, $\ker 1_R = \langle 1 - e \rangle$. Similarly the cosieve $\{R \to 1\}$ correspond to the idempotent 0. □

Corollary 6.5. *For an object A of $\mathcal{A}_K$, $\mathbf{\Omega}(A)$ is isomorphic to the set of idempotents of A and the Heyting algebra structure of $\mathbf{\Omega}(A)$ is the boolean algebra of idempotents of A.*

Proof. Classically an étale algebra over K is isomorphic to a product of field extensions of K. Let A be an object of $\mathcal{A}_K$, then $A \cong F_1 \times ... \times F_n$ where F_i is a finite field extension of K. The set of idempotents of A is

$\{(d_1, ..., d_n) \mid 1 \le j \le n, d_j \in F_j, d_j = 0 \text{ or } d_j = 1\}$. But this is exactly the set $\mathbf{\Omega}(F_1) \times ... \times \mathbf{\Omega}(F_n) \cong \mathbf{\Omega}(A)$. It is obvious that since $\mathbf{\Omega}(A)$ is isomorphic to a product of boolean algebras, it is a boolean algebra with the operators defined pointwise. □

Corollary 6.6. *The topos* $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ *is boolean.*

Proof. The subobject classifier of $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ is $1 \xrightarrow{\text{true}} \mathbf{\Omega}$ where for an object A of $\mathcal{A}_K$ one has $\mathrm{true}_A(*) = 1 \in A$. □

It is not possible to show that the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ is boolean in an intuitionistic metatheory as we shall demonstrate here. First we recall the definition of the *Limited principle of omniscience* (LPO for short).

Definition 6.7 (LPO)**.** For any binary sequence α the following statement holds

$$\forall n[\alpha(n) = 0] \vee \exists n[\alpha(n) = 1]$$

LPO cannot be shown to hold intuitionistically. One can, nevertheless, show that it is weaker than the law of excluded middle [Bridges and Richman, 1987].

Theorem 6.8. *Intuitionistically, if* $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ *is boolean then* LPO *holds.*

Proof. Let $\alpha \in K[[X]]$ be a binary sequence. By Lemma 4.7 one has an isomorphism $\Lambda : \mathbf{F}[[X]] \xrightarrow{\sim} \mathbf{F}^{\widetilde{\mathbf{N}}}$. Let $\Lambda_K(\alpha) = \beta \in \mathbf{F}^{\widetilde{\mathbf{N}}}(K)$. Assume the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ is boolean. Then one has $K \Vdash \forall n[\beta(n) = 0] \vee \exists n[\beta(n) = 1]$. By $\boxed{\vee}$ this holds only if there exist a cocover of K

$$\{\vartheta_i : K \to A_i \mid i \in I\} \cup \{\xi_j : K \to B_j \mid j \in J\}$$

such that $B_j \Vdash \forall n[(\beta\xi_j)(n) = 0]$ for all $j \in J$ and $A_i \Vdash \exists n[(\beta\vartheta_i)(n) = 1]$ for all $i \in I$. Note that at least one of I or J is nonempty since K is not covered by the empty cover.

For each $i \in I$ there exist a cocover $\{\eta_\ell : A_i \to D_\ell\}_{\ell \in L}$ of A_i such that for all $\ell \in L$, we have $D_\ell \Vdash (\beta\vartheta_i\eta_\ell)(m) = 1$ for some $m \in \widetilde{\mathbf{N}}(D_\ell)$. Let $m = \sum_{t \in T} e_t n_t$ then we have a cocover $\{\xi_t : D_\ell \to C_t = D_\ell/\langle 1 - e_t\rangle\}_{t \in T}$ such that $C_t \Vdash (\beta\vartheta_i\eta_\ell\xi_t)(n_t) = 1$ which implies $\xi_t\eta_\ell\vartheta_i(\alpha(n_t)) = 1$. For each t we can check whether $\alpha(n_t) = 1$. If $\alpha(n_t) = 1$ then we have witness for $\exists n[\alpha(n) = 1]$. Otherwise, we have $\alpha(n_t) = 0$ and $\xi_t\eta_\ell\vartheta_i(0) = 1$. Thus the map $\xi_t\eta_\ell\vartheta_i : K \to C_t$ from the field K cannot be injective, which leaves us with the conclusion that C_t is trivial. If for all $t \in T$, C_t is trivial then D_ℓ is trivial as well. Similarly, if for every

$\ell \in L$, D_ℓ is trivial then A_i is trivial as well. At this point one either have either (i) a natural number m such that $\alpha(m) = 1$ in which case we have a witness for $\exists n[\alpha(n) = 0]$. Or (ii) we have shown that for all $i \in I$, A_i is trivial in which case we have a cocover $\{\xi_j : K \to B_j \mid j \in J\}$ such that $B_j \Vdash \forall n[(\beta\xi_j)(n) = 0]$ for all $j \in J$. Which by $\boxed{\text{LC}}$ means $K \Vdash \forall n[\beta(n) = 0]$ which by $\boxed{\forall}$ means that for all arrows $K \to R$ and elements $d \in \widetilde{\mathbf{N}}(R)$, $R \Vdash \beta(d) = 0$. In particular for the arrow $K \xrightarrow{1_K} K$ and every natural number m one has $K \Vdash \beta(m) = 0$ which implies $K \Vdash \alpha(m) = 0$. By $\boxed{=}$ we get that $\forall m \in \mathbb{N}[\alpha(m) = 0]$. Thus we have shown that LPO holds. □

Corollary 6.9. *It cannot be shown in an intuitionistic metatheory that the topos* $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ *is boolean.*

7 Eliminating the assumption of algebraic closure

Let K be a field of characteristic 0. We consider a typed language $\mathcal{L}[N, F]_K$ of the form described in Section I.3.2 with two basic types N and F and the elements of the field K as its set of constants. Consider a theory T in the language $\mathcal{L}[N, F]_K$, such that T has as an axiom every atomic formula or the negation of one valid in the field K, T equips N with the (Peano) axioms of natural numbers and equips F with the axioms of a field containing K. If we interpret the types N and F by the objects $\widetilde{\mathbf{N}}$ and $\mathbf{F}$, respectively, in the topos $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ then we have, by the results proved earlier, a model of T in $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$.

Let AlgCl be the axiom schema of separable algebraic closure with quantification over the type F, then one has that $T + \mathrm{AlgCl}$ has a model in $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$ with the same interpretation. Let ϕ be a sentence in the language such that $T + \mathrm{AlgCl} \vdash \phi$ in IHOL deduction system. By soundness (See I.3.2) one has that $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J}) \Vdash \phi$, i.e. for all étale algebras R over K, $R \Vdash \phi$ which can be seen as a constructive interpretation of the existence of the separable algebraic closure of K.

In the next Chapter we will give an example of the application of this model to Newton–Puiseux theorem. Here we discuss another example briefly. Suppose one want to show that

"For a discrete field K of characteristic 0, if $f \in K[X, Y]$ is smooth, i.e. $1 \in \langle f, f_X, f_Y \rangle$, then $K[X, Y]/\langle f \rangle$ is a Prüfer ring."

To prove that a ring is Prüfer one needs to prove that it is arithmetical, that is $\forall x, y \exists u, v, w[yu = vx \wedge yw = (1 - u)x]$. Proving that $K[X, Y]/\langle f \rangle$

is arithmetical is easier in the case where K is algebraically closed [Coquand et al., 2010]. Let $\mathbf{F}$ be the algebraic closure of K in $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$. Now $\mathbf{F}[X, Y]/\langle f \rangle$ being arithmetical amounts to having a solution u,v, and w to a linear system $yu = vx$, $yw = (1 - u)x$. Having obtained such solution, by Rouché–Capelli–Fontené theorem we can then conclude that the system have a solution in $K[X, Y]/\langle f \rangle$.

IV

Dynamic Newton–Puiseux Theorem

1 Dynamic Newton–Puiseux Theorem

The proof of Newton–Puiseux theorem in Chapter II depended on the assumption that we have an algebraically closed field at our disposal. In Chapter III we have shown that if assuming the existence of an algebraic closure of fields of characteristic 0 one has a sentence ϕ valid in the system of higher order intuitionistic logic then $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J}) \Vdash \phi$. The statement of Newton–Puiseux theorem of Lemma II.2.3 is one such sentence. Thus we have:

Theorem 1.1. *Let* $\mathbf{F}$ *be the algebraically closed field object of characteristic* 0 *as described in Section III.3.*
Let $G(X,Y) = Y^n + \sum_{i=1}^{n} \alpha_i(X) Y^{n-i} \in \mathbf{F}[[X]][Y]$ *be a monic non-constant polynomial separable over* $\mathbf{F}((X))$. *Then there exist a positive integer* m *and factorization*

$$G(T^m, Y) = \prod_{i=1}^{n} (Y - \eta_i) \qquad \eta_i \in \mathbf{F}[[T]]$$

If we consider only polynomials over the base field, we get the simpler statement:

Theorem 1.2. *In the topos* $\mathrm{Sh}(\mathcal{A}_K^{op}, \mathbf{J})$, *let* $G(X,Y) \in K[[X]][Y]$ *be a monic non-constant polynomial of degree* n *separable over* $K((X))$. *Then there exist* *a positive integer m and a factorization*

$$G(T^m, Y) = \prod_{i=1}^{n} (Y - \eta_i) \qquad \eta_i \in \mathbf{F}[[T]]$$

One surprising aspect of Newton–Puiseux algorithm is that one needs only to find a finite number of roots during the execution of the algorithm. Classically, if one starts with a monic polynomial $G(X,Y) \in K[[X]][Y]$ of degree n, where K, a field of characteristic 0, is not algebraically closed then one can find a finite algebraic extension L/K and a factorization $G(T^m, Y) = \prod_{i=1}^{n}(Y - \eta_i)$ with $\eta_i \in L[[T]]$. This aspect of the algorithm becomes clearer in the sheaf model. We have seen for instance that, in the model, a power series over the algebraically closed field $\mathbf{F}$ is given at each object A of $\mathcal{A}_K^{op}$ as a power series over A. By the forcing conditions, the meaning of the existential quantifier is given locally, that is to say by existence on a finite cover. Thus if a statement asserting the existence of a power series with certain properties is forced, e.g. $K \Vdash \exists \alpha \; \phi(\alpha)$, then the witness of this existential quantifier is a power series $\alpha \in A[[X]]$ where the algebra A is a finite extension of K. The failure of the axiom of dependent choice clarifies the matter even more, by showing that one cannot form power series with infinitely increasing order of roots.

The Newton–Puiseux theorem (Theorem 1.2) has the following computational content.

Theorem 1.3. *Let K be a field of characteristic* 0 *and let Let $G(X,Y) = Y^n + \sum_{i=1}^{n} \alpha_i(X)Y^{n-i} \in K[[X]][Y]$ be a monic non-constant polynomial separable over $K((X))$. Then there exist an étale algebra R over K and a positive integer m such that*

$$G(T^m, Y) = \prod_{i=1}^{n} (Y - \eta_i) \qquad \eta_i \in R[[T]]$$

2 Analysis of the algorithm

Theorem 1.3, as will become apparent from the examples at the end of this section, is not deterministic, in the sense that the étale algebra R is one of several over which the polynomial $G(X,Y)$ factors linearly. On one hand this is not surprising since one can easily see that for any $R \to A$ one has a factorization of $G(X,Y)$ over $A[[X^{1/r}]]$ for some r. On the other hand, one can postulate the existence of a minimal étale algebra(s) B such that $G(X,Y)$ factors linearly over $B[[X^{1/m}]]$ and if $G(X,Y)$ factors linearly over $A[[X^{1/r}]]$, where A is étale, then one has a morphism $B \to A$[1]. In this section we will see that such an algebra indeed exists. We will also show that the morphism $B \to A$ satisfies a stronger condition. Since the examples from the Haskell program will make it clear that there is more than one such minimal algebra. We will also look at the relation between two such algebras.

In order to achieve our task we will consider an arbitrary regular K-algebra A such that the polynomial under consideration $G(X,Y)$ factors linearly over $A[[X^{1/r}]]$ for some r. Then starting from the base field K we will build an algebra R by repetitive extensions and quotients. At each new algebra we obtain by extension or quotient we will show we have a morphism from this algebra to A satisfying certain property. We note that while we can state the dynamic version of Newton–Puiseux directly with the aid of the sheaf model as was done in the previous section, the model is of no help to us here. The reason is that, when presented with a forcing $R \Vdash \phi$ we have no way of knowing how it was constructed and certainly we cannot, without inspecting the algorithm, know whether R is minimal in the above sense. In the following we will be working solely in characteristic 0.

We consider a polynomial $G(X,Y) \in K[[X]][Y]$. Since we start from the base field, we will only need to adjoin roots of monic polynomials. Thus we need not consider the full class of étale K-algebras. It will be sufficient to consider only the triangular separable ones, which we define here:

A *triangular separable K-algebra*

$$R = K[a_1, \ldots, a_n],\ p_1(a_1) = 0,\ p_2(a_1, a_2) = 0, \ldots$$

is a sequence of separable extension starting from a field K, with p_1 in $K[X]$, p_2 in $K[a_1][X]$, ... all monic and separable polynomials. It follows immediately by Lemma III.1.8 that a triangular separable algebra is étale, hence regular. In this case however, the idempotent elements have a simpler direct description. If we have a decomposition $p_l(a_1, \ldots, a_{l-1}, X) = g(X)q(X)$ with g, q in $K[a_1, \ldots, a_{l-1}, X]$ then since p_l is separable, we have a relation $rg + sq = 1$ and $e = r(a_l)g(a_l)$, $1 - e = s(a_l)q(a_l)$ are then idempotent element. We then have a decomposition of R in two triangular separable algebras $p_1, \ldots, p_{l-1}, g, p_{l+1}, \ldots$ and $p_1, \ldots, p_{l-1}, q, p_{l+1}, \ldots$. If we iterate this process we obtain the notion of *decomposition* of a triangular separable algebra R into finitely many triangular separable algebra $R_1, \ldots, R_n$.

This decomposition stops when all the polynomials $p_1, \ldots, p_l$ are irre-

[1] One can say that in some sense B is initial among the étale algebras forcing the Newton–Puiseux statement.

ducible, i.e. when R is a field. For a triangular separable K-algebra R and an ideal I of R, if R/I is a triangular separable K-algebra then we describe R/I as being a *refinement* of R. Thus a refinement of $K[a_1, ..., a_n], p_1, ..., p_n$ is of the form $K[b_1, ..., b_n], q_1, ..., q_n$ with $q_i \mid p_i$.

Now we are ready to begin our analysis. In the following we refer to the elementary symmetric polynomials in n variables by $\sigma_1, ..., \sigma_n$ taking $\sigma_i(X_1, ..., X_n) = \sum_{1 \leq j_1 < ... j_i \leq n} X_{j_1} ... X_{j_i}$.

Lemma 2.1. *Let R be a reduced ring. Given $a_1, ..., a_n \in R$, if $\sigma_i(a_1, ..., a_n) = 0$ for $0 < i \leqslant n$ then $a_1 = a_2 = ... = a_n = 0$.*

Proof. We have $\prod_{i=1}^{n}(X - a_i) = X^n$. Hence, $a_i^n = 0$ for $0 < i \leqslant n$ and since R is reduced, $a_i = 0$. □

Lemma 2.2. *Let R be a reduced ring. Given $\alpha_1, ..., \alpha_n \in R[[X]]$ such that for some positive rational number d we have $\operatorname{ord}(\sigma_i(\alpha_1, ..., \alpha_n)) \geqslant di$ for $0 < i \leqslant n$. Then $\operatorname{ord}(\alpha_i) \geqslant d$ for $0 < i \leqslant n$.*

Proof. Let $\alpha_i = \sum_{j=0}^{\infty} \alpha_i(j) X^j$. We show that $\alpha_i(j) = 0$ if $j < d$. Assume that we have $\alpha_i(j) = 0$ for $j < m < d$. We show then $\alpha_i(m) = 0$ for $i = 1, \ldots, n$. The coefficient of X^{im} in $\sigma_i(\alpha_1, ..., \alpha_n)$ is $\sigma_i(\alpha_1(m), \ldots, \alpha_n(m))$. Since $\operatorname{ord}(\sigma_i(\alpha_1, ..., \alpha_n)) > mi$ we get that $\sigma_i(\alpha_1(m), \ldots, \alpha_n(m)) = 0$ and hence by Lemma 2.1 we get that $\alpha_i(m) = 0$ for $i = 1, \ldots, n$. □

Lemma 2.3. *For a ring R and a reduced extension $R \to A$, let $F = Y^n + \sum_{i=1}^{n} \alpha_i Y^{n-i}$ be an element of $R[[X]][Y]$ such that $F(T^q, T^p Z) = T^{np} F_1(T, Z)$ with F_1 in $R[[T]][Z]$ for some $q > 0, p$. If $F(U^m, Y)$ factors linearly over $A[[U]]$ for some $m > 0$ then $F_1(0, Z)$ factors linearly over A.*

Proof. We have $F(U^m, Y) = \prod_{i=1}^{n}(Y - \eta_i)$, $\eta_i \in A[[U]]$
and hence we have $F(V^{mq}, V^{mp} Z) = \prod_{i=1}^{n}(V^{mp} Z - \eta_i(V^q))$, $\eta_i(U) \in A[[U]]$ and

$$F_1(V^m, Z) = \prod_{i=1}^{n}(Z - V^{-mp}\eta_i(V^q)) = Z^n + \sum_{i=1}^{n} V^{-imp} \beta_i(V^q) Z^{n-i}$$

Since $F_1(T, Z)$ is in $R[[T]][Z]$ we have $imp \leqslant \operatorname{ord} \beta_i(V^q)$.
Since $\beta_i(V^q) = \sigma_i(\eta_1(V^q), \ldots, \eta_n(V^q))$, Lemma 2.2 shows that $mp \leqslant \operatorname{ord} \eta_i(V^q)$ for $0 < i \leqslant n$. Hence $\mu_i(V) = V^{-mp}\eta_i(V^q)$ is in $A[[V]]$ and since $F_1(V, Z) = \prod_{i=1}^{n}(Z - \mu_i(V))$, we have that $F_1(0, Z)$ factors linearly over A, of roots $\mu_i(0)$. □

Definition 2.4. Let $R = K[b_1, ..., b_n], p_1, ..., p_n$ be a triangular separable algebra with p_i of degree m_i and A an algebra over K. Then A splits R if there exist a family of elements $\{a_{i_1,...,i_l} \in A \mid 0 < l \leq n, 0 < i_j \leq m_j\}$ such that

$$p_1 = \prod_{d=0}^{m_1}(X - a_{d_1})$$

$$p_{l+1}(a_{i_1}, a_{i_1,i_2}, ..., a_{i_1,...,i_l}, X) = \prod_{d=0}^{m_{l+1}}(X - a_{i_1,...,i_l,d})$$

for $0 < l < n$

We can view the previous definition as that of a tree of homomorphisms from the subalgebras of R to A. At the root we have the identity homomorphism from K to A under which p_1 factors linearly, i.e. $p_1 = \prod_{j_1=0}^{m_1}(X - a_{j_1})$. From this we obtain m_1 homomorphisms $\varphi_1, ..., \varphi_{m_1}$ from $K[b_1]$ to A each taking b_1 to a different a_{j_1}. If p_2 factors linearly under say φ_1, i.e. $\varphi_1(p_2) = \prod_{j_2=0}^{m_2}(X - a_{1_1,j_2})$ then we obtain m_2 different (since p_2 is separable) homomorphisms $\varphi_{11}, ..., \varphi_{1m_2}$ from $K[b_1, b_2]$ to A. Similarly we obtain m_2 different homomorphisms from $K[b_1, b_2]$ to A by extending $\varphi_2, \varphi_3, ...etc$, thus having $m_1 m_2$ homomorphism in total. Continuing in this fashion we obtain the m different homomorphisms of the family $\mathcal{S}$.

We note that if an K-algebra A splits a triangular separable K-algebra Rthen $A \otimes_K R \cong A^{[R:K]}$. If A is a field then the converse is also true as the following lemma shows.

Lemma 2.5. *Let L/K be a field and $R = K[a_1, ..., a_n], p_1, ..., p_n$ a triangular separable algebra. Then $L \otimes_K R \cong L^{[R:K]}$ only if L splits R.*

Proof. Let $\deg p_i = m_i$, $[R : K] = m = \prod_{i=1}^{n} m_i$ and let $L \otimes_K R \cong L^{[R:K]}$. Then there exist a system of orthogonal idempotents $e_1, ..., e_m$ such that $A = L \otimes_K R \cong A/(1 - e_1) \times .... \times A/(1 - e_m) = L^m$. Let a_{ij} be the image of a_i in $A/(1 - e_j)$. Then we have $(a_{11}, ..., a_{n1}) \neq (a_{12}, ..., a_{n2}) \neq ... \neq (a_{1m}, ..., a_{nm})$ since otherwise we will have the ideals $\langle 1 - e_i \rangle = \langle 1 - e_j \rangle$ for some $i \neq j$. Since p_1 is separable there are up to m_1 different images a_{1j} of a_1. Thus the size of the set $\{a_{1j} \mid 0 < j \leq m\}$ is equal to m_1 only if p_1 factors linearly over L. Similarly, for each different image $\bar{a}_1$ of a_1 there are up to m_2 possible images of a_2 in L since the polynomial $p_2(\bar{a}_1, X)$ is separable. Thus the size of the set $\{(a_{1j}, a_{2j}) \mid 0 < j \leq m\}$ is equal $m_1 m_2$ only if p_1 factors linearly over L and for each root $\bar{a}_1$ of p_1 the polynomial $p_2(\bar{a}_1, X)$ factors linearly over L. Continuing in this fashion we find that the size of the set $\{(a_{1j}, ..., a_{nj}) \mid 0 < j \leq m\}$ is equal to $m_1 ... m_n = m$ only if L splits R. ☐

Lemma 2.6. *Let A be a regular algebra over a field K and let p be a monic non-constant polynomial of degree m in $A[X]$ such that $p = \prod_{i=1}^{m}(X - a_i)$ with $a_i \in A$. If g is a monic non-constant polynomial of degree n such that $g \mid p$ then we have a decomposition $A \cong R_1 \times ... \times R_l$ such that for any R_j in the product $g = \prod_{i=1}^{n}(X - \bar{a}_i)$ with $\bar{a}_i \in R_j$ the image in R_j of some $a_k, 0 < k \leq m$.*

Proof. Let $p = (X - a_1)...(X - a_n)$ for $a_1, ..., a_n \in A$. Let $p = gq$. Then $p(a_1) = g(a_1)q(a_1) = 0$. We can find a decomposition of A into regular algebras $A_1 \times ... \times A_t \times B_1 \times B_s$ such that $g(a_1) = 0$ in $A_i, 0 < i \leq t$ and $g(a_1)$ is a unit in $B_i, 0 < i \leq s$ in which case $q(a_1) = 0$ in B_i. By induction we can find a decomposition of A into a product of regular algebras $R_1, \ldots, R_l$ such that g factors linearly over R_i. □

From Definition 2.4 it is obvious that if an algebra A splits a triangular separable algebra R then A/I splits R for any ideal I of A.

Lemma 2.7. *Let A be a regular K-algebra and R a triangular separable K-algebra such that A splits R. Let B be a refinement of R. Then we can find a decomposition $A \cong A_1 \times ... \times A_m$ such that A_i splits B for $0 < i \leq m$.*

Proof. Let $R = K[a_1, ..., a_n], p_1, ..., p_n$. Then $B = K[\bar{a}_1, ..., \bar{a}_n], g_1, ..., g_n$ where $g_j \mid p_j$ for $0 < j \leq n$. Let $\deg p_j = m_j$ and $\deg g_j = \ell_j$ for $0 < j \leq n$. Since A splits R we have a family of elements $\{a_{i_1,...,i_l} \in A \mid 0 < l \leq n, 0 < i_j \leq m_j\}$ satisfying the condition of Definition 2.4. we have $p_1 = \prod_{i=1}^{m_1}(X - a_{i_1})$. By Lemma 2.6 we decompose A into the product $A_1 \times ... \times A_t$ such that for any given A_k in the product we have $p_1 = \prod_{i=1}^{m_1}(X - \bar{a}_{i_1})$ and $g_1 = \prod_{i=1}^{\ell_1}(X - \bar{a}_{i_1})$ with $\bar{a}_{i_1} \in A_k$ for $0 < i \leq m_1$. Since each $\bar{a}_{i_1}$ is an image of some a_{j_1} and $p_2(a_{j_1}, X)$ factors linearly over A we have that $p_2(\bar{a}_{i_1}, X)$ factors linearly over A_k but then $g_2(\bar{a}_{i_1}, X)$ divides $p_2(\bar{a}_{i_1}, X)$ and thus by Lemma 2.6 we can decompose A_k into the product $B_1 \times ... \times B_s$ such that for a given B_r in the product we have $p_2(\bar{a}_{i_1}, X) = \prod_{j=1}^{m_2}(X - \bar{a}_{i_1,j_2})$ and $g_2(\bar{a}_{i_1}, X) = \prod_{j=1}^{\ell_2}(X - \bar{a}_{i_1,j_2})$. By induction on the m_1 values of $\bar{a}_{i_1}$ we can find a decomposition $D_1 \times ... \times D_l$ such that in each D_i we have $g_1(X) = \prod_{i=1}^{\ell_1}(X - \bar{a}_{i_1})$ and $g_2(\bar{a}_{i_1}, X) = \prod_{j=1}^{\ell_2}(X - \bar{a}_{i_1,j_2})$ for $0 < i \leq \ell_1$. Continuing in this fashion we can find a decomposition of A such that each algebra in the decomposition splits B. □

Lemma 2.8. *Let A be a regular K-algebra with decomposition $A \cong A_1 \times ... \times A_t$. Let B a triangular separable algebra. If A_i splits B for all $1 \leq i \leq t$ then A splits B.*

Proof. Let $B = K[a_1, ..., a_n], g_1, ..., g_n$ with $\deg g_i = m_i$. Then we have a family of elements $\{a^{(i)}_{k_1,...,k_l} \mid 0 < k_j \leq m_j, 0 < j \leq n\}$ in A_i satisfying the conditions of Definition 2.4. We claim that the family

$$\mathcal{S} = \{a_{k_1,...,k_l} \mid a_{k_1,...,k_l} = (a^{(1)}_{k_1,...,k_l}, ..., a^{(t)}_{k_1,...,k_l}), 0 < k_j \leq m_j, 0 < j \leq n\}$$

of A elements satisfy the conditions of Definition 2.4. Since we have a factorization $g_1 = \prod_{l=1}^{m_1}(X - a_l^{(i)})$ over A_i, we have a factorization $g_1 = \prod_{l=1}^{m_1}(X - (a_l^{(1)}, ..., a_l^{(t)})) = \prod_{l=1}^{m_1}(X - a_l)$ over A. Since for $0 < l \leq m_1$ we have a factorization $g_2(a_l^{(i)}, X) = \prod_{j=1}^{m_2}(X - a_{l,j}^{(i)})$ of in A_i, we have a factorization $g_2(a_l, X) = \prod_{j=1}^{m_2}(X - (a_{l,j}^{(1)}, ..., a_{l,j}^{(t)})) = \prod_{j=1}^{m_2}(X - a_{l,j})$. Continuing in this fashion we verify that the family $\mathcal{S}$ satisfy the requirements of Definition 2.4. □

Corollary 2.9. *Let A be a regular K-algebra and B be a triangular separable K-algebra such that A splits B. Then A splits any refinement of B.*

Lemma 2.10. *Let R be a regular ring and let $a_1, ..., a_n \in R$ such that $1 \in \langle a_1, ..., a_n \rangle$. Then we can find a decomposition $R \cong R_1 \times ... \times R_m$ such that for each R_i we have a_j a unit in R_i for some $1 \leq j \leq n$.*

Proof. We have a decomposition $R \cong A \times B$ with a_n unit in A and zero in B. We have $1 \in \langle a_1, ..., a_{n-1} \rangle$ in B. The statement follows by induction. □

Going back to Newton–Puiseux theorem.

Lemma 2.11. *Let R be a triangular separable algebra over a field K. Let $F(X,Y) = \sum_{i=0}^{n} \alpha_i(X)Y^{n-i} \in R[[X]][Y]$ be a monic polynomial such that $PF + QF_Y = \gamma$ for some $P, Q \in R[[X]][Y]$ and $\gamma \# 0$ in $K[[X]]$. Then we can find a decomposition $R_1, ...$ of R such that in each R_i we have $\alpha_k(m)$ a unit for some m and $k = n$ or $k = n-1$.*

Proof. Since $\gamma \# 0 \in K[[X]]$ we have $\gamma(\ell)$ a unit for some ℓ. Since $PF + QF_Y = \gamma$, we have $\eta\alpha_n + \theta\alpha_{n-1} = \gamma$ with $\eta = P(0)$ and $\theta = Q(0)$. Then we have $\sum_{i+j=\ell} \eta(i)\alpha_n(j) + \theta(i)\alpha_{n-1}(j) = \gamma(\ell)$. By Lemma 2.10 we have a decomposition $R_1, ...$ of R such that in R_i we have $\alpha_k(m)$ is a unit for some m and $k = n \vee k = n-1$. □

With the help of Lemma 2.3, Lemma 2.8 and Corollary 2.9 we have the following result.

Lemma 2.12. *Let $R = K[a_1, ..., a_n], p_1, ..., p_n$ be a triangular separable algebra with $\deg p_i = m_i$. Let $F(a_1, ..., a_n, X, Y) = Y^n + \sum_{i=1}^{n} \alpha_i(X)Y^{n-i} \in R[[X]][Y]$ be a monic non-constant polynomial of degree $n \geq 2$ such that $PF + QF_Y = \gamma$ for some $P, Q \in R[[X]][Y]$, $\gamma \in R[[X]]$ with $\gamma \# 0$. There exists then a decomposition $R_1, \ldots$ of R and for each i there exist $m > 0$ and a proper factorization $F(T^m, Y) = G(T, Y)H(T, Y)$ with G and H in $S_i[[T]][Y]$ where $S_i = R_i[b], q$ is a separable extension of R_i.*
Moreover, Let A be a regular K-algebra such that A splits R and let $\{a_{i_1,...,i_l} \mid 0 < l \leq n, 0 < i \leq m_i\}$ be the family of elements in A satisfying the conditions in Definition 2.4. If $F(a_{i_1}, ..., a_{i_1,...,i_n}, X, Y)$ factors linearly over $A[[U]]$ for $0 < i \leq m_i$ where $U^v = X$ for some positive integer v then A splits S_i.

Proof. By Lemma 2.11 we have a decomposition $A_1, ...$ of R such that in each A_i we have $\alpha_k(m)$ a unit for some m and $k = n$ or $k = n - 1$. The rest of the proof proceeds as the proof of Lemma II.2.2, assuming w.l.o.g. $\alpha_1 = 0$. We first find a decomposition $R_1, \ldots$ of R and for each l we can then find m and p such that $\alpha_m(p)$ is invertible and $\alpha_i(j) = 0$ whenever $j/i < p/m$ in R_l. We can then write

$$F(T^m, T^p Z) = T^{np}(Z^n + c_2(T)Z^{n-2} + \cdots + c_n(T))$$

with $\operatorname{ord} c_m = 0$. Since A splits R then by Lemma 2.7 we can find a decomposition $A_1, \ldots$ of A such that each A_i splits R_l for each l. We then find a further decomposition $R_{l1}, R_{l2}, \ldots$ of R_l and for each t a number s and a separable extension $R_{lt}[a]$ of R_{lt} such that

$$q = Z^n + c_2(0)Z^{n-2} + \cdots + c_n(0) = (Z - a)^s L(Z)$$

with $L(a)$ invertible. Similarly, we can decompose each A_i further into $B_1, \ldots$ such that each B_i splits each R_{lt} for all l, t. Let the family $\mathcal{F} = \{b_{i_1,...,i_l} \mid 0 < l \leq m, 0 < i \leq m_i\}$ be the image of the family $\{a_{i_1,...,i_l} \mid 0 < l \leq n, 0 < i \leq m_i\}$ in B_i. Then B_i splits R with $\mathcal{F}$ as the family of elements of B_i satisfying Definition 2.4. But then $F(b_{i_1}, ..., b_{i_1,...,i_n}, X, Y)$ factors linearly over B_i. For some subfamily $\{c_{i_1}, ..., c_{i_1,...,i_l} \mid 0 < l \leq n, 0 < i_j \leq \bar{m}_j \leq m_j\} \subset \mathcal{F}$ of elements in B_i we have that B_i splits R_{lt}. Thus $F(c_{i_1}, ..., c_{i_1,...,i_n}, X, Y)$ factors linearly over B_i for all $c_{i_1}, ..., c_{i_1,...,i_n}$ in the family. By Lemma 2.3 we have that $q(c_{i_1}, ..., c_{i_1,...,i_n}, Z)$ factors linearly over B_i for all $c_{i_1}, ..., c_{i_1,...,i_n}$. Thus B_i splits the extension $R_{lt}[a]$. But then by Lemma 2.8 we have that A splits $R_{lt}[a]$. Using Hensel's Lemma II.2.1, we can lift this to a proper decomposition $Z^n + c_2(T)Z^{n-2} + \cdots + c_n(T) = G_1(T, Z)H_1(T, Z)$ with $G_1(T, Z)$ monic of degree t and $H_1(T, Z)$ monic of degree u. We take $G(T, Y) = T^{tp}G_1(T, Y/T^p)$ and $H(T, Y) = T^{up}H_1(T, Y/T^p)$. □

As a corollary we get the following version of Newton–Puiseux theorem which follows by induction from Lemma 2.12.

Theorem 2.13. *Let $F(X,Y) = Y^n + \sum_{i=1}^{n} \alpha_i(X)Y^{n-i} \in K[[X]][Y]$ be a monic non-constant polynomial separable over $K((X))$. There exists then a triangular separable algebra R over K and $m > 0$ and a factorization*

$$F(T^m, Y) = \prod_{i=1}^{n} (Y - \eta_i) \qquad \eta_i \in R[[T]]$$

Moreover, if A is a regular algebra over K such that $F(X,Y)$ factors linearly over $A[[X^{1/s}]]$ for some positive integer s then A splits R.

We note that the algebra R above is not unique.

Corollary 2.14. *Let A and B be two triangular separable algebras obtained by the algorithm of Theorem 2.13, i.e. minimal in the sense expressed in the theorem. Then A splits B and B splits A. Consequently, a triangular separable algebra obtained by this algorithm splits itself.*

Thus given any two algebras $R_1$and R_2 obtained by the algorithm and two prime ideals $P_1 \in \text{Spec}(R_1)$ and $P_2 \in \text{Spec}(R_2)$ we have a field isomorphism $R_1/P_1 \cong R_2/P_2$. Therefore all the algebras obtained are approximations of the same field L. Since L splits all the algebras and itself is a refinement, L splits itself, i.e. $L \otimes_K L \cong L^{[L:K]}$ and L is a normal, in fact a Galois extension of K.

Classically, this field L is the field of constants generated over K by the set of coefficients of the Puiseux expansions of F. The set of Puiseux expansions of F is closed under the action of $\text{Gal}(\bar{K}/K)$, where $\bar{K}$ is the algebraic closure of K. Thus the field of constants generated by the coefficients of the expansions of F is a Galois extension. The algebras generated by our algorithm are powers of this field of constants, hence are in some sense minimal extensions.

Even without the notion of prime ideals we can still show interesting relationship between the algebras produced by the algorithm of Theorem 2.13. The plan is to show that any two such algebras A and B are essentially isomorphic in the sense that each of them is equal to the power of some common triangular separable algebra R, i.e. $A \cong R^m$ and $B \cong R^n$ for some positive integers m, n. To show that $A \cong R^m$ we have to be able to decompose A. To do this we need to constructively obtain a system of orthogonal nontrivial (unless $A \cong R$ already) idempotents $e_1, ..., e_m$. Since A and B split each other, the composition of these maps gives a homomorphism from A to itself. We know that a homomorphism between a field and itself is an automorphism thus as we would expect if there is a homomorphism from a triangular separable algebra A to itself that is not an automorphism we can decompose this algebra non trivially. We use the composition of the split maps from A to B and vice versa as our homomorphism this will enable us to repeat the process after the initial decomposition, that is if $A/e_1, B/e_2$ are algebras in the decompositions of A and B, respectively, we know that they split each other. This process of decomposition stops once we reach the common algebra R.

Lemma 2.15. *Let A be a triangular separable algebra over a field K and let $\pi : A \to A$ be K-homomorphism. Then π is either an automorphism of A or we can find a non-trivial decomposition $A \cong A_1 \times ... \times A_t$.*

Proof. Let $A = K[a_1, ..., a_l], p_1, ..., p_l$ with $\deg p_i = n_i$. Let π map a_i to $\bar{a}_i$, for $0 < i \leq l$. Then $\bar{a}_i$ is a root of $\pi(p_i) = p_i(\bar{a}_1, ..., \bar{a}_{i-1}, X)$. The set of vectors $\mathcal{S} = \{a_1^{i_1} ... a_l^{i_l} \mid 0 \leq i_j < n_j, 0 < j \leq l\}$ is a basis for the vector space A over K. If the image $\pi(\mathcal{S}) = \{\bar{a}_1^{i_1} ... \bar{a}_l^{i_l} \mid 0 \leq i_j < n_j, 0 < j \leq l\}$ is a basis for A, i.e. $\pi(\mathcal{S})$ is a linearly independent set then π is surjective and thus an automorphism.
Assuming π is not an automorphism, then the kernel of π is non-trivial, i.e. we have a non-zero non-unit element in $\ker \pi$, thus we have a non-trivial decomposition of A. □

Theorem 2.16. *Let A and B be triangular separable algebras over a field K such that A splits B and B splits A. Then there exist a triangular separable algebra R over K and two positive integers m, n such that $A \cong R^n$ and $B \cong R^m$.*

Proof. First we note that by Corollary 2.9 if A splits B then A splits any refinement of B. Trivially if A splits B then any refinement of A splits B. Since A and B split each other then there is K-homomorphisms $\vartheta : B \to A$ and $\varphi : A \to B$. The maps $\pi = \vartheta \circ \varphi$ and $\varepsilon = \varphi \circ \vartheta$ are K-homomorphisms from A to A and B to B respectively. If both π and ε are automorphisms then we are done. Otherwise, by Lemma 2.15 we can find a decomposition of either A or B. By induction on $\dim(A) + \dim(B)$ the statement follows. □

Theorems 2.16 and 2.13 show that the algebras obtained by the algorithm of Theorem 2.13 are equal to the power of some common algebra. This common triangular separable algebra is an approximation, for lack of irreducibility test for polynomials, of the normal field extension of K generated by the coefficients of the Puiseux expansions $\eta_i \in \bar{K}[[X^{1/m}]]$ of F, where $\bar{K}$ is the algebraic closure of K.

The following are examples from a Haskell implementation of the algorithm. We truncate the different factors unevenly for readability.

Example 2.1. Applying the algorithm to $F(X,Y) = Y^4 - 3Y^2 + XY + X^2 \in Q[X][Y]$ we get.

- $Q[a,b,c], a = 0,\ b^2 - 13/36 = 0,\ c^2 - 3 = 0$
  $F(X,Y) =$
  $(Y + (-b - \frac{1}{6})X + (-\frac{31}{351}b - \frac{7}{162})X^3 + (-\frac{1415}{41067}b - \frac{29}{1458})X^5 + ...)$
  $(Y + (b - \frac{1}{6})X + (\frac{31}{351}b - \frac{7}{162})X^3 + (\frac{1415}{41067}b - \frac{29}{1458})X^5 + ...)$
  $(Y - c + \frac{1}{6}X + \frac{5}{72}cX^2 + \frac{7}{162}X^3 + \frac{185}{10368}cX^4 + \frac{29}{1458}X^5 + ...)$
  $(Y + c + \frac{1}{6}X - \frac{5}{72}cX^2 + \frac{7}{162}X^3 - \frac{185}{10368}cX^4 + \frac{29}{1458}X^5 + ...)$

- $Q[a,b,c], a^2 - 3 = 0, b - a/3 = 0, c^2 - 13/36 = 0$
  $F(X,Y) =$
  $(Y - a + \frac{1}{6}X + \frac{5}{72}aX^2 + \frac{7}{162}X^3 + \frac{185}{10368}aX^4 + \frac{29}{1458}X^5 + ...)$
  $(Y + (-c - \frac{1}{6})X + (-\frac{31}{351}c - \frac{7}{162})X^3 + (-\frac{1415}{41067}c - \frac{29}{1458})X^5 + ...)$
  $(Y + (c - \frac{1}{6})X + (\frac{31}{351}c - \frac{7}{162})X^3 + (\frac{1415}{41067}c - \frac{29}{1458})X^5 + ...)$
  $(Y + a + \frac{1}{6}X - \frac{5}{72}aX^2 + \frac{7}{162}X^3 - \frac{185}{10368}aX^4 + \frac{29}{1458}X^5 + ...)$

- $Q[a,b,c], a^2 - 3 = 0, b + 2a/3 = 0, c^2 - 13/36 = 0$
  $F(X,Y) =$
  $(Y - a + \frac{1}{6}X + \frac{5}{72}aX^2 + \frac{7}{162}X^3 + \frac{185}{10368}aX^4 + \frac{29}{1458}X^5 + ...)$
  $(Y + a + \frac{1}{6}X - \frac{5}{72}aX^2 + \frac{7}{162}X^3 - \frac{185}{10368}aX^4 + \frac{29}{1458}X^5 + ...)$
  $(Y + (-c - \frac{1}{6})X + (-\frac{31}{351}c - \frac{7}{162})X^3 + (-\frac{1415}{41067}c - \frac{29}{1458})X^5 + ...)$
  $(Y + (c - \frac{1}{6})X + (\frac{31}{351}c - \frac{7}{162})X^3 + (\frac{1415}{41067}c - \frac{29}{1458})X^5 + ...)$

The algebras in the above example can be readily seen to be isomorphic. However, as we will show next, this is not always the case.

Example 2.2. To illustrate Theorem 2.16 we show how it works in the context of an example computation. The polynomial is $F(X,Y) = Y^6 + X^6 + 3X^2Y^4 + 3X^4Y^2 - 4X^2Y^2$. The following are two of the several triangular separable algebras obtained by our algorithm along with their respective factorization of $F(X,Y)$.

$$A = Q[a,b,c,d,e], p_1, p_2, p_3, p_5$$
$$p_1 = Y^4 - 4,\ p_2 = Y - \tfrac{1}{5}a,\ p_3 = Y^2 - \tfrac{1}{4},$$
$$p_4 = Y^3 + \tfrac{2}{3}a^2Y + \tfrac{20}{27}a^3,\ p_5 = Y^2 + \tfrac{3}{4}d^2 + \tfrac{2}{3}a^2$$
$$F(X,Y) =$$
$$\big(Y - aX^{\frac{1}{2}} + \frac{3}{16}a^3X^{\frac{3}{2}} + ...\big)\big(Y - cX^2 + ...\big)\big(Y + cX^2 + ...\big)$$
$$\big(Y + (-d + \tfrac{1}{3}a)X^{\frac{1}{2}} + (-\tfrac{3}{16}ad^2 - \tfrac{1}{16}a^2d - \tfrac{7}{48}a^3)X^{\frac{3}{2}} + ...\big)$$
$$\big(Y + (-e + \tfrac{1}{2}d + a/3)X^{\frac{1}{2}} +$$
$$(\tfrac{3}{16}ade - \tfrac{1}{16}a^2e + \tfrac{3}{32}ad^2 + \tfrac{1}{32}a^2d - \tfrac{1}{48}a^3)X^{\frac{3}{2}} + ...\big)$$
$$\big(Y + (e + \tfrac{1}{2}d + \tfrac{1}{3}a)X^{\frac{1}{2}} +$$
$$(-\tfrac{3}{16}ade + \tfrac{1}{16}a^2e + \tfrac{3}{32}ad^2 + \tfrac{1}{32}a^2d - \tfrac{1}{48}a^3)X^{\frac{3}{2}} + ...\big)$$

$$B = Q[r,t,u,v,w], q_1, q_2, q_3, q_5$$
$$q_1 = Y^4 - 4,\ q_2 = Y + \tfrac{4}{5}r,\ q_3 = Y,\ q_4 = Y^2 - \tfrac{1}{4},\ q_5 = Y^2 + r^2$$
$$F(X,Y) = (Y - rX^{\frac{1}{2}} + \tfrac{3}{16}r^3X^{\frac{3}{2}} + ...)(Y + rX^{\frac{1}{2}} - \tfrac{3}{16}r^3X^{\frac{3}{2}} + ...)$$
$$(Y - vX^2 + ...)(Y + vX^2 + ...)$$
$$(Y - wX^{\frac{1}{2}} - \tfrac{3}{16}r^2wX^{\frac{3}{2}} + ...)(Y + wX^{\frac{1}{2}} + \tfrac{3}{16}r^2wX^{\frac{3}{2}} + ...)$$

We now show that the two algebras indeed split each other. Over B the polynomial p_1 factors as $p_1 = (Y - r)(Y + r)(Y - w)(Y + w)$. Each of these factors partly specify a homomorphism taking a to a zero of p_1 in B. For each we get a factorization of p_4 over B.

- $a \mapsto r$
  $p_4 = (Y + 2r/3)(Y - w - r/3)(Y + w - r/3)$
- $a \mapsto -r$
  $p_4 = (Y - 2r/3)(Y - w + r/3)(Y + w + r/3)$
- $a \mapsto w$
  $p_4 = (Y - r - w/3)(Y + r - w/3)(Y + 2w/3)$
- $a \mapsto -w$
  $p_4 = (Y - r + w/3)(Y + r + w/3)(Y - 2w/3)$

For each of the 4 mappings of a we get 3 mappings of d. Now we see we have 12 different mappings arising from the different mappings of a and d. Each of these 12 mappings will give rise to 2 different mappings of e (factorization of p_5)...etc. Thus we have a number of homomorphisms equal to the dimension of the algebra, that is 48 homomorphisms. We avoid listing all these homomorphisms here. In conclusion, we see that B splits A. Similarly, we have that A splits B. We show only one of the 16 homomorphisms below. The polynomial q_1 factors linearly over A as $q_1 = (Y-a)(Y-d+a/3)(Y-e+d/2+a/3)(Y+e+d/2+a/3)$. Under the map $r \mapsto a$ we get a factorization of q_5 over A as

$$
\begin{aligned}
&q_5 = Y^2 + a^2 = \\
&(Y - a^2d^2e/8 + a^3de/12 - 5e/9 - a^3d^2/8 - 2d/3 - 2a/9) \\
&(Y + a^2d^2e/8 - a^3de/12 + 5e/9 + a^3d^2/8 + 2d/3 + 2a/9)
\end{aligned}
$$

Now to the application of Theorem 2.16. Under the map above we have an endomorphism $a \mapsto r \mapsto a$ and $d \mapsto -2r/3 \mapsto -2a/3$. Thus in the kernel we have the non-zero element $d + 2a/3$ and as expected $Y + 2a/3$ divides p_4. Using this we obtain a decomposition of $A \cong A_1 \times A_2$. We have $A_1 = Q[a,b,c,d,e], p_1, p_2, p_3, g_4, p_5$ with $g_4 = Y + 2a/3$ and $A_2 = Q[a,b,c,d,e], p_1, p_2, p_3, h_4, p_5$ with $h_4 = Y^2 - 2aY/3 + 10a^2/9$.
With $d + 2a/3 = 0$ in A_1, $p_5 = Y^2 + 3d^2/4 + 2a^2/3 = Y^2 + a^2$ and we can see immediately that $A_1 \cong B$. Similarly, we can decompose the algebra $A_2 \cong C_1 \times C_2$, where $C_1 = Q[a,b,c,d,e], p_1, p_2, p_3, h_4, g_5$ with $g_5 = Y - d/2 + 2a/3$ and $C_2 = Q[a,b,c,d,e], p_1, p_2, p_3, h_4, h_5$ with $h_5 = Y + d/2 - 2a/3$. The polynomial q_5 factors linearly over both C_1 and C_2 as $q_5 = (Y - d + a/3)(Y + d - a/3)$. We can readily see that both C_1 and C_2 are isomorphic to B, through the C_1 automorphism $a \mapsto r \mapsto a, d \mapsto w + r/3 \mapsto d$. Thus proving $A \cong B^3$.

Part B

Type Theory: The Independence of Markov's Principle

Introduction

Markov's principle has a special status in constructive mathematics. One way to formulate this principle is that if it is impossible that a given algorithm does not terminate, then it does terminate. It is equivalent to the fact that if a set of natural number and its complement are both computably enumerable, then this set is decidable. This form is often used in recursivity theory. This principle was first formulated by Markov, who called it "Leningrad's principle", and founded a branch of constructive mathematics around this principle [Margenstern, 1995].

This principle is also equivalent to the fact that if a given real number is *not* equal to 0 then this number is *apart* from 0 (that is this number is $< -r$ or $> r$ for some rational number $r > 0$). On this form, it was explicitly *refuted* by Brouwer in intuitionistic mathematics, who gave an example of a real number (well defined intuitionistically) which is not equal to 0, but also not apart from 0. (The motivation of Brouwer for this example was to show the necessity of using *negation* in intuitionistic mathematics [Brouwer, 1975].) The idea of Brouwer can be represented formally using topological models [van Dalen, 1978].

In a neutral approach to mathematics, such as Bishop's [Bishop, 1967], Markov's principle is simply left undecided. We also expect to be able to prove that Markov's principle is *not* provable in formal system in which we can express Bishop's mathematics. For instance, Kreisel [Kreisel, 1959] introduced *modified realizability* to show that Markov's principle is not derivable in the formal system HA^{ω}. Similarly, one would expect that Markov's principle is *not* derivable in Martin-Löf type theory [Martin-Löf, 1972], but, as far as we know, such a result has not been established yet. [2]

We say that a statement A is *independent* of some formal system if A

[2] The paper [Hyland and Ong, 1993] presents a model of the calculus of constructions using the idea of modified realizability, and it seems possible to use also this technique to interpret the type theory we consider and prove in this way the independence of Markov's principle.

cannot be derived in that system. A statement in the formal system of Martin-Löf type theory (MLTT) is represented by a closed type. A statement/type A is derivable if it is inhabited by some term t (written MLTT $\vdash t : A$). This is the so-called propositions-as-types principle. Correspondingly we say that a statement A (represented as a type) is independent of MLTT if there is no term t such that MLTT $\vdash t : A$.

The main result of this paper is to show that Markov's principle is independent of Martin-Löf type theory.[3]

The main idea for proving this independence is to follow Brouwer's argument. We want to extend type theory with a "generic" infinite sequence of 0 and 1 and establish that it is both absurd that this generic sequence is never 0, but also that we cannot show that it *has to* take the value 0. To add such a generic sequence is exactly like adding a *Cohen real* [Cohen, 1963] in forcing extension of set theory. A natural attempt for doing this will be to consider a *topological model* of type theory (sheaf model over Cantor space), extending the work [van Dalen, 1978] to type theory. However, while it is well understood how to represent universes in *presheaf* model [Hofmann and Streicher, 199?], it has turned out to be surprisingly difficult to represent universes in *sheaf* models, see [Xu and Escardó, 2016] and [Streicher, 2005]. Our approach is here instead a purely *syntactical* description of a forcing extension of type theory (refining previous work of [Coquand and Jaber, 2010]), which contains a formal symbol for the generic sequence and a proof that it is absurd that this generic sequence is never 0, together with a *normalization* theorem, from which we can deduce that we *cannot* prove that this generic sequence has to take the value 0. Since this formal system is an extension of type theory, the independence of Markov's principle follows[4].

As stated in [Kopylov and Nogin, 2001], which describes an elegant generalization of this principle in type theory, Markov's principle is an important technical tool for proving termination of computations, and thus can play a crucial role if type theory is extended with general recursion as in [Constable and Smith, 1987].

This part is organized as follows. In Chapter V we first describe the

[3]Some authors define independence in the stronger sense "A statement is independent of a formal system if neither the statement nor its negation is provable in the system", e.g. [Kunen, 1980]. We will also establish the independence of Markov's principle in this stronger sense.

[4]In [Coste-Roy et al., 1980] it is shown that Markov's principle is independent from HA^ω in a different topos. Namely, the topos of sheaves over the category $\mathcal{P}(N)$ of subsets of natural numbers with the finite partitions topology. It seems quite likely that one also use this site similarly to Cantor space to show the independence of Markov's principle from type theory.

rules of the version of type theory we are considering. This version can be seen as a simplified version of type theory as represented in the system Agda [Norell, 2007], and in particular, contrary to the work [Coquand and Jaber, 2010], we allow η-conversion, and we express conversion as *judgment*. Markov's principle can be formulated in a natural way in this formal system. We describe then the forcing extension of type theory, where we add a Cohen real. For proving normalization, we follow Tait's computability method [Tait, 1967; Martin-Löf, 1972], but we have to consider an extension of this with a computability *relation* in order to interpret the conversion judgment. Using this computability argument, it is then possible to show that we cannot show that the generic sequence has to take the value 0. We end by a refinement of this method, giving a consistent extension of type theory where the *negation* of Markov's principle is provable

Chapter V is dedicated to an informal discussion of the problems with finding a general sheaf model of type theory. First of these problems is the problem of interpretation of the universe as mentioned earlier. Second of these is the problem of interpretation of Σ types in a model without choice (we have seen such topos in Chapter III).

V

The Independence of Markov's Principle in Type Theory

1 Type theory and forcing extension

The syntax of our type theory is given by the grammar:

$$\begin{aligned} t, u, A, B := x \mid & \perp\mathsf{rec}\ (\lambda x.A) \mid \mathsf{unitrec}\ (\lambda x.A)\ t \\ & \mid \mathsf{boolrec}\ (\lambda x.A)\ t\ u \mid \mathsf{natrec}\ (\lambda x.A)\ t\ u \\ & \mid U \mid N \mid N_0 \mid N_1 \mid N_2 \mid 0 \mid 1 \mid \mathsf{S}\ t \\ & \mid \Pi(x : A)B \mid \lambda x.t \mid t\ u \mid \Sigma(x : A)B \mid (t, u) \mid t.1 \mid t.2 \end{aligned}$$

The terms N_0, N_1, N_2, and N will denote , respectively, the empty type, the unit type, the type of booleans, and the type of natural numbers. The term U will denote the universe, i.e. the type of small types. We use the notation $\bar{n}$ as a short hand for the term $\mathsf{S}^n\, 0$, where S is the successor constructor of natural numbers.

1.1 Type system

We describe a type theory with one universe à la Russell, natural numbers, functional extensionality and surjective pairing, hereafter referred to as MLTT.[1] The type theory has the following judgment forms 1. $\Gamma \vdash$

[1]This is a type system similar to Martin-löf's [Martin-Löf, 1972] except that we have η-conversion and surjective pairing.

2. $\Gamma \vdash A$ 3. $\Gamma \vdash t : A$ 4. $\Gamma \vdash A = B$ 5. $\Gamma \vdash t = u : A$ The first expresses that Γ is a well-formed context, the second that A is a type in the context Γ, and the third that t is a term of type A in the context Γ. The fourth and fifth express type and term equality respectively. Below we outline the inference rules of this type theory. We use the notation $F \to G$ for $\Pi(x : F)G$ when G doesn't depend on F and $\neg A$ for $A \to N_0$.

Natural numbers:

$$\frac{\Gamma \vdash}{\Gamma \vdash N} \qquad \frac{\Gamma \vdash}{\Gamma \vdash 0 : N} \qquad \frac{\Gamma \vdash n : N}{\Gamma \vdash \mathsf{S}\, n : N}$$

$$\frac{\Gamma, x : N \vdash F \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash g : \Pi(x : N)(F[x] \to F[\mathsf{S}\, x])}{\Gamma \vdash \mathsf{natrec}\ (\lambda x.F)\, a_0\, g : \Pi(x : N)F}$$

$$\frac{\Gamma, x : N \vdash F \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash g : \Pi(x : N)(F[x] \to F[\mathsf{S}\, x])}{\Gamma \vdash \mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, 0 = a_0 : F[0]}$$

$$\frac{\Gamma, x : N \vdash F \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash n : N \quad \Gamma \vdash g : \Pi(x : N)(F[x] \to F[\mathsf{S}\, x])}{\Gamma \vdash \mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, (\mathsf{S}\, n) = g\, n\, (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, n) : F[\mathsf{S}\, n]}$$

$$\frac{\Gamma, x : N \vdash F = G \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash g : \Pi(x : N)(F[x] \to F[\mathsf{S}\, x])}{\Gamma \vdash \mathsf{natrec}\ (\lambda x.F)\, a_0\, g = \mathsf{natrec}\ (\lambda x.G)\, a_0\, g : \Pi(x : N)F}$$

Booleans:

$$\frac{\Gamma \vdash}{\Gamma \vdash N_2} \qquad \frac{\Gamma \vdash}{\Gamma \vdash 0 : N_2} \qquad \frac{\Gamma \vdash}{\Gamma \vdash 1 : N_2}$$

$$\frac{\Gamma, x : N_2 \vdash F \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash a_1 : F[1]}{\Gamma \vdash \mathsf{boolrec}\ (\lambda x.F)\, a_0\, a_1 : \Pi(x : N_2)F}$$

$$\frac{\Gamma, x : N_2 \vdash F \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash a_1 : F[1]}{\Gamma \vdash \mathsf{boolrec}\ (\lambda x.F)\, a_0\, a_1\, 0 = a_0 : F[0]}$$

$$\frac{\Gamma, x : N_2 \vdash F \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash a_1 : F[1]}{\Gamma \vdash \mathsf{boolrec}\ (\lambda x.F)\, a_0\, a_1\, 1 = a_1 : F[1]}$$

$$\frac{\Gamma, x : N_2 \vdash F = G \quad \Gamma \vdash a_0 : F[0] \quad \Gamma \vdash a_1 : F[1]}{\Gamma \vdash \mathsf{boolrec}\ (\lambda x.F)\, a_0\, a_1 = \mathsf{boolrec}\ (\lambda x.G)\, a_0\, a_1 : \Pi(x : N_2)F}$$

Unit Type:

$$\frac{\Gamma \vdash}{\Gamma \vdash N_1} \qquad \frac{\Gamma \vdash}{\Gamma \vdash 0 : N_1}$$

$$\frac{\Gamma, x : N_1 \vdash F \quad \Gamma \vdash a : F[0]}{\Gamma \vdash \mathsf{unitrec}\ (\lambda x.F)\, a : \Pi(x : N_1)F} \qquad \frac{\Gamma, x : N_1 \vdash F \quad \Gamma \vdash a : F[0]}{\Gamma \vdash \mathsf{unitrec}\ (\lambda x.F)\, a\, 0 = a : F[0]}$$

$$\frac{\Gamma, x\!:\!N_1 \vdash F = G \quad \Gamma \vdash a\!:\!F[0]}{\Gamma \vdash \mathsf{unitrec}\ (\lambda x.F)\, a = \mathsf{unitrec}\ (\lambda x.G)\, a\!:\!\Pi(x\!:\!N_1)F}$$

Empty type:

$$\frac{\Gamma \vdash}{\Gamma \vdash N_0} \qquad \frac{\Gamma, x\!:\!N_0 \vdash F}{\Gamma \vdash \bot\mathsf{rec}\ (\lambda x.F)\!:\!\Pi(x\!:\!N_0)F}$$

$$\frac{\Gamma, x\!:\!N_0 \vDash_p F = G}{\Gamma \vDash_p \bot\mathsf{rec}\ (\lambda x.F) = \bot\mathsf{rec}\ (\lambda x.G)\!:\!\Pi(x\!:\!N_0)F}$$

Dependent functions:

$$\frac{\Gamma \vdash F \quad \Gamma, x\!:\!F \vdash G}{\Gamma \vdash \Pi(x\!:\!F)G} \qquad \frac{\Gamma \vdash F = H \quad \Gamma, x\!:\!F \vdash G = E}{\Gamma \vdash \Pi(x\!:\!F)G = \Pi(x\!:\!H)E}$$

$$\frac{\Gamma, x\!:\!F \vdash t\!:\!G}{\Gamma \vdash \lambda x.t\!:\!\Pi(x\!:\!F)G} \qquad \frac{\Gamma \vdash g\!:\!\Pi(x\!:\!F)G \quad \Gamma \vdash a\!:\!F}{\Gamma \vdash g\, a\!:\!G[a]}$$

$$\frac{\Gamma, x\!:\!F \vdash t\!:\!G \quad \Gamma \vdash a\!:\!F}{\Gamma \vdash (\lambda x.t)a = t[a]\!:\!G[a]}$$

$$\frac{\Gamma \vdash g\!:\!\Pi(x\!:\!F)G \quad \Gamma \vdash u = v\!:\!F}{\Gamma \vdash g\, u = g\, v\!:\!G[u]} \qquad \frac{\Gamma \vdash h = g\!:\!\Pi(x\!:\!F)G \quad \Gamma \vdash u\!:\!F}{\Gamma \vdash h\, u = g\, u\!:\!G[u]}$$

$$\frac{\Gamma \vdash h\!:\!\Pi(x\!:\!F)G \quad \Gamma \vdash g\!:\!\Pi(x\!:\!F)G \quad \Gamma, x\!:\!F \vdash h\, x = g\, x\!:\!G[x]}{\Gamma \vdash h = g\!:\!\Pi(x\!:\!F)G}$$

Dependent pairs:

$$\frac{\Gamma \vdash F \quad \Gamma, x\!:\!F \vdash G}{\Gamma \vdash \Sigma(x\!:\!F)G} \qquad \frac{\Gamma \vdash F = H \quad \Gamma, x\!:\!F \vdash G = E}{\Gamma \vdash \Sigma(x\!:\!F)G = \Sigma(x\!:\!H)E}$$

$$\frac{\Gamma, x\!:\!F \vdash G \quad \Gamma \vdash a\!:\!F \quad \Gamma \vdash b\!:\!G[a]}{\Gamma \vdash (a,b)\!:\!\Sigma(x\!:\!F)G} \qquad \frac{\Gamma \vdash t\!:\!\Sigma(x\!:\!F)G}{\Gamma \vdash t.1\!:\!F} \qquad \frac{\Gamma \vdash t\!:\!\Sigma(x\!:\!F)G}{\Gamma \vdash t.2\!:\!G[t.1]}$$

$$\frac{\Gamma, x\!:\!F \vdash G \quad \Gamma \vdash t\!:\!F \quad \Gamma \vdash u\!:\!G[t]}{\Gamma \vdash (t,u).1 = t\!:\!F} \qquad \frac{\Gamma, x\!:\!F \vdash G \quad \Gamma \vdash t\!:\!F \quad \Gamma \vdash u\!:\!G[t]}{\Gamma \vdash (t,u).2 = u\!:\!G[t]}$$

$$\frac{\Gamma \vdash t = u\!:\!\Sigma(x\!:\!F)G}{\Gamma \vdash t.1 = u.1\!:\!F} \qquad \frac{\Gamma \vdash t = u\!:\!\Sigma(x\!:\!F)G}{\Gamma \vdash t.2 = u.2\!:\!G[t.1]}$$

$$\frac{\Gamma \vdash t\!:\!\Sigma(x\!:\!F)G \quad \Gamma \vdash u\!:\!\Sigma(x\!:\!F)G \quad \Gamma \vdash t.1 = u.1\!:\!F \quad \Gamma \vdash t.2 = u.2\!:\!G[t.1]}{\Gamma \vdash t = u\!:\!\Sigma(x\!:\!F)G}$$

Universe:

$$\frac{\Gamma \vdash}{\Gamma \vdash U} \quad \frac{\Gamma \vdash F : U}{\Gamma \vdash F} \quad \frac{\Gamma \vdash F = G : U}{\Gamma \vdash F = G} \quad \frac{\Gamma \vdash}{\Gamma \vdash N : U} \quad \frac{\Gamma \vdash}{\Gamma \vdash N_2 : U}$$

$$\frac{\Gamma \vdash F : U \quad \Gamma, x : F \vdash G : U}{\Gamma \vdash \Pi(x : F) G : U} \quad \frac{\Gamma \vdash F = H : U \quad \Gamma, x : F \vdash G = E : U}{\Gamma \vdash \Pi(x : F) G = \Pi(x : H) E : U}$$

$$\frac{\Gamma \vdash F : U \quad \Gamma, x : F \vdash G : U}{\Gamma \vdash \Sigma(x : F) G : U} \quad \frac{\Gamma \vdash F = H : U \quad \Gamma, x : F \vdash G = E : U}{\Gamma \vdash \Sigma(x : F) G = \Sigma(x : H) E : U}$$

Congruence:

$$\frac{\Gamma \vdash t : F \quad \Gamma \vdash F = G}{\Gamma \vdash t : G} \quad \frac{\Gamma \vdash t = u : F \quad \Gamma \vdash F = G}{\Gamma \vdash t = u : G}$$

$$\frac{\Gamma \vdash F}{\Gamma \vdash F = F} \quad \frac{\Gamma \vdash F = G}{\Gamma \vdash G = F} \quad \frac{\Gamma \vdash F = G \quad \Gamma \vdash G = H}{\Gamma \vdash F = H}$$

$$\frac{\Gamma \vdash t : F}{\Gamma \vdash t = t : F} \quad \frac{\Gamma \vdash t = u : F}{\Gamma \vdash u = t : F} \quad \frac{\Gamma \vdash t = u : F \quad \Gamma \vdash u = v : F}{\Gamma \vdash t = v : F}$$

The following four rules are admissible in the this type system [Abel and Scherer, 2012]:

$$\frac{\Gamma \vdash a : A}{\Gamma \vdash A} \quad \frac{\Gamma \vdash a = b : A}{\Gamma \vdash a : A}$$

$$\frac{\Gamma, x : F \vdash G \quad \Gamma \vdash a = b : F}{\Gamma \vdash G[a] = G[b]} \quad \frac{\Gamma, x : F \vdash t : G \quad \Gamma \vdash a = b : F}{\Gamma \vdash t[a] = t[b] : G[a]}$$

1.2 Markov's principle

Markov's principle can be represented in type theory by the type

$$\mathrm{MP} := \Pi(h : N \to N_2)[\neg\neg(\Sigma(x : N)\, \mathsf{IsZero}\,(h\,x)) \to \Sigma(x : N)\, \mathsf{IsZero}\,(h\,x)]$$

where $\mathsf{IsZero} : N_2 \to U$ is defined by $\mathsf{IsZero} := \lambda y.\mathsf{boolrec}\ (\lambda x.U)\ N_1\ N_0\ y$. Note that $\mathsf{IsZero}\,(h\,n)$ is inhabited when $h\,n = 0$ and empty when $h\,n = 1$. Thus $\Sigma(x : N)\, \mathsf{IsZero}\,(h\,x)$ is inhabited if there is n such that $h\,n = 0$.

The main result of this paper is the following:

Theorem 1.1. *There is no term t such that* $\mathrm{MLTT} \vdash t : \mathrm{MP}$.

An *extension* of MLTT is given by introducing new objects, judgment forms and derivation rules. This means in particular that any judgment valid in MLTT is valid in the extension. A *consistent* extension is one in which the type N_0 is uninhabited.

To show Theorem 1.1 we will form a consistent extension of MLTT with a new constant f where $\vdash \mathsf{f} : N \to N_2$. We will then show that $\neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x))$ is derivable while $\Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x)$ is not derivable. Thus showing that MP is not derivable in this extension and consequently not derivable in MLTT.

While this is sufficient to establish independence in the sense of non-derivability of MP. To establish the independence of MP in the stronger sense one also needs to show that ¬MP is not derivable in MLTT. This can achieved by reference to the work of Aczel [Aczel, 1999] where it is shown that MLTT extended with $\vdash \mathsf{dne} : \Pi(A : U)(\neg\neg A \to A)$ is consistent. Since $h : N \to N_2, x : N \vdash \mathsf{IsZero}\,(h\,x) : U$ we have $h : N \to N_2 \vdash \Sigma(x : N)\,\mathsf{IsZero}\,(h\,x) : U$. If we let $T(h) := \Sigma(x : N)\,\mathsf{IsZero}\,(h\,x)$ we get that

$$h : N \to N_2 \vdash \mathsf{dne}\,T(h) : \neg\neg T(h) \to T(h)$$

By λ abstraction we have $\vdash \lambda h.\mathsf{dne}\,T(h) : \mathrm{MP}$. We can then conclude that there is no term t such that MLTT $\vdash t : \neg\mathrm{MP}$.

Finally, we will refine the result of Theorem 1.1 by building a consistent extension of MLTT where ¬MP is derivable.

1.3 Forcing extension

A *condition* p is a graph of a partial finite function from $\mathbb{N}$ to $\{0,1\}$. We denote by $\langle\rangle$ the empty condition. We write $p(n) = b$ when $(n, b) \in p$. We say q *extends* p (written $q \leqslant p$) if p is a subset of q. A condition can be thought of as a compact open in Cantor space $2^{\mathbb{N}}$. Two conditions p and q are *compatible* if $p \cup q$ is a condition and we write pq for $p \cup q$, otherwise they are *incompatible*. If $n \notin \mathrm{dom}(p)$ we write $p(n, 0)$ for $p \cup \{(n, 0)\}$ and $p(n, 1)$ for $p \cup \{(n, 1)\}$. We define the notion of *partition* corresponding to the notion of finite covering of a compact open in Cantor space.

Definition 1.2 (Partition)**.** We write $p \lhd p_1, \ldots, p_n$ to say that $p_1, \ldots, p_n$ is a partition of p and we define it as follows:

1. $p \lhd p$.
2. If $n \notin \mathrm{dom}(p)$ and $p(n, 0) \lhd \ldots, q_i, \ldots$ and $p(n, 1) \lhd \ldots, r_j, \ldots$ then $p \lhd \ldots, q_i, \ldots, r_j, \ldots$.

Note that if $p \lhd p_1, \ldots, p_n$ then p_i and p_j are incompatible whenever $i \neq j$. If moreover $q \leqslant p$ then $q \lhd \ldots, qp_j, \ldots$ where p_j is compatible with q.

We extend the given type theory by annotating the judgments with conditions, i.e. replacing each judgment $\Gamma \vdash J$ in the given type system with a judgment $\Gamma \vdash_p J$.

In addition we add the locality rule:

$$\textsc{loc}\,\frac{\Gamma \vdash_{p_1} J \quad \dots \quad \Gamma \vdash_{p_n} J}{\Gamma \vdash_p J}\, p \lhd p_1 \dots p_n$$

We add a term f for the generic point along with the introduction and conversion rules:

$$\textsf{f-I}\,\frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{f} : N \to N_2} \qquad \textsf{f-}\textsc{eval}\,\frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{f}\,\bar{n} = p(n) : N_2}\, n \in \mathrm{dom}(p)$$

We add a term w and the rule:

$$\textsc{w-term}\,\frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{w} : \neg\neg(\Sigma(x : N)\, \mathsf{IsZero}\,(\mathsf{f}\, x))}$$

Since w inhabits $\neg\neg(\Sigma(x : N)\, \mathsf{IsZero}\,(\mathsf{f}\, x))$, our goal is then to show that no term inhabits $\Sigma(x : N)\, \mathsf{IsZero}(\mathsf{f}\, x)$.

It follows directly from the description of the forcing extension that:

Lemma 1.3. *If $\Gamma \vdash J$ then $\Gamma \vdash_p J$ for all p. In particular, if $\vdash t : A$ then $\vdash_p t : A$ for all p.*

Note that if $q \leqslant p$ and $\Gamma \vdash_p J$ then $\Gamma \vdash_q J$ (monotonicity). A statement A (represented as a closed type) is derivable in this extension if $\vdash_{\langle\rangle} t : A$ for some t, which in turn implies $\vdash_p t : A$ for all p.

Similarly to [Coquand and Jaber, 2010] we can state a conservativity result for this extension. Let $\vdash g : N \to N_2$ and $\vdash v : \neg\neg(\Sigma(x : N)\, \mathsf{IsZero}\,(g\, x))$ be two terms of standard type theory. We say that g is compatible with a condition p if g is such that $\vdash g\, \bar{n} = b : N_2$ whenever $(n, b) \in p$ and $\vdash g\, \bar{n} = 0 : N_2$ otherwise. We say that v is compatible with a condition p if g is compatible with p and v is given by $v := \lambda x.x\, (\bar{n}_p, 0)$ where n_p is the smallest natural number such that $n_p \notin \mathrm{dom}(p)$. To see that v is well typed, note that by design $\Gamma \vdash g\, \bar{n}_p = 0 : N_2$ thus $\Gamma \vdash \mathsf{IsZero}\,(g\, \bar{n}_p) = N_1$ and $\Gamma \vdash (\bar{n}_p, 0) : \Sigma(x : N) \mathsf{IsZero}\,(g\, x)$. We have then

$$\Gamma, x : \neg(\Sigma(y : N)\, \mathsf{IsZero}\,(g\, y)) \vdash x\, (\bar{n}_p, 0) : N_0$$

Thus $\Gamma \vdash \lambda x.x\, (\bar{n}_p, 0) : \neg\neg(\Sigma(y : N)\, \mathsf{IsZero}\,(g\, y))$.

Lemma 1.4 (Conservativity)**.** *Let p be a condition and let $\vdash g : N \to N_2$ and $\vdash v : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(g\,x))$ be compatible with p. If $\Gamma \vdash_p J$ then $\Gamma[g/\mathsf{f}, v/\mathsf{w}] \vdash J[g/\mathsf{f}, v/\mathsf{w}]$, i.e. replacing f with g then w with v we obtain a valid judgment in standard type theory. In particular, if $\Gamma \vdash_{\langle\rangle} J$ where neither f nor w occur in Γ or J then $\Gamma \vdash J$ is a valid judgment in standard type theory.*

Proof. The proof is by induction on the type system and it is straightforward for all the standard rules.

For (f-EVAL) we have $(\mathsf{f}\,\overline{n})[g/\mathsf{f}, v/\mathsf{w}] := g\,\overline{n}$ and since g is compatible with p we have $\Gamma[g/\mathsf{f}, v/\mathsf{w}] \vdash g\,\overline{n} = p(n) : N_2$ whenever $n \in \mathrm{dom}(p)$.

For (w-TERM) we have

$$\begin{aligned}(\mathsf{w} : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x)))[g/\mathsf{f}, v/\mathsf{w}]&\\ &:= (\mathsf{w} : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(g\,x)))[v/\mathsf{w}]\\ &:= v : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(g\,x)).\end{aligned}$$

For (LOC) the statement follows from the observation that when g is compatible with p and $p \lhd p_1, \ldots, p_n$ then g is compatible with exactly one p_i for $1 \leqslant i \leqslant n$. □

2 A Semantics of the forcing extension

In this section we outline a semantics for the forcing extension given in the previous section. We will interpret the judgments of type theory by computability predicates and relations defined by reducibility to computable weak head normal forms.

2.1 Reduction rules

We extend the β, ι conversion with $\mathsf{f}\,\overline{n} \Rightarrow_p b$ whenever $(n, b) \in p$. In order to ease the presentation of the proofs and definitions we introduce *evaluation contexts* following [Wright and Felleisen, 1994].

$$\begin{aligned}\mathbb{E} ::= [\,] \mid \mathbb{E}\,u \mid \mathbb{E}.1 \mid \mathbb{E}.2 \mid \mathsf{S}\,\mathbb{E} \mid \mathsf{f}\,\mathbb{E}\\ \bot\mathsf{rec}\ (\lambda x.C)\,\mathbb{E} \mid \mathsf{unitrec}\ (\lambda x.C)\,a\,\mathbb{E}\\ \mid \mathsf{boolrec}\ (\lambda x.C)\,a_0\,a_1\,\mathbb{E} \mid \mathsf{natrec}\ (\lambda x.C)\,c_z\,g\,\mathbb{E}\end{aligned}$$

An expression $\mathbb{E}[e]$ is then the expression resulting from replacing the hole $[\,]$ by e. We reserve the symbols $\mathbb{E}$ and $\mathbb{C}$ for evaluation contexts. We have the following reduction rules:

$$\frac{}{\mathsf{unitrec}\ (\lambda x.C)\ c\ 0 \to c}$$

$$\frac{}{\mathsf{boolrec}\ (\lambda x.C)\ c_0\ c_1\ 0 \to c_0} \qquad \frac{}{\mathsf{boolrec}\ (\lambda x.C)\ c_0\ c_1\ 1 \to c_1}$$

$$\frac{}{\mathsf{natrec}\ (\lambda x.C)\ c_z\ g\ 0 \to c_z}$$

$$\frac{}{\mathsf{natrec}\ (\lambda x.C)\ c_z\ g\ (\mathsf{S}\ \bar{k}) \to g\ \bar{k}\ (\mathsf{natrec}\ (\lambda x.C)\ c_z\ g\ \bar{k})}$$

$$\frac{}{(\lambda x.t)\ a \to t[a/x]} \qquad \frac{}{(u,v).1 \to u} \qquad \frac{}{(u,v).2 \to v}$$

$$\frac{e \to e'}{e \to_p e'} \qquad \textsc{f-red}\ \frac{k \in \mathsf{dom}(p)}{\mathsf{f}\ \bar{k} \to_p p(k)} \qquad \frac{e \to_p e'}{\mathbb{E}[e] \Rightarrow_p \mathbb{E}[e']}$$

Note that we reduce under S.

The relation $\Rightarrow$ is monotone, that is if $q \leqslant p$ and $t \Rightarrow_p u$ then $t \Rightarrow_q u$. We will also need to show that the reduction is local, i.e. if $p \lhd p_1, \dots, p_n$ and $t \Rightarrow_{p_i} u$ then $t \Rightarrow_p u$.

Lemma 2.1. *If $m \notin \mathsf{dom}(p)$ and $t \to_{p(m,0)} u$ and $t \to_{p(m,1)} u$ then $t \to_p u$.*

Proof. By induction on the derivation of $t \to_{p(m,0)} u$. If $t \to_{p(m,0)} u$ is derived by (f-RED) then $t := \mathsf{f}\ \bar{k}$ and $u := p(m,0)(k)$ for some $k \in \mathsf{dom}(p(m,0))$. But since we also have a reduction $\mathsf{f}\ \bar{k} \to_{p(m,1)} u$, we have $p(m,1)(k) := u := p(m,0)(k)$ which could only be the case if $k \in \mathsf{dom}(p)$. Thus we have a reduction $\mathsf{f}\ \bar{k} \to_p u := p(k)$. Alternatively, we have a derivation $t \to u$, in which case we have $t \to_p u$ directly. □

Lemma 2.2. *If $m \notin \mathsf{dom}(p)$ and $t \Rightarrow_{p(m,0)} u$ and $t \Rightarrow_{p(m,1)} u$ then $t \Rightarrow_p u$.*

Proof. From the reduction $t \Rightarrow_{p(k,0)} u$ we have $t := \mathbb{E}[e]$, $u := \mathbb{E}[e']$ and $e \to_{p(m,0)} e'$ for some context $\mathbb{E}$. But then we also have a reduction $\mathbb{E}[e] \Rightarrow_{p(m,1)} \mathbb{E}[e']$, thus $e \to_{p(m,1)} e'$. By Lemma 2.1 we have $e \to_p e'$ and thus $\mathbb{E}[e] \Rightarrow_p \mathbb{E}[e']$. □

Lemma 2.3. *Let $q \leqslant p$. If $t \to_q u$ then either $t \to_p u$ or $t := \mathsf{f}\ m$ for some $m \in \mathsf{dom}(q) \setminus \mathsf{dom}(p)$.*

Proof. By induction on the derivation of $t \to_q u$. If the reduction $t \to_q u$ has the form $\mathsf{f}\ \bar{k} \to_q q(k)$ then either $k \notin \mathsf{dom}(p)$ and the statement follows or $k \in \mathsf{dom}(p)$ and we have $t \to_p u$. Alternatively, we have $t \to u$ and immediately $t \to_p u$. □

Lemma 2.4. *Let $q \leqslant p$. If $t \Rightarrow_q u$ then either $t \Rightarrow_p u$ or t has the form $\mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \in \mathrm{dom}(q) \setminus \mathrm{dom}(p)$.*

Proof. If $t \Rightarrow_q u$ then $t := \mathbb{E}[e]$, $u := \mathbb{E}[e']$ and $e \rightarrow_q e'$ for some context $\mathbb{E}$. By Lemma 2.3 either $e := \mathsf{f}\,\overline{m}$ for $m \notin \mathrm{dom}(p)$ and the statement follows or $e \rightarrow_p e'$ in which case we have $t \Rightarrow_p u$. □

Corollary 2.5. *For any condition p and $m \notin \mathrm{dom}(p)$. Let $t \Rightarrow_{p(m,0)} u$ and $t \Rightarrow_{p(m,1)} v$. If $u := v$ then $t \Rightarrow_p u$; otherwise, t has the form $\mathbb{E}[\mathsf{f}\,\overline{m}]$.*

Proof. Follows by Lemma 2.2 and Lemma 2.4. □

Next we define the relation $p \vdash t \Rightarrow u : A$ to mean $t \Rightarrow_p u$ and $\vdash_p t = u : A$ and we write $p \vdash A \Rightarrow B$ for $p \vdash A \Rightarrow B : U$. We note that it holds that if $p \vdash t \Rightarrow u : \Pi(x:F)G$ and $\vdash a : F$ then $p \vdash t\,a \Rightarrow u\,a : G[a]$ and if $p \vdash t \Rightarrow u : \Sigma(x:F)G$ then $p \vdash t.1 \Rightarrow u.1 : F$ and $p \vdash t.2 \Rightarrow u.2 : G[t.1]$. We define a closure for this relation as follows:

$$\frac{\vdash_p t : A}{p \vdash t \Rightarrow^* t : A} \qquad \frac{p \vdash t \Rightarrow u : A}{p \vdash t \Rightarrow^* u : A} \qquad \frac{p \vdash t \Rightarrow u : A \quad p \vdash u \Rightarrow^* v : A}{p \vdash t \Rightarrow^* v : A}$$

$$\frac{\vdash_p A}{p \vdash A \Rightarrow^* A} \qquad \frac{p \vdash A \Rightarrow B}{p \vdash A \Rightarrow^* B} \qquad \frac{p \vdash A \Rightarrow B \quad p \vdash B \Rightarrow^* C}{p \vdash A \Rightarrow^* C}$$

A term t is *in* p-whnf if whenever $t \Rightarrow_p u$ then $t := u$. A whnf is canonical if it has the form:

$$0, 1, \overline{n}, \lambda x.t, (a, b), \mathsf{f}, \mathsf{w},$$
$$\bot\mathsf{rec}\ (\lambda x.C), \mathsf{unitrec}\ (\lambda x.C)\ a, \mathsf{boolrec}\ (\lambda x.C)\ a_0\ a_1, \mathsf{natrec}\ (\lambda x.C)\ c_z\ g,$$
$$N_0, N_1, N_2, N, U, \Pi(x:F)G, \Sigma(x:F)G.$$

A p-whnf is proper if it is canonical or it is of the form $\mathbb{E}[\mathsf{f}\,\bar{k}]$ for $k \notin \mathrm{dom}(p)$.

We have the following corollaries to Lemma 2.2 and Corollary 2.5.

Corollary 2.6. *Let $m \notin \mathrm{dom}(p)$. Let $p(m,0) \vdash t \Rightarrow_{p(m,0)} u : A$ and $p(m,1) \vdash t \Rightarrow_{p(m,1)} v : A$. If $u := v$ then $p \vdash t \Rightarrow u : A$; otherwise t has the form $\mathbb{E}[\mathsf{f}\,\overline{m}]$.*

Corollary 2.7. *If $p \vdash t \Rightarrow u : A$ and $q \leqslant p$ then $q \vdash t \Rightarrow u : A$. If $p \lhd p_1, \ldots, p_n$ and $p_i \vdash t \Rightarrow u : A$ for all i then $p \vdash t \Rightarrow u : A$.*

Proof. Let $q \leqslant p$. If $t \Rightarrow_p u$ we have $t \Rightarrow_q u$ and if $\vdash_p t = u : A$ then $\vdash_q t = u : A$. Thus $q \vdash t \Rightarrow u : A$ whenever $p \vdash t \Rightarrow u : A$. Let $p \lhd p_1, \ldots, p_n$. If for all i, $t \Rightarrow_{p_i} u : A$ then from Lemma 2.2, by induction on the partition, we have $t \Rightarrow_p u : A$. If $\vdash_{p_i} t = u : A$ for all i, then $\vdash_p t = u : A$. Thus we have $p \vdash t \Rightarrow u : A$ whenever $p_i \vdash t \Rightarrow u : A$ for all i. □

Note that if $q \leqslant p$ and $p \vdash t \Rightarrow^* u : A$ then $q \vdash t \Rightarrow^* u : A$ and similarly if $p \vdash A \Rightarrow^* B$ then $q \vdash A \Rightarrow^* B$.

Note also that if $m \notin \mathrm{dom}(p)$ and $p(m,0) \vdash t \Rightarrow^* u : A$ and $p(m,1) \vdash t \Rightarrow^* u : A$ it is not necessarily the case that $p \vdash t \Rightarrow^* u : A$. For example we have that $\{(m,0)\} \vdash \mathsf{boolrec}\ (\lambda x.N)\ \bar{n}\ \bar{n}\ (\mathsf{f}\ \bar{m}) \Rightarrow^* \bar{n} : N$ and $\{(m,1)\} \vdash \mathsf{boolrec}\ (\lambda x.N)\ \bar{n}\ \bar{n}\ (\mathsf{f}\ \bar{m}) \Rightarrow^* \bar{n} : N$ but it is *not* true that $\langle\rangle \vdash \mathsf{boolrec}\ (\lambda x.N)\ \bar{n}\ \bar{n}\ (\mathsf{f}\ \bar{m}) \Rightarrow^* \bar{n} : N$

For a closed term $\vdash_p t : A$, we say that t *has* a p-whnf if $p \vdash t \Rightarrow^* u : A$ and u is in p-whnf. If moreover u is canonical, respectively proper, we say that t has a canonical, respectively proper, p-whnf. Note that a canonical p-whnf has no further reduction at any $q \leqslant p$. A proper p-whnf that is not canonical, i.e. of the form $\mathbb{E}[\mathsf{f}\ \bar{k}]$ for $k \notin \mathrm{dom}(p)$, could have further reduction at some $q \leqslant p$, namely at any $q \leqslant p(k,0)$ or $q \leqslant p(k,1)$. Since the reduction relation is deterministic we have

Lemma 2.8. *A term $\vdash_p t : A$ has at most one p-whnf.*

Corollary 2.9. *Let $\vdash_p t : A$ and $m \notin \mathrm{dom}(p)$. If t has proper $p(m,0)$-whnf and a proper $p(m,1)$-whnf then t has a proper p-whnf.*

Proof. Let $p(m,0) \vdash t \Rightarrow^* u : A$ and $p(m,1) \vdash t \Rightarrow^* v : A$ with u in proper $p(m,0)$-whnf and v in proper $p(m,1)$-whnf. If $t := u$ or $t := v$ then t is already in proper p-whnf. Alternatively we have reductions $p(m,0) \vdash t \Rightarrow u_1 : A$ and $p(m,1) \vdash t \Rightarrow v_1 : A$. By Corollary 2.6 either t is in proper p-whnf or $u_1 := v_1$ and $p \vdash t \Rightarrow u_1 : A$. It then follows by induction that u_1, and thus t, has a proper p-whnf. □

2.2 Computability predicate and relation

We define inductively a forcing relation $p \Vdash A$ to express that a type A is computable at p. Mutually by recursion we define relations $p \Vdash a : A$, $p \Vdash A = B$, and $p \Vdash a = b : A$. The definition fits the generalized mutual induction-recursion schema [Dybjer, 2000][2].

Definition 2.10 (Computability predicate and relation)**.**

($\mathbf{F_{N_0}}$) If $p \vdash A \Rightarrow^* N_0$ then $p \Vdash A$.

[2] However, for the canonical proof below we actually need something weaker than an inductive-recursive definition (arbitrary fixed-point instead of *least* fixed-point), reflecting the fact that the universe is defined in an open way [Martin-Löf, 1972].

1. $p \Vdash t : A$ does not hold for all t.
2. $p \Vdash t = u : A$ does not hold for all t and u.
3. If $p \Vdash B$ then $p \Vdash A = B$ if
   i. $p \vdash B \Rightarrow^* N_0$.
   ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash A = B$.

$(\mathbf{F_{N_1}})$ If $p \vdash A \Rightarrow^* N_1$ then $p \Vdash A$.

1. $p \Vdash t : A$ if
   i. $p \vdash t \Rightarrow^* 0 : A$.
   ii. $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t : A$.
2. If $p \Vdash t : A$ and $p \Vdash u : A$ then $p \Vdash t = u : A$ if
   i. $p \vdash t \Rightarrow^* 0 : A$ and $p \vdash u \Rightarrow^* 0 : A$.
   ii. $p \vdash t \Rightarrow^* 0 : A$ and $p \vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t = u : A$.
   iii. $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t = u : A$.
3. If $p \Vdash B$ then $p \Vdash A = B$ if
   i. $p \vdash B \Rightarrow^* N_1$.
   ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash A = B$.

$(\mathbf{F_{N_2}})$ If $p \vdash A \Rightarrow^* N_2$ then $p \Vdash A$.

1. $p \Vdash t : A$ if
   i. $p \vdash t \Rightarrow^* b : A$ for some $b \in \{0,1\}$.
   ii. $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t : A$.
2. If $p \Vdash t : A$ and $p \Vdash u : A$ then $p \Vdash t = u : A$ if
   i. $p \vdash t \Rightarrow^* b : A$ and $p \vdash u \Rightarrow^* b : A$ for some $b \in \{0,1\}$.
   ii. $p \vdash t \Rightarrow^* b : A$ for some $b \in \{0,1\}$ and $p \vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t = u : A$.
   iii. $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t = u : A$.
3. If $p \Vdash B$ then $p \Vdash A = B$ if
   i. $p \vdash B \Rightarrow^* N_2$.

ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash A = B$.

($\mathbf{F_N}$) If $p \vdash A \Rightarrow^* N$ then $p \Vdash A$.

1. $p \Vdash t : A$ if
   i. $p \vdash t \Rightarrow^* \bar{n} : A$ for some $n \in \mathbb{N}$.
   ii. $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t : A$.
2. If $p \Vdash t : A$ and $p \Vdash u : A$ then $p \Vdash t = u : A$ if
   i. $p \vdash t \Rightarrow^* \bar{n} : A$ and $p \vdash u \Rightarrow^* \bar{n} : A$ for some $n \in \mathbb{N}$.
   ii. $p \vdash t \Rightarrow^* \bar{n} : A$ for some $n \in \mathbb{N}$ and $p \vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t = u : A$.
   iii. $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : A$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash t = u : A$.
3. If $p \Vdash B$ then $p \Vdash A = B$ if
   i. $p \vdash B \Rightarrow^* N$.
   ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash A = B$.

($\mathbf{F_\Pi}$) If $p \vdash A \Rightarrow^* \Pi(x : F)G$ then $p \Vdash A$ if $p \Vdash F$ and for all $q \leqslant p$, $q \Vdash G[a]$ whenever $q \Vdash a : F$ and $q \Vdash G[a] = G[b]$ whenever $q \Vdash a = b : F$.

1. If $\vdash_p f : A$ then $p \Vdash f : A$ if for all $q \leqslant p$, $q \Vdash f\,a : G[a]$ whenever $q \Vdash a : F$ and $q \Vdash f\,a = f\,b : G[a]$ whenever $q \Vdash a = b : F$.
2. If $p \Vdash f : A$ and $p \Vdash g : A$ then $p \Vdash f = g : A$ if $\vdash_p f = g : A$ and for all $q \leqslant p$, $q \Vdash f\,a = g\,a : G[a]$ whenever $q \Vdash a : F$.
3. If $p \Vdash B$ then $p \Vdash A = B$ if
   i. $\vdash_p A = B$ and $p \vdash B \Rightarrow^* \Pi(x : H)E$ and $p \Vdash F = H$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a]$ whenever $q \Vdash a : F$.
   ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash A = B$.

($\mathbf{F_\Sigma}$) If $p \vdash A \Rightarrow^* \Sigma(x : F)G$ then $p \Vdash A$ if $p \Vdash F$ and for all $q \leqslant p$, $q \Vdash G[a]$ whenever $q \Vdash a : F$ and $q \Vdash G[a] = G[b]$ whenever $q \Vdash a = b : F$.

1. If $\vdash_p t : A$ then $p \Vdash t : A$ if $p \Vdash t.1 : F$ and $p \Vdash t.2 : G[t.1]$.

2. If $p \Vdash t : A$ and $p \Vdash u : A$ then $p \Vdash t = u : A$ if $\vdash_p t = u : A$ and $p \Vdash t.1 = u.1 : F$ and $p \Vdash t.2 = u.2 : G[t.1]$.
3. If $p \Vdash B$ then $p \Vdash A = B$ if
   i. $\vdash_p A = B$ and $p \vdash B \Rightarrow^* \Sigma(x : H)E$ and $p \Vdash F = H$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a]$ whenever $q \Vdash a : F$.
   ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash A = B$.

($\mathbf{F_U}$) $p \Vdash U$.

1. $p \Vdash C : U$ if
   i. $p \vdash C \Rightarrow^* M : U$ for some $M \in \{N_0, N_1, N_2, N\}$.
   ii. $p \vdash C \Rightarrow^* \Pi(x : F)G : U$ and $p \Vdash F : U$ and for all $q \leqslant p$, $q \Vdash G[a] : U$ whenever $q \Vdash a : F$ and $q \Vdash G[a] = G[b] : U$ whenever $q \Vdash a = b : F$.
   iii. $p \vdash C \Rightarrow^* \Sigma(x : F)G : U$ and $p \Vdash F : U$ and for all $q \leqslant p$, $q \Vdash G[a] : U$ whenever $q \Vdash a : F$ and $q \Vdash G[a] = G[b] : U$ whenever $q \Vdash a = b : F$.
   iv. $p \vdash C \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : U$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash C : U$.
2. If $p \Vdash C : U$ and $p \Vdash D : U$ then $p \Vdash C = D : U$ if
   i. $p \vdash C \Rightarrow^* M : U$ and $D \Rightarrow^* M : U$ for some $M \in \{N_0, N_1, N_2, N\}$.
   ii. $p \vdash C \Rightarrow^* \Pi(x : F)G : U$ and $p \vdash D \Rightarrow^* \Pi(x : H)E : U$ and $p \Vdash F = H : U$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a] : U$ whenever $q \Vdash a : F$.
   iii. $p \vdash C \Rightarrow^* \Sigma(x : F)G : U$ and $p \vdash D \Rightarrow^* \Sigma(x : H)E : U$ and $p \Vdash F = H : U$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a] : U$ whenever $q \Vdash a : F$.
   iv. $p \vdash C \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : U$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash C = D : U$.
   v. $p \vdash D \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}] : U$ for some $m \notin \mathrm{dom}(p)$ and for all $i \in \{0,1\}$, $p(m,i) \Vdash C = D : U$.
3. $p \Vdash U = B$ iff $B := U$.

($\mathbf{F_{Loc}}$) If $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and $p(m,i) \Vdash A$ for all $i \in \{0,1\}$ then $p \Vdash A$.

1. If $p(m,i) \Vdash t : A$ for all $i \in \{0,1\}$ then $p \Vdash t : A$.
2. If $p \Vdash t : A$ and $p \Vdash u : A$ and $p(m,i) \Vdash t : A$ for all $i \in \{0,1\}$ then $p \Vdash t = u : A$.

3. If $p \Vdash B$ then $p \Vdash A = B$ if $p(m,i) \Vdash A = B$ for all $i \in \{0,1\}$.

We note from the definition that when $p \Vdash A = B$ then $p \Vdash A$ and $p \Vdash B$, when $p \Vdash a : A$ then $p \Vdash A$ and when $p \Vdash a = b : A$ then $p \Vdash a : A$ and $p \Vdash b : A$. We remark also if $p \vdash A \Rightarrow^* U$ then $A := U$ since we have only one universe.

The clause (F_{Loc}) gives semantics to *variable types*. For example, if $p := \{(0,0)\}$ and $q := \{(0,1)\}$ the type $R := \mathsf{boolrec}\ (\lambda x.U)\ N_1\ N\ (\mathsf{f}\ 0)$ has reductions $p \vdash R \Rightarrow^* N_1$ and $q \vdash R \Rightarrow^* N$. Thus $p \Vdash R$ and $q \Vdash R$ and since $\langle\rangle \lhd p, q$ we have $\langle\rangle \Vdash R$.

Immediately from Definition 2.10 we get:

Lemma 2.11. *If $p \Vdash A$ then $\vdash_p A$. If $p \Vdash a : A$ then $\vdash_p a : A$.*
If $p \Vdash A = B$ then $\vdash_p A = B$. If $p \Vdash a = b : A$ then $\vdash_p a = b : A$.

Lemma 2.12. *If $p \Vdash A$ then there is a partition $p \lhd p_1, \dots, p_n$ where A has a canonical p_i-whnf for all i.*

Proof. The statement follows from the definition by induction on the derivation of $p \Vdash A$ □

Corollary 2.13. *Let $p \lhd p_1, \dots, p_n$. If $p_i \Vdash A$ for all i then A has a proper p-whnf.*

Proof. Follows from Lemma 2.12 and Corollary 2.9 by induction on the partition. □

Lemma 2.14. *If $p \Vdash A$ and $q \leqslant p$ then $q \Vdash A$.*

Proof. Let $p \Vdash A$ and $q \leqslant p$. By induction on the derivation of $p \Vdash A$

($\mathbf{F_N}$) Since $p \vdash A \Rightarrow^* N$ and the reduction relation is monotone we have $q \vdash A \Rightarrow^* N$, thus $q \Vdash A$. The statement follows similarly for (F_{N_0}), (F_{N_1}), (F_{N_2}) and (F_U).

($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. Since $p \Vdash F$, by the induction $q \Vdash F$. Let $s \leqslant q$, we have then $s \leqslant p$. It then follows from $p \Vdash A$ that $s \Vdash G[a]$ whenever $s \Vdash a : F$ and $s \Vdash G[a] = G[b]$ whenever $s \Vdash a = b : F$. Thus $q \Vdash A$. The statement follows similarly for (F_Σ).

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\ \overline{m}]$. If $m \in \mathrm{dom}(q)$ then $q \leqslant p(m,0)$ or $q \leqslant p(m,1)$ and since $p(m,i) \Vdash A$, by the induction $q \Vdash A$. Alternatively, $q \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\ \overline{m}]$. But $q \lhd q(m,0), q(m,1)$ and $q(m,i) \leqslant p(m,i)$. By the induction $q(m,i) \Vdash A$ for all $i \in \{0,1\}$ and thus $q \Vdash A$. □

Lemma 2.15. *If $p \Vdash t : A$ and $q \leqslant p$ then $q \Vdash t : A$.*

Proof. Let $p \Vdash t : A$ and $q \leqslant p$. By induction on the derivation of $p \Vdash A$.

($\mathbf{F_N}$) Since $p \vdash A \Rightarrow^* N$ then $q \vdash A \Rightarrow^* N$. By induction on the derivation of $p \Vdash t : A$. If $p \vdash t \Rightarrow^* \bar{n} : A$ for $n \in \mathbb{N}$ then $q \vdash t \Rightarrow^* \bar{n} : A$, hence, $q \Vdash t : A$. Alternatively, $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$ for some $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash t : A$ for all $b \in \{0,1\}$. If $k \in \mathrm{dom}(q)$ then $q \leqslant p(k,1)$ or $q \leqslant p(k,0)$ and in either case, by the induction, $q \Vdash t : A$. Otherwise, we have $q(k,b) \leqslant p(k,b)$ and by the induction $q(k,b) \Vdash t : A$ for all $b \in \{0,1\}$. By the definition $q \Vdash t : A$. The statement follows similarly for ($\mathrm{F}_{\mathrm{N}_0}$), ($\mathrm{F}_{\mathrm{N}_1}$), and ($\mathrm{F}_{\mathrm{N}_2}$).

($\mathbf{F_U}$) We can show the statement by a proof similar to that of Lemma 2.14.

($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. We have $q \vdash A \Rightarrow^* \Pi(x : F)G$. From $\vdash_p t : A$ we have $\vdash_q t : A$. Let $r \leqslant q$. If $r \Vdash a : F$ then since $r \leqslant p$ we have $r \Vdash t\,a : G[a]$. Similarly if $r \Vdash a = b : F$ then $r \Vdash t\,a = t\,b : G[a]$. Thus $q \Vdash t : A$.

($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. We have $q \vdash A \Rightarrow^* \Sigma(x : F)G$. From $\vdash_p t : A$ we have $\vdash_q t : A$. Since $p \Vdash t : A$ we have $p \Vdash t.1 : F$ and $p \Vdash t.2 : G[t.1]$. By the induction $q \Vdash t.1 : F$ and $q \Vdash t.2 : G[t.1]$, thus $q \Vdash t : A$.

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. Since $p \Vdash t : A$ we have $p(k,b) \Vdash t : A$ for all $b \in \{0,1\}$. If $k \in \mathrm{dom}(q)$ then $q \leqslant p(k,0)$ or $q \leqslant p(k,1)$ and by the induction $q \Vdash t : A$. Otherwise, $q \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ and since $q(k,b) \leqslant p(k,b)$, by the induction, $q(k,b) \Vdash t : A$ for all $b \in \{0,1\}$. By definition $q \Vdash t : A$. □

Lemma 2.16. *If $p \Vdash A = B$ and $q \leqslant p$ then $q \Vdash A = B$.*

Proof. Let $p \Vdash A = B$ and $q \leqslant p$. We have then that $p \Vdash A$ and $p \Vdash B$. By Lemma2.14 we have that $q \Vdash A$ and $q \Vdash B$. By induction on the derivation of $p \Vdash A$.

($\mathbf{F_N}$) We have then $p \vdash A \Rightarrow^* N$. Since $p \Vdash A = B$ either

i. $p \vdash B \Rightarrow^* N$. In this case $q \vdash B \Rightarrow^* N$ and $q \vdash A \Rightarrow^* N$ and we have $q \Vdash A = B$.

ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. If $k \in \mathrm{dom}(q)$ then $q \leqslant p(k,1)$ or $q \leqslant p(k,0)$;

in either case, by the induction, $q \Vdash A = B$. Otherwise, $q \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$. Since $q(k,b) \leqslant p(k,b)$, by the induction $q(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. By the definition we have $q \Vdash A = B$.

The statement follows similarly for $(\mathrm{F}_{\mathrm{N}_0})$, $(\mathrm{F}_{\mathrm{N}_1})$, and $(\mathrm{F}_{\mathrm{N}_2})$.

$(\mathbf{F_U})$ We have then that $B := U$ and thus $q \Vdash A = B$.

$(\mathbf{F_\Pi})$ Let $p \vdash A \Rightarrow^* \Pi(x:F)G$. Either i. $p \vdash B \Rightarrow^* \Pi(x:H)E$ and $p \Vdash F = H$ and for all $r \leqslant p$, $r \Vdash G[a] = E[a]$ whenever $r \Vdash a:F$. In this case $q \vdash B \Rightarrow^* \Pi(x:H)E$. By the induction $q \Vdash F = H$. If $s \leqslant q$ and $s \Vdash u:F$ then since $s \leqslant p$ we have that $s \Vdash G[u] = E[u]$. Thus $q \Vdash A = B$. ii. $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ and $p(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. If $k \in \mathrm{dom}(q)$ then $q \leqslant p(k,0)$ or $q \leqslant p(k,1)$; in either case, by the induction, $q \Vdash A = B$. Otherwise, $q \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$. But $q(k,b) \leqslant p(k,b)$ and by the induction $q(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. By the definition $q \Vdash A = B$.

The statement follows similarly for (F_Σ).

$(\mathbf{F_{Loc}})$ Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for $k \notin \mathrm{dom}(p)$. We have that $p(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. If $k \in \mathrm{dom}(q)$ then $q \leqslant p(k,0)$ or $q \leqslant p(k,1)$ and by the induction $q \Vdash A = B$. Otherwise, $q \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ and since $q(k,b) \leqslant p(k,b)$, by the induction, $q(k,b) \Vdash A = B$ for all b. By the definition $q \Vdash A = B$. ☐

Lemma 2.17. *If $p \Vdash t = u:A$ and $q \leqslant p$ then $q \Vdash t = u:A$.*

Proof. Let $p \Vdash t = u:A$ and $q \leqslant p$. We have then that $p \Vdash A$, $p \Vdash t:A$, and $p \Vdash u:A$. By Lemma 2.14 $q \Vdash A$. By Lemma 2.15 $q \Vdash t:A$ and $q \Vdash u:A$. By induction on the derivation $p \Vdash A$.

$(\mathbf{F_N})$ Since $p \vdash A \Rightarrow^* N$ then $q \vdash A \Rightarrow^* N$. By induction on the derivation of $p \Vdash t = u:A$.

i. Let $p \vdash t \Rightarrow^* \bar{n}:A$ and $p \vdash u \Rightarrow^* \bar{n}:A$ for $n \in \mathbb{N}$. We have $q \vdash t \Rightarrow^* \bar{n}:A$ and $q \vdash u \Rightarrow^* \bar{n}:A$, hence, $q \Vdash t = u:A$.

ii. Let $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]:A$ for some $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash t = u:A$ for all $b \in \{0,1\}$. If $k \in \mathrm{dom}(q)$ then $q \leqslant p(k,1)$ or $q \leqslant p(k,0)$ and in either case, by the induction, $q \Vdash t = u:A$. Otherwise, we have $q(k,b) \leqslant p(k,b)$ and by the induction $q(k,b) \Vdash t = u:A$ for all $b \in \{0,1\}$. By the definition $q \Vdash t = u:A$.

iii. Let $p \vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$ for some $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash t = u : A$ for all $b \in \{0,1\}$. The statement follows similarly to (ii).

The statement follows similarly for $(\mathrm{F}_{\mathrm{N}_0})$, $(\mathrm{F}_{\mathrm{N}_1})$, and $(\mathrm{F}_{\mathrm{N}_2})$.

$(\mathbf{F_U})$ The statement follows by a proof similar to that of Lemma 2.16.

$(\mathbf{F_\Pi})$ Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. We have $q \vdash A \Rightarrow^* \Pi(x : F)G$. Let $r \leqslant q$. If $r \Vdash a : F$ then since $r \leqslant p$ we have $r \Vdash t\,a = u\,a : G[a]$. Thus $q \Vdash t = u : A$.

$(\mathbf{F_\Sigma})$ Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. We have $q \vdash A \Rightarrow^* \Sigma(x : F)G$. Since $p \Vdash t = u : A$ we have $p \Vdash t.1 = u.1 : F$ and $p \Vdash t.2 = u.2 : G[t.1]$. By the induction $q \Vdash t.1 = u.1 : F$ and $q \Vdash t.2 = u.2 : G[t.1]$, thus $q \Vdash t = u : A$.

$(\mathbf{F_{Loc}})$ Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. Since $p \Vdash t = u : A$ we have $p(k,b) \Vdash t = u : A$ for all $b \in \{0,1\}$. If $k \in \mathrm{dom}(q)$ then $q \leqslant p(k,0)$ or $q \leqslant p(k,1)$ and by the induction $q \Vdash t = u : A$. Otherwise, $q \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ and since $q(k,b) \leqslant p(k,b)$, by the induction, $q(k,b) \Vdash t = u : A$ for all b. By definition $q \Vdash t = u : A$. □

We collect the results of Lemmas 2.14, Lemma 2.15, Lemma 2.17, and Lemma 2.16 in the following corollary.

Corollary 2.18 (Monotonicity). *If $p \Vdash J$ and $q \leqslant p$ then $q \Vdash J$.*

We write $\Vdash J$ when $\langle\rangle \Vdash J$. By monotonicity $\Vdash J$ iff $p \Vdash J$ for all p.

Lemma 2.19. *If $p(m,0) \Vdash A$ and $p(m,1) \Vdash A$ for some $m \notin \mathrm{dom}(p)$ then $p \Vdash A$.*

Proof. By Corollary 2.13, either A has a canonical p-whnf or $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for $k \notin \mathrm{dom}(p)$.

- Let $p \vdash A \Rightarrow^* M$ with $M \in \{N_0, N_1, N_2, N, U\}$. We have immediately that $p \Vdash A$.

- Let $p \vdash A \Rightarrow^* M$ with M of the form $\Pi(x : F)G$ or $\Sigma(x : F)G$. We have then that $p(m,b) \vdash A \Rightarrow^* M$ for all $b \in \{0,1\}$. Since $p(m,b) \Vdash A$ we have $p(m,b) \Vdash F$ for all b and by the induction $p \Vdash F$. Let $q \leqslant p$ and $q \Vdash a : F$. If $m \in \mathrm{dom}(q)$ then $q \leqslant p(m,b)$ for some $b \in \{0,1\}$. Assume, w.l.o.g, $q \leqslant p(m,0)$. Since $p(m,0) \Vdash A$ we have by the definition that $q \Vdash G[a]$. Alternatively, if $m \notin$

dom(q) we have a partition $q \lhd q(m,0), q(m,1)$. By monotonicity $q(m,b) \Vdash a : F$, and since $q(m,b) \leqslant p(m,b)$, we have $q(m,b) \Vdash G[a]$ for all $b \in \{0,1\}$. By the induction $q \Vdash G[a]$. Similarly we can show $q \Vdash G[t] = G[u]$ whenever $q \Vdash t = u : F$.

- Alternatively, let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. If $k = m$ then by the definition $p \Vdash A$. Otherwise, $p(m,0) \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ and by the definition $p(m,0)(k,b) \Vdash A$ for all $b \in \{0,1\}$. Similarly, $p(m,1)(k,b) \Vdash A$ for all $b \in \{0,1\}$. But $p(k,b) \lhd p(m,0)(k,b), p(m,1)(k,b)$. By the induction $p(k,b) \Vdash A$ for all $b \in \{0,1\}$ and by the definition $p \Vdash A$. □

Lemma 2.20. *If $p(m,0) \Vdash A = B$ and $p(m,1) \Vdash A = B$ for some $m \notin$ dom(p) then $p \Vdash A = B$.*

Proof. If $p(m,b) \Vdash A = B$ then $p(m,b) \Vdash A$ and $p(m,b) \Vdash B$. By Lemma 2.19 we have $p \Vdash A$ and $p \Vdash B$. By induction on the derivation of $p \Vdash A$.

- Let $p \vdash A \Rightarrow^* M$ with $M \in \{N_0, N_1, N_2, N\}$. Since $p \Vdash B$ then B has a proper p-whnf. If $p \vdash B \Rightarrow^* C$ and C is canonical then $p(m,b) \vdash B \Rightarrow^* C$ for all $b \in \{0,1\}$. Since $p(m,b) \Vdash A = B$, by the definition, we have that $C := M$. Thus $p \Vdash A = B$. Alternatively, $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for $k \notin \mathrm{dom}(p)$. If $k = m$ then by the definition we have immediately that $p \Vdash A = B$. Otherwise, we have $p(m,b) \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ and by the definition $p(m,b)(k,i) \Vdash A = B$ for all $i, b \in \{0,1\}$. But $p(k,i) \lhd p(m,0)(k,i), p(m,1)(k,i)$ and by the induction $p(k,i) \Vdash A = B$ for all $i \in \{0,1\}$. By the definition $p \Vdash A = B$.

- Let $A := U$. Since $p(m,0) \Vdash A = B$ we have $B := U$ and thus $p \Vdash A = B$.

- Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. We have then that $p(m,b) \vdash A \Rightarrow^* \Pi(x : F)G$ for all $b \in \{0,1\}$. Since $p \Vdash B$ we have that B has a proper p-whnf. If $p \vdash B \Rightarrow^* C$ where C is canonical then $p(m,b) \vdash B \Rightarrow^* C$ for all $b \in \{0,1\}$. By the definition $C := \Pi(x : H)E$ for some H and E. Since $p(m,b) \Vdash A = B$ we have that $p(m,b) \Vdash F = H$ for all $b \in \{0,1\}$. By the induction $p \Vdash F = H$. Let $q \leqslant p$ and $q \Vdash a : F$. If $q \leqslant p(m,0)$ or $q \leqslant p(m,1)$ then we have that $q \Vdash G[a] = E[a]$. Otherwise, we have $q(m,b) \leqslant p(m,b)$ for all $b \in \{0,1\}$. By monotonicity $q(m,b) \Vdash a : F$ and thus $q(m,b) \Vdash G[a] = E[a]$ for all $b \in \{0,1\}$. By the induction $q \Vdash G[a] = E[a]$. Thus $p \Vdash A = B$.

- Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. The statement follows similarly to the above.

- Alternatively, let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. If $k = m$ then by the definition $p \Vdash A = B$. Otherwise, $p(m,0) \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ and by the definition $p(m,b)(k,i) \Vdash A = B$ for all $i, b \in \{0,1\}$. But $p(k,i) \lhd p(m,0)(k,i), p(m,1)(k,i)$. By the induction $p(k,i) \Vdash A = B$ for all $i \in \{0,1\}$ and by the definition $p \Vdash A = B$. □

Lemma 2.21. *1. If $p(m,0) \Vdash t : A$ and $p(m,1) \Vdash t : A$ for some $m \notin \mathrm{dom}(p)$ then $p \Vdash t : A$.*

2. If $p(m,0) \Vdash t = u : A$ and $p(m,1) \Vdash t = u : A$ for some $m \notin \mathrm{dom}(p)$ then $p \Vdash t = u : A$.

Proof. We prove the two statements mutually by induction.

1. Let $p(m,0) \Vdash t : A$ and $p(m,1) \Vdash t : A$ for some $m \notin \mathrm{dom}(p)$. Then $p(m,b) \Vdash A$ for all $b \in \{0,1\}$. By Lemma 2.19 we have $p \Vdash A$. By induction on the derivation of $p \Vdash A$.

   ($\mathbf{F_N}$) From the definition since $p(m,b) \Vdash t : A$ for all $b \in \{0,1\}$ we have that t has a proper $p(m,b)$-whnf. By Lemma 2.9 t has a proper p-whnf. If $p \vdash t \Rightarrow^* u : A$ where u is canonical then $p(m,b) \vdash t \Rightarrow^* u : A$ for all b. Since $p(m,b) \Vdash t : A$ we have by the definition that $u := \bar{n}$ for some $n \in \mathbb{N}$. Thus $p \Vdash t : A$. Otherwise, $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$ for some $k \notin \mathrm{dom}(p)$. If $k = m$ then by the definition $p \Vdash t : A$. Alternatively $k \neq m$ and $p(m,b) \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$. From the definition $p(m,b)(k,i) \Vdash t : A$ for all $b, i \in \{0,1\}$. Since $p(k,i) \lhd p(m,0)(k,i), p(m,1)(k,i)$, by the induction that $p(k,i) \Vdash t : A$ for all $i \in \{0,1\}$. By the definition $p \Vdash t : A$.
   The statement follows similarly for ($\mathrm{F}_{\mathrm{N}_0}$), ($\mathrm{F}_{\mathrm{N}_1}$), and ($\mathrm{F}_{\mathrm{N}_2}$).

   ($\mathbf{F_U}$) Follows similarly to Lemma 2.19.

   ($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x:F)G$. Since $p(m,b) \Vdash t : A$ we have that $\vdash_{p(m,b)} t : A$ for all $b \in \{0,1\}$. Thus $\vdash_p t : A$. Let $q \leqslant p$, $q \Vdash a : F$ and $q \Vdash a = c : F$. If $q \leqslant p(m,0)$ or $q \leqslant p(m,1)$ then $q \Vdash t\,a : G[a]$ and $q \Vdash t\,a = t\,c : G[a]$. Otherwise, $q(m,b) \leqslant p(m,b)$ for all $b \in \{0,1\}$. By monotonicity $q(m,b) \Vdash a : F$ and $q(m,b) \Vdash a = c : F$. We have then $q(m,b) \Vdash t\,a : G[a]$ and $q(m,b) \Vdash t\,a = t\,c : G[a]$ for all $b \in \{0,1\}$. By the induction $q \Vdash t\,a : G[a]$ and $q \Vdash t\,a = t\,c : G[a]$. Thus $p \Vdash t : A$.

   ($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x:F)G$. Since $p(m,b) \Vdash t : A$ we have $\vdash_{p(m,b)} t : A$ for all $b \in \{0,1\}$. Thus $\vdash_p t : A$. For all $b \in \{0,1\}$ we have $p(m,b) \Vdash t.1 : F$ and $p(m,b) \Vdash t.2 : G[t.1]$. By the induction $p \Vdash t.1 : F$ and $p \Vdash t.2 : G[t.1]$. Thus $p \Vdash t : \Sigma(x:F)G$.

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$. If $m = k$ and since $p(m,b) \Vdash t : A$ for all $b \in \{0,1\}$ then by the definition $p \Vdash t : A$. Otherwise, we have $p(m,b) \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$. Since $p(m,b) \Vdash t : A$ we have $p(m,b)(k,i) \Vdash t : A$ for all $i, b \in \{0,1\}$. But $p(k,i) \lhd p(m,0)(k,i), p(m,1)(k,i)$ and by the induction $p(k,i) \Vdash t : A$ for all $i \in \{0,1\}$. By the definition $p \Vdash t : A$.

2. Let $p(m,0) \Vdash t = u : A$ and $p(m,1) \Vdash t = u : A$ for some $m \notin \mathrm{dom}(p)$. Then $p(m,b) \Vdash A$ for all $b \in \{0,1\}$. By Lemma 2.19 we have $p \Vdash A$. Since $p(m,b) \Vdash t = u : A$ we have that $p(m,b) \Vdash t : A$ and $p(m,b) \Vdash u : A$ for all $b \in \{0,1\}$. By induction on the derivation of $p \Vdash A$.

   ($\mathbf{F_N}$) By the induction $p \Vdash t : A$ and $p \Vdash u : A$.

   i. Let $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$ for $k \notin \mathrm{dom}(p)$. If $k = m$ the we immediately have $p \Vdash t = u : A$. Otherwise, we have that $p(m,b) \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$ for all $b \in \{0,1\}$. Since $p(m,b) \Vdash t = u : A$ we have that $p(m,b)(k,i) \Vdash t = u : A$ for all $b, i \in \{0,1\}$. But $p(k,i) \lhd p(m,0)(k,i), p(m,1)(k,i)$ and by the induction $p(k,i) \Vdash t = u : A$ for all $i \in \{0,1\}$. By the definition $p \Vdash t = u : A$.

   ii. Let $p \vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$ for $k \notin \mathrm{dom}(p)$. The statement follows similarly.

   iii. Alternatively, let $p \vdash t \Rightarrow^* t' : A$ and $p \vdash u \Rightarrow^* u' : A$ where t' and u' are canonical. Thus $p(m,b) \vdash t \Rightarrow^* t' : A$ and $p(m,b) \vdash u \Rightarrow^* u' : A$ for all $b \in \{0,1\}$. Since $p(m,b) \vdash t = u : A$ and t' and u' are canonical, by the definition, $t' := u' := \bar{n}$ for some $n \in \mathbb{N}$. Thus $p \Vdash t = u : A$.

   The statement follows similarly for ($\mathrm{F}_{\mathrm{N}_0}$), ($\mathrm{F}_{\mathrm{N}_1}$), and ($\mathrm{F}_{\mathrm{N}_2}$).

   ($\mathbf{F_U}$) Follows similarly to Lemma 2.20.

   ($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. Since $p(m,b) \Vdash t = u : A$ we have that $p(m,b) \Vdash t : A$ and $p(m,b) \Vdash u : A$ for all $b \in \{0,1\}$. By Lemma 2.21 $p \Vdash t : A$ and $p \Vdash u : A$. Let $q \leqslant p$ and $q \Vdash a : F$. If $q \leqslant p(m,0)$ or $q \leqslant p(m,1)$ then $q \Vdash t\,a = u\,a : G[a]$. Otherwise, $q(m,b) \leqslant p(m,b)$ and $q(m,b) \Vdash a : F$, thus $q(m,b) \Vdash t\,a = u\,a : G[a]$ for all $b \in \{0,1\}$. By the induction $q \Vdash t\,a = u\,a : G[a]$. Thus $p \Vdash t = u : A$.

   ($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. We have that $p(m,b) \Vdash t.1 = u.1 : F$ and $p(m,b) \Vdash t.2 = u.2 : G[t.1]$ for all $b \in \{0,1\}$. By the induction $p \Vdash t.1 = u.1 : F$ and $p \Vdash t.2 = u.2 : G[t.1]$. Thus $p \Vdash t = u : A$.

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$. If $m = k$ and since $p(m,b) \Vdash t = u : A$ for all $b \in \{0,1\}$ then by the definition $p \Vdash t : A$. Otherwise, we have $p(m,b) \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$. Since $p(m,b) \Vdash t = u : A$ we have $p(m,b)(k,i) \Vdash t = u : A$ for all $i, b \in \{0,1\}$. But $p(k,i) \lhd p(m,0)(k,i), p(m,1)(k,i)$ and by the induction $p(k,i) \Vdash t = u : A$ for all $i \in \{0,1\}$. By the definition $p \Vdash t = u : A$. □

Corollary 2.22 (Local character)**.** *If $p \lhd p_1, \ldots, p_n$ and $p_i \Vdash J$ for all i then $p \Vdash J$.*

Proof. Follows from Lemma 2.19, Lemma 2.21, and Lemma 2.20 by induction. □

Lemma 2.23. *Let $p \vdash A \Rightarrow^* M$ where $M \in \{N_1, N_2, N\}$. If $p \Vdash a : A$ then there is a partition $p \lhd p_1, \ldots, p_n$ where a has a canonical p_i-whnf for all i. If $p \Vdash a = b : A$ then there is a partition $p \lhd q_1, \ldots, q_m$ where a and b have the same canonical q_j-whnf for each j.*

In particular, for $p \vdash A \Rightarrow^* N$. If $p \Vdash a : A$ then there is a partition $p \lhd p_1, \ldots, p_n$ where for each i, $p_i \vdash a \Rightarrow^* \bar{n}_i : N$ for some $n_i \in \mathbb{N}$. If $p \vdash a = b : A$ then there is a partition $p \lhd q_1, \ldots, q_m$ where for each j, $q_j \vdash a \Rightarrow^* \overline{m}_j$ and $q \vdash b \Rightarrow^* \overline{m}_j$ for some $m_j \in \mathbb{N}$.

Proof. Follows by induction from the definition. □

Lemma 2.24. *Let $p \Vdash A = B$.*

1. *If $p \Vdash t : A$ then $p \Vdash t : B$ and if $p \Vdash u : B$ then $p \Vdash u : A$.*
2. *If $p \Vdash t = u : A$ then $p \Vdash t = u : B$ and if $p \Vdash v = w : B$ then $p \Vdash v = w : A$.*

Proof. We will prove the two statements mutually by induction on the derivation of $p \Vdash A$.

1. Let $p \Vdash t : A$ and $p \Vdash u : B$.

   ($\mathbf{F_N}$) By induction on the derivation of $p \Vdash A = B$.

   i. Let $p \vdash B \Rightarrow^* N$. The statement then follows directly.

   ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for $m \notin \mathrm{dom}(p)$ and $p(m,b) \Vdash A = B$ for all $b \in \{0,1\}$. Since $p \Vdash t : A$, by monotonicity $p(m,b) \Vdash t : A$ and by the induction $p(m,b) \Vdash t : B$ for all $b \in \{0,1\}$. By the definition $p \Vdash t : B$.
   Since $p \Vdash u : B$, by monotonicity $p(m,b) \Vdash u : B$ and $p(m,b) \Vdash A = B$. By the induction $p(m,b) \Vdash u : A$ for all b. By local character $p \Vdash u : A$.

($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. By induction on the derivation of $p \Vdash A = B$.

i. Let $\vdash_p A = B$ and $p \vdash B \Rightarrow^* \Pi(x : H)E$ and $p \Vdash F = H$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a]$ whenever $q \Vdash a : F$. Since $p \Vdash t : A$ then $\vdash_p t : A$, thus $\vdash_p t : B$. Let $q \leqslant p$ and $q \Vdash a : H$. By monotonicity $q \Vdash F = H$. By the induction $q \Vdash a : F$, hence, $q \Vdash t\,a : G[a]$ and by the induction $q \Vdash t\,a : E[a]$. Similarly, $q \Vdash t\,a = t\,b : E[a]$ whenever $q \Vdash a = b : H$. Thus $p \Vdash t : B$. Similarly, we get $p \Vdash u : A$ from $p \Vdash u : B$.

ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$ and $p(k, b) \Vdash A = B$ for all $b \in \{0, 1\}$. Since $p \Vdash t : A$ then by monotonicity $p(k, b) \Vdash t : A$ and by the induction $p(k, b) \Vdash t : B$ for all b. By the definition $p \Vdash t : B$.
Since $p \Vdash u : B$ then by definition $p(k, b) \Vdash u : B$ and by the induction $p(k, b) \Vdash u : A$ for all $b \in \{0, 1\}$. By local character $p \Vdash u : A$.

($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. By induction on the derivation of $p \Vdash A = B$.

i. Let $\vdash_p A = B$ and $p \vdash B \Rightarrow^* \Sigma(x : H)E$ and $p \Vdash F = H$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a]$ whenever $q \Vdash a : F$. Since $p \Vdash t : A$ then $\vdash_p t : A$, thus $\vdash_p t : B$. Since $p \Vdash t.1 : F$, by the induction $p \Vdash t.1 : H$. Since $p \Vdash t.2 : H[t.1]$, by the induction $p \Vdash t.2 : E[t.1]$. Thus $p \Vdash t : B$. Similarly, we get $p \Vdash u : A$ from $p \Vdash u : B$.

ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$ and $p(k, b) \Vdash A = B$ for all $b \in \{0, 1\}$. Since $p \Vdash t : A$ then by monotonicity $p(k, b) \Vdash t : A$ and by the induction $p(k, b) \Vdash t : B$ for all b. By the definition $p \Vdash t : B$. Since $p \Vdash u : B$ then by definition $p(k, b) \Vdash g : B$ and by the induction $p(k, b) \Vdash g : A$ for all $b \in \{0, 1\}$. By local character $p \Vdash u : A$.

($\mathbf{F_U}$) Since $p \Vdash A = B$, we have $B := U$ and the statements follow directly.

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. Since $p \Vdash A = B$, we have $p(k, b) \Vdash A = B$ for all $b \in \{0, 1\}$. Since $p \Vdash t : A$ then $p(k, b) \Vdash t : A$ and by the induction $p(k, b) \Vdash t : B$ for all $b \in \{0, 1\}$. By local character $p \Vdash t : B$. Since $p \Vdash u : B$ then $p(k, b) \Vdash u : B$ and by the induction $p(k, b) \Vdash u : A$ for all $b \in \{0, 1\}$. By the definition $p \Vdash u : A$.

2. Let $p \Vdash t = u : A$ and $p \Vdash v = w : B$. By induction on the derivation of $p \Vdash A$.

   ($\mathbf{F_N}$) By induction on the derivation of $p \Vdash A = B$.

   i. Let $p \vdash B \Rightarrow^* N$. The statement then follows directly.

   ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for $m \notin \mathrm{dom}(p)$ and $p(m,b) \Vdash A = B$ for all $b \in \{0,1\}$. Since $p \Vdash t = u : A$, by monotonicity $p(m,b) \Vdash t = u : A$ and by the induction $p(m,b) \Vdash t = u : B$ for all b. By the definition $p \Vdash t = u : B$. Since $p \Vdash v = w : A$, by monotonicity $p(m,b) \Vdash v = w : B$ and $p(m,b) \Vdash A = B$. By the induction $p(m,b) \Vdash v = w : A$ for all b. By local character $p \Vdash v = w : A$.

   ($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. By induction on the derivation of $p \Vdash A = B$.

   i. Let $\vdash_p A = B$ and $p \vdash B \Rightarrow^* \Pi(x : H)E$ and $p \Vdash F = H$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a]$ whenever $q \Vdash a : F$. Since $p \Vdash t = u : A$ then $p \Vdash t : A$ and $p \Vdash u : A$. By the above $p \Vdash t : B$ and $p \Vdash u : B$. Let $q \leqslant p$ and $q \Vdash a : H$. By monotonicity $q \Vdash F = H$. By the induction $q \Vdash a : F$, hence, $q \Vdash t\,a = u\,a : G[a]$ and by the induction $q \Vdash t\,a = u\,a : E[a]$. Thus $p \Vdash t = u : B$. Similarly, Since $p \Vdash v = w : B$ we get $p \Vdash v = w : A$.

   ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. Since $p \Vdash t = u : A$ then by monotonicity $p(k,b) \Vdash t = u : A$ and by the induction $p(k,b) \Vdash t = u : B$ for all $b \in \{0,1\}$. By the definition $p \Vdash t = u : B$.
   Since $p \Vdash v = w : B$ then by definition $p(k,b) \Vdash v = w : B$ and by the induction $p(k,b) \Vdash v = w : A$ for all $b \in \{0,1\}$. By local character $p \Vdash v = w : A$.

   ($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. By induction on the derivation of $p \Vdash A = B$.

   i. Let $\vdash_p A = B$ and $p \vdash B \Rightarrow^* \Sigma(x : H)E$ and $p \Vdash F = H$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a]$ whenever $q \Vdash a : F$. Since $p \Vdash t = u : A$ then $p \Vdash t : A$ and $p \Vdash u : A$. By the above $p \Vdash t : B$ and $p \Vdash u : B$. Since $p \Vdash t.1 = u.1 : F$, by the induction $p \Vdash t.1 = u.1 : H$. Since $p \Vdash t.2 = u.2 : G[t.1]$, by the induction $p \Vdash t.2 = u.2 : E[t.1]$. Thus $p \Vdash t = u : B$. Similarly we get $p \Vdash v = w : A$ since $p \Vdash v = w : B$.

   ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. Since $p \Vdash t = u : A$ we have

$p \Vdash t : A$ and $p \Vdash u : A$. By the above we get $p \Vdash t : B$ and $p \Vdash u : B$. By monotonicity $p(k, b) \Vdash t = u : A$ and by the induction $p(k, b) \Vdash t = u : B$ for all $b \in \{0, 1\}$. By the definition $p \Vdash t = u : B$.
Since $p \Vdash v = w : B$ then by definition $p(k, b) \Vdash v = w : B$ and by the induction $p(k, b) \Vdash v = w : A$ for all $b \in \{0, 1\}$. By local character $p \Vdash v = w : A$.

($\mathbf{F_U}$) Since $p \Vdash A = B$, we have $B := U$ and the statements follow directly.

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. Since $p \Vdash A = B$, we have $p(k, b) \Vdash A = B$ for all $b \in \{0, 1\}$. Since $p \Vdash t = u : A$ then $p(k, b) \Vdash t = u : A$ and by the induction $p(k, b) \Vdash t = u : B$ for all $b \in \{0, 1\}$. By local character $p \Vdash t : B$.

Since $p \Vdash v = w : B$, we have $p \Vdash v : B$ and $p \Vdash w : B$. By the above $p \Vdash v : A$ and $p \Vdash w : A$. By monotonicity $p(k, b) \Vdash v = w : B$ and by the induction $p(k, b) \Vdash v = w : A$ for all $b \in \{0, 1\}$. By definition $p \Vdash v = w : A$. □

Lemma 2.25. *The relation $p \Vdash - = -$ is an equivalence relation. That is,*

(Reflexivity) If $p \Vdash A$ then $p \Vdash A = A$.

(Symmetry) If $p \Vdash A = B$ then $p \Vdash B = A$.

(Transitivity) If $p \Vdash A = B$ and $p \Vdash B = C$ then $p \Vdash A = C$.

Proof.

(Reflexivity) Reflexivity follows directly from the definition.

(Symmetry) Let $p \Vdash A = B$. We have then that $p \Vdash A$ and $p \Vdash B$. By induction on the derivation of $p \Vdash A$

- Let $p \vdash A \Rightarrow^* M$ where M is canonical. If B have canonical p-whnf then the statement follows in straight forward fashion from the definition. Otherwise, let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for $k \notin \mathrm{dom}(p)$. From $p \Vdash A = B$ we have by the definition that $p(k, b) \Vdash A = B$ for all $k \in \{0, 1\}$. By the induction $p(k, b) \Vdash B = A$ for all $b \in \{0, 1\}$. By the definition (i.e. ($\mathsf{F_{Loc}}$)) we have that $p \Vdash B = A$.
- Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for $m \notin \mathrm{dom}(p)$ we have then from $p \Vdash A = B$ that $p(m, b) \Vdash A = B$ for all $b \in \{0, 1\}$. By the induction $p(m, b) \Vdash B = A$ for all $b \in \{0, 1\}$. If B has a canonical p-whnf then from the definition (i.e. ($\mathrm{F_N}$),($\mathrm{F_{N_0}}$),($\mathrm{F_{N_1}}$), ($\mathrm{F_{N_2}}$),

(F_Π), (F_Σ)) we have that $p \Vdash B = A$. If $p \vdash B \Rightarrow^* \mathbb{E}[f\,\bar{k}]$ for $k \notin \mathrm{dom}(p)$ then by monotonicity $p(m,b)(k,i) \Vdash A = B$ for all $b, i \in \{0,1\}$. By the induction $p(m,b)(k,i) \Vdash B = A$ for all $b, i \in \{0,1\}$. But $p(k,i) \lhd p(m,0)(k,i), p(m,1)(k,i)$ and by local character $p(k,i) \Vdash B = A$ for all $i \in \{0,1\}$. By the definition we have that $p \Vdash B = A$.

(Transitivity) Let $p \Vdash A = B$ and $p \Vdash B = C$. We then have that $p \Vdash A$, $p \Vdash B$ and $p \Vdash C$. Thus A,B and C have proper p-whnf. By induction on the derivation of $p \Vdash A$.

$(\mathbf{F_N})$ By induction on the derivation of $p \Vdash A = B$.

i. Let $p \vdash B \Rightarrow^* N$ then the statement follows directly.

ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for $m \notin \mathrm{dom}(p)$. Then from $p \Vdash A = B$ we have that $p(m,b) \Vdash A = B$ for all $b \in \{0,1\}$. From $p \Vdash B = C$ we have that $p(m,b) \Vdash B = C$ for all $b \in \{0,1\}$. By the induction $p(m,b) \Vdash A = C$ for all $b \in \{0,1\}$. By local character $p \Vdash A = C$.

$(\mathbf{F_\Pi})$ Let $p \vdash A \Rightarrow^* \Pi(x\!:\!F)G$. By induction on the derivation of $p \Vdash A = B$.

i. Let $p \vdash B \Rightarrow^* \Pi(x : H)E$ and $p \Vdash F = H$ and for all $q \leqslant p$, $q \Vdash G[a] = E[a]$ whenever $q \Vdash a\!:\!F$. By induction on the derivation of $p \Vdash B = C$.

- Let $p \vdash C \Rightarrow^* \Pi(x : T)R$ and $p \Vdash H = T$ and for all $q \leqslant p$, $q \Vdash E[b] = R[b]$ whenever $q \Vdash b : H$. By the induction $p \Vdash F = T$. Let $q \leqslant p$ and $q \Vdash a : F$. By monotonicity $q \Vdash F = H$ and by Lemma 2.24 $q \Vdash a\!:\!H$. Thus $q \Vdash E[a] = R[a]$. But $q \Vdash G[a] = E[a]$. By the induction $q \Vdash G[a] = R[a]$. Thus $p \Vdash A = C$.
- Let $p \vdash C \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and $p(m,b) \Vdash B = C$ for all $b \in \{0,1\}$. By monotonicity $p(m,b) \Vdash A = B$, and by the induction $p(m,b) \Vdash A = C$ for all $b \in \{0,1\}$. By the definition $p \Vdash A = C$.

ii. Let $p \vdash B \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and $p(m,b) \Vdash A = B$ for all $b \in \{0,1\}$. By monotonicity $p(m,b) \Vdash B = C$, and by the induction $p(m,b) \Vdash A = C$ for all $b \in \{0,1\}$. By local character $p \Vdash A = C$.

$(\mathbf{F_\Sigma})$ Follows similarly to the above.

$(\mathbf{F_U})$ Since $p \Vdash A = B$ and $p \Vdash B = C$, we have $B := U$ and $C := U$ and the statements follow directly.

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. Since $p \Vdash A = B$, we have $p(k,b) \Vdash A = B$ for all $b \in \{0,1\}$. By monotonicity $p(k,b) \Vdash B = C$ and by the induction $p(k,b) \Vdash A = C$ for all $b \in \{0,1\}$. By the definition $p \Vdash A = C$. □

Lemma 2.26. *The relation $p \Vdash - = - : A$ is an equivalence relation. That is*

(Reflexivity) If $p \Vdash t : A$ then $p \Vdash t = t : A$.

(Symmetry) If $p \Vdash t = u : A$ then $p \Vdash u = t : A$.

(Transitivity) If $p \Vdash t = u : A$ and $p \Vdash u = v : A$ then $p \Vdash t = v : A$.

Proof. (Reflexivity) Let $p \Vdash t : A$. We have that $p \Vdash A$. Reflexivity follows from the definition in a straightforward fashion by induction on the derivation of $p \Vdash A$

(Symmetry) Let $p \Vdash t = u : A$. We have that $p \Vdash t : A$, $p \Vdash u : A$ and $p \Vdash A$. By induction on the derivation of $p \Vdash A$

($\mathbf{F_N}$) By induction on the derivation of $p \Vdash t : A$.

i. Let $p \vdash t \Rightarrow^* \bar{n} : A$ for some $n \in \mathbb{N}$. By induction on the derivation of $p \Vdash t = u : A$. If $p \vdash u \Rightarrow^* \bar{n} : A$ then we have immediately that $p \Vdash u = t : A$. Alternatively, $p \vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for some $m \notin \mathrm{dom}(p)$ and $p(m,i) \Vdash t = u : A$ for all $i \in \{0,1\}$. By the induction $p(m,i) \Vdash u = t : A$ for all $i \in \{0,1\}$ and by the definition $p \Vdash u = t : A$.

ii. Let $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\overline{m}]$ for $m \notin \mathrm{dom}(p)$. Since $p \Vdash t = u : A$ we have that $p(m,b) \Vdash t = u : A$ for all $b \in \{0,1\}$. By the induction $p(m,b) \Vdash u = t : A$ for all $b \in \{0,1\}$. By local character $p \Vdash u = t : A$.

The statement follows similarly for ($\mathsf{F}_{\mathsf{N}_1}$) and ($\mathsf{F}_{\mathsf{N}_2}$)

($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. Since $p \Vdash t = u : A$ we have $\vdash_p t = u : A$ and thus $\vdash_p u = t : A$. Let $q \leqslant p$ and $q \Vdash a : F$ we then have $q \Vdash t\,a = u\,a : G[a]$ and by the induction $q \Vdash u\,a = t\,a : G[a]$. Thus $p \Vdash u = t : A$.

($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. Since $p \Vdash t = u : A$ then $\vdash_p t = u : A$ and we have $\vdash_p u = t : A$. By the definition, $p \Vdash t.1 = u.1 : F$ and $p \Vdash t.2 = u.2 : G[t.1]$. By the induction $p \Vdash u.1 = t.1 : F$. Since $p \Vdash A$ we have that $p \Vdash G[t.1] = G[u.1]$. By Lemma 2.24 $p \Vdash t.2 = u.2 : G[u.1]$. By the induction $p \Vdash u.2 = t.2 : G[u.1]$. Thus $p \Vdash u = t : A$.

($\mathbf{F_U}$) Follows similarly to Lemma 2.25

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. Since $p \Vdash t = u : A$ we have that $p(k,b) \Vdash t = u : A$ for all $b \in \{0,1\}$. By the induction $p(k,b) \Vdash u = t : A$ for all $b \in \{0,1\}$. By the definition $p \Vdash u = t : A$.

(Transitivity) Let $p \Vdash t = u : A$ and $p \Vdash u = v : A$. We have that $p \Vdash A$, $p \Vdash t : A$, $p \Vdash u : A$ and $p \Vdash v : A$. By induction on the derivation of $p \Vdash A$.

($\mathbf{F_N}$) By induction on the derivation of $p \Vdash t = u : A$.

i. Let $p \vdash t \Rightarrow^* \bar{n} : A$ and $p \vdash u \Rightarrow^* \bar{n} : A$. By induction on the derivation of $p \Vdash u = v : A$. If $p \vdash v \Rightarrow^* \bar{n} : A$ we have immediately that $p \Vdash t = v : A$. If $p \vdash v \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash u = v : A$ for all $b \in \{0,1\}$. By the induction $p(k,b) \Vdash t = v : A$ for all $b \in \{0,1\}$. By the definition $p \Vdash t = v : A$.

ii. Let $p \vdash t \Rightarrow^* \bar{n} : A$ and $p \Vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{m}]$ for some $m \notin \mathrm{dom}(p)$ where $p(m,b) \Vdash t = u : A$ for all $b \in \{0,1\}$. Since $p \Vdash u = v : A$ we have that $p(m,b) \Vdash u = v : A$ and by the induction $p(m,b) \Vdash t = v : A$. By local character $p \Vdash t = v : A$.

iii. Let $p \vdash t \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$ where $p(k,b) \Vdash t = u : A$ for all $b \in \{0,1\}$. Since $p \Vdash u = v : A$, by monotonicity $p(k,b) \Vdash u = v : A$. By the induction $p(k,b) \Vdash t = v : A$ for all $b \in \{0,1\}$. By the definition $p \Vdash t = v : A$.

The statement follows similarly for $(\mathsf{F}_{\mathsf{N}_1})$ and $(\mathsf{F}_{\mathsf{N}_2})$

($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. Since $p \Vdash t = u : A$ we have that $\vdash_p t = u : A$. Similarly we have $\vdash_p u = v : A$ and thus $\vdash_p t = v : A$. Let $q \leqslant p$ and $q \Vdash a : F$. We have then $q \Vdash t\,a = u\,a : G[a]$ and $q \Vdash u\,a = v\,a : G[a]$. By the induction $q \Vdash t\,a = v\,a : G[a]$. Thus $p \Vdash t = v : A$.

($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. Since $p \Vdash t = u : A$ we have that $p \Vdash t.2 = u.1 : F$ and $p \Vdash t.2 = u.2 : G[t.1]$. Similarly we have that $p \Vdash u.1 = v.1 : F$ and $p \Vdash u.2 = v.2 : G[u.1]$. Since $p \Vdash A$ we have that $p \Vdash G[t.1] = G[u.1]$ and by Lemma 2.24 $p \Vdash u.2 = v.2 : G[t.1]$. By the induction then we have $p \Vdash t.1 = v.1 : F$ and $p \Vdash t.2 = v.2 : G[t.1]$. We have then that $p \Vdash t = v : A$.

($\mathbf{F_U}$) Follows similarly to Lemma 2.25.

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for some $k \notin \mathrm{dom}(p)$. We have then that $p(k,b) \Vdash t = u : A$ and $p(k,b) \Vdash u = v : A$ for all $b \in$

$\{0,1\}$. By the induction $p(k,b) \Vdash t = v : A$ for all $b \in \{0,1\}$. By the definition we have then that $p \Vdash t = v : A$. □

3 Soundness

In this section we show that the type theory described in Section 1 is sound with respect to the semantics described in Section 2. That is, we aim to show that for any judgment J whenever $\vdash_p J$ then $p \Vdash J$.

Lemma 3.1. *If $p \vdash A \Rightarrow^* B$ and $p \Vdash B$ then $p \Vdash A$ and $p \Vdash A = B$.*

Proof. Follows from the definition. □

Lemma 3.2. *Let $p \Vdash A$. If $p \vdash t \Rightarrow u : A$ and $p \Vdash u : A$ then $p \Vdash t : A$ and $p \Vdash t = u : A$.*

Proof. Let $p \vdash t \Rightarrow u : A$ and $p \Vdash u : A$. By induction on the derivation of $p \Vdash A$.

($\mathbf{F_U}$) The statement follows similarly to Lemma 3.1.

($\mathbf{F_N}$) By induction on the derivation of $p \Vdash u : A$. If $p \vdash u \Rightarrow^* \bar{n} : N$ for some $n \in \mathbb{N}$ then $p \vdash t \Rightarrow^* \bar{n} : N$ and the statement follows by the definition. If $p \vdash u \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}] : A$ for $k \notin \mathrm{dom}(p)$ and $p(k,b) \Vdash u : A$ for all $b \in \{0,1\}$ then since $p(k,b) \vdash t \Rightarrow u : A$, by the induction, $p(k,b) \Vdash t : A$ and $p(k,b) \Vdash t = u : A$. By the definition $p \Vdash t : A$ and $p \Vdash t = u : A$. The statement follows similarly for $(\mathrm{F}_{\mathrm{N}_1})$, $(\mathrm{F}_{\mathrm{N}_2})$.

($\mathbf{F_\Pi}$) Let $p \vdash A \Rightarrow^* \Pi(x : F)G$. Since $p \vdash t \Rightarrow u : A$ we have $\vdash_p t : A$. Let $q \leqslant p$ and $q \Vdash a : F$. We have $q \vdash t\,a \Rightarrow u\,a : G[a]$. By the induction $q \Vdash t\,a : G[a]$ and $q \Vdash t\,a = u\,a : G[a]$. If $q \Vdash a = b : F$ we similarly get $q \Vdash t\,b : G[b]$ and $q \Vdash t\,b = u\,b : G[b]$. Since $q \Vdash G[a] = G[b]$, by Lemma 2.24 $q \Vdash t\,b = u\,b : G[a]$. But $q \Vdash u\,a = u\,b : G[a]$. By symmetry and transitivity $q \Vdash t\,a = t\,b : G[a]$. Thus $p \Vdash t : A$ and $p \Vdash t = u : A$.

($\mathbf{F_\Sigma}$) Let $p \vdash A \Rightarrow^* \Sigma(x : F)G$. From $p \vdash t \Rightarrow u : A$ we have $\vdash_p t : A$ and we have $p \vdash t.1 \Rightarrow u.1 : F$ and $p \vdash t.2 \Rightarrow u.2 : G[u.1]$. By the induction $p \Vdash t.1 : F$ and $p \Vdash t.1 = u.1 : F$. By the induction $p \Vdash t.2 : G[u.1]$ and $p \Vdash t.2 = u.2 : G[u.1]$. But since $p \Vdash A$ and we have shown $p \Vdash t.1 = u.1 : F$ we get $p \Vdash G[t.1] = G[u.1]$. By Lemma 2.24 $p \Vdash t.2 : G[t.1]$ and $p \Vdash t.2 = u.2 : G[t.1]$. Thus $p \Vdash t : A$ and $p \Vdash t = u : A$

($\mathbf{F_{Loc}}$) Let $p \vdash A \Rightarrow^* \mathbb{E}[\mathsf{f}\,\bar{k}]$ for $k \notin \mathrm{dom}(p)$. Since $p \Vdash u : A$ we have $p(k,b) \Vdash u : A$ for all $b \in \{0,1\}$. But we have $p(k,b) \vdash t \Rightarrow u : A$. By the induction $p(k,b) \Vdash t : A$ and $p(k,b) \Vdash t = u : A$. By the definition $p \Vdash t : A$ and $p \Vdash t = u : A$. □

Corollary 3.3. *Let $p \vdash t \Rightarrow^* u : A$ and $p \Vdash A$. If $p \Vdash u : A$ then $p \Vdash t : A$ and $p \Vdash t = u : A$.*

Corollary 3.4. $\Vdash \mathsf{f} : N \to N_2$.

Proof. It's direct to see that $\Vdash N \to N_2$. For an arbitrary condition p let $p \Vdash n : N$. By Lemma 2.23 we have a partition $p \lhd p_1, \ldots, p_m$ where for each i, $p_i \vdash n \Rightarrow^* \overline{m}_i : N$ for some $m_i \in \mathbb{N}$. We have thus a reduction $p_i \vdash \mathsf{f}\, n \Rightarrow^* \mathsf{f}\, \overline{m}_i : N_2$. If $m_i \in \mathrm{dom}(p_i)$ then $p_i \vdash \mathsf{f}\, n \Rightarrow^* \mathsf{f}\, \overline{m}_i \Rightarrow p_i(m_i) : N_2$ and by definition $p_i \Vdash \mathsf{f}\, n : N_2$. If for any j, $m_j \notin \mathrm{dom}(p_j)$ then $p_j(m_j, 0) \vdash \mathsf{f}\, n \Rightarrow^* \mathsf{f}\overline{m}_j \Rightarrow 0 : N_2$ and $p_j(m_j, 1) \vdash \mathsf{f}\, n \Rightarrow^* \mathsf{f}\overline{m}_j \Rightarrow 1 : N_2$. Thus $p_j(m_j, 0) \Vdash \mathsf{f}\, n : N_2$ and $p_j(m_j, 1) \Vdash \mathsf{f}\, n : N_2$. By the definition $p_j \Vdash \mathsf{f}\, n : N_2$. We thus have that $p_i \Vdash \mathsf{f}\, n : N_2$ for all i and by local character $p \Vdash \mathsf{f}\, n : N_2$.

Let $p \Vdash n_1 = n_2 : N$. By Lemma 2.23 there is a partition $p \lhd p_1, \ldots, p_m$ where for each i, $p_i \vdash n_1 \Rightarrow^* \overline{m}_i : N$ and $p_i \vdash n_2 \Rightarrow^* \overline{m}_i : N$ for some $m_i \in \mathbb{N}$. We then have that $p_i \vdash \mathsf{f}\, n_1 \Rightarrow^* \mathsf{f}\, \overline{m}_i : N_2$ and $p_i \vdash \mathsf{f}\, n_2 \Rightarrow^* \mathsf{f}\, \overline{m}_i : N_2$. If $m_i \in \mathrm{dom}(p_i)$ then $p_i \vdash \mathsf{f}\, n_1 \Rightarrow^* p_i(m_i) : N_2$ and $p_i \vdash \mathsf{f}\, n_2 \Rightarrow^* p_i(m_i) : N_2$. By Corollary 3.3, symmetry and transitivity $p_i \Vdash \mathsf{f}\, n_1 = \mathsf{f}\, n_2 : N_2$. If on the other hand $m_j \notin \mathrm{dom}(p_j)$ for some j then similarly $p_j(m_j, 0) \Vdash \mathsf{f}\, n_1 = \mathsf{f}\, n_2 : N_2$ and $p_j(m_j, 1) \Vdash \mathsf{f}\, n_1 = f\, n_2 : N_2$. By the definition $p_j \Vdash \mathsf{f}\, n_1 = \mathsf{f}\, n_2 : N_2$. Thus we have that $p_i \Vdash \mathsf{f}\, n_1 = \mathsf{f}\, n_2 : N$ for all i. By local character $p \Vdash \mathsf{f}\, n_1 = \mathsf{f}\, n_2 : N_2$. Hence $\Vdash \mathsf{f} : N \to N_2$. □

Lemma 3.5. *If $\vdash_p t : \neg A$ and $p \Vdash A$ then $p \Vdash t : \neg A$ iff for all $q \leqslant p$ there is no term u such that $q \Vdash u : A$.*

Proof. Let $p \Vdash A$ and $\vdash_p t : \neg A$. We have directly that $p \Vdash \neg A$. Assume $p \Vdash t : \neg A$. If $q \Vdash u : A$ for some $q \leqslant p$, then $q \Vdash t\, u : N_0$ which is impossible. Conversely, assume it is the case that for all $q \leqslant p$ there is no u for which $q \Vdash u : A$. Since $r \Vdash a : A$ and $r \Vdash a = b : A$ never hold for any $r \leqslant p$, the statement "$r \Vdash t\, a : N_0$ whenever $r \Vdash a : A$ and $r \Vdash t\, a = t\, b : N_0$ whenever $r \Vdash a = b : A$" holds trivially. □

Lemma 3.6. $\Vdash \mathsf{w} : \neg\neg(\Sigma(x : N)\mathsf{IsZero}(\mathsf{f}\, x))$.

Proof. By Lemma 3.5 it is enough to show that for all q there is no term u for which $q \Vdash u : \neg(\Sigma(x : N)\mathsf{IsZero}(\mathsf{f}\, x))$. Assume $q \Vdash u : \neg(\Sigma(x : N)\mathsf{IsZero}(\mathsf{f}\, x))$ for some u. Let $m \notin \mathrm{dom}(q)$ we have then

$q(m,0) \Vdash (\overline{m},0) : \Sigma(x:N)\mathsf{IsZero}(\mathsf{f}\ x)$ thus $q(m,0) \Vdash u\ (\overline{m},0) : N_0$ which is impossible. □

Let $\Gamma := x_1 : A_1 \ldots, x_n : A_n[x_1, \ldots, x_{n-1}]$ and $\rho := a_1, \ldots, a_n$. We say that $\Delta \vdash_p \rho : \Gamma$ if $\Delta \vdash_p a_1 : A_1, \ldots, \Delta \vdash_p a_n : A_n[a_1, \ldots, a_{n-1}]$.

If moreover, $\sigma = b_1, \ldots, b_n$, we say that $\Delta \vdash_p \rho = \sigma : \Gamma$ if $\Delta \vdash_p a_1 = b_1 : A_1, \ldots, \Delta \vdash_p a_n = b_n : A_n[a_1, \ldots, a_{n-1}]$.

Let $\Gamma := x_1 : A_1 \ldots, x_n : A_n[x_1, \ldots, x_{n-1}]$ and $\rho := a_1, \ldots, a_n$. We say $p \Vdash \rho : \Gamma$ if $p \Vdash a_1 : A, \ldots, p \Vdash a_n : A_n[a_1, \ldots, a_{n-1}]$.

If moreover $\sigma := b_1, \ldots, b_n$ and $p \Vdash \sigma : \Gamma$, we say $p \Vdash \rho = \sigma : \Gamma$ if $p \Vdash a_1 = b_1 : A_1, \ldots, p \Vdash a_n = b_n : A_n[a_1, \ldots, a_{n-1}]$.

Lemma 3.7. *If $p \Vdash \rho : \Gamma$ then $\vdash_p \rho : \Gamma$. If $p \Vdash \rho = \sigma : \Gamma$ then $\vdash_p \rho = \sigma : \Gamma$.*

Proof. Follows by induction from Lemma 2.11. □

Definition 3.8.

1. We write $\Gamma \vDash_p A$ if $\Gamma \vdash_p A$ and for all $q \leqslant p$ whenever $q \Vdash \rho : \Gamma$ then $q \Vdash A\rho$ and whenever $q \Vdash \rho = \sigma : \Gamma$ then $q \Vdash A\rho = A\sigma$.

2. We write $\Gamma \vDash_p t : A$ if $\Gamma \vdash_p t : A$, $\Gamma \vDash_p A$ and for all $q \leqslant p$ whenever $q \Vdash \rho : \Gamma$ then $q \Vdash t\rho : A\rho$ and whenever $q \Vdash \rho = \sigma : \Gamma$ then $q \Vdash t\rho = t\sigma : A\rho$.

3. We write $\Gamma \vDash_p A = B$ if $\Gamma \vdash_p A = B$, $\Gamma \vDash_p A$, $\Gamma \vDash_p B$ and for all $q \leqslant p$ whenever $q \Vdash \rho : \Gamma$ then $q \Vdash A\rho = B\rho$.

4. We write $\Gamma \vDash_p t = u : A$ if $\Gamma \vdash_p t = u : A$, $\Gamma \vDash_p t : A$, $\Gamma \vDash_p u : A$ and for all $q \leqslant p$ whenever $q \Vdash \rho : \Gamma$ then $q \Vdash t\rho = u\rho : A\rho$.

In the following we will show that whenever we have a rule $\dfrac{\Gamma_1 \vdash_p J_1 \quad \ldots \quad \Gamma_\ell \vdash_p J_\ell}{\Gamma \vdash_p J}$ in the type system then it holds that $\dfrac{\Gamma_1 \vDash_p J_1 \quad \ldots \quad \Gamma_\ell \vDash_p J_\ell}{\Gamma \vDash_p J}$. Which is sufficient to show soundness.

Lemma 3.9. $\dfrac{\Gamma \vDash_p F \quad \Gamma, x : F \vDash_p G}{\Gamma \vDash_p \Pi(x:F)G} \quad \dfrac{\Gamma \vDash_p F \quad \Gamma, x : F \vDash_p G}{\Gamma \vDash_p \Sigma(x:F)G}$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. Let $r \leqslant q$ and $r \Vdash a : F\rho$. We have $r \Vdash (\rho, a) : (\Gamma, x : F)$. Since $\Gamma \vDash_r F$ we have $r \Vdash F\rho = F\sigma$ and by Lemma 2.24 $r \Vdash a : F\sigma$. Thus $r \Vdash (\rho, a) = (\sigma, a) : (\Gamma, x : F)$. We have $r \Vdash G\rho[a]$ and $r \Vdash G\rho[a] = G\sigma[a]$. If moreover $r \Vdash a = b : F\rho$ then

$r \Vdash (\rho, a) = (\rho, b) : (\Gamma, x : F)$. Thus $r \Vdash G\rho[a] = G\rho[b]$. By the definition $q \Vdash (\Pi(x : F)G)\rho = (\Pi(x : F)G)\sigma$ and $q \Vdash (\Sigma(x : F)G)\rho = (\Sigma(x : F)G)\sigma$. □

Lemma 3.10.

$$\frac{\Gamma \vDash_p F = H \quad \Gamma, x{:}F \vDash_p G = E}{\Gamma \vDash_p \Pi(x{:}F)G = \Pi(x{:}H)E} \qquad \frac{\Gamma \vDash_p F = H \quad \Gamma, x{:}F \vDash_p G = E}{\Gamma \vDash_p \Sigma(x{:}F)G = \Sigma(x{:}H)E}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. Similarly to Lemma 3.9, we can show $q \Vdash (\Sigma(x{:}F)G)\rho$, $q \Vdash (\Sigma(x{:}H)E)\rho$, $q \Vdash (\Pi(x{:}F)G)\rho$, and $q \Vdash (\Pi(x{:}H)E)\rho$.

From the premise $q \Vdash F\rho = H\rho$. Let $r \leqslant q$ and $r \Vdash a : F\rho$. We have then $r \Vdash (\rho, a) : (\Gamma, x : F)$. Thus $r \Vdash G\rho[a] = E\rho[a]$. By the definition $q \Vdash (\Pi(x{:}F)G)\rho = (\Pi(x{:}H)E)\rho$ and $q \Vdash (\Sigma(x{:}F)G)\rho = (\Sigma(x{:}H)E)\rho$. □

Lemma 3.11. $\dfrac{\Gamma, x{:}F \vDash_p t : G}{\Gamma \vDash_p \lambda x.t : \Pi(x{:}F)G}$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. By Lemma 3.7 $\vdash_q \rho : \Gamma$. Let $r \leqslant q$ and $r \Vdash d : F\rho$. Since $\Gamma, x : F \vdash_p t : G$ we have that $x : F\rho \vdash_r t\rho : G\rho$. Since $\vdash_r d : F\rho$ we have $\vdash_r (\lambda x.t\rho)d = t\rho[d] : G\rho[d]$. By the reduction rules $(\lambda x.t\rho)d \Rightarrow t\rho[d]$. Thus $r \vdash (\lambda x.t\rho)d \Rightarrow t\rho[d] : G\rho[d]$. But $r \Vdash (\rho, d) : (\Gamma, x : F)$, hence, $r \Vdash t\rho[d] : G\rho[d]$. By Lemma 3.2 we have that $r \Vdash (\lambda x.t\rho)\, d : G\rho[d]$ and $r \Vdash (\lambda x.t\rho)\, d = t\rho[d] : G\rho[d]$.

Let $r \Vdash e = d : F\rho$ we have similarly that $r \Vdash (\lambda x.t\rho)\, e = t\rho[e] : G\rho[e]$. We have also that $r \Vdash (\rho, d) = (\rho, e) : G\rho[d]$, thus $r \Vdash t\rho[d] = t\rho[e] : G\rho[d]$ and $r \Vdash G\rho[d] = G\rho[e]$. By Lemma 2.24 we have $r \Vdash (\lambda x.t\rho)\, e = t\rho[e] : G\rho[d]$. By symmetry and transitivity we have $r \Vdash (\lambda x.t\rho)\, d = (\lambda x.t\rho)\, e : G\rho[d]$. Thus $q \Vdash (\lambda x.t)\rho : (\Pi(x{:}F)G)\rho$.

Let $q \Vdash \rho = \sigma : \Gamma$. We get $q \Vdash F\rho = F\sigma$. Similarly to the above we can show $q \Vdash (\lambda x.t)\sigma : (\Pi(x{:}F)G)\sigma$. Let $r \leqslant q$ and $r \Vdash a : F\rho$. By Lemma 2.24 $r \Vdash a : F\sigma$. We then have $r \Vdash (\rho, a) = (\sigma, a) : (\Gamma, x : F)$. Thus we have $r \Vdash G\rho[a] = G\sigma[a]$. Thus $q \Vdash (\Pi(x : F)G)\rho = (\Pi(x : F)G)\sigma$ and by Lemma 2.24 $q \Vdash (\lambda x.t)\sigma : (\Pi(x{:}F)G)\rho$. We have $r \Vdash t\rho[a] = t\sigma[a] : G\rho[a]$. But $r \Vdash (\lambda x.t\rho)\, a = t\rho[a] : G\rho[a]$ and $r \Vdash (\lambda x.t\sigma)a = t\sigma[a] : G\sigma[a]$. By Lemma 2.24 $r \Vdash (\lambda x.t\sigma)a = t\sigma[a] : G\rho[a]$. By Symmetry and transitivity $r \Vdash (\lambda x.t\rho)\, a = (\lambda x.t\sigma)\, a : G\rho[a]$. Thus $q \Vdash (\lambda x.t)\rho = (\lambda x.t)\sigma : (\Pi(x : F)G)\rho$. □

Lemma 3.12. $\dfrac{\Gamma, x{:}F \vDash_p t : G \quad \Gamma \vDash_p a : F}{\Gamma \vDash_p (\lambda x.t)\, a = t[a] : G[a]}$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. We have $q \Vdash a\rho : F\rho$. As in Lemma 3.11 $q \vdash ((\lambda x.t)\, a)\rho \Rightarrow t[a]\rho : G[a]\rho$ which by Lemma 3.2 imply that $q \Vdash ((\lambda x.t)\, a)\rho = t[a]\rho : G\rho[a]$. □

Lemma 3.13. $$\frac{\Gamma \vDash_p g : \Pi(x:F)G \quad \Gamma \vDash_p a : F}{\Gamma \vDash_p g\, a : G[a]}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. We have $q \Vdash g\rho : (\Pi(x:F)G)\rho$ and $q \Vdash a\rho : F\rho$. By the definition $q \Vdash (g\, a)\rho : G[a]\rho$.

Let $q \Vdash \rho = \sigma : \Gamma$. We have then $q \Vdash g\rho = g\sigma : (\Pi(x:F)G)\rho$ and $q \Vdash a\rho = a\sigma : F\rho$. From the definition $q \Vdash g\rho\, a\rho = g\sigma\, a\rho : G[a]\rho$. From the definition $q \Vdash g\sigma\, a\rho = g\sigma\, a\sigma : G[a]\rho$. By transitivity $q \Vdash (g\, a)\rho = (g\, a)\sigma : G[a]\rho$. □

Lemma 3.14.

1. $$\frac{\Gamma \vDash_p g : \Pi(x:F)G \quad \Gamma \vDash_p u = v : F}{\Gamma \vDash_p g\, u = g\, v : G[u]}$$
2. $$\frac{\Gamma \vDash_p h = g : \Pi(x:F)G \quad \Gamma \vDash_p u : F}{\Gamma \vDash_p h\, u = g\, u : G[u]}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$.

1. We have $q \Vdash g\rho : (\Pi(x:F)G)\rho$ and $q \Vdash u\rho = v\rho : F\rho$. From the definition get $q \Vdash (g\, u)\rho = (g\, v)\rho : G[u]\rho$.
2. We have $q \Vdash h\rho = g\rho : (\Pi(x:F)G)\rho$ and $q \Vdash u\rho : F\rho$. From the definition we get $q \Vdash (h\, u)\rho = (g\, u)\rho : G[u]\rho$. □

Lemma 3.15.

$$\frac{\Gamma \vDash_p h : \Pi(x:F)G \quad \Gamma \vDash_p g : \Pi(x:F)G \quad \Gamma, x:F \vDash_p h\, x = g\, x : G[x]}{\Gamma \vDash_p h = g : \Pi(x:F)G}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. We have $q \Vdash h\rho : (\Pi(x:F)G)\rho$ and $q \Vdash g\rho : (\Pi(x:F)G)\rho$. Let $r \leqslant q$ and $r \Vdash a : F\rho$. We have then that $r \Vdash (\rho, a) : \Gamma, x : F$. Thus $r \Vdash h\rho\, a = g\rho\, a : G\rho[a]$. By the definition $q \Vdash h\rho = g\rho : (\Pi(x:F)G)\rho$. □

Lemma 3.16. $$\frac{\Gamma, x:F \vDash_p G \quad \Gamma \vDash_p a : F \quad \Gamma \vDash_p b : G[a]}{\Gamma \vDash_p (a, b) : \Sigma(x:F)G}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. By the typing rules $\Gamma \vdash_q (a, b).1 = a : F$ and $\Gamma \vdash_q (a, b).2 = b : F[a]$. But $\vdash_q \rho : \Gamma$. By substitution we have $\vdash_q ((a, b).1)\rho = a\rho : F\rho$ and $\vdash_q ((a, b).2)\rho = b\rho : G[a]\rho$. But $((a, b).1)\rho \Rightarrow_q a\rho$ and $((a, b).2)\rho \Rightarrow_q b\rho$. Thus $q \vdash ((a, b).1)\rho \Rightarrow a\rho : F\rho$ and $q \vdash ((a, b).2)\rho \Rightarrow b\rho : G[a]\rho$. From the premise $q \Vdash a\rho : F\rho$

and $q \Vdash b\rho : G[a]\rho$. By Lemma 3.2 $q \Vdash ((a,b).1)\rho : F\rho$ and $q \Vdash ((a,b).2)\rho : G[a]\rho$. By Lemma 3.2 $q \Vdash ((a,b).1)\rho = a\rho : F\rho$, thus $q \Vdash (\rho, a\rho) = (\rho, ((a,b).1)\rho) : (\Gamma, x:F)$. Hence $q \Vdash G[a]\rho = G[(a,b).1]\rho$. By Lemma 2.24 $q \Vdash ((a,b).2)\rho : G[(a,b).1]\rho$. By the definition we have then that $q \Vdash (a,b)\rho : (\Sigma(x:F)G)\rho$.

Let $q \Vdash \rho = \sigma : \Gamma$. Similarly we can show $q \Vdash (a,b)\sigma : (\Sigma(x:F)G)\sigma$. We have that $q \Vdash a\rho = a\sigma : F\rho$ and $q \Vdash b\rho = b\sigma : G[a]\rho$. We have also $q \Vdash (\rho, a\rho) = (\sigma, a\sigma) : (\Gamma, x : F)$ we thus have $q \Vdash G[a]\rho = G[a]\sigma$. By Lemma 3.2 $q \Vdash ((a,b).2)\sigma = b\sigma : G[a]\sigma$. By Lemma 2.24 $q \Vdash ((a,b).2)\sigma = b\sigma : G[a]\rho$. But we also have by Lemma 3.2 that $q \Vdash ((a,b).2)\sigma = a\sigma : F\sigma$. Hence, by Lemma 2.24, we have $q \Vdash ((a,b).2)\sigma = a\sigma : F\rho$. By symmetry and transitivity we then have that $q \Vdash ((a,b).1)\rho = ((a,b).1)\sigma : F\rho$ and $q \Vdash ((a,b).2)\rho = ((a,b).2)\sigma : G[(a,b).1]\rho$. Thus we have that $q \Vdash (a,b)\rho = (a,b)\sigma : (\Sigma(x:F)G)\rho$. ☐

Lemma 3.17.

1. $$\frac{\Gamma, x:F \vDash_p G \quad \Gamma \vDash_p t:F \quad \Gamma \vDash_p u:G[t]}{\Gamma \vDash_p (t,u).1 = t:F}$$

2. $$\frac{\Gamma, x:F \vDash_p G \quad \Gamma \vDash_p t:F \quad \Gamma \vDash_p u:G[t]}{\Gamma \vDash_p (t,u).2 = u:G[t]}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho:\Gamma$.

1. We have $\vdash_q t\rho : F\rho$ and $\vdash_q u\rho : G[t]\rho$. By substitution we get $\vdash_q ((t,u).1)\rho = t\rho : F\rho$. But $((t,u).1)\rho \Rightarrow_q t\rho$, thus $q \vdash ((t,u).1)\rho \Rightarrow_q t\rho : F\rho$. We have that $q \Vdash t\rho : F\rho$. Thus by Lemma 3.2 $q \Vdash ((t,u).1)\rho : F\rho$ and $q \Vdash ((t,u).1)\rho = t\rho : F\rho$.

2. Similarly we have $q \Vdash (t,u)\rho.2 \Rightarrow u\rho : G[t\rho]$. Since $q \Vdash u\rho : G[t]\rho$, by Lemma 3.2, we have that $q \Vdash ((t,u).2)\rho : G[t]\rho$ and $q \Vdash ((t,u).2)\rho = u\rho : G[t\rho]$. ☐

Lemma 3.18.

1. $$\frac{\Gamma \vDash_p t:\Sigma(x:F)G}{\Gamma \vDash_p t.1:F} \qquad \frac{\Gamma \vDash_p t:\Sigma(x:F)G}{\Gamma \vDash_p t.2:G[t.1]}$$

2. $$\frac{\Gamma \vDash_p t = u:\Sigma(x:F)G}{\Gamma \vDash_p t.1 = u.1:F} \qquad \frac{\Gamma \vDash_p t = u:\Sigma(x:F)G}{\Gamma \vDash_p t.2 = u.2:G[t.1]}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho:\Gamma$.

1. We have $q \Vdash t\rho : (\Sigma(x : F)G)\rho$. By the definition we have $q \Vdash (t.1)\rho : F\rho$ and $q \Vdash (t.2)\rho : G[t.1]\rho$. Let $q \Vdash \rho = \sigma : \Gamma$. We have that $q \Vdash t\rho = t\sigma : (\Sigma(x : F)G)\rho$. By the definition $q \Vdash (t.1)\rho = (t.1)\sigma : F\rho$ and $q \Vdash (t.2)\rho = (t.2)\sigma : G[t.1]\rho$.

2. We have $q \Vdash t\rho = u\rho : (\Sigma(x : F)G))\rho$. By the definition $q \Vdash (t.1)\rho = (u.1)\rho : F\rho$ and $q \Vdash (t.2)\rho = (u.2)\rho : G[t.1]\rho$. □

Lemma 3.19.

$$\frac{\begin{array}{l}\Gamma \vDash_p t : \Sigma(x : F)G \\ \Gamma \vDash_p u : \Sigma(x : F)G \quad \Gamma \vDash_p t.1 = u.1 : F \quad \Gamma \vDash_p t.2 = u.2 : G[t.1]\end{array}}{\Gamma \vDash_p t = u : \Sigma(x : F)G}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. We have $q \Vdash t\rho : (\Sigma(x : F)G)\rho$ and $q \Vdash u\rho : (\Sigma(x : F)G)\rho$. We also have $q \Vdash (t.1)\rho = (u.1)\rho : F\rho$ and $q \Vdash (t.2)\rho = (u.2)\rho : G[t.1]\rho$. By the definition $q \Vdash t\rho = u\rho : (\Sigma(x : F)G)\rho$. □

Lemma 3.20. *1.* $\dfrac{\Gamma \vdash_p}{\Gamma \vDash_p N}$ 2. $\dfrac{\Gamma \vdash_p}{\Gamma \vDash_p 0 : N}$ 3. $\dfrac{\Gamma \vDash_p n : N}{\Gamma \vDash_p \mathsf{S}\, n : N}$

Proof. 1 and 2 follow directly from the definition while 3 follows from Lemma 2.23. □

Lemma 3.21.

$$\frac{\Gamma, x : N \vDash_p F \quad \Gamma \vDash_p a_0 : F[0] \quad \Gamma \vDash_p g : \Pi(x : N)(F[x] \to F[\mathsf{S}\, x])}{\Gamma \vDash_p \mathsf{natrec}\ (\lambda x.F)\, a_0\, g : \Pi(x : N)F}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. We have then that $q \vdash \rho : \Gamma$, hence, $\vdash_q (\mathsf{natrec}\ (\lambda x.F)\rho\, a_0\, g)\rho : (\Pi(x : N)F)\rho$. Let $r \leqslant q$. Let $r \Vdash a : N$, $r \Vdash b : N$ and $r \Vdash a = b : N$. By Lemma 2.23 there is a partition $r \lhd r_1, \ldots, r_m$ such that for each i, $r_i \vdash a \Rightarrow^* \overline{n}_i : N$ and $r_i \vdash b \Rightarrow^* \overline{n}_i : N$ for some $n_i \in \mathbb{N}$. In order to show that $q \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho : (\Pi(x : N)F)\rho$ we need to show that $r \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, a : F\rho[a]$, $r \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, b : F\rho[b]$, and $r \Vdash \mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, a = (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, b : F\rho[a]$. By local character it will however be sufficient to show that for each i we have $r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, a : F\rho[a]$, $r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, b : F\rho[b]$, and $r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, a = (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, b\rho : F\rho[a]$.

We have that

$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, a \Rightarrow^* (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, \overline{n}_i : F\rho[a]$$
$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, b \Rightarrow^* (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, \overline{n}_i : F\rho[b]$$

Let $\bar{n}_i := \mathsf{S}^{k_i}0$. By induction on k_i. If $k_i = 0$ then

$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ a \Rightarrow^* (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ 0 \Rightarrow a_0\rho : F\rho[a]$$
$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ b \Rightarrow^* (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ 0 \Rightarrow a_0\rho : F\rho[b]$$

By Lemma 2.24 we have then that

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ a = a_0\rho : F\rho[a]$$
$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ b = a_0\rho : F\rho[b]$$

Since $r_i \Vdash a = b : N$ we have $r_i \Vdash (\rho, a) = (\rho, b) : (\Gamma, x : N)$ and thus $r_i \Vdash F\rho[a] = F\rho[b]$. By Lemma 2.24, symmetry and transitivity we have

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ a = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ b : F\rho[a]$$

Assume the statement holds for $k_i \leq \ell$. Let $\bar{n}_i = \mathsf{S}\,\bar{\ell}$. We have then

$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ a \Rightarrow^* g\rho\,\bar{\ell}\,((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\,\bar{\ell}) : F\rho[\mathsf{S}\,\bar{\ell}]$$
$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ b \Rightarrow^* g\rho\,\bar{\ell}\,((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\,\bar{\ell}) : F\rho[\mathsf{S}\,\bar{\ell}]$$

By the induction hypothesis $r_i \Vdash ((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\,\bar{\ell}) : F\rho[\bar{\ell}]$. But $\Gamma \vDash_p g : \Pi(x:N)(F[x] \to F[\mathsf{S}\,x])$ and thus

$$r_i \Vdash g\rho(\bar{\ell})\,((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\,\bar{\ell}) : F\rho[\mathsf{S}\,\bar{\ell}]$$

By Corollary 3.3, symmetry and transitivity we get that

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ a : F\rho[\mathsf{S}\,\bar{\ell}]$$
$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ b : F\rho[\mathsf{S}\,\bar{\ell}]$$
$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ a = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ b : F\rho[\mathsf{S}\,\bar{\ell}]$$

But $r_i \Vdash a = \mathsf{S}\,\bar{\ell} : N$, thus, $r_i \Vdash F\rho[a] = F\rho[\mathsf{S}\,\bar{\ell}]$. By Lemma 2.24 we get then that

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ a = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\ b : F\rho[a]$$

As indicated above, this is sufficient to show $q \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho : \Pi(x : N)F\rho$. Similarly we can show $q \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma : \Pi(x :$

$N)F\sigma$. To show that $q \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma : \Pi(x:N)F\rho$ we need to show that whenever $r \Vdash a : F\rho$ for some $r \leqslant q$ we have $r \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\, a = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\, a : F\rho[a]$. Let $r \Vdash a : F$ for $r \leqslant q$. By Lemma 2.23 we have a partition $r \lhd r_1, \ldots, r_m$ where $r_i \vdash a \Rightarrow^* \bar{n}_i : N$. Similarly to the above it will be sufficient to show $r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\, a = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\, a : F\rho[a]$ for each i.

Let $\bar{n}_i := \mathsf{S}^{k_i}\, 0$. By induction on k_i. If $k_i = 0$ then as above

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\, a = a_0\rho : F\rho[a]$$
$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\, a = a_0\sigma : F\sigma[b]$$

Since $r \Vdash \rho = \sigma : \Gamma$ we have $r_i \Vdash (\rho, a) = (\sigma, a) : (\Gamma, x : N)$. We have then that $r_i \Vdash F\rho[a] = F\sigma[a]$. But we also have that $r_i \Vdash a_0\rho = a_0\sigma : F\rho[a]$. By Lemma 2.24, symmetry and transitivity it then follows that $r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\, a = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\, a : F\rho[a]$.

Assume the statement holds for $k_i \leq \ell$. Let $\bar{n}_i = \mathsf{S}\,\bar{\ell}$. As before we have that

$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\, a \Rightarrow^* g\rho\,\bar{\ell}\,((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\,\bar{\ell}) : F\rho[\mathsf{S}\,\bar{\ell}]$$
$$r_i \vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\, a \Rightarrow^* g\sigma\,\bar{\ell}\,((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\,\bar{\ell}) : F\sigma[\mathsf{S}\,\bar{\ell}]$$

By the induction hypothesis

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\,\bar{\ell} = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\,\bar{\ell} : F\rho[\bar{\ell}]$$

But $r_i \Vdash g\rho = g\sigma : (\Pi(x:N)(F[x] \to F[\mathsf{S}\,x]))\rho$, thus

$$r_i \Vdash g\rho\,\bar{\ell}\,((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\,\bar{\ell}) = g\sigma\,\bar{\ell}\,((\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\,\bar{\ell}) : F\rho[\mathsf{S}\,\bar{\ell}]$$

But $r_i \Vdash F\rho[\mathsf{S}\,\bar{\ell}] = F\sigma[\mathsf{S}\,\bar{\ell}]$ and $r_i \Vdash F\rho[a] = F\rho[\mathsf{S}\,\bar{\ell}]$. By Lemma 2.24, symmetry and transitivity we have then that

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho\, a = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma\, a : F\rho[a]$$

Which is sufficient to show

$$q \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\rho = (\mathsf{natrec}\ (\lambda x.F)\ a_0\ g)\sigma : \Pi(x:N)F\rho$$

□

Lemma 3.22.

$$\frac{\Gamma, x:N \vdash F \quad \Gamma \models_p a_0 : F[0] \quad \Gamma \models_p g : \Pi(x:N)(F[x] \to F[\mathsf{S}\,x])}{\Gamma \models_p \mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, 0 = a_0 : F[0]}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. We have $\vdash_q \rho : \Gamma$ and thus we get that $\vdash_q (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, 0)\rho = a_0\rho : F\rho[0]$. But $(\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, 0)\rho \Rightarrow a_0\rho$. Thus $q \vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, 0)\rho \Rightarrow a_0\rho : F\rho[0]$. But $q \Vdash a_0\rho : F\rho[0]$. By Lemma 3.2 we have $q \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, 0)\rho = a_0\rho : F\rho[0]$. □

Lemma 3.23.

$$\frac{\Gamma, x:N \models_p F \quad \Gamma \models_p a_0 : F[0] \quad \Gamma \models_p n : N \quad \Gamma \models_p g : \Pi(x:N)(F[x] \to F[\mathsf{S}\,x])}{\Gamma \models_p \mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, (\mathsf{S}\,n) = g\, n\, (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, n) : F[\mathsf{S}\,n]}$$

Proof. Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. We have $q \Vdash n\rho : N$. By Lemma 2.23 there is a partition $q \lhd q_1, \ldots, q_m$ where $q_i \vdash n\rho \Rightarrow^* \bar{n}_i : N$. Thus $q_i \vdash \mathsf{S}\, n\rho \Rightarrow^* \mathsf{S}\, \bar{n}_i : N$. We have then that

$$\begin{aligned} q_i \vdash & (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, (\mathsf{S}\,n))\rho \Rightarrow^* (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, (\mathsf{S}\,\bar{n}_i) \\ & \Rightarrow^* g\rho\, \bar{n}_i\, ((\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, \bar{n}_i) : F\rho[\mathsf{S}\,\bar{n}_i] \end{aligned}$$

But

$$\begin{aligned} q_i \vdash & (g\, n\, (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, n))\rho \\ & \Rightarrow^* g\rho\, \bar{n}_i\, ((\mathsf{natrec}\ (\lambda x.F)\, a_0\, g)\rho\, \bar{n}_i) : F\rho[\mathsf{S}\,\bar{n}_i] \end{aligned}$$

By Corollary 3.3, symmetry and transitivity we have

$$q_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, (\mathsf{S}\,n))\rho = (g\, n\, (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, n))\rho : F\rho[\mathsf{S}\,\bar{n}_i]$$

Since $q_i \Vdash n\rho = \bar{n}_i : N$ we have that $q_i \Vdash F\rho[\mathsf{S}\,\bar{n}_i] = F\rho[\mathsf{S}\,n\rho]$. By Lemma 2.24 we then have

$$q_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, (\mathsf{S}\,n))\rho = (g\, n\, (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, n))\rho : F[\mathsf{S}\,n]\rho$$

By local character we have

$$q \Vdash (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, (\mathsf{S}\,n))\rho = (g\, n\, (\mathsf{natrec}\ (\lambda x.F)\, a_0\, g\, n))\rho : F[\mathsf{S}\,n]\rho$$

□

Lemma 3.24.

$$\frac{\Gamma, x{:}N \vDash_p F = G \quad \Gamma \vDash_p a_0{:}F[0] \quad \Gamma \vDash_p g{:}\Pi(x{:}N)(F[x] \to F[\mathsf{S}\,x])}{\Gamma \vDash_p \mathsf{natrec}\ (\lambda x.F)\ a_0\, g = \mathsf{natrec}\ (\lambda x.G)\ a_0\, g{:}\Pi(x{:}N)F}$$

Proof. Since the proof is very similar to that of Lemma 3.21 we only sketch the idea below.

Let $q \leqslant p$ and $q \Vdash \rho : \Gamma$. Let $r \leqslant q$ and $r \Vdash n : N$. By Lemma 2.23 there is a partition $r \lhd r_1, \ldots, r_m$ such that for each i, $r_i \vdash n \Rightarrow^* \overline{n}_i : N$. But then we have $r_i \Vdash (\rho, \overline{n}_i) : (\Gamma, x : N)$ and thus $r_i \Vdash F\rho[\overline{n}_i] = G\rho[\overline{n}_i]$. From this observation, in a manner similar to that used in the proof of Lemma 3.21, one can show by induction on $\overline{n}_i$ that

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\, g)\rho\, \overline{n}_i = \mathsf{natrec}\ (\lambda x.G)\ a_0\, g\, \overline{n}_i{:}F\rho[\overline{n}_i]$$

Having shown the above forcing for each i, it will then follow from Corollary 3.3 and Lemma 2.24 that

$$r_i \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\, g)\rho\, n = (\mathsf{natrec}\ (\lambda x.G)\ a_0\, g)\rho\, n{:}F\rho[n]$$

By local character it will then follow that

$$r \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\, g)\rho\, n = (\mathsf{natrec}\ (\lambda x.G)\ a_0\, g)\rho\, n{:}F\rho[n]$$

Thus proving

$$q \Vdash (\mathsf{natrec}\ (\lambda x.F)\ a_0\, g)\rho = (\mathsf{natrec}\ (\lambda x.G)\ a_0\, g)\rho{:}(\Pi(x{:}N)F)\rho$$

□

Lemma 3.25. $\dfrac{\Gamma \vdash_p}{\Gamma \vDash_p N_2} \quad \dfrac{\Gamma \vdash_p}{\Gamma \vDash_p 0{:}N_2} \quad \dfrac{\Gamma \vdash_p}{\Gamma \vDash_p 1{:}N_2}$

$$\frac{\Gamma, x{:}N_2 \vDash_p F \quad \Gamma \vDash_p a_0{:}F[0] \quad \Gamma \vDash_p a_1{:}F[1]}{\Gamma \vDash_p \mathsf{boolrec}\ (\lambda x.F)\ a_0\, a_1{:}\Pi(x{:}N_2)F}$$

$$\frac{\Gamma, x{:}N_2 \vDash_p F \quad \Gamma \vDash_p a_0{:}F[0] \quad \Gamma \vDash_p a_1{:}F[1]}{\Gamma \vDash_p \mathsf{boolrec}\ (\lambda x.F)\ a_0\, a_1\, 0 = a_0{:}F[0]}$$

$$\frac{\Gamma, x{:}N_2 \vdash_p F \quad \Gamma \vDash_p a_0{:}F[0] \quad \Gamma \vDash_p a_1{:}F[1]}{\Gamma \vDash_p \mathsf{boolrec}\ (\lambda x.F)\ a_0\, a_1\, 1 = a_1{:}F[1]}$$

$$\frac{\Gamma, x{:}N_2 \vDash_p F = G \quad \Gamma \vDash_p a_0{:}F[0] \quad \Gamma \vDash_p a_1{:}F[1]}{\Gamma \vDash_p \mathsf{boolrec}\ (\lambda x.F)\ a_0\, a_1 = \mathsf{boolrec}\ (\lambda x.G)\ a_0\, a_1{:}\Pi(x{:}N_2)F}$$

Proof. Follows similarly to Lemma 3.20, Lemma 3.21, Lemma 3.22, Lemma 3.23, and Lemma 3.24. □

Lemma 3.26. $\dfrac{\Gamma \vdash_p}{\Gamma \vDash_p N_1} \quad \dfrac{\Gamma \vdash_p}{\Gamma \vDash_p 0 : N_1}$

$$\dfrac{\Gamma, x : N_1 \vDash_p F \quad \Gamma \vDash_p a : F[0]}{\Gamma \vDash_p \mathsf{unitrec}\ (\lambda x.F)\ a : \Pi(x : N_1)F} \quad \dfrac{\Gamma, x : N_1 \vDash F \quad \Gamma \vDash_p a : F[0]}{\Gamma \vDash_p \mathsf{unitrec}\ (\lambda x.F)\ a\ 0 = a : F[0]}$$

$$\dfrac{\Gamma, x : N_1 \vDash_p F = G \quad \Gamma \vDash_p a : F[0]}{\Gamma \vDash_p \mathsf{unitrec}\ (\lambda x.F)\ a = \mathsf{unitrec}\ (\lambda x.G)\ a : \Pi(x : N_1)F}$$

Proof. Follows similarly to Lemma 3.20, Lemma 3.21, Lemma 3.22, and Lemma 3.24. □

Lemma 3.27. $\dfrac{\Gamma \vdash_p}{\Gamma \vDash_p N_0} \quad \dfrac{\Gamma, x : N_0 \vDash_p F}{\Gamma \vDash_p \bot\mathsf{rec}\ (\lambda x.F) : \Pi(x : N_0)F}$

$$\dfrac{\Gamma, x : N_0 \vDash_p F = G}{\Gamma \vDash_p \bot\mathsf{rec}\ (\lambda x.F) = \bot\mathsf{rec}\ (\lambda x.G) : \Pi(x : N_0)F}$$

Proof. Follows similarly to Lemma 3.20, Lemma 3.21, and Lemma 3.24. □

Lemma 3.28.

$$\dfrac{\Gamma \vdash_p}{\Gamma \vDash_p U} \quad \dfrac{\Gamma \vDash_p F : U}{\Gamma \vDash_p F} \quad \dfrac{\Gamma \vDash_p F = G : U}{\Gamma \vDash_p F = G} \quad \dfrac{\Gamma \vdash_p}{\Gamma \vDash_p N : U} \quad \dfrac{\Gamma \vdash_p}{\Gamma \vDash_p N_2 : U}$$

$$\dfrac{\Gamma \vDash_p F : U \quad \Gamma, x : F \vDash_p G : U}{\Gamma \vDash_p \Pi(x : F)G : U} \quad \dfrac{\Gamma \vDash_p F = H : U \quad \Gamma, x : F \vDash_p G = E : U}{\Gamma \vDash_p \Pi(x : F)G = \Pi(x : H)E : U}$$

$$\dfrac{\Gamma \vDash_p F : U \quad \Gamma, x : F \vDash_p G : U}{\Gamma \vDash_p \Sigma(x : F)G : U} \quad \dfrac{\Gamma \vDash_p F = H : U \quad \Gamma, x : F \vDash_p G = E : U}{\Gamma \vDash_p \Sigma(x : F)G = \Sigma(x : H)E : U}$$

Proof. Follows similarly to the above. □

Lemma 3.29.

$$\dfrac{\Gamma \vDash_p t : F \quad \Gamma \vDash_p F = G}{\Gamma \vDash_p t : G} \quad \dfrac{\Gamma \vDash_p t = u : F \quad \Gamma \vDash_p F = G}{\Gamma \vDash_p t = u : G}$$

$$\dfrac{\Gamma \vDash_p F}{\Gamma \vDash_p F = F} \quad \dfrac{\Gamma \vDash_p F = G}{\Gamma \vDash_p G = F} \quad \dfrac{\Gamma \vDash_p F = G \quad \Gamma \vDash_p G = H}{\Gamma \vDash_p F = H}$$

$$\frac{\Gamma \vDash_p t : F}{\Gamma \vDash_p t = t : F} \qquad \frac{\Gamma \vDash_p t = u : F}{\Gamma \vDash_p u = t : F} \qquad \frac{\Gamma \vDash_p t = u : F \quad \Gamma \vDash_p u = v : F}{\Gamma \vDash_p t = v : F}$$

Proof. Follows from Lemma 2.24, Lemma 2.25 and Lemma 2.26. □

We have then the following corollary.

Corollary 3.30 (Soundness). *If $\Gamma \vdash_p J$ then $\Gamma \vDash_p J$*

Theorem 3.31 (Fundamental Theorem). *If $\vdash_p J$ then $p \Vdash J$.*

Proof. Follows from Corollary 3.30. □

4 Markov's principle

Now we have enough machinery to show the independence of MP from type theory. The idea is that if a judgment J is derivable in type theory (i.e. $\vdash J$) then it is derivable in the forcing extension (i.e. $\vdash_{\langle\rangle} J$) and by Theorem 3.31 it holds in the interpretation (i.e. $\Vdash J$). It thus suffices to show that there no t such that $\Vdash t : \mathrm{MP}$ to establish the independence of MP from type theory. First we recall the formulation of MP.

$$\mathrm{MP} := \Pi(h : N \to N_2)[\neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(h\,x)) \to \Sigma(x : N)\,\mathsf{IsZero}\,(h\,x)]$$

where $\mathsf{IsZero} : N_2 \to U$ is given by $\mathsf{IsZero} := \lambda y.\mathsf{boolrec}\ (\lambda x.U)\ N_1\ N_0\ y$.

Lemma 4.1. *There is no term t such that $\Vdash t : \Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x)$.*

Proof. Assume $\Vdash t : \Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x)$ for some t. We then have $\Vdash t.1 : N$ and $\Vdash t.2 : \mathsf{IsZero}\,(\mathsf{f}\,t.1)$. By Lemma 2.23 one has a partition $\langle\rangle \lhd p_1, \ldots, p_n$ where for each i, $p_i \vdash t.1 \Rightarrow^* \overline{m}_i$ for some $\overline{m}_i \in \mathbb{N}$. Hence $p_i \vdash \mathsf{IsZero}\,(\mathsf{f}\,t.1) \Rightarrow^* \mathsf{IsZero}\,(\mathsf{f}\,\overline{m}_i)$ and by Lemma 3.1 $p_i \Vdash \mathsf{IsZero}\,(\mathsf{f}\,t.1) = \mathsf{IsZero}\,(\mathsf{f}\,\overline{m}_i)$. But, by definition, a partition of $\langle\rangle$ must contain a condition, say p_j, such that $p_j(k) = 1$ whenever $k \in \mathrm{dom}(p_j)$ (this holds vacuously for $\langle\rangle \lhd \langle\rangle$). Assume $m_j \in \mathrm{dom}(p_j)$, then $p_j \vdash \mathsf{IsZero}\,(\mathsf{f}\,t.1) \Rightarrow^* \mathsf{IsZero}\,(\mathsf{f}\,m_j) \Rightarrow^* N_0$. By monotonicity, from $\Vdash t.2 : \mathsf{IsZero}\,(\mathsf{f}\,t.1)$ we get $p_j \Vdash t.2 : \mathsf{IsZero}\,(\mathsf{f}\,t.1)$. But $p_j \vdash \mathsf{IsZero}\,(\mathsf{f}\,t.1) \Rightarrow^* N_0$ thus $p_j \Vdash \mathsf{IsZero}\,(\mathsf{f}\,t.1) = N_0$. Hence, by Lemma 2.24, $p_j \Vdash t.2 : N_0$ which is impossible, thus contradicting our assumption. If on the other hand $m_j \notin \mathrm{dom}(p_j)$ then since $p_j \lhd p_j(m_j, 0), p_j(m_j, 1)$ we can apply the above argument with $p_j(m_j, 1)$ instead of p_j. □

Lemma 4.2. *There is no term t such that $\Vdash t : \mathrm{MP}$.*

Proof. Assume $\Vdash t : \mathrm{MP}$ for some t. From the definition, whenever $\Vdash g : N \to N_2$ we have $\Vdash t\,g : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(g\,x)) \to \Sigma(x : N)\,\mathsf{IsZero}\,(g\,x)$. Since by Corollary 3.4, $\Vdash \mathsf{f} : N \to N_2$ we have $\Vdash t\,\mathsf{f} : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x)) \to \Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x)$. Since by Lemma 3.6 $\Vdash \mathsf{w} : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x))$ we have $\Vdash (t\,\mathsf{f})\,\mathsf{w} : \Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}\,x)$ which is impossible by Lemma 4.1. □

From Theorem 3.31, Lemma 4.2, and Lemma 1.3 we can then conclude:

Theorem 1.1. *There is no term t such that* MLTT $\vdash t : \mathrm{MP}$.

4.1 Many Cohen reals

We extend the type system in Section 1 further by adding a generic point f_q for each condition q. The introduction and conversion rules for f_q are given by:

$$\frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{f}_q : N \to N_2} \qquad \frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{f}_q\,\overline{n} = 1}\; n \in \mathrm{dom}(q)$$

$$\frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{f}_q\,\overline{n} = p(n)}\; n \notin \mathrm{dom}(q), n \in \mathrm{dom}(p)$$

With the reduction rules:

$$\frac{n \in \mathrm{dom}(q)}{\mathsf{f}_q\,\overline{n} \to 1} \qquad \frac{n \notin \mathrm{dom}(q), n \in \mathrm{dom}(p)}{\mathsf{f}_q\,\overline{n} \to_p p(n)}$$

We observe that with these added rules the reduction relation is still monotone.

For each f_q we add a term:

$$\frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{w}_q : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}_q\,x))}$$

Finally we add a term mw witnessing the negation of MP

$$\frac{\Gamma \vdash_p}{\Gamma \vdash_p \mathsf{mw} : \neg\mathrm{MP}}$$

By analogy to Corollary 3.4 we have

Lemma 4.3. $\Vdash \mathsf{f}_q : N \to N_2$ *for all* q.

Lemma 4.4. $\Vdash \mathsf{w}_q : \neg\neg(\Sigma(x:N)\,\mathsf{IsZero}\,(\mathsf{f}_q\,x))$ *for all* q.

Proof. Assume $p \Vdash t : \neg(\Sigma(x:N)\mathsf{IsZero}\,(\mathsf{f}_q\,x))$ for some p and t. Let $m \notin \mathrm{dom}(q) \cup \mathrm{dom}(p)$, we have $p(\overline{m},0) \Vdash \mathsf{f}_q\,m = 0$. Thus $p(\overline{m},0) \Vdash (\overline{m},0) : \Sigma(x:N)\,\mathsf{IsZero}\,(\mathsf{f}_q\,x)$ and $p(\overline{m},0) \Vdash t\,(\overline{m},0) : N_0$ which is impossible. □

Lemma 4.5. *There is no term* t *for which* $q \Vdash t : \Sigma(x:N)\,\mathsf{IsZero}\,(\mathsf{f}_q\,x)$.

Proof. Assume $q \Vdash t : \Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}_q\,x)$ for some t. We then have $q \Vdash t.1 : N$ and $q \Vdash t.2 : \mathsf{IsZero}\,(\mathsf{f}_q\,t.1)$. By Lemma 2.23 one has a partition $q \lhd q_1, \ldots, q_n$ where for each i, $t.1 \Rightarrow^*_{q_i} \overline{m}_i$ for some $\overline{m}_i \in \mathbb{N}$. Hence $q_i \vdash \mathsf{IsZero}\,(\mathsf{f}_q\,t.1) \Rightarrow^* \mathsf{IsZero}\,(\mathsf{f}_q\,\overline{m}_i)$. But any partition of q contain a condition, say q_j, where $q_j(k) = 1$ whenever $k \notin \mathrm{dom}(q)$ and $k \in \mathrm{dom}(q_j)$. Assume $m_j \in \mathrm{dom}(q_j)$. If $m_j \in \mathrm{dom}(q)$ then $q_j \vdash \mathsf{f}_q\,m_j \Rightarrow 1 : N_2$ and if $m_j \notin \mathrm{dom}(q)$ then $q_j \vdash \mathsf{f}_q\,\overline{m}_j \Rightarrow q_j(k) := 1 : N_2$. Thus $q_j \vdash \mathsf{IsZero}\,(\mathsf{f}_q\,t.1) \Rightarrow^* N_0$ and by Lemma 3.1 $q_j \Vdash \mathsf{IsZero}\,(\mathsf{f}\,t.1) = N_0$. From $\Vdash t.2 : \mathsf{IsZero}\,(\mathsf{f}\,t.1)$ by monotonicity and Lemma 2.24 we have $q_j \Vdash t.2 : N_0$ which is impossible. If on the other hand $m_j \notin \mathrm{dom}(q_j)$ then since $q_j \lhd q_j(m_j,0), q_j(m_j,1)$ we can apply the above argument with $q_j(m_j,1)$ instead of q_j. □

Lemma 4.6. $\Vdash \mathsf{mw} : \neg\mathrm{MP}$

Proof. Assume $p \Vdash t : \mathrm{MP}$ for some p and t. Thus whenever $q \leqslant p$ and $q \Vdash u : N \to N_2$ then $q \Vdash t\,u : \neg\neg(\Sigma(x : N)\,\mathsf{IsZero}\,(u\,x)) \to (\Sigma(x : N)\,\mathsf{IsZero}\,(u\,x))$. But we have $q \Vdash \mathsf{f}_q : N \to N_2$ by Lemma 4.3. Hence $q \Vdash t\,\mathsf{f}_q : \neg\neg(\Sigma(x : N)\mathsf{IsZero}\,(\mathsf{f}_q\,x)) \to (\Sigma(x : N)\mathsf{IsZero}\,(\mathsf{f}_q\,x))$. But $q \Vdash \mathsf{w}_q : \neg\neg(\Sigma(x : N)\mathsf{IsZero}\,(\mathsf{f}_q\,x))$ by Lemma 4.4. Thus $q \Vdash (t\,\mathsf{f}_q)\,\mathsf{w}_q : \Sigma(x : N)\,\mathsf{IsZero}\,(\mathsf{f}_q\,x)$ which is impossible by Lemma 4.5. □

We have then that this extension is sound with respect to the interpretation. Hence we have shown the following statement.

Theorem 4.7. *There is a consistent extension of* MLTT *where* ¬MP *is derivable.*

Conclusion and Future Work: Interpretation of the Universe in Sheaf Models

In this chapter we discuss, rather informally, the problem of the interpretation of the universe in sheaf models of type theory.

1 The universe in type theory

The type theory we are considering here is similar to the one presented in Section V.1 except that we have a smallness judgment $\Gamma \vdash A$ small along with the rule

$$\frac{\Gamma \vdash A \text{ small}}{\Gamma \vdash A}$$

and change the rules for the universe to:

$$\frac{}{U \vdash \mathsf{El} \text{ small}} \qquad \frac{\Gamma \vdash a : U}{\Gamma \vdash \mathsf{El}[a] \text{ small}}$$

$$\frac{\Gamma \vdash A \text{ small}}{\Gamma \vdash \ulcorner A \urcorner : U} \qquad \frac{\Gamma \vdash A}{\Gamma \vdash \mathsf{El}[\ulcorner A \urcorner] = A} \qquad \frac{\Gamma \vdash a : U}{\Gamma \vdash \ulcorner \mathsf{El}[a] \urcorner = a : U}$$

Where $[a] = (\mathrm{id}, a) : \Gamma \to \Gamma.U$.

We start by recalling briefly the definition of a presheaf model of type theory by Hofmann [1997], Coquand [2013] and others.

For readability we use the same symbols for types, contexts, terms and their interpretation.

1.1 Presheaf models of type theory

Let $\mathcal{C}$ be a small category. A context Γ is interpreted by a presheaf over $\mathcal{C}$. A substitution is interpreted by a presheaf map $\sigma : \Delta \to \Gamma$.

A type $\Gamma \vdash A$ is interpreted by a presheaf over the category of elements of Γ. That is to say A is a family of sets $A(\rho)$ for each X and $\rho \in \Gamma(X)$ with restrictions $A(\rho) \ni u \mapsto uf \in A(\rho f)$ such that $(uf)g = u(fg)$. Let $\sigma : \Delta \to \Gamma$, the type $\Delta \vdash A\sigma$ is then given by $(A\sigma)(\delta) = A(\sigma\delta)$ for any X and $\delta \in \Delta(X)$.

A comprehension $\Gamma.A$ is the presheaf defined by $\Gamma.A(X) = \{(\rho, d) \mid u \in A(\rho), \rho \in \Gamma(X)\}$ with restriction $(\rho, u) \mapsto (\rho f, uf)$ along $f : Y \to X$. We then have a projection $\Gamma.A \to \Gamma$ given by $(\rho, u) \mapsto \rho$.

A term $\Gamma \vdash a : A$ is interpreted by a section of the projection $\Gamma.A \to \Gamma$. That is to say a term is a family of elements $a(\rho)$ where $a(\rho) \in A(\rho)$ for each X and $\rho \in \Gamma(X)$ such that $a(\rho)f = a(\rho f)$. Let $\sigma : \Delta \to \Gamma$ and $\delta \in \Delta(X)$ the term $\Delta \vdash a\sigma : A\sigma$ is given by $(a\sigma)\delta = a(\sigma\delta)$.

Given $\Gamma \vdash A$ and $\Gamma.A \vdash B$ we interpret type formers as follows:

For a type $\Gamma \vdash \Pi(x : A)B$ an element of $(\Pi(x : A)B)(\rho)$ is given by a family ω indexed by arrows $Y \to X$ where $\omega_h \in \prod_{a \in A\rho h} B(\rho h, a)$ for $h : Y \to X$, $\rho \in \Gamma(X)$. The restriction ωf along $f : Y \to X$ is given by $(\omega f)_g = \omega_{fg}$ for $g : Z \to Y$.

For $\Gamma \vdash A$ we say that A is small if $A(\rho)$ is small for all X and $\rho \in \Gamma(X)$. An element of $U(X)$ is a presheaf over (the Yoneda of) X, i.e. a family $(A(h))_{h:W \to X}$ with restrictions Af along $f : Y \to X$ given by $Af(g) = A(fg)$ for $g : Z \to Y$.

For $\Gamma \vdash A$ small we define $\ulcorner A \urcorner(\rho)$ to be the presheaf over X given by $(\ulcorner A \urcorner(\rho))(f) = A(\rho f)$ for $f : Y \to X$ and $\rho \in \Gamma(X)$.

The classifying (display/decoding) family El is a small presheaf over U given by $\mathsf{El}(A) = A(1_X)$ for $A \in U(X)$.

For $\Gamma \vdash A$ small we have then $\Gamma \vdash \ulcorner A \urcorner : U$ and $\Gamma \vdash \mathsf{El}[\ulcorner A \urcorner] = A$. Moreover we have $\Gamma \vdash \ulcorner \mathsf{El}[a] \urcorner = a : U$ for $\Gamma \vdash a : U$. This is the commonly called Hofmann-Streicher interpretation of the universe [Hofmann and Streicher, 199?].

One can then verify that type theory is sound for this model, see for example [Bezem et al., 2014].

1.2 Sheaf models of type theory

Let $(\mathcal{C}, \mathbf{J})$ be a site. A natural starting point for finding a sheaf interpretation would be to replicate the above interpretation only replacing presheaves with sheaves. That is we interpret a context Γ by a sheaf over $(\mathcal{C}, \mathbf{J})$ and a substitution by a sheaf morphism. The topology $\mathbf{J}$ induces a topology $\mathbf{J}_\Gamma$ on the category $\mathcal{C}_\Gamma$ (the category of elements of Γ) given by covering families $\{f_i : (\rho f_i, X_i) \to (\rho, X)\}_{i \in I} \in \mathbf{J}_\Gamma(\rho, X)$ whenever $\{f_i : X_i \to X\}_{i \in I} \in \mathbf{J}(X)$. A type $\Gamma \vdash A$ is then interpreted as a sheaf over the site $(\mathcal{C}_\Gamma, \mathbf{J}_\Gamma)$. The interpretation of terms, type formers, and context comprehension remain the same. One can then verify that the sheaf property is stable for the interpretation of type formers and context comprehension.

This interpretation fails when it comes to the universe though. If we take $U(X)$ to be the set of sheaves over X with the same restrictions as described before then the U so defined is only a presheaf but not a sheaf. Given two elements A and B in $U(X)$ and a coverage $\{f_i : X_i \to X\}_{i \in I}$ it could very well be that A and B restrict to the same elements on the coverage while they are not identical, that is $A_{f_i} = B_{f_i}$ for all $i \in I$ while $A(1_X) \neq B(1_X)$. The sheaf property does not hold for this interpretation of the universe then.

But what happens if we sheafify U? To answer that we note first that the presheaf U satisfy the *gluing* part of the sheaf axiom; that is, for a coverage $\{f_i : X_i \to X\}_{i \in I}$ whenever there is a matching family $C_i \in U(X_i)$ one has an element $C \in U(X)$ such that C restricts to C_i along f_i. However, U does not satisfy the *uniqueness* part of the sheaf axiom, i.e. the presheaf U is not separated. The sheafification $\widetilde{U}$ of U then amounts to a quotient, that is we take as elements of $\widetilde{U}(X)$ the equivalence classes defined by the equivalence relation $\sim$ on $U(X)$ given by: $A \sim B$ whenever there is a coverage $\{f_i : X \to X\}_{i \in I}$ such that $A_{f_i} = B_{f_i}$ for all i. But then it is not clear how to define the display $\mathsf{El}[[A]_\sim]$, where $[A]_\sim$ is the equivalence class of some $A \in U(X)$.

See [Xu and Escardó, 2016] for more thoughts on the issue of universes in sheaf models of type theory.

2 Stack models of type theory

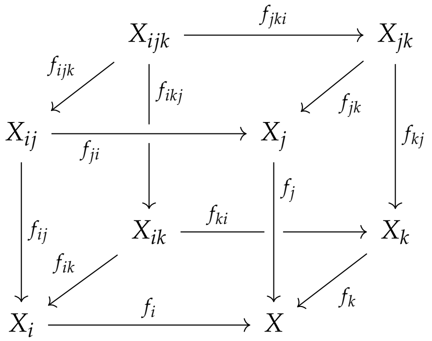
A possible solution to the problem of interpretation of the universe in sheaves is to interpret type theory in stacks. One can think of stacks as higher order sheaves of groupoids.[3]

Let $(\mathcal{C}, \mathbf{J})$ be a small site where $\mathcal{C}$ has pullbacks.

Let **Gpd** be the 1-category of groupoids and groupoid morphisms.

[3]This section outlines a work in progress in collaboration with Thierry Coquand and Fabian Ruch.

Consider a coverage $\{f_i : X_i \to X\}_{i \in I} \in \mathbf{J}(X)$ and the pullback cube



Assume we are given a functor $F : \mathcal{C}^{op} \to \mathbf{Gpd}$.

Given an object X of $\mathcal{C}$ let $\mathcal{C}_X$ be the slice of $\mathcal{C}$ over X and $\mathbf{J}_X$ the comma topology. Let $\mathrm{Hom}_{/X}(x, y) \in \mathbf{Set}^{\mathcal{C}^{op}}$ be the presheaf given by $\mathrm{Hom}_{/X}(x, y)(Y \xrightarrow{f} X) = \mathrm{Hom}_{F(Y)}(xf, yf)$ and restrictions along $h : g \to f$ given by the action of the groupoid morphism $F(h) : F(Y) \to F(Z)$, where $g : Z \to X$, $f : Y \to X$, $h : Z \to Y$ and $f \circ h = g$.[4]

Given a family of elements $x_i \in F(X_i)$ and a family of isomorphisms $\mu_{ij} : x_i f_{ij} \xrightarrow{\sim} x_j f_{ji}$ such that for each triple indices $i, j, k \in I$ it holds that $\mu_{jk} f_{jki} \circ \mu_{ij} f_{ijk} = \mu_{ik} f_{ikj}$, where $\mu_{jk} f_{jki}$, $\mu_{ij} f_{ijk}$, and $\mu_{ik} f_{ikj}$ are the restrictions of μ_{jk}, μ_{ij}, and μ_{ik} along f_{jki}, f_{ijk}, and f_{ikj}. The family $(\{x_i\}, \{\mu_{ij}\})$ is called descent datum (for the coverage $\{f_i : X_i \to X\}_{i \in I}$).

The functor F is a *prestack* if it satisfies:

1. Given $x, y \in F(X)$ the presheaf $\mathrm{Hom}_{/X}(x, y)$ is a sheaf on the site $(\mathcal{C}_X, \mathbf{J}_X)$.

F is a *stack* if it moreover satisfies

2. Given a descent datum $(\{x_i\}, \{\mu_{ij}\})$ as described above there exist $x \in F(X)$ and isomorphisms $\nu_i : x f_i \xrightarrow{\sim} x_i$ for each i such that $\mu_{ij} \circ \nu_i f_{ij} = \nu_j f_{ji}$. We say that $(x, \{\nu_i\})$ is a glue of the descent datum $(\{x_i\}, \{\mu_{ij}\})$.

2.1 Interpretation of type theory in stacks

Assume we have a site $(\mathcal{C}, \mathbf{J})$ as described above. We outline an interpretation of type theory in stacks.

[4] We introduce $\circ$ for composition in this section to avoid confusion with restriction of groupoid morphism.

Contexts, Grothendieck construction and dependent types: A Context Γ is interpreted by a stack over $(\mathcal{C}, \mathbf{J})$. A substitution $\sigma : \Delta \to \Gamma$ is interpreted by a stack morphism, that is, by a family of groupoid morphisms $\sigma_X : \Delta(X) \to \Gamma(X)$ natural in X.

Given a context $\Gamma : \mathcal{C}^{op} \to \mathbf{Gpd}$ let $\mathcal{C}_\Gamma$ be the Grothendieck construction for Γ, that is to say, the objects of $\mathcal{C}_\Gamma$ are pairs (X, γ) where $X \in \mathcal{C}$ and $\gamma \in \Gamma(X)$ and morphisms are pairs $(f, p) : (Y, \alpha) \to (X, \gamma)$ where $f : Y \to X$ is a morphism in $\mathcal{C}$ and $p : \alpha \to \gamma f$ is an isomorphism in the groupoid $\Gamma(Y)$.

The topology $\mathbf{J}$ induces a topology $\mathbf{J}_\Gamma$ on $\mathcal{C}_\Gamma$ given by families $\{(f_i, \mathrm{id}) : (X_i, \gamma f_i) \to (X, \gamma)\}_{i \in I} \in \mathbf{J}_\Gamma((X, \gamma))$ whenever $\{f_i : X_i \to X\}_{i \in I} \in \mathbf{J}(X)$.

A type $\Gamma \vdash A$ is interpreted by a stack $A : \mathcal{C}_\Gamma^{op} \to \mathbf{Gpd}$. Let $\sigma : \Delta \to \Gamma$ be a substitution. As in the presheaf interpretation the substitution $A\sigma$ is interpreted by precomposition with the induced $\sigma : \mathcal{C}_\Delta^{op} \to \mathcal{C}_\Gamma^{op}$.

Context comprehension: Given $\Gamma \vdash A$ the comprehension $\Gamma.A$ is given by the stack where $\Gamma.A(X)$ is the groupoid with elements (γ, a) where $\gamma \in \Gamma(X)$, $a \in A(\gamma)$ and morphisms $(p, \nu) : (\alpha, b) \to (\gamma, a)$ where $p : \alpha \xrightarrow{\sim} \gamma$ in $\Gamma(X)$ and $\nu : b \xrightarrow{\sim} a(\mathrm{id}, p)$ in $A(\alpha)$. For $f : Y \to X$ we have $(\gamma, a)f = (\gamma f, a(f, \mathrm{id}))$ and $(p, \nu)f = (pf, \nu(f, \mathrm{id}))$.

Terms: A term $\Gamma \vdash a : A$ is interpreted by a section of the projection $\mathsf{p} : \Gamma.A \to \Gamma$. That is to say by a family of elements $a(\gamma) \in A(\gamma)$ for each X and $\gamma \in \Gamma(X)$ and a family of arrows $a(q) : a(\gamma) \to a(\beta)(\mathrm{id}, q)$ for each $q : \gamma \xrightarrow{\sim} \beta$ in $\Gamma(X)$. Moreover, if $(f, p) : (Y, \alpha) \to (X, \gamma)$ then $a(q)(f, p) = a(qf)(\mathrm{id}, p)$, $a(\gamma)(f, p) = a(\gamma f)(\mathrm{id}, p)$ and if $r : \beta \xrightarrow{\sim} \delta$ in $\Gamma(X)$ then $a(r \circ q) = a(r)(\mathrm{id}, q) \circ a(q)$.

A substitution $\Delta \vdash a\sigma : A\sigma$ for some $\sigma : \Delta \to \Gamma$ is interpreted by precomposition, i.e. $a\sigma(\delta) = a(\sigma\delta)$ and $a\sigma(q) = a(\sigma q)$, for $\delta \in \Delta(X)$ and $q : \delta \to \delta' \in \Delta(X)$.

One can then define an arrow between terms $\Gamma \vdash a : A$ and $\Gamma \vdash b : A$ as a family of arrows $a(\gamma) \xrightarrow{\sim} b(\gamma)$ for each X and $\gamma \in \Gamma(X)$ natural in γ. One then can show that $\mathrm{Tm}_\Gamma(A)$ is groupoid.

Dependent functions and dependent pairs: Let $\Gamma \vdash A$ and $\Gamma.A \vdash B$.

The type $\Gamma \vdash \Sigma(x : A)B$ is interpreted by the stack where for X and $\gamma \in \Gamma(X)$ the elements of $(\Sigma(x : A)B)(\gamma)$ are pairs (a, b) with $a \in A(\gamma)$ and $b \in B(\gamma, a)$. The morphisms are pairs $(\nu, \varrho) : (a, b) \to (a', b')$ where $\nu : a \to a'$ in $A(\gamma)$ and $\varrho : b \to b'(\mathrm{id}, \nu)$ in $B(\gamma, a)$. For (f, p) :

$(Y, \alpha) \to (X, \gamma)$ we have $(a, b)(f, p) = (a(f, p), b(f, p, \mathrm{id}))$ and similarly $(\nu, \varrho)(f, p) = (\nu(f, p), \varrho(f, p, \mathrm{id}))$.

Assuming a subcanonical topology (i.e. the Yoneda $\mathbf{y}X$ is a sheaf for all X).

Let X be an object of $\mathcal{C}$ and $\gamma \in \Gamma(X)$. By the Yoneda Lemma γ induces a substitution $\gamma : \mathbf{y}X \to \Gamma$. The type $\Gamma \vdash \Pi(x\!:\!A)B$ is interpreted as:

$(\Pi(x\!:\!A)B)(\gamma) = \mathrm{Tm}_{\mathbf{y}X.A\gamma}B(\gamma \circ \mathsf{p}, \mathsf{q})$

where $\mathsf{p} : \mathbf{y}X.A\gamma \to \mathbf{y}X$ and $\mathsf{q} \in \mathrm{Tm}_{\mathbf{y}X.A\gamma}(A\gamma)\mathsf{p}$ are the familiar first and second projections.

A map $(f, \mathrm{id}) : (Y, \gamma f) \to (X, \gamma)$ induces a substitution $\mathbf{y}Y.A\gamma f \to \mathbf{y}X.A\gamma$. Let $(f, p) : (Y, \alpha) \to (X, \gamma)$ and $\omega \in (\Pi(x : A)B)(\gamma)$. We describe the restriction $\omega(f, p)$ of ω along (f, p). Given $g : Z \to Y$ and $a \in A(\alpha g)$ then $b = (\omega(f, p))(g, a)$ is given by transporting a along the isomorphism $(\mathrm{id}, p^{-1}g) : A(\alpha g) \xrightarrow{\sim} A(\gamma f g)$ followed by an application $b' = \omega(fg, a(\mathrm{id}, p^{-1}g)) \in B(\gamma f g, a(\mathrm{id}, p^{-1}g))$ then a transport along the isomorphism $(\mathrm{id}, pg, \mathrm{id}) : B(\gamma f g, a(\mathrm{id}, p^{-1}g)) \xrightarrow{\sim} B(\alpha g, a)$, that is, $b = b'(\mathrm{id}, pg, \mathrm{id})$. The description of $\omega(f, p)(g, d)$ where $d : a \xrightarrow{\sim} a' \in A(\alpha g)$ is similar.

The restrictions of elements and morphisms in $(\Pi(x : A)B)(\gamma)$ along (f, p) is then given by the action of the substitution induced by (f, id)on $\mathrm{Tm}_{\mathbf{y}X.A\gamma}B(\gamma \circ \mathsf{p}, \mathsf{q})$.

In order to show that the $\Pi(x\!:\!A)B$ so described is a stack one needs to impose extra conditions on the objects of our model. Namely we require that the stacks in this model come equipped with a uniform glue operation, that is an operation that produces glues for descent data and behaves well with restrictions. We omit these details here.

Small types and the universe: We interpret the judgment $\Gamma \vdash A$ small as saying that the stack $A : \mathcal{C}_\Gamma^{op} \to \mathbf{Gpd}$ is a small sheaf, or rather more accurately A is a stack of small discrete groupoids.

The universe U is interpreted as follows: $U(X)$ is the groupoid of small sheaves over $(\mathcal{C}_X, \mathbf{J}_X)$ and sheaf isomorphisms. That is to say an element of $U(X)$ is a type $\mathbf{y}X \vdash A$ where the judgment $\mathbf{y}X \vdash A$ small holds. The U so described is indeed a stack [Vistoli, 2004, Example 3.20].

The coding $\ulcorner\ \urcorner$ and decoding El are interpreted as described in the presheaf interpretation.

Natural numbers and booleans As is usual in sheaf models, the interpretation of natural numbers is given by a constant sheaf N, i.e. by the sheafification of the constant presheaf of natural numbers. Customarily this sheafification is described in terms of equivalence classes (the plus construction). The eliminator natrec has then to be well defined with respect to these equivalence classes. When it comes to elimination of natural numbers into small types (i.e. sheaves) this is quite straightforward. The situation is not so obvious when it comes to elimination of natural numbers into stacks. To properly define the elimination of natural numbers in stacks we require that an element in $N(X)$ can be uniquely represented.

The situation is similar for the type N_2 of booleans.

Identity types and univalence We can furthermore interpret identity types. For $\Gamma \vdash a : A$ and $\Gamma \vdash b : A$ we interpret $\mathrm{Id}_A(a,b)$ by the the set (discrete groupoid) $\mathrm{Id}_A(a,b)(\gamma) = \mathrm{Hom}_{A(\gamma)}(a(\gamma), b(\gamma))$.

We say that two types A and B are equivalent if there is a a contractible map $\varphi : A \to B$, i.e. the fibers of φ over b are contractible for all $b : B$. If A and B are small (i.e. sheaves) then φ being contractible means that its fiber over any $b : B$ contains exactly one element. Which in turn implies that A and B are isomorphic as sheaves. We can then transform an equivalence between two small types to a path (isomorphism) in the universe U.

Cantor space and the independence of Markov's principle We now turn our attention to Cantor space. The stacks over this site have a relatively simple description.

Since we have at most one arrow between any two objects in this site, whenever $q \leqslant p$ and $b \in A(p)$ we will write $b|q$ for the restriction of b to $A(q)$.

For any object/condition p the objects in the coverage/partition $p \lhd p_1, \dots, p_n$ are disjoint which means that the isomorphism part of a descent datum is always trivial and any family $\{x_i \in F(p_i)\}$ forms a descent datum.

Thus a stack over Cantor space is defined as a prestack F with an operation glue that satisfies:
Whenever $a_0 \in F(p(m,0))$ and $a_1 \in F(p(m,1))$ for some $m \notin \mathrm{dom}(p)$ then $\mathsf{glue}(a_0, a_1) \in F(p)$.

The restrictions of $\mathsf{glue}(a_0, a_1)$ are given by:

$$\mathsf{glue}(a_0, a_1)|q = \begin{cases} a_0|q, & q \leqslant p(m,0) \\ a_1|q, & q \leqslant p(m,1) \\ \mathsf{glue}(a_0|q(m,0), a_1|q(m,1)), & \textit{Otherwise (i.e. } m \notin \mathrm{dom}(p)) \end{cases}$$

Now let $r \leqslant q \leqslant p$. If $r \leqslant p(m,0)$ then $\mathsf{glue}(a_0, a_1)|r = a_0|r$. If $q \leqslant p(m,0)$ then $\mathsf{glue}(a_0, a_1)|q = a_0|q$ and $(\mathsf{glue}(a_0, a_1)|q)|r = (a_0|q)|r = a_0|r$. Otherwise, $\mathsf{glue}(a_0, a_1)|q = \mathsf{glue}(a_0|q(m,0), a_1|q(m,1))$. But since $r \leqslant p(m,0)$ we have $r \leqslant q(m,0)$ and $(\mathsf{glue}(a_0, a_1)|q)|r = (a_0|q(m,0))|r = a_0|r$. Similarly if $r \leqslant p(m,1)$. If on the other hand, $m \notin \mathrm{dom}(r)$ we have

$$\begin{aligned}(\mathsf{glue}(a_0, a_1)|q)|r &= \mathsf{glue}(a_0|q(m,0), a_1|q(m,1))|r \\ &= \mathsf{glue}(a_0|r(m,0), a_1|r(m,1)) = \mathsf{glue}(a_0, a_1)|r\end{aligned}$$

Thus stacks over Cantor space satisfy the uniform glue condition mentioned earlier.

We thus have the following strong conjecture:

Conjecture 2.1. *Type theory with one universe, natural numbers, Π and Σ types, and one level of univalence has a sound interpretation in stacks over Cantor space.*

A Corollary of the above conjecture would be that Markov's principle is independent from type theory. Thus giving another proof of the result in Part B.

Considering the forcing extension of type theory presented in Part B. We remark that the locality rule:

$$\frac{\Gamma \vdash_{p_1} J \quad \dots \quad \Gamma \vdash_{p_n} J}{\Gamma \vdash_p J}\ p \lhd p_1 \dots p_n$$

cannot hold in general in the stack interpretation. For example if $p \lhd p_1, \dots, p_n$ and $\Gamma \vdash_{p_i} A = B : U$ for all i it is not necessarily the case that $\Gamma \vdash_p A = B : U$.

However this property still holds for small types. We can then replace the locality rule with:

$$\frac{\Gamma \vdash_{p_1} a : A \quad \dots \quad \Gamma \vdash_{p_n} a : A}{\Gamma \vdash_p a : A}\ p \lhd p_1 \dots p_n$$

$$\frac{\Gamma \vdash_{p_1} a = b : N_1 \quad \dots \quad \Gamma \vdash_{p_n} a = b : N_1}{\Gamma \vdash_p a = b : N_1}\ p \lhd p_1 \dots p_n$$

$$\frac{\Gamma \vdash_{p_1} a = b : N_2 \quad \dots \quad \Gamma \vdash_{p_n} a = b : N_2}{\Gamma \vdash_p a : N_2}\ p \lhd p_1 \dots p_n$$

$$\frac{\Gamma \vdash_{p_1} a = b : N \quad \ldots \quad \Gamma \vdash_{p_n} a = b : N}{\Gamma \vdash_p a = b : N}\ p \lhd p_1 \ldots p_n$$

The site of étale *K*-algebras and the separable closure Whether type theory has an interpretation in stacks over the site of étale K-algebras is still an open problem. In light of the work presented in Part A, if such interpretation exists one can then extend type theory with the type of separable algebraic closure of a field K. This should be quite useful for formalization of algebra in type theory.

Bibliography

Abel, A. and Scherer, G. [2012], 'On irrelevance and algorithmic equality in predicative type theory', *Logical Methods in Computer Science* **8**(1), 1–36. TYPES'10 special issue.

Abhyankar, S. S. [1976], 'Historical ramblings in algebraic geometry and related algebra', *Amer. Math. Monthly* **83**(6), 409–448.

Abhyankar, S. S. [1990], *Algebraic Geometry for Scientists and Engineers*, American Mathematical Society.

Aczel, P. [1999], *Types for Proofs and Programs: International Workshop, TYPES' 98 Kloster Irsee, Germany, March 27–31, 1998 Selected Papers*, Springer Berlin Heidelberg, Berlin, Heidelberg, chapter On Relating Type Theories and Set Theories, pp. 1–18.

Artin, M., Grothendieck, A. and Verdier, J. [1972], *Théorie des topos et cohomologie étalé des schémas*, Théorie des topos et cohomologie étale des schémas, Springer-Verlag.

Awodey, S. [1997], Logic in topoi: Functorial Semantics for Higher-Order Logic, PhD thesis, The University of Chicago.

Basu, S., Pollack, R. and Roy, M.-F. [2006], *Algorithms in Real Algebraic Geometry (Algorithms and Computation in Mathematics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Bell, J. [2011], *Set Theory: Boolean-Valued Models and Independence Proofs*, Oxford Logic Guides, OUP Oxford.

Bell, J. L. [2005], The Development of Categorical Logic, Vol. 12 of *Handbook of Philosophical Logic*, Springer Netherlands, pp. 279–361.

Beth, E. W. [1956], *Semantic Construction of Intuitionistic Logic*, Noord-Hollandsche Uitg. Mij.

Bezem, M., Coquand, T. and Huber, S. [2014], A model of type theory in cubical sets, *in* '19th International Conference on Types for Proofs and Programs (TYPES 2013)', Vol. 26, pp. 107–128.

Bishop, E. [1967], *Foundations of Constructive Analysis*, McGraw-Hill.

Bridges, D. and Richman, F. [1987], *Varieties of Constructive Mathematics*, Lecture note series, Cambridge University Press.

Brouwer, L. E. J. [1975], Essentially negative properties, *in* A. Heyting, ed., 'Collected Works', Vol. I, North-Holland, pp. 478 – 479.

Cohen, P. J. [1963], 'The independence of the continuum hypothesis', *Proceedings of the National Academy of Sciences of the United States of America* **50**(6), 1143–1148.

Constable, R. L. and Smith, S. F. [1987], Partial objects in constructive type theory, *in* 'Proceedings of Second IEEE Symposium on Logic in Computer Science', pp. 183–193.

Coquand, T. [2013], What is a model of type theory?

Coquand, T. and Jaber, G. [2010], 'A note on forcing and type theory', *Fundamenta Informaticae* **100**(1-4), 43–52.

Coquand, T., Lombardi, H. and Quitté, C. [2010], 'Curves and coherent prüfer rings', *J. Symb. Comput.* **45**(12), 1378–1390.

Coquand, T. and Mannaa, B. [2016], The independence of markov's principle in type theory, *in* '1st International Conference on Formal Structures for Computation and Deduction, FSCD 2016, June 22-26, 2016, Porto, Portugal', pp. 17:1–17:18.

Coste, M. [1972], 'Language interne d'un topos', *Sém. Bénabou, Université Paris-Nord* .

Coste, M., Lombardi, H. and Roy, M.-F. [2001], 'Dynamical method in algebra: effective nullstellensätze', *Annals of Pure and Applied Logic* **111**(3), 203 – 256.

Coste-Roy, M.-F., Coste, M. and Mahé, L. [1980], 'Contribution to the study of the natural number object in elementary topoi', *Journal of Pure and Applied Algebra* **17**(1), 35 – 68.

Della Dora, J., Dicrescenzo, C. and Duval, D. [1985], About a new method for computing in algebraic number fields, *in* B. Caviness, ed., 'EUROCAL '85', Vol. 204 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 289–290.

Descartes, R. [1637], *Discours de la méthode*, Jan Maire, chapter La Géométrie, pp. 376–493.

Duval, D. [1989], 'Rational Puiseux expansions', *Compos. Math.* **70**(2), 119–154.

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

Edwards, H. M. [2005], *Essays in constructive mathematics*, Springer.

Fröhlich, A. and Shepherdson, J. C. [1956], 'Effective procedures in field theory', *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **248**(950), 407–432.

Grothendieck, A. and Dieudonné, J. [1960], *Éléments de géométrie algébrique: I. Le langage des schémas*, Institut des hautes études scientifiques.

Hofmann, M. [1997], Syntax and semantics of dependent types, *in* 'Semantics and Logics of Computation', Cambridge University Press, pp. 79–130.

Hofmann, M. and Streicher, T. [199?], Lifting grothendieck universes. unpublished note.

Hofmann, M. and Streicher, T. [1998], The groupoid interpretation of type theory, *in* 'Twenty-five years of constructive type theory (Venice, 1995)', Vol. 36 of *Oxford Logic Guides*, Oxford Univ. Press, New York, pp. 83–111.

Huber, S. [2015], A model of type theory in cubical sets, Licentiate thesis, Chalmers University of Technology.

Hyland, J. and Ong, C.-H. [1993], Modified realizability toposes and strong normalization proofs (extended abstract), *in* 'Typed Lambda Calculi and Applications, LNCS 664', Springer-Verlag, pp. 179–194.

Interview: F. William Lawvere -part 2 [2008], *Bulletin of the International Center for Mathematics* pp. 21–28.

Johnstone, P. T. [2002*a*], *Sketches of an Elephant: A Topos Theory Compendium - Volume 2*, number 44 *in* 'Oxford Logic Guides', Oxford University Press.

Johnstone, P. T. [2002*b*], *Sketches of an Elephant: A Topos Theory Compendiumm - Volume 1*, number 43 *in* 'Oxford Logic Guides', Oxford University Press.

Joyal, A. [1975], 'Le théorème de chevalley-tarski et remarques sur l'algèbre constructive', *Cah. Topol. Géom. Différ.* **16**, 256–258.

Kopylov, A. and Nogin, A. [2001], *Computer Science Logic: 15th International Workshop, CSL 2001 10th Annual Conference of the EACSL Paris, France, September 10–13, 2001, Proceedings*, Springer Berlin Heidelberg, Berlin, Heidelberg, chapter Markov's Principle for Propositional Type Theory, pp. 570–584.

Kreisel, G. [1959], Interpretation of analysis by means of constructive functionals of finite types, *in* A. Heyting, ed., 'Constructivity in Mathematics', Amsterdam, North-Holland Pub. Co., pp. 101–128.

Kunen, K. [1980], *Set Theory: An Introduction to Independence Proofs*, Vol. 102 of *Studies in Logic and the Foundations of Mathematics*, Elsevier.

Lawvere, F. W. [1970], Quantifiers and sheaves, *in* 'Actes du Congres International des Mathématiciens, Nice', Vol. 1, pp. 329–334.

Lombardi, H. and Quitté, C. [2011], *Algèbre Commutative, Méthodes Constructives*, Mathématiques en devenir, Calvage et Mounet.

MacLane, S. and Moerdijk, I. [1992], *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, corrected edn, Springer.

Makkai, M. and Reyes, G. E. [1977], *First order categorical logic: model-theoretical methods in the theory of topoi and related categories*, Vol. 611 of *Lecture notes in mathematics*, Springer-Verlag.

Mannaa, B. and Coquand, T. [2013], 'Dynamic newton-puiseux theorem', *J. Logic & Analysis* **5**.

Mannaa, B. and Coquand, T. [2014], A sheaf model of the algebraic closure, *in* P. Oliva, ed., 'Proceedings Fifth International Workshop on Classical Logic and Computation, Vienna, Austria, July 13, 2014', Vol. 164 of *Electronic Proceedings in Theoretical Computer Science*, Open Publishing Association, pp. 18–32.

Margenstern, M. [1995], 'L'école constructive de Markov', *Revue d'histoire des mathématiques* **1**(2), 271–305.

Martin-Löf, P. [1972], An intuitionistic theory of types. reprinted in Twenty-five years of constructive type theory, Oxford University Press, 1998, 127–172.

McLarty, C. [1990], 'The uses and abuses of the history of topos theory', *The British Journal for the Philosophy of Science* **41**(3), 351–375.

Miller, H. [2000], 'Leray in oflag xviia: The origins of sheaf theory, sheaf cohomology, and spectral sequences', *Gazette des Mathematicens* **84 suppl**, 17–34.

Mines, R., Richman, F. and Ruitenburg, W. [1988], *A course in constructive algebra*, Universitext (1979), Springer-Verlag.

Mitchell, W. [1972], 'Boolean topoi and the theory of sets', *Journal of Pure and Applied Algebra* **2**(3), 261–274.

Newton, S. I. [1736], *The method of fluxions and infinite series: with its application to the geometry of curve-lines*, printed by Henry Woodfall; and sold by John Nourse.

Norell, U. [2007], Towards a practical programming language based on dependent type theory, PhD thesis, Department of Computer Science and Engineering, Chalmers University of Technology, SE-412 96 Göteborg, Sweden.

Osius, G. [1975], A note on kripke-joyal semantics for the internal language of topoi, Model Theory and Topoi: A Collection of Lectures by Various Authors, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 349–354.

Palmgren, E. [1997], 'Constructive sheaf semantics', *Mathematical Logic Quarterly* **43**(3), 321–327.

Puiseux, V. [1850], 'Recherches sur les fonctions algébriques', *J. Math. Pures Appl* (15), 365–480.

Ščedrov, A. [1984], *Forcing and classifying topoi*, Vol. 48 of *Memoirs of the AMS*, American Mathematical Society (AMS).

Solovay, R. M. [1965], $2^{\aleph_0}$ can be anything it ought to be, *in* J. Addison, L. Henkin and A. Tarski, eds, 'The Theory of Models: Proceedings of the 1963 International Symposium at Berkeley', Studies in logic and the foundations of mathematics, Elsevier Science.

Streicher, T. [2005], Universes in toposes, *in* L. Crosilla and P. Schuster, eds, 'From Sets and Types to Topology and Analysis: Towards practicable foundations for constructive mathematics', Oxford University Press, pp. 78–90.

Tait, W. W. [1967], 'Intensional interpretations of functionals of finite type I', *Journal of Symbolic Logic* **32**(2), 198–212.

Tierney, M. [1972], *Toposes, Algebraic Geometry and Logic: Dalhousie University, Halifax, January 16–19, 1971*, Springer Berlin Heidelberg, Berlin, Heidelberg, chapter Sheaf theory and the continuum hypothesis, pp. 13–42.

Troelstra, A. S. and van Dalen, D. [1988], *Constructivism in Mathematics: An Introduction*, Vol. I and II of *Studies in Logic and the Foundations of Mathematics*, North-Holland.

Troelstra, A. S. and van Ulsen, P. [1999], The discovery of ew beth's semantics for intuitionistic logic, JFAK. Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday.

van Dalen, D. [1978], 'An interpretation of intuitionistic analysis', *Annals of Mathematical Logic* **13**(1), 1 – 43.

Vistoli, A. [2004], 'Notes on Grothendieck topologies, fibered categories and descent theory', *ArXiv Mathematics e-prints* .

von Tschirnhaus, E. W. and Green, R. F. [2003], 'A method for removing all intermediate terms from a given equation', *SIGSAM Bull.* **37**(1), 1–3.

Vopěnka, P. [1965], 'The limits of sheaves and applications on construction of models', *Bull. Acad. Plon. Sci.* **13**, 189–192.

Waerden, V. D. [1930], 'Eine bemerkung über die unzerlegbarkeit von polynomen', *Math. Annalen* **102**, 738–739.

Walker, R. J. [1978], *Algebraic curves*, Springer-Verlag.

Wright, A. K. and Felleisen, M. [1994], 'A syntactic approach to type soundness', *Information and Computation* **115**(1), 38 – 94.

Xu, C. and Escardó, M. [2016], Universes in sheaf models. University of Birmingham.