Polytime Functions in Two-Sorted Bounded Arithmetic

Masterarbeit
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Remo Goetschi

2008

Leiter der Arbeit:

Prof. Dr. Thomas Strahm
Institut für Informatik und angewandte Mathematik

Contents

1 Introduction

In this master thesis we examine the theory $\mathbf{V^1}$, a two-sorted (first-order) theory of bounded arithmetic which was introduced by Cook. Theories of bounded arithmetic are weak subsystems of Peano Arithmetic, where the induction axiom scheme is restricted to formulas with only bounded quantifiers, i.e. quantifiers of the form $\exists x \leq t$ or $\forall x \leq t$. "Two-sorted" means that there are two sorts of individuals, namely "numbers" and "finite sets of numbers". This two-sorted approach is due to Zambella ([11]) and Cook ([6]) and turned out to be useful for studying especially weak complexity classes such as AC^0 (but also P).

The theory $\mathbf{V^1}$ characterises the polynomial time computable functions FP in the sense that the functions definable in $\mathbf{V^1}$ (called the *provably total* functions of $\mathbf{V^1}$), correspond precisely to the functions in FP. $\mathbf{V^1}$ is the second element of a whole hierarchy $\mathbf{V^0} \subset \mathbf{V^1} \subseteq \mathbf{V^2} \subseteq \ldots$ of theories, where, for $i \geq 1$, $\mathbf{V^i}$ characterises the $(i-1)$-th level of the polynomial hierarchy.

One leading motivation for relating bounded arithmetic and complexity classes is to gain insight into the numerous open problems in theoretical computer science, namely the question of whether the polynomial hierarchy is proper and as a consequence whether the classes P and NP are distinct. The PhD thesis of Buss ([2]) was a milestone in relating complexity classes and bounded arithmetic. The theories $\mathbf{V^0} \subset \mathbf{V^1} \subseteq \mathbf{V^2} \subseteq \ldots$ correspond to Buss' theories $\mathbf{S^1_2} \subseteq \mathbf{S^2_2} \subseteq \ldots$.

The main references for this thesis are the lecture notes of Prof. Stephen Cook and the yet unpublished[i] book ([6]) "Foundations of Proof Complexity: Bounded Arithmetic and Propositional Translations" by Stephen Cook and Phuong Nguyen. [11, 12, 3, 2] are good sources for advanced study of the subject of bounded arithmetic and complexity classes.

This thesis is intended to be easily accessible to students and novice readers of logic texts. Therefore it contains more "prose" than the advanced logician is used to. Section 2 is rather elementary and contains an introduction into bounded arithmetic as well as a brief introduction into first-order logic. Experienced readers might just want to skim through in order to look at the notations. In section 3, two-sorted first-order logic, together with a two-sorted version of the sequent calculus, is introduced. We also prove basic results such as soundness, completeness and "free-cut" elimination. In section 4, the theory $\mathbf{V^1}$ is introduced and many of its properties are proved, especially a two-sorted version of Parikh's theorem ([8]). At the end, section

[i] in April 2008

4 then contains a lower and an upper bound theorem for $\mathbf{V^1}$. These theorems imply that $\mathbf{V^1}$ characterises FP.

All but very few of the theorems and lemmas are proved in detail. Familiarity with basic complexity theory and some experience with first-order predicate logic are helpful. The reader should also know about elementary set theoretic notions such as functions and relations.

List of Abbreviations

- $\equiv$ stands for syntactical equivalence.
- $\mathcal{P}(X)$ stands for the powerset of X.
- $\mathcal{P}_{\mathbf{f}}(X)$ stands for the *finite power set* of X, that is all finite subsets of X.
- $\Rightarrow$ stands for implication in the meta-language.
- s.t.: such that
- w.r.t.: with respect to
- w.l.o.g.: without loss of generality
- RHS: right hand side
- LHS: left hand side

Acknowledgements

I am grateful to Prof. Dr. Thomas Strahm and Prof. Dr. Gerhard Jäger for their advice and supervision, and for giving me the possibility to study the subject of mathematical logic. I further thank the famous physicist Patrick Liniger for proofreading, the staff of Open Innovation GmbH for letting me use their source code repository and for paying my salary and Muriel Riesen for everything else.

Remo Goetschi
City of Bern, May 2008

2 Preliminaries

2.1 First-order Predicate Logic

This section provides a brief introduction into first-order logic together with basic definitions that are used later. For simplicity, first-order predicate logic will be called just first-order logic.

2.1.1 Syntax

First-order logic is an extension of first-order propositional logic and allows predications over individuals (e.g. natural numbers) using functions and relations over those individuals. In the following, we define first-order languages, terms and formulas. A first-order language $\mathcal{L}$ consists of:

1. a (possibly empty) set of n-ary function symbols, for each $n \in \mathbb{N}$. A 0-ary function symbol is called a constant symbol,
2. a (possibly empty) set of n-ary predicate symbols, for each $n \in \mathbb{N}$,
3. an infinite set of variables,
4. logical connectives $\vee, \wedge, \neg$,
5. quantifiers $\forall, \exists$,
6. parentheses $(,)$.

We use $f, g, h, \ldots$ as meta-symbols to denote function symbols[i], $P, Q, R, \ldots$ for predicate symbols and $x, y, z, \ldots, a, b, c, \ldots$ for variables. Further we require that each language contains at least one predicate symbol. Terms over a language $\mathcal{L}$ are intended to range over a set of individuals, called the universe (cf. section 2.1.2). $\mathcal{L}$-terms are inductively defined in the following way:

1. Every variable of $\mathcal{L}$ is an $\mathcal{L}$-term.
2. If f is an n-ary function symbol of $\mathcal{L}$ and $t_1, \ldots, t_n$ are $\mathcal{L}$-terms, then $f(t_1, \ldots, t_n)$ is an $\mathcal{L}$-term.

Using terms, we can now build $\mathcal{L}$-formulas inductively as follows:

1. If P is an n-ary predicate symbol in $\mathcal{L}$ and $t_1, \ldots, t_n$ are $\mathcal{L}$-terms, then $P(t_1, \ldots, t_n)$ is an (atomic) $\mathcal{L}$-formula.

[i] sometimes with sub- or superscripts

2. If A and B are $\mathcal{L}$-formulas, then $(A \vee B), (A \wedge B), (\neg A)$ are $\mathcal{L}$-formulas.

3. If A is an $\mathcal{L}$-formula and x is a variable, then $(\forall x)A$ and $(\exists x)A$ are $\mathcal{L}$-formulas.

We use $r, s, t, \ldots$ as meta-symbols to denote terms, capital letters $A, B, C, \ldots$ for formulas and capital Greek letters $\Gamma, \Delta, \Phi, \ldots$ to denote sets of formulas. In addition, we use the following abbreviations: $A \to B$ stands for $(\neg A \vee B)$ and $A \leftrightarrow B$ stands for $(A \to B \wedge B \to A)$. We often omit parentheses and follow the convention that precedence follows the order $\neg, \wedge, \vee, \to$. For example, $\neg A \vee B \to C$ stands for $\big((\neg A) \vee B\big) \to C$. All languages in this thesis are extensions of the first-order *language of arithmetic* (or a two-sorted version thereof) which we will call $\mathcal{L}_{\mathcal{A}}$.

Definition 2.1 ($\mathcal{L}_{\mathcal{A}}$)**.** $\mathcal{L}_{\mathcal{A}} = (0, 1, +, \times, =, \leq)$

In the above definition $0, 1$ are 0-ary function symbols (constants), $+, \times$ are binary function symbols and $=, \leq$ are binary predicate symbols. We will use $t \neq s$ as an abbreviation for $\neg(t = s)$ and $t < s$ for $(t \leq s \wedge t \neq s)$.

Definition 2.2 (Free/Bound Variables)**.** *For terms and formulas, the set of free variables is inductively defined as follows*

$$
\begin{aligned}
&free(x) = \{x\}, \quad free(f(t_1, \ldots, t_n)) = free(t_1) \cup \cdots \cup free(t_n),\\
&free(P(t_1, \ldots, t_n)) = free(t_1) \cup \cdots \cup free(t_n),\\
&free(A \wedge B) = free(A \vee B) = free(A) \cup free(B),\\
&free(\neg A) = free(A),\\
&free(\forall x A) = free(\exists x A) = free(A) \setminus \{x\}
\end{aligned}
$$

The set of bound variables is defined accordingly with $bound(\forall x A) = bound(\exists x A) = bound(A) \cup \{x\}$.

Obviously, a variable x can occur both free and bound in a formula A. A formula without free occurrences of variables is called a *closed* formula or a *sentence*. A term that does not contain variables is called a closed term. We adopt common syntactic conventions and write $A(t/x)$ for the formula obtained by replacing all *free* occurrences of x in A by the term t. In general, we demand that t is *freely substitutable* for x in A, which means that no (free) variable of t becomes bound in $A(t/x)$. This is to prevent unwanted semantic side effects[i]. Further we write $A(x)$ and mean that the variable x might occur free in A and then $A(t)$ means $A(t/x)$ in the same context. Note that $A(x)$ does not necessarily mean that x actually occurs in A. Out of context, $A(t)$ stands for $A(t/x)$ for some x. This notation applies accordingly for terms (e.g. $t(x)$).

[i]For example, the variable x is *not* freely substitutable for y in $\exists x(y < x)$.

Notation We often write $\vec{x}$ instead of $x_1, \ldots, x_k$, for some $k \geq 0$. Let $A(\vec{x})$ be a formula with all free variables indicated. Then $\forall A(\vec{x})$ stands for

$$\forall x_1 \ldots \forall x_k A(x_1, \ldots, x_k)$$

and is called the *universal closure* of the formula A. If Φ is a set of formulas, $\forall\Phi$ denotes the set of universal closures of the formulas in Φ.

2.1.2 Semantics

In order to assign a *meaning* (i.e. a truth value) to first-order formulas, function and relation symbols must obtain interpretations in a set of individuals. The concept of an $\mathcal{L}$-structure defines such interpretations. An $\mathcal{L}$-structure consists of

1. a nonempty set M of individuals (the *universe*),
2. for each n-ary function symbol f of $\mathcal{L}$ an interpretation $f^{\mathcal{M}} : \mathsf{M}^n \to \mathsf{M}$,
3. for each n-ary relation symbol P of $\mathcal{L}$ an interpretation $P^{\mathcal{M}} \subseteq \mathsf{M}^n$. If $\mathcal{L}$ contains the relation symbol $=$, then $=^{\mathcal{M}}$ must be the equality relation on M.

To obtain interpretations for terms with variables (i.e. non-closed or open terms) we introduce *variable assignments*. A variable assignment (or just assignment) σ for an $\mathcal{L}$-structure $\mathcal{M}$ is a mapping from the set of variables in $\mathcal{L}$ to M and gives meaning to the free variables of a formula A. Now the interpretation $t^{\mathcal{M}}[\sigma]$ of an $\mathcal{L}$-term t in an $\mathcal{L}$-structure $\mathcal{M}$ with respect to an assignment σ can be defined inductively:

1. If t is a variable x, then $t^{\mathcal{M}}[\sigma]$ is just $\sigma(x)$.
2. If t is of the form $f(t_1, \ldots, t_n)$ then $t^{\mathcal{M}}[\sigma]$ is $f^{\mathcal{M}}(t_1^{\mathcal{M}}[\sigma], \ldots, t_n^{\mathcal{M}}[\sigma])$.

If t is closed, then it is of course independent of σ and its interpretation is simply denoted $t^{\mathcal{M}}$. In the following $\sigma(\mathsf{m}/x)$ denotes an assignment which is identical to σ with the exception that it maps x to $\mathsf{m} \in \mathsf{M}$. We can now define the *truth value* of a formula A[i] with respect to a structure $\mathcal{M}$ and an assignment σ. We write $\mathcal{M}[\sigma] \models A$ if A is true in $\mathcal{M}$ with respect to σ. We define $\mathcal{M}[\sigma] \models A$ by structural induction on the built-up of A as follows:

1. If A is $P(t_1, \ldots, t_n)$, then $\mathcal{M}[\sigma] \models A$ iff $(t_1^{\mathcal{M}}[\sigma], \ldots, t_n^{\mathcal{M}}[\sigma]) \in P^{\mathcal{M}}$.

[i] We will often not explicitly mention the underlying language $\mathcal{L}$ if it is clear from the context or irrelevant.

2. If A is $\neg B$, then $\mathcal{M}[\sigma] \models A$ iff $\mathcal{M}[\sigma] \models B$ does not hold (written $\mathcal{M}[\sigma] \not\models A$).

3. If A is $B \wedge C$, then $\mathcal{M}[\sigma] \models A$ iff $\mathcal{M}[\sigma] \models B$ and $\mathcal{M}[\sigma] \models C$.

4. If A is $B \vee C$, then $\mathcal{M}[\sigma] \models A$ iff $\mathcal{M}[\sigma] \models B$ or $\mathcal{M}[\sigma] \models C$ (or both).

5. If A is $\forall x B$, then $\mathcal{M}[\sigma] \models A$ iff $\mathcal{M}[\sigma(\mathsf{m}/x)] \models B$ for all $\mathsf{m} \in \mathsf{M}$.

6. If A is $\exists x B$, then $\mathcal{M}[\sigma] \models A$ iff $\mathcal{M}[\sigma(\mathsf{m}/x)] \models B$ for at least one $\mathsf{m} \in \mathsf{M}$.

Definition 2.3 (Validity)**.** *We say that a formula A is* valid *in a structure $\mathcal{M}$, formally $\mathcal{M} \models A$, iff $\mathcal{M}[\sigma] \models A$ for all assignments σ and we then call $\mathcal{M}$ a* model of A*. We call a formula valid iff it is valid in all structures. A set Γ of formulas is valid in $\mathcal{M}$, formally $\mathcal{M} \models \Gamma$, iff every formula in Γ is valid in $\mathcal{M}$ (and we analogously call $\mathcal{M}$ a model of Γ).*

Definition 2.4 (Satisfiability)**.** *We say that a formula A is satisfiable if there exists a structure $\mathcal{M}$ and $\mathcal{M} \models A$.*

Definition 2.5 (Logical Consequence)**.** *We say* Γ logically implies A *(or A is a* logical consequence *of Γ), formally $\Gamma \models A$, iff every model of Γ is also a model of A. That is iff for all $\mathcal{M}$, $\mathcal{M} \models \Gamma \Rightarrow \mathcal{M} \models A$.*

As a convention, we write $\models A$ instead of $\emptyset \models A$ and $A \models B$ instead of $\{A\} \models B$. In the case of $\models A$, we say that A is *valid*. Note that the relation $\models$ is transitive, i.e. $A \models B$ and $B \models C$ implies $A \models C$. If both $A \models B$ and $B \models A$ we say that A and B are *equivalent* and we sometimes write $A \Leftrightarrow B$.

Remark 2.6. We follow Buss ([3]) and others and define logical consequence in such a way that free variables are implicitly *universally quantified*[i]. The following lemma is an obvious consequence thereof.

Lemma 2.7 (Universal Closure Property)**.** *For every formula A*

$$A(x) \Leftrightarrow \forall x A(x).$$

Note that the statement "$A \leftrightarrow B$ is valid" is stronger than $A \Leftrightarrow B$. We will make implicit use of the following lemma later in the text.

Lemma 2.8 (Formula Replacement Lemma)**.** *Let A and A' be formulas and assume that $A \leftrightarrow A'$ is valid. Assume further that a formula B' is obtained from a formula B by substituting A' for all occurrences of the subformula A in B. Then $B \leftrightarrow B'$ is valid as well.*

[i]This is just a convention and makes some arguments simpler. Note that our definition differs from the one in [6].

Proof. Straightforward by structural induction on B. See for example theorem 2.5.8 in [9]. □

Next we define the term of a *theory*. We often use bold capital letters ($\mathbf{T}, \mathbf{S}, \dots$) to denote theories.

Definition 2.9 (Theory)**.** *An $\mathcal{L}$-theory is a set of $\mathcal{L}$-formulas closed under logical consequence. An axiomatisation of an $\mathcal{L}$-theory $\mathbf{T}$ is a set Γ of $\mathcal{L}$-formulas, called axioms, such that $\mathbf{T}$ is exactly the set of $\mathcal{L}$-formulas logically implied by Γ.*

We call the formulas of $\mathbf{T}$ *theorems of* $\mathbf{T}$. It is obvious that $A \in \mathbf{T}$ and $\mathbf{T} \models A$ are equivalent (we use the second notation). Note that by the above definitions, theories are also closed under universal quantification. That is, $\mathbf{T} \models A \Leftrightarrow \mathbf{T} \models \forall x A$ for every formula A. Given a set Γ of axioms, the corresponding theory is precisely the set $\{A \mid \Gamma \models A\}$.

2.2 Bounded Arithmetic

This short section serves as a brief introduction into the topic of bounded arithmetic. Some important notions and techniques are introduced by means of the well-known one-sorted theory $\mathbf{I\Delta_0}$. It is shown what it means for functions and predicates to be *definable* and *provably total* in a theory and we will see how complexity classes can be *characterised* by theories. Many of the definitions involved will reappear in chapter 3 in a two-sorted context. The underlying language is always the language $\mathcal{L}_\mathcal{A}$ or an extension thereof (written $\mathcal{L} \supseteq \mathcal{L}_\mathcal{A}$). We will use *infix notation* for the function and relation symbols of $\mathcal{L}_\mathcal{A}$ and follow the standard precedence rules ($\times$ binds stronger than $+$). For example, $(x+1) \times x = x \times x + x$ stands for $\times(+(x,1),x) = +(\times(x,x),x)$. Our theories are intended to prove facts about the natural numbers $\mathbb{N}$.

Definition 2.10 (standard model $\underline{\mathbb{N}}$)**.** *The $\mathcal{L}_\mathcal{A}$-structure $\underline{\mathbb{N}}$ has universe $\mathbb{N}$ and the function and relations symbols $0, 1, +, \times, =, \leq$ get their standard interpretations in the natural numbers. We call $\underline{\mathbb{N}}$ the standard model (of the natural numbers).*

2.2.1 Setting up $\mathbf{I\Delta_0}$

We begin by defining a slightly modified version **BA** (for "Basic Arithmetic") of the theory **Q**, called *Robinson's Arithmetic*[i]. **BA** is has the following axioms:

B1. $x + 1 \neq 0$
B2. $x + 1 = y + 1 \rightarrow x = y$
B3. $x + 0 = x$
B4. $x + (y + 1) = (x + y) + 1$
B5. $x \times 0 = 0$
B6. $x \times (y + 1) = (x \times y) + x$
B7. $(x \leq y \wedge y \leq x) \rightarrow x = y$
B8. $x \leq x + y$
C. $0 + 1 = 1$

Note that by addition of 1s we can build terms that represent arbitrary natural numbers, e.g. $1+1+1+1$ for 4. This is cumbersome to write down.

[i] Our theory **BA** is slightly weaker than Robinson's **Q**.

Therefore, we define $\underline{0} \equiv 0, \underline{1} \equiv 1$ and $\underline{n+1} \equiv (\underline{n}+1)$ for $n \geq 1$ and we call $\underline{n}$ a *numeral*. Lacking induction, the theory **BA** already contains all true quantifier-free sentences over $\mathcal{L}_{\mathcal{A}}$. In order to gain some "confidence" in the axioms involved, we prove this theorem in full detail. In subsequent proofs, we will leave out more details.

Theorem 2.11. *Let A be a sentence over $\mathcal{L}_{\mathcal{A}}$ not containing any quantifiers. Then*

$$\underline{\mathbb{N}} \models A \quad \Leftrightarrow \quad \mathbf{BA} \models A.$$

Proof. Axioms **B1** to **B8** and **C** are valid in $\underline{\mathbb{N}}$. Therefore $\underline{\mathbb{N}} \models \mathbf{BA}$ and $\underline{\mathbb{N}} \models A$. For the direction $\Rightarrow$ we need to establish step by step theorems of **BA**.

Claim 1: For $m, n \in \mathbb{N}$, if $m < n$, then $\mathbf{BA} \models \underline{m} \neq \underline{n}$.
Proof: By induction on m. We first show the case $m = 0$. Since free variables are implicitly universally quantified in the context of logical consequence (see definition 2.5), we have $\mathbf{B1} \models \underline{n-1}+1 \neq 0 \equiv \underline{n} \neq 0$ and because $=$ is always interpreted as true equality, we have $\mathbf{BA} \models \underline{0} \neq \underline{n}$. However, the case where $n = 1$ needs special attention because it does not follow from **B1** alone[i]. But **C** ensures that the terms $0+1$ and 1 are mapped to the same individuals in models of **BA**. Then it follows from $\mathbf{B1} \models 0+1 \neq 0$ that $\mathbf{BA} \models 1 \neq 0$ [ii]. For the case $m > 0$ note that $\mathbf{B2} \models \underline{m-1} \neq \underline{n-1} \rightarrow \underline{m} \neq \underline{n}$ (contraposition). By the induction hypothesis we have $\mathbf{BA} \models \underline{m-1} \neq \underline{n-1}$. It follows that $\mathbf{BA} \models \underline{m} \neq \underline{n}$. [iii]

Claim 2: For all $m, n \in \mathbb{N}$, $\mathbf{BA} \models \underline{m} + \underline{n} = \underline{m+n}$. [iv]
Proof: By induction on n. The case $n = 0$ follows immediately from **B3**. For $n > 0$ we have

$$\mathbf{B4} \models \underline{m} + \underbrace{(\underline{n-1}+1)}_{\equiv \underline{n}} = (\underline{m} + \underline{n-1}) + 1. \tag{2.1}$$

By induction hypothesis we obtain $\mathbf{BA} \models \underline{m} + \underline{n-1} = \underline{m+n-1}$. Hence we can make a substitution in (2.1) and obtain $\mathbf{BA} \models \underline{m} + \underline{n} = \underline{m+n-1} + 1 \equiv \underline{m} + \underline{n} = \underline{m+n}$. But again, this only holds for $n > 1$. For the case $n = 1$, what we claimed follows from axiom $\mathbf{C} \equiv 0 + 1 = 1$ together with $\mathbf{B4} \models \underline{m} + (0+1) = (\underline{m}+0) + 1$ (again apply the induction hypothesis).

[i] Because $0 + 1 \not\equiv \underline{1}$

[ii] Note that terms that are interpreted the same way by models of **BA** are interchangeable.

[iii] If the reader does not feel comfortable with this step, she is invited to check that from $\mathcal{M} \models A$ and $\mathcal{M} \models A \rightarrow B \equiv \neg A \vee B$ it follows that $\mathcal{M} \models B$.

[iv] Note that on the right hand side '+' denotes "real world" addition, whereas on the left hand side it is the function symbol of our language $\mathcal{L}_{\mathcal{A}}$.

Claim 3: For all $m, n \in \mathbb{N}$, $\mathbf{BA} \models \underline{m} \times \underline{n} = \underline{m \times n}$.
Proof: By induction on n. The case $n = 0$ follows immediately from **B5**. For $n > 0$ we first threat the case where $n = 1$. We have $\mathbf{B6} \models \underline{m} \times (0 + 1) = (\underline{m} \times 0) + \underline{m}$. From **C** we conclude $\mathbf{BA} \models \underline{m} \times 1 = (\underline{m} \times 0) + \underline{m}$ and with **B4** we obtain $\mathbf{BA} \models \underline{m} \times 1 = 0 + \underline{m}$. By claim 2 we have $\mathbf{BA} \models 0 + \underline{m} = \underline{m}$ and thus $\mathbf{BA} \models \underline{m} \times 1 = \underline{m}$. If $n > 1$ we have

$$\mathbf{B6} \models \underline{m} \times \underbrace{(\underline{n-1} + 1)}_{\equiv \underline{n}} = (\underline{m} \times \underline{n-1}) + \underline{m}. \tag{2.2}$$

By the induction hypothesis we obtain $\mathbf{BA} \models \underline{m} \times \underline{n} = \underline{m \times n - m} + \underline{m}$ and by claim 2 we can replace the right hand side and obtain $\mathbf{BA} \models \underline{m} \times \underline{n} = \underline{m \times n}$.

Claim 4: If a closed term t is interpreted as n in the standard model $\underline{\mathbb{N}}$, then $\mathbf{BA} \models t = \underline{n}$.
Proof: By structural induction on closed terms t. If t is a 0-ary function symbol, i.e. $t \equiv 0$ ($t \equiv 1$), then $\mathbf{BA} \models 0 = 0$ ($\mathbf{BA} \models 1 = 1$) because $=$ is always the true equality relation. Case $t \equiv t_1 + t_2$: Assume t_1 (t_2) is interpreted as n_1 (n_2) in $\underline{\mathbb{N}}$. By the induction hypothesis we have $\mathbf{BA} \models t_1 = \underline{n_1}$ and $\mathbf{BA} \models t_2 = \underline{n_2}$. By claim 2 we have $\mathbf{BA} \models \underline{n_1} + \underline{n_2} = \underline{n_1 + n_2}$ and thus $\mathbf{BA} \models t_1 + t_2 = \underline{n_1 + n_2}$. Case $t \equiv t_1 \times t_2$: Analogously using claim 3.

From claim 4 it follows that if t and s are closed terms, then $\underline{\mathbb{N}} \models t = s$ implies $\mathbf{BA} \models t = s$. Given $\underline{\mathbb{N}} \models t \neq s$, assume $\underline{\mathbb{N}} \models t = \underline{m}$ and $\underline{\mathbb{N}} \models s = \underline{n}$. Then by claim 4 $\mathbf{BA} \models t = \underline{m}$ and $\mathbf{BA} \models s = \underline{n}$. We assume $m < n$ (without loss of generality) and conclude from claim 1 that $\mathbf{BA} \models \underline{m} \neq \underline{n}$ and hence $\mathbf{BA} \models t \neq s$. Therefore $\underline{\mathbb{N}} \models t \neq s$ implies $\mathbf{BA} \models t \neq s$.

Next we show that for any $m \leq n$, $\mathbf{BA} \models \underline{m} \leq \underline{n}$. By claim 2 we have $\mathbf{BA} \models \underline{n} = \underline{m} + \underline{k}$ for $k = n - m$. Then we can conclude from $\mathbf{B8} \models \underline{m} \leq \underline{m} + \underline{k}$ that $\mathbf{BA} \models \underline{m} \leq \underline{n}$. If not $m \leq n$, then $n < m$ and $\mathbf{BA} \models \underline{n} \neq \underline{m}$ by claim 1. By the above we obtain $\mathbf{BA} \models \underline{n} \leq \underline{m}$ [i] The contraposition of **B7** yields $\mathbf{BA} \models \underline{m} \neq \underline{n} \rightarrow \neg(\underline{m} \leq \underline{n} \wedge \underline{n} \leq \underline{m})$. Then it is obvious that $\mathbf{BA} \models \neg \underline{m} \leq \underline{n}$ holds.

Now we are ready to prove by structural induction on quantifier-free $\mathcal{L}_{\mathcal{A}}$-sentences A that $\underline{\mathbb{N}} \models A \Rightarrow \mathbf{BA} \models A$ and $\underline{\mathbb{N}} \models \neg A \Rightarrow \mathbf{BA} \models \neg A$. If A is atomic it has either the form $t = s$ or $t \leq s$. Both cases follow from the above.

Case $A \equiv A_1 \wedge A_2$: We we have $\underline{\mathbb{N}} \models A_1$ and $\underline{\mathbb{N}} \models A_2$. By the induction hypothesis we obtain $\mathbf{BA} \models A_1$ and $\mathbf{BA} \models A_2$ and hence $\mathbf{BA} \models A_1 \wedge A_2$ (because every model of **BA** makes A_1 *and* A_2 true). If $\underline{\mathbb{N}} \models \neg(A_1 \wedge A_2)$, then $\underline{\mathbb{N}} \not\models (A_1 \wedge A_2)$ (by definition). Assume (without loss of generality)

[i] Since $n < m \Rightarrow n \leq m$.

$\underline{\mathbb{N}} \not\models A_1$, then $\underline{\mathbb{N}} \models \neg A_1$ and $\mathbf{BA} \models \neg A_1$ by induction hypothesis. Since $\neg A_1 \models \neg(A_1 \wedge A_2)$ we have $\mathbf{BA} \models \neg(A_1 \wedge A_2)$ [i].

Case $A \equiv A_1 \vee A_2$: We have $\underline{\mathbb{N}} \models A_1$ or $N \models A_2$. Assume w.l.o.g. $\underline{\mathbb{N}} \models A_1$. By the induction hypothesis we get $\mathbf{BA} \models A_1$ and therefore $\mathbf{BA} \models A_1 \vee A_2$. If $\underline{\mathbb{N}} \models \neg(A_1 \vee A_2)$, then $\underline{\mathbb{N}} \models \neg A_1$ and $\underline{\mathbb{N}} \models \neg A_2$ [ii]. By induction hypothesis we obtain $\mathbf{BA} \models \neg A_1$ and $\mathbf{BA} \models \neg A_2$ and thus $\mathbf{BA} \models \neg A_1 \wedge \neg A_2$ and $\mathbf{BA} \models \neg(A_1 \vee A_2)$.

The case $A \equiv \neg A_1$ follows directly from the induction hypothesis. If $\underline{\mathbb{N}} \models \neg\neg A_1$ then $\underline{\mathbb{N}} \models A_1$ and $\mathbf{BA} \models A_1$ (induction hypothesis) and thus $\mathbf{BA} \models \neg\neg A_1$ (logical consequence). □

Consider the theory $\{A \in \mathcal{L}_{\mathcal{A}} \mid \underline{\mathbb{N}} \models A\}$ of all true arithmetical formulas. By Gödel's Incompleteness Theorem, this theory has no recursive set of axioms (otherwise it were incomplete or inconsistent). Also note that it is not even recursively enumerable.

BA is a very weak fragment of arithmetic (in the sense that it is only a very small subset of the theory of all true formulas). In order to obtain stronger theories we need the concept of *induction*, which we formulate in terms of a set of axioms, called an *axiom scheme*.

Definition 2.12 (Induction axiom scheme)**.** *Let Φ be a set of formulas.* **Φ-IND** *is the set of formulas of the form*

$$\Big(A(0) \wedge \forall x\big(A(x) \to A(x+1)\big)\Big) \to \forall x A(x)$$

where $A \in \Phi$. Note that $A(x)$ can have free variables other than x.

We extend **BA** by induction to obtain the famous theory of *Peano Arithmetic*, denoted **PA**.

Definition 2.13 (**PA**)**. PA** *is the theory axiomatised by* **B1** *to* **B8** *and the* **Φ-IND** *axioms, where* **Φ** *is the set of all $\mathcal{L}_{\mathcal{A}}$-formulas.*

Remark 2.14. Axiom **C** follows from the other axioms and induction.

Peano Arithmetic is a very strong theory. By restricting the induction axiom scheme to so-called *bounded formulas*, we obtain theories of *bounded arithmetic* which are subtheories of **PA**.

[i] Recall that theories are closed under logical consequence

[ii] By using De Morgan's laws. E.g. it is easy to check that $\mathcal{M} \models \neg(A \vee B) \Rightarrow \mathcal{M} \models \neg A \wedge \neg B$.

Definition 2.15 (Bounded Formula)**.** *Assume that the variable x does not occur in the term t. Then $\exists x \leq tA$ is an abbreviation for $\exists x(x \leq t \wedge A)$, and $\forall x \leq t$ stands for $\forall x(x \leq t \rightarrow A)$. Quantifiers of this form are called bounded and a formula in which every quantifier is bounded is called a bounded formula.*

As a convention, we write $\exists \vec{x}$ instead of $\exists x_1 \exists x_2 \ldots \exists x_k$ and $\forall \vec{x}$ instead of $\forall x_1 \forall x_2 \ldots \forall x_k$ (for some $k \geq 0$).

Definition 2.16 (Δ_0, Σ_1)**.** *Δ_0 is the set of all bounded formulas. Σ_1 is the set of formulas of the form $\exists x A$, where $A \in \Delta_0$.*

Definition 2.17 $(\mathbf{I\Delta_0})$**.** $\mathbf{I\Delta_0}$ *is the theory specified by the same axioms as* **PA** *with the exception of* $\mathbf{\Phi}$ *being the set Δ_0 of bounded formulas.*

$\mathbf{I\Delta_0}$ is much stronger than **BA** and contains already all true Σ_1-sentences over $\mathcal{L}_{\mathcal{A}}$. The proof of the next theorem is very similar to the proof of theorem 2.11 and is omitted.

Theorem 2.18. *Let A be a Σ_1-sentence over $\mathcal{L}_{\mathcal{A}}$. Then*

$$\underline{\mathbb{N}} \models A \quad \Leftrightarrow \quad \mathbf{I\Delta_0} \models A.$$

2.2.2 Introducing new Functions and Predicates

We are interested in what functions and predicates are *definable* in a given theory. In particular, we try to find theories of which the definable functions are precisely the functions of a specific complexity class, for example the polynomial time computable functions. To this end we introduce the concept of definability of functions and predicates. For convenience (and without loss of generality) we talk only about functions and relations in the natural numbers and we often use the same symbol for a function (relation) in the real world and for the function (relation) in our language.

Definition 2.19 (Unique Existence $\exists!$)**.** *The notation $\exists! x A(x)$ stands for $\exists x\big(A(x) \wedge \forall y(A(y) \rightarrow x = y)\big)$.*

Definition 2.20 (Definable Predicates and Functions)**.** *Let $\mathcal{L} \supseteq \mathcal{L}_{\mathcal{A}}$ be a language and let Φ be a set of $\mathcal{L}$-formulas.*
(a) Let $R(\vec{x}) \subseteq \mathbb{N}^n$ be a (real world) n-ary relation and assume that the symbol R is not in $\mathcal{L}$. Further assume that $\underline{\mathbb{N}}'$ is an expansion of the standard model $\underline{\mathbb{N}}$ with $R^{\underline{\mathbb{N}}'} = R$ and the extra symbols in $\mathcal{L} \setminus \mathcal{L}_{\mathcal{A}}$ get their intended

interpretations. We call R $\mathbf{\Phi}$-definable if there is a formula $A(\vec{x}) \in \mathbf{\Phi}$ such that

$$\underline{\mathbb{N}}' \models R(\vec{x}) \leftrightarrow A(\vec{x}).$$

We then call $R(\vec{x}) \leftrightarrow A(\vec{x})$ the defining axiom *for R.*

(b) Let $\mathbf{T}$ be a theory over $\mathcal{L} \supseteq \mathcal{L}_{\mathcal{A}}$. Let $f(\vec{x}) : \mathbb{N}^n \to \mathbb{N}$ be a (real world) n-ary function and assume that the function symbol f is not in $\mathcal{L}$. Further assume that $\underline{\mathbb{N}}'$ is an expansion of $\underline{\mathbb{N}}$ with $f^{\underline{\mathbb{N}}'} = f \subseteq \mathbb{N}^n \times \mathbb{N}$ and the extra symbols in $\mathcal{L} \setminus \mathcal{L}_{\mathcal{A}}$ get their intended interpretations. We call f $\mathbf{\Phi}$-definable in $\mathbf{T}$ if its graph (relation) is $\mathbf{\Phi}$-definable and $\mathbf{T} \models \forall \vec{x} \exists! y A(\vec{x}, y)$ (which we call the totality condition*). That is if there is a formula $A(\vec{x}, y) \in \mathbf{\Phi}$ such that*

$$\underline{\mathbb{N}}' \models y = f(\vec{x}) \leftrightarrow A(\vec{x}, y).$$

We then call $y = f(\vec{x}) \leftrightarrow A(\vec{x}, y)$ the defining axiom *for f.*

Note that the definability of relations does not depend on a particular theory but only on the underlying language and the standard model (or an expansion thereof). Also note that when Φ is a set of $\mathcal{L}_{\mathcal{A}}$-formulas, then a k-ary relation over $\mathbb{N}$ is Φ-definable iff there is a formula $A(x_1, \ldots, x_k) \in \Phi$ s.t. for all $(n_1, \ldots, n_k) \in \mathbb{N}^k$:

$$(n_1, \ldots, n_k) \in R \quad \text{iff} \quad \underline{\mathbb{N}} \models A(\underline{n_1}, \ldots, \underline{n_k}).$$

As an example we will proof that:

Lemma 2.21. *The relation* pow2(x) $\Leftrightarrow$ *"x is a power of 2" is Δ_0-definable.*

Proof. The defining axiom[i] of pow2(x) is

$$\text{pow2}(x) \quad \leftrightarrow \quad x \neq 0 \ \wedge \ \forall y \leq x \big((1 < y \wedge y | x) \rightarrow 2 | y \big) \tag{2.3}$$

where $|$ is the divisibility relation (in infix notation) that has the Δ_0-defining axiom $x|y \leftrightarrow \exists z \leq y (x \times z = y)$. Hence every subformula of the form $t|s$ in (2.3) can be replaced with $\exists z \leq s (t \times z = s)$ in order to obtain a Δ_0-defining axiom of pow2. □

Definition 2.22 (provably total function)**.** *A function that is Σ_1-definable in $\mathbf{T}$ is called* provably total *(or provably recursive) in $\mathbf{T}$.*

[i]There might be more than one defining axiom, but we generally speak of *the* defining axiom.

It turns out ([10]) that the Δ_0-definable relations coincide with the relations in the linear time hierarchy. Further, the provably total functions of $\mathbf{I\Delta_0}$ are precisely the functions whose graph relation is in the linear time hierarchy. Later, we will show that the theory $\mathbf{V^1}$ characterises the polynomial time computable functions in the same way as $\mathbf{I\Delta_0}$ characterises the functions in the linear time hierarchy.

3 Two-sorted First-order Logic

In section 4 we will introduce the theory $\mathbf{V^1}$, a theory of so-called two-sorted first-order logic, which is an extension of (ordinary) first-order logic that contains two "sorts" of terms. In our case one sort is for numbers and the other sort is for finite sets of numbers. In this section we define the syntax and the semantics of this logic and present a two-sorted version LK^2 of Gentzen's ([7]) sequent calculus.

3.1 Syntax

Two-sorted first-order logic is an extension of (one-sorted) first-order logic. The difference is that in the two-sorted context we have variables of two sorts. The variables of the first sort are denoted by small letters $(x, y, z, \dots)$ and are called *number variables*. The variables of the second sort are denoted by capital letters $(X, Y, Z, \dots)$ and are called *set variables* (sometimes string variables). Number variables are intended to range over the natural numbers and set variables over sets of natural numbers (which can be represented by binary strings, as we will see later). Function and predicate symbols can also take arguments of both sorts and there are two kinds of function symbols (number and set function symbols).

Of course, in general, the domain of number and set variables can be *any* set. But since our domains of interest are natural numbers and finite sets of natural numbers we call them number and set variables, respectively.

Definition 3.1 (Two-sorted Language)**.** *For each $n, m \in \mathbb{N}$, a two-sorted (first-order) language is like a one-sorted language with the exception that it contains in addition an infinite set of set variables, a set of (n, m)-ary* number function symbols*, a set of (n, m)-ary* set function symbols *and a set of (n, m)-ary* predicate symbols.

The idea behind this is that functions and predicates take n arguments of the first sort (numbers) and m arguments of the second sort (sets). Number functions are into the natural numbers and set functions into finite sets of natural numbers. We use $f, g, h, \dots$ as meta-symbols for number function symbols, $F, G, H, \dots$ for set function symbols and $P, Q, R, \dots$ for predicate symbols. As in the one-sorted case, we define terms and formulas. Note that from now on, we assume that our languages are two-sorted without explicitly mentioning it. Also, we omit to mention the underlying language $\mathcal{L}$ if it is clear from the context.

Although We use capital letters for formulas $(A, B, C, \dots)$, set variables $(X, Y, Z, \dots)$, set terms $(T, S, \dots)$, relations $(P, Q, R, \dots)$ and set functions

$(F, G, H, \dots)$, there should be no confusion. The meaning of the meta variables is always evident from the context.

Definition 3.2 ($\mathcal{L}$-term)**.** *1. Every number variable is a number term.*

2. Every set variable is a set term.

3. If f is an (n,m)-ary number function symbol, $t_1, \dots, t_n$ are n number terms and $T_1, \dots, T_m$ are m set terms, then $f(t_1, \dots, t_n, T_1, \dots, T_m)$ is a number term.

4. If F is an (n,m)-ary set function symbol, $t_1, \dots, t_n$ and $T_1, \dots, T_m$ are as above, then $F(t_1, \dots, t_n, T_1, \dots, T_m)$ is a set term.

Definition 3.3 ($\mathcal{L}$-formula)**.** *1. If P is an (n,m)-ary predicate symbol, $t_1, \dots, t_n$ are number terms and $T_1, \dots, T_m$ are set terms, then $P(t_1, \dots, t_n, T_1, \dots, T_m)$ is an* atomic *formula.*

2. If A, B are formulas, so are $A \vee B, A \wedge B, \neg A$.

3. If A is a formula, x is a number variable and X is a set variable, then $\forall x A, \exists x A, \forall X A, \exists X A$ are formulas.

We extend our language of arithmetic $\mathcal{L}_{\mathcal{A}}$ to a two-sorted version $\mathcal{L}^2_{\mathcal{A}}$.

Definition 3.4 ($\mathcal{L}^2_{\mathcal{A}}$)**.** $\mathcal{L}^2_{\mathcal{A}} = (0, 1, +, \times, ||, =_1, =_2, \leq, \in)$

$0, 1, +, \times$ are the (number) function symbols from $\mathcal{L}_{\mathcal{A}}$ [i], $\leq$ and $=_1$ are $(2,0)$-ary predicate symbols (where $=_1$ is the original $=$). $||$ is an $(0,1)$-ary number function symbol and its interpretation will be the *least upperbound function* of a set X (roughly speaking the length of the binary representation of X). We will write $|X|$ instead of $||(X)$. The $(1,1)$-ary predicate symbol $\in$ will be interpreted as *set membership* and we write $x \in X$ instead of $\in (x, X)$. Finally, the $(0,2)$-ary predicate symbol $=_2$ will be interpreted as *set equality*. Whenever the meaning is clear from the context we will simply write $=$ instead of $=_1$ or $=_2$, respectively. Note that $\mathcal{L}^2_{\mathcal{A}}$ has no set terms except set variables. In the following we will work exclusively with extensions of $\mathcal{L}^2_{\mathcal{A}}$. We now generalise the notion of *universal closure* of formulas and sets of formulas.

Definition 3.5 (Universal Closure)**.** *Let $A(\vec{x}, \vec{X})$ be a formula with all free variables indicated. Then $\forall A(\vec{x}, \vec{X})$ stands for* $\forall x_1 \dots \forall x_m \forall X_1 \dots X_n A(x_1, \dots, x_m, X_1, \dots, X_n)$ *and is called* universal closure *of the formula A. If Φ is a set of formulas, $\forall \Phi$ denotes the set of universal closures of the formulas in Φ.*

[i] 0 and 1 are $(0,0)$-ary, $+$ and $\times$ are $(2,0)$-ary.

3.2 Semantics

Semantics is a generalisation of the one-sorted case. The main difference is that the universe of a structure consists of two sets instead of one.

Definition 3.6 ($\mathcal{L}$-structure)**.** *An $\mathcal{L}$-structure $\mathcal{M}$ consists of:*

1. *a pair of nonempty sets $(\mathsf{M}_1, \mathsf{M}_2)$ (the* universe*),*
2. *for each (n, m)-ary number function symbol f an interpretation $f^{\mathcal{M}} : \mathsf{M}_1^n \times \mathsf{M}_2^m \to \mathsf{M}_1$,*
3. *for each (n, m)-ary set function symbol F an interpretation $F^{\mathcal{M}} : \mathsf{M}_1^n \times \mathsf{M}_2^m \to \mathsf{M}_2$,*
4. *for each (n, m)-ary predicate symbol P an interpretation $P^{\mathcal{M}} \subseteq \mathsf{M}_1^n \times \mathsf{M}_2^m$. Further, $=_1^{\mathcal{M}}$ and $=_2^{\mathcal{M}}$ are the "true" equality relations on numbers and sets, respectively.*

To obtain interpretations for terms with free variables (possibly of both sorts) we generalise the notion of a variable assignment. A variable assignment σ is a mapping from the set of number variables to M_1 and from the set of set variables to M_2. As in the one-sorted case, the assignment $\sigma(\mathsf{m}/x)$ is the same as σ with the exception that it maps x to $\mathsf{m} \in \mathsf{M}_1$. The same applies for $\sigma(\mathsf{M}/X)$.

The interpretation $t^{\mathcal{M}}[\sigma]$ of a term t (T) in a two-sorted structure $\mathcal{M}$ with respect to an assignment σ generalises as follows:

1. If t is a number variable x, then $t^{\mathcal{M}}[\sigma]$ is $\sigma(x)$.
2. If T is a set variable X, then $T^{\mathcal{M}}[\sigma]$ is $\sigma(X)$.
3. If t is of the form $f(t_1, \ldots, t_n, T_1, \ldots, T_m)$ then $t^{\mathcal{M}}[\sigma]$ is $f^{\mathcal{M}}(t_1^{\mathcal{M}}[\sigma], \ldots, t_n^{\mathcal{M}}[\sigma], T_1^{\mathcal{M}}[\sigma], \ldots, T_m^{\mathcal{M}}[\sigma])$.
4. If T is of the form $F(t_1, \ldots, t_n, T_1, \ldots, T_m)$ then $T^{\mathcal{M}}[\sigma]$ is $F^{\mathcal{M}}(t_1^{\mathcal{M}}[\sigma], \ldots, t_n^{\mathcal{M}}[\sigma], T_1^{\mathcal{M}}[\sigma], \ldots, T_m^{\mathcal{M}}[\sigma])$.

The *truth value* of a (two-sorted) $\mathcal{L}$-formula A with respect to a structure $\mathcal{M}$ and an assignment σ is defined accordingly on the built up of formulas:

1. $\mathcal{M} \models P(t_1, \ldots, t_n, T_1, \ldots, T_m)[\sigma]$ iff $(t_1^{\mathcal{M}}[\sigma], \ldots, t_n^{\mathcal{M}}[\sigma], T_1^{\mathcal{M}}[\sigma], \ldots, T_m^{\mathcal{M}}[\sigma]) \in P^{\mathcal{M}}$.
2. If $\mathcal{L}$ contains $=_1$, then $\mathcal{M} \models (s =_1 t)[\sigma]$ iff $s^{\mathcal{M}}[\sigma] = t^{\mathcal{M}}[\sigma]$.

3. If $\mathcal{L}$ contains $=_2$, then $\mathcal{M} \models (S =_2 T)[\sigma]$ iff $S^{\mathcal{M}}[\sigma] = T^{\mathcal{M}}[\sigma]$.
4. $\mathcal{M} \models \neg A[\sigma]$ iff $\mathcal{M} \models A[\sigma]$ does not hold (written $\mathcal{M} \not\models A[\sigma]$).
5. $\mathcal{M} \models (A \wedge B)[\sigma]$ iff $\mathcal{M} \models A[\sigma]$ and $\mathcal{M} \models B[\sigma]$.
6. $\mathcal{M} \models (A \vee B)[\sigma]$ iff $\mathcal{M} \models A[\sigma]$ or $\mathcal{M} \models B[\sigma]$.
7. $\mathcal{M} \models \forall x A[\sigma]$ iff $\mathcal{M} \models A[\sigma(\mathsf{m}/x)]$ for all $\mathsf{m} \in \mathsf{M}_1$.
8. $\mathcal{M} \models \forall X A[\sigma]$ iff $\mathcal{M} \models A[\sigma(\mathsf{M}/X)]$ for all $\mathsf{M} \in \mathsf{M}_2$.
9. $\mathcal{M} \models \exists x A[\sigma]$ iff $\mathcal{M} \models B[\sigma(\mathsf{m}/x)]$ for at least one $\mathsf{m} \in \mathsf{M}_1$.
10. $\mathcal{M} \models \exists X A[\sigma]$ iff $\mathcal{M} \models B[\sigma(\mathsf{M}/X)]$ for at least one $\mathsf{M} \in \mathsf{M}_2$.

Definition 3.7 (Two-sorted Standard Model $\underline{\mathbb{N}_2}$)**.** *The $\mathcal{L}^2_{\mathcal{A}}$-structure $\underline{\mathbb{N}_2}$ has universe $(\mathbb{N}, \mathcal{P}_{\mathbf{f}}(\mathbb{N}))$. The symbols $0, 1, +, \times, =, \leq$ have the same interpretation as in $\underline{\mathbb{N}}$. $\in$ is interpreted as the set membership relation. $||$ is interpreted as the "least upper bound" function on finite subsets of $\mathbb{N}$* [i]*. We call $\underline{\mathbb{N}_2}$ the (two-sorted) standard model (of the natural numbers).*

The notions of *model, validity, satisfiability, logical consequence and theory* generalise in the obvious way to the two-sorted case. Also, all results from section 2.1 continue to hold.

In two sorted first-order logic we have "second order" objects (sets of natural numbers). Nevertheless, two-sorted first-order logic is equivalent to one-sorted first-order logic because one can merge the two sorts by introducing additional unary predicate symbols FS and SS, together with some appropriate axioms, to identify the two sorts. For details, see [6].

[i] That is 1 plus the greatest element of a set S or 0 if S is empty.

3.3 The Two-sorted Sequent Calculus LK^2

3.3.1 Rules and proofs

The sequent calculus is a deduction formalism that does not derive formulas but so-called sequents. A sequent is an expression of the form $\Gamma \vdash \Delta$, where Γ and Δ are finite (possibly empty) sequences of (two-sorted first-order) formulas, called *cedents*. Γ is called the *antecedent* and Δ is called the *succedent*. Cedents need to be given a *meaning*. Informally, the conjunction of the formulas in Γ implies the disjunction of the formulas in Δ. Formally, for $\Gamma = A_1, \ldots, A_m$ and $\Delta = B_1, \ldots, B_n$ ($\emptyset$ denotes the empty sequence)

$$\begin{aligned}
(\Gamma \vdash \Delta)^\star &\equiv (A_1 \wedge \cdots \wedge A_m) \rightarrow (B_1 \vee \cdots \vee B_n),\\
(\Gamma \vdash \emptyset)^\star &\equiv (A_1 \wedge \cdots \wedge A_m) \rightarrow \bot,\\
(\emptyset \vdash \Delta)^\star &\equiv (B_1 \vee \cdots \vee B_n),\\
(\emptyset \vdash \emptyset)^\star &\equiv \bot.
\end{aligned}$$

The symbol $\bot$ (falsity) is not part of our languages. But since each language contains at least one predicate symbol P, we can define $\bot$ as an abbreviation for $\forall \vec{x} \forall \vec{X} \big(P(\vec{x}, \vec{X}) \wedge \neg P(\vec{x}, \vec{X}) \big)$. It is obvious that $\mathcal{M}[\sigma] \not\models \bot$ for every structure $\mathcal{M}$ and every assignment σ. The notions of validity, logical consequence etc. generalise from formulas to sequents in the obvious way. If no confusion arises, then we may write $\Gamma \vdash \Delta$ instead of $(\Gamma \vdash \Delta)^\star$. We now describe the axioms and rules of the sequent calculus LK^2. The only (logical) *axiom* of LK^2 is

$$A \ \vdash \ A$$

where A is any formula. In the following Γ, Δ (with superscripts) denote cedents. LK^2 consists of the following *structural rules*:

$$\text{(exchange-left)} \ \frac{\Gamma', A, B, \Gamma'' \ \vdash \ \Delta}{\Gamma', B, A, \Gamma'' \ \vdash \ \Delta} \qquad \text{(exchange-right)} \ \frac{\Gamma \ \vdash \ \Delta', A, B, \Delta''}{\Gamma \ \vdash \ \Delta', B, A, \Delta''}$$

$$\text{(weakening-left)} \ \frac{\Gamma \ \vdash \ \Delta}{\Gamma, A \ \vdash \ \Delta} \qquad \text{(weakening-right)} \ \frac{\Gamma \ \vdash \ \Delta}{\Gamma \ \vdash \ \Delta, A}$$

$$\text{(contraction-left)} \ \frac{\Gamma, A, A \ \vdash \ \Delta}{\Gamma, A \ \vdash \ \Delta} \qquad \text{(contraction-right)} \ \frac{\Gamma \ \vdash \ \Delta, A, A}{\Gamma \ \vdash \ \Delta, A}$$

LK^2 has the following *propositional rules*:

$$(\neg\text{-left}) \ \frac{\Gamma \ \vdash \ \Delta, A}{\Gamma, \neg A \ \vdash \ \Delta} \qquad (\neg\text{-right}) \ \frac{\Gamma, A \ \vdash \ \Delta}{\Gamma \ \vdash \ \Delta, \neg A}$$

$$(\vee\text{-left})\ \frac{\Gamma, A \vdash \Delta \qquad \Gamma, B \vdash \Delta}{\Gamma, A \vee B \vdash \Delta} \qquad (\vee\text{-right})\ \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \vee B}$$

$$(\wedge\text{-left})\ \frac{\Gamma, A, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta} \qquad (\wedge\text{-right})\ \frac{\Gamma \vdash \Delta, A \qquad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B}$$

LK^2 has the following *quantifier rules*:

$$(\exists\text{-left})\ \frac{\Gamma, A(b) \vdash \Delta}{\Gamma, \exists x A(x) \vdash \Delta} \qquad (\exists\text{-right})\ \frac{\Gamma, \vdash \Delta, A(t)}{\Gamma \vdash \Delta, \exists x A(x)}$$

$$(\text{set } \exists\text{-left})\ \frac{\Gamma, A(M) \vdash \Delta}{\Gamma, \exists X A(X) \vdash \Delta} \qquad (\text{set } \exists\text{-right})\ \frac{\Gamma, \vdash \Delta, A(T)}{\Gamma \vdash \Delta, \exists X A(X)}$$

$$(\forall\text{-left})\ \frac{\Gamma, A(t) \vdash \Delta}{\Gamma, \forall x A(x) \vdash \Delta} \qquad (\forall\text{-right})\ \frac{\Gamma, \vdash \Delta, A(b)}{\Gamma \vdash \Delta, \forall x A(x)}$$

$$(\text{set } \forall\text{-left})\ \frac{\Gamma, A(T) \vdash \Delta}{\Gamma, \forall X A(X) \vdash \Delta} \qquad (\text{set } \forall\text{-right})\ \frac{\Gamma, \vdash \Delta, A(M)}{\Gamma \vdash \Delta, \forall X A(X)}$$

where t is any number term and T is any set term. The free variables b and M are called *eigenvariables* and must not occur in $\Gamma \cup \Delta$. Otherwise the calculus were not sound.[i] Finally, LK^2 contains the *cut rule*:

$$(\text{cut})\ \frac{\Gamma \vdash \Delta, A \qquad \Gamma, A \vdash \Delta}{\Gamma \vdash \Delta}$$

Definition 3.8 (LK^2 proof)**.** *An LK^2 proof of a sequent S is a finite tree where the vertices are labelled with sequents, the root is labelled with S, the leaves are labelled with (logical) axioms of LK^2 and every edge of the tree corresponds to a rule of LK^2 in the obvious way. An LK^2 proof of a formula A is a proof of the sequent $\vdash A$.*

Definition 3.9 (LK^2-Φ proof)**.** *An LK^2-Φ proof of a sequent S an LK^2 proof of S in which sequents at leaves are either (logical) axioms or (non-logical) axioms of the form $\vdash A$ where $A \in \Phi$.*

[i] To see this, consider the (valid) axiom $A(b) \vdash A(b)$. By the $\forall$-right rule we obtain the sequent $A(b) \vdash \forall x A(x)$, which is not valid. Accordingly, $\exists x A(x) \vdash A(b)$ is not valid.

3.3.2 Soundness and Completeness

Theorem 3.10 (Derivational Soundness of LK^2 with Equality)**.** *Let $\mathcal{L}$ be a two-sorted first-order language (it may contain $=_1$ and $=_2$) and let Φ be a set of $\mathcal{L}$-formulas. If a sequent $\Gamma \vdash \Delta$ has an LK^2-Φ proof, then $(\Gamma \vdash \Delta)^\star$ is a logical consequence of Φ.*

Proof. The theorem is proved by induction on the number of sequents in a proof. Assuming that there is an LK^2-Φ proof of a sequent S, we need to show that $\Phi \models S^\star$.
Base case: If S is a logical axiom, then $S^\star$ is obviously valid. If S is a nonlogical axiom of the form $\vdash A$, where $A \in \Phi$, then $(\vdash A)^\star \equiv A$ is obviously a logical consequence of Φ.
Induction step: It is easy to see that for every rule of LK^2, the bottom sequent is a logical consequence of the top sequent. Then the lemma follows immediately from the transitivity of the logical consequence relation $\models$. □

Variable convention In the context of LK^2, it is convenient to distinguish between free and bound variables. We denote free number variables by $a, b, c, \ldots$, free set variables by $M, N, \ldots$[i], bound number variables by $x, y, z, \ldots$ and bound set variables by $X, Y, Z, \ldots$. In the following, sequents satisfy the restriction that no free variables occur as bound variables and vice versa. Note that with this restriction, any term t (T) is always *freely substitutable* for a (M) in $A(a)$ ($A(M)$).

Notation A double line in a derivation tree means that there are implicit applications of structural rules, as in the following example (three applications of the exchange-left rule):

$$\dfrac{\dfrac{A, B, C \vdash \Delta}{C, B, A \vdash \Delta}}{}$$

We will use the following well-known special case of König's Lemma[ii] in subsequent completeness proofs.

Lemma 3.11 (König's Lemma)**.** *Every tree that contains infinitely many vertices, each having finite degree, has at least one infinite path.*

In order to prove derivational completeness of LK^2 we first establish the following completeness lemma.

[i] Since $A, B, C, \ldots$ are already reserved as meta-variables for formulas.
[ii] The World Wide Web is full of proofs thereof.

Lemma 3.12 (Completeness Lemma)**.** *Let $\mathcal{L}$ be a countable two-sorted first-order language not containing $=_1$ and $=_2$ and let Φ be a (possibly infinite) set of $\mathcal{L}$-sentences. If $(\Gamma \vdash \Delta)^\bigstar$ is a logical consequence of Φ, then there exists a finite subset $\{C_1, \ldots, C_n\} \subseteq \Phi$ such that the sequent*

$$\Gamma, C_1, \ldots, C_n \vdash \Delta$$

has a cut-free LK^2 proof.

Proof. We adapt the proof in [3] to the two-sorted case and assume that Φ is countable. Assume $\Phi \models (\Gamma \vdash \Delta)^\bigstar$. We assign a (distinct) binary string to each function symbol, predicate symbol, variable and to all logical connectives, quantifiers and parentheses. Hence every $\mathcal{L}$-formula and every $\mathcal{L}$-term can be assigned a unique binary string. This allows us to build lists of all formulas and of all set and number terms (which, by our variable convention, do not contain bound variables). We further build these lists in such a way that each formula and each term occurs infinitely often in the list, e.g. by enumerating them in the order $1., 1., 2., 1., 2., 3., 1., 2., 3., 4., \ldots$. Let $A_1, A_2, \ldots$, $t_1, t_2, \ldots$ and $T_1, T_2, \ldots$ denote these lists, respectively. We then enumerate tuples (A_i, t_j, T_k) in such a way that all combinations occur infinitely often. We need this property later in the construction of the countermodel. A possible enumeration is

$$\begin{aligned}
&(A_1, t_1, T_1), (A_2, t_1, T_1), (A_2, t_1, T_2), (A_2, t_2, T_1), (A_2, t_2, T_2), (A_1, t_1, T_2),\\
&(A_1, t_2, T_1), (A_1, t_2, T_2), (A_3, t_1, T_1), (A_3, t_1, T_2), (A_3, t_1, T_3), (A_3, t_2, T_1),\\
&(A_3, t_2, T_2), (A_3, t_2, T_3), (A_3, t_3, T_1), (A_3, t_3, T_2), (A_3, t_3, T_3), (A_2, t_1, T_3),\\
&(A_2, t_2, T_3), (A_2, t_3, T_1), (A_2, t_3, T_2), (A_2, t_3, T_3), (A_2, t_1, T_3), (A_2, t_2, T_3),\\
&(A_2, t_3, T_1), (A_2, t_3, T_2), (A_2, t_3, T_3), \ldots
\end{aligned}$$

We now define a procedure that delivers a cut-free proof Π of some sequent

$$\Gamma, C_1, \ldots, C_n \vdash \Delta$$

with $C_1, \ldots, C_n \in \Phi$ for some n. We prove that the procedure always terminates by showing that, if it doesn't, then $\Phi \not\models (\Gamma \vdash \Delta)^\bigstar$. In the following, we call a sequent *active* if it is a leaf of the proof tree and is not directly derivable from an axiom of LK^2 using only weakening and exchange rules. We begin with an end sequent $\Gamma \vdash \Delta$ and work upwards by subsequently modifying the proof Π. We use the same name Π for the initial proof and its modifications.

Loop: Let (A_i, t_j, T_k) be the next tuple in the enumeration.

1. *Step:* If $A_i \in \Phi$, then replace *every* sequent $\Pi \vdash \Omega$ in $\mathsf{\Pi}$ with the sequent $\Pi, A_i \vdash \Omega$ (also the end sequent!). Continue with step 2.

2. *Step:* If A_i is atomic, do nothing and continue the loop.

3. *Step:* If A_i is not atomic, then do the following:

   Case (3a): If $A_i \equiv \neg B$, then every active sequent in $\mathsf{\Pi}$ of the form $\Pi', \neg B, \Pi'' \vdash \Omega$ is replaced with the following derivation:

   $$\dfrac{\Pi', \neg B, \Pi'' \ \vdash \ \Omega, B}{\Pi', \neg B, \Pi'' \ \vdash \ \Omega}$$

   and every active sequent in $\mathsf{\Pi}$ of the form $\Pi \vdash \Omega', \neg B, \Omega''$ is replaced by the derivation

   $$\dfrac{\Pi, B \ \vdash \ \Omega', \neg B, \Omega''}{\Pi \ \vdash \ \Omega', \neg B, \Omega''}$$

   Case (3b): If $A_i \equiv B \vee C$, then every active sequent in $\mathsf{\Pi}$ of the form $\Pi', B \vee C, \Pi'' \vdash \Omega$ is replaced with the derivation

   $$\dfrac{\Pi', B \vee C, \Pi'', B \ \vdash \ \Omega \qquad \Pi', B \vee C, \Pi'', C \ \vdash \ \Omega}{\Pi', B \vee C, \Pi'' \ \vdash \ \Omega}$$

   and every active sequent in $\mathsf{\Pi}$ of the form $\Pi \vdash \Omega', B \vee C, \Omega''$ is replaced with the derivation

   $$\dfrac{\Pi \ \vdash \ \Omega', B \vee C, \Omega'', B, C}{\Pi \ \vdash \ \Omega', B \vee C, \Omega''}$$

   Case (3c): The case $A_i \equiv B \wedge C$ is dual to the case $A_i \equiv B \vee C$.

   Case (3d): If $A_i \equiv (\exists x)B(x)$, then every active sequent in $\mathsf{\Pi}$ of the form $\Pi', (\exists x)B(x), \Pi'' \vdash \Omega$ is replaced with the derivation

   $$\dfrac{B(c), \Pi', (\exists x)B(x), \Pi'' \ \vdash \ \Omega}{\Pi', (\exists x)B(x), \Pi'' \ \vdash \ \Omega}$$

   where c is a "fresh" free number variable not yet used in $\mathsf{\Pi}$ [i]. And every active sequent in $\mathsf{\Pi}$ of the form $\Pi \vdash \Omega', (\exists x)B(x), \Omega''$ is replaced with the derivation

   $$\dfrac{\Pi \ \vdash \ \Omega', (\exists x)B(x), \Omega'', B(t_j)}{\Pi \ \vdash \ \Omega', (\exists x)B(x), \Omega''}$$

[i] Note that since Φ contains no free variables, c does not occur in Φ.

Case (3e): If $A_i \equiv (\exists X)B(X)$, then every active sequent in Π of the form $\Pi', (\exists X)B(X), \Pi'' \vdash \Omega$ is replaced with the derivation

$$\frac{B(M), \Pi', (\exists X)B(X), \Pi'' \ \vdash \ \Omega}{\Pi', (\exists X)B(X), \Pi'' \ \vdash \ \Omega}$$

where M is a "fresh" free set variable not yet used in Π. And every active sequent in Π of the form $\Pi \vdash \Omega', (\exists X)B(X), \Omega''$ is replaced with the derivation

$$\frac{\Pi \ \vdash \ \Omega', (\exists X)B(X), \Omega'', B(T_k)}{\Pi \ \vdash \ \Omega', (\exists X)B(X), \Omega''}$$

Cases (3f) and (3g): The cases $A_i \equiv (\forall x)B(x)$ and $A_i \equiv (\forall X)B(X)$ are dual to the cases 3d and 3e, respectively.

4. *Step:* If there are no active sequents remaining in Π, then exit. Otherwise, continue with the next iteration.

End of the loop.

If the above procedure terminates, then Π is a proof of $\Gamma, C_1, \ldots, C_n \vdash \Delta$ for some $C_1, \ldots, C_n \in \Phi$ (we have to use the contraction rule to eliminate double occurrences of the C_is). To prove that it does terminate, we assume that it does not terminate and then show that this contradicts our assumption that $\Phi \models (\Gamma \vdash \Delta)^\bigstar$.

So assume that the above procedure runs forever. Then, in general, it builds an infinite tree. And this tree will give us a structure that disproves $\Phi \models (\Gamma \vdash \Delta)^\bigstar$. Let Π denote this tree. In general, Π is infinite. If Φ is nonempty, each vertice in Π is a *generalised* sequent of the form $\Pi, A_1, A_2, \cdots \vdash \Omega$ having an infinite number of formulas in its antecedent (eventually, all sequents of Π contain all formulas $A_1, A_2, \ldots$ of Φ since these are "thrown in" at step 1). Note, however, that at each step of the infinite construction process, every sequent of Π is finite. In the special case where Π is finite, Π has some active sequent containing only atomic formulas (otherwise the procedure would terminate). In this case let π be the branch going from the root of Π up to this active sequent. If Π is infinite, then it has infinitely many vertices and each vertice has finite degree (at most 2). Therefore Π contains an infinite path π (starting at the root) by König's lemma 3.11. Note that π defines a sequence of sequents.

We use π to construct a structure $\mathcal{M}$ and an assignment σ s.t. $\mathcal{M}[\sigma] \not\models (\Gamma \vdash \Delta)^\bigstar$ and $\mathcal{M} \models \Phi$, i.e. $\Phi \not\models (\Gamma \vdash \Delta)^\bigstar$. Let the universe of $\mathcal{M}$ be $(\mathsf{M}_1, \mathsf{M}_2)$ where M_1 is the set of all $\mathcal{L}$-number terms and M_2 is the set of all

$\mathcal{L}$-set terms. σ just maps a variable a (or A) to itself. A number function symbol f is interpreted s.t. $f^{\mathcal{M}}(\vec{t},\vec{T})$ is the term $f(\vec{t},\vec{T})$ (analogously for set function symbols F). So $\mathcal{M}$ and σ just map terms to themselves[i]. For any predicate symbol P, let $(\vec{t},\vec{T}) \in P^{\mathcal{M}}$ hold iff the formula $P(\vec{t},\vec{T})$ occurs in the antecedent of some sequent in π.

We claim that every formula A occurring in an antecedent along π is true in $\mathcal{M}[\sigma]$ and that every formula A occurring in a succedent along π is false in $\mathcal{M}[\sigma]$. We show this by structural induction on A. If A is atomic, the claim is true by the above definition of $\mathcal{M}$. Consider the case $A \equiv (\exists x)B(x)$. If A appears in an antecedent of π, then, according to our procedure, also a formula $B(c)$ appears in some antecedent of π. By the induction hypothesis, $\mathcal{M}[\sigma] \models B(c)$ and hence $\mathcal{M}[\sigma] \models (\exists x)B(x)$. If A appears in a succedent of π, then, for every number term t, $B(t)$ eventually occurs in a succedent (since in case (3d) we always keep a copy of $(\exists x)B(x)$ [ii] and since every term t_i appears infinitely often in our enumeration!). Therefore, for all t, $\mathcal{M}[\sigma] \not\models B(t)$ by induction hypothesis. This implies that $\mathcal{M}[\sigma] \not\models A$. The cases $A \equiv (\exists X)B(X)$, $A \equiv (\forall x)B(x)$ and $A \equiv (\forall X)B(X)$ are dual and the other cases are straight-forward. Note that A cannot occur in both an antecedent and a succedent of π, since then these formulas would persist upward in π s.t. some particular sequent S in π would have A occurring both in its antecedent and in its succedent. But then S would not be active and the branch π therefore not infinite.

Note that the end sequent of π is $\Gamma, C_1, C_2, \cdots \vdash \Delta$ where the infinite sequence $C_1, C_2, \ldots$ contains all formulas of Φ (even infinitely often). From the above claim we conclude that $\mathcal{M}[\sigma] \models \Gamma$, $\mathcal{M}[\sigma] \not\models \Delta$ and hence $\mathcal{M}[\sigma] \not\models (\Gamma \vdash \Delta)^{\bigstar}$. Since Φ contains no free variables[iii] we also have $\mathcal{M} \models \Phi$ and are done with the proof. □

Using the completeness lemma it is easy to show that the following derivational completeness theorem holds.

Theorem 3.13 (Derivational Completeness of LK^2 without Equality)**.** *Let $\mathcal{L}$ be a two-sorted first-order language not containing $=_1$ and $=_2$ and let Φ be a set of $\mathcal{L}$-formulas. If $(\Gamma \vdash \Delta)^{\bigstar}$ is a logical consequence of Φ, then $\Gamma \vdash \Delta$ has an LK^2-Φ proof.*

Proof. Let Φ be a set of $\mathcal{L}$-formulas such that $(\Gamma \vdash \Delta)^{\bigstar}$ is a logical consequence of Φ. Because Φ and $\forall\Phi$ are equivalent (i.e. have exactly the same

[i] A common technique; in German it is often called "Terminterpretation".

[ii] The cases where the $\exists$-right and $\forall$-left rules are applied are the only ones where it is really necessary to keep a copy of the "active" formula.

[iii] Note that the proof does not work when Φ is a set of formulas instead of sentences.

models), $(\Gamma \vdash \Delta)^\bigstar$ is also a logical consequence of the set of sentences $\forall\Phi$. By the completeness lemma 3.12, there are sentences $\forall C_1, \ldots, \forall C_n \in \forall\Phi$ such that

$$\Gamma, \forall C_1, \ldots, \forall C_n \vdash \Delta$$

has a cut-free LK^2 proof. Since every $\forall C_1, \ldots, \forall C_n$ has an LK^2-Φ proof (using the $\forall$-rules) we can use the cut rule n times (plus the weakening rules) to obtain an LK^2-Φ proof of $\Gamma \vdash \Delta$. □

It is evident that the above proofs can be easily generalised to the case of n-sorted first-order logic (for $n > 2$). Note that if the underlying language contains $=$, then the above completeness theorem does not hold because all our structures interpret $=$ as the true equality relation. For example, the valid formula $x = x$ does not have an LK^2 proof. We will soon resolve this grievance by adding special equality axioms to LK^2.

3.3.3 Eliminating Free Cuts

For the proof of the witnessing theorem of $\mathbf{V^1}$ (4.95) it is important that we can restrict applications of the cut rule to formulas in Φ. However, in order to do this, Φ needs to be closed under substitution of terms for free variables.

Definition 3.14. *A set Φ of formulas is called* closed under substitution of terms for free variables *if it satisfies the following condition: If $A(b) \in \Phi$ ($A(B) \in \Phi$), then also $A(t) \in \Phi$ ($A(T) \in \Phi$) where t (T) is any number (set) term.*

Definition 3.15 (Anchored LK^2-Φ Proof)**.** *An application of the cut rule in an LK^2-Φ proof Π is called anchored if its cut formula is in Φ. Π is called anchored if it contains only anchored applications of the cut rule.*

The term "anchored" is taken from [3] and [6]. Note that Buss' definition of anchored is slightly more complicated than the one of Cook we use here.

Theorem 3.16 (Anchored Completeness of LK^2 without Equality)**.** *Let $\mathcal{L}$ be a two-sorted first-order language not containing $=_1$ and $=_2$ and let Φ be a set of $\mathcal{L}$-formulas closed under substitution of terms for free variables. If $(\Gamma \vdash \Delta)^\bigstar$ is a logical consequence of Φ, then $\Gamma \vdash \Delta$ has an anchored LK^2-Φ proof.*

Proof. Note that if Φ are sentences, then the above follows from theorem 3.13 since then $\forall\Phi$ is the same as Φ. We slightly modify the proof of the completeness lemma 3.12. Now we try to find a proof of the sequent $\Gamma \vdash \Delta$ from the non-logical axioms Φ involving only cut formulas in Φ. We call a

sequent *active* if it is a leaf of the proof tree and it is not directly derivable from a logical axiom of LK^2 or from a non-logical axiom of Φ using only weakening and exchange rules. We begin with an end sequent $\Gamma \vdash \Delta$ and work upwards by subsequently modifying the proof Π.

Loop: Let (A_i, t_j, T_k) be the next tuple in the enumeration.

1. *Step:* If $A_i \in \Phi$, then every active sequent $\Pi \vdash \Omega$ is replaced with the derivation

$$\dfrac{\dfrac{\vdash\ A_i}{\Pi\ \vdash\ \Omega, A_i} \qquad \Pi, A_i\ \vdash\ \Omega}{\Pi\ \vdash\ \Omega}$$

   Then continue with step 2.

   (For the other steps, proceed as in steps 2 to 4 of the completeness lemma 3.12.)

End of the loop.

If the above procedure terminates, then Π is an anchored LK^2-Φ proof of $\Gamma \vdash \Delta$. To prove that it does terminate, we assume that it does not terminate and then show that this contradicts our assumption that $\Phi \models (\Gamma \vdash \Delta)^\bigstar$.

So assume that the above procedure runs forever and let Π be the result thereof. Then Π is an infinite tree. In general, we deal again with *generalised* sequents containing infinitely many formulas. Again, Π contains an infinite path π (starting at the root) by König's Lemma (3.11) and we use π to construct a structure $\mathcal{M}$ and an assignment σ in the same way as in the proof of the completeness lemma.

We claim again that every formula A occurring in an antecedent along π is true in $\mathcal{M}[\sigma]$ and that every formula A occurring in a succedent along π is false in $\mathcal{M}[\sigma]$. The argument is the same as in the proof of the completeness lemma 3.12. Therefore $\mathcal{M}[\sigma] \not\models (\Gamma \vdash \Delta)^\bigstar$. Note that the A_is in the application of the cut rule above do not occur in a succedent along the infinite path π (otherwise it were not infinite).

It remains to show that $\mathcal{M} \models \Phi$. Since π is an infinite path and every $A_i \in \Phi$ occurs infinitely often in the loop, the first step is applied infinitely often. Thus every $A_i(\vec{a}, \vec{M}) \in \Phi$ (with all free variables indicated) occurs in some antecedent in π and hence $\mathcal{M}[\sigma] \models A_i(\vec{a}, \vec{M})$, by the above claim. But since Φ is closed under substitution of terms for free variables, we also have $\mathcal{M}[\sigma] \models A_i(\vec{t}, \vec{T})$, for all terms $\vec{t}, \vec{T}$. Since the universes of $\mathcal{M}$ consist

precisely of all terms and since σ maps terms to themselves, it follows that $\mathcal{M}[\sigma] \models \forall\vec{x}\forall X A_i(\vec{x}, \vec{X})$ and hence $\mathcal{M}[\sigma] \models \forall\Phi$ and $\mathcal{M} \models \Phi$. Therefore we have showed that $\Phi \not\models (\Gamma \vdash \Delta)^\star$. □

3.3.4 The Special Treatment of the Equality Symbol(s)

If a language contains the equality symbols $=_1$ and $=_2$, it is natural to consider only structures that interpret these symbols as equality (section 3.2). Our definition of logical consequence is subject to this restriction. However, up to now, the calculus LK^2 is not. If we want to allow $=_1$ and $=_2$ in the underlying language $\mathcal{L}$, we need to formulate new axioms for LK^2, so-called equality axioms. For convenience we will often write $=$ instead of $=_1$ or $=_2$. It is always clear which symbol is meant. For the proof of lemma 3.21 below, we need the common notions of *equivalence relation* and *equivalence class*.

Definition 3.17 (Equivalence Relation)**.** *A relation $R \subseteq \mathsf{S} \times \mathsf{S}$ over a set S is called an equivalence relation if*

(a) $(x, x) \in R$ *for all* $x \in \mathsf{S}$ *(reflexivity),*

(b) $(x, y) \in R \Rightarrow (y, x) \in R$ *for all* $x, y \in \mathsf{S}$ *(symmetry),*

(c) $(x, y) \in R$ *and* $(y, z) \in R \Rightarrow (x, z) \in R$ *for all* $x, y, z \in \mathsf{S}$ *(transitivity).*

Definition 3.18 (Equivalence Class)**.** *Let R be an equivalence relation on some set S. Then*

$$[a]_R = \{x \in \mathsf{S} \mid (a, x) \in R\} \subseteq \mathsf{S}$$

is called the equivalence class of a under R.

Definition 3.19 (Equality Axioms $\varepsilon_\mathcal{L}$)**.** *Let $\vec{x} = \vec{y}$ stand for $x_1 = y_1 \wedge \cdots \wedge x_n = y_n$ (accordingly for $\vec{X} = \vec{Y}$). The set $\varepsilon_\mathcal{L}$ of equality axioms of $\mathcal{L}$ contains the axioms below and is closed under replacement of terms for free variables.*

E1′. $x = x$
E1″. $X = X$
E2′. $x = y \rightarrow y = x$
E2″. $X = Y \rightarrow Y = X$
E3′. $x = y \wedge y = z \rightarrow x = z$
E3″. $X = Y \wedge Y = Z \rightarrow X = Z$
E4′. $\vec{x} = \vec{y} \wedge \vec{X} = \vec{Y} \rightarrow f(\vec{x}, \vec{X}) = f(\vec{y}, \vec{Y})$
E4″. $\vec{x} = \vec{y} \wedge \vec{X} = \vec{Y} \rightarrow F(\vec{x}, \vec{X}) = F(\vec{y}, \vec{Y})$
E5. $\vec{x} = \vec{y} \wedge \vec{X} = \vec{Y} \wedge P(\vec{x}, \vec{X}) \rightarrow P(\vec{y}, \vec{Y})$

Definition 3.20 (Weak Structure)**.** *A weak $\mathcal{L}$-structure is like a (proper) $\mathcal{L}$-structure with the exception that $=_1, =_2$ can be interpreted as any relation.*

Note that every (proper) structure is a weak structure.

Lemma 3.21. *For every weak model $\mathcal{M}$ of $\varepsilon_{\mathcal{L}}$, there exists a proper $\mathcal{L}$-structure $\mathcal{M}'$ s.t. $\mathcal{M}$ and $\mathcal{M}'$ satisfy the same formulas.*

Proof. Let $\mathcal{M}$ be a weak structure with universe $(\mathsf{M}_1, \mathsf{M}_2)$ and $\mathcal{M} \models \varepsilon_{\mathcal{L}}$. We have to construct a proper structure $\mathcal{M}'$ with $\mathcal{M}' \models A \Leftrightarrow \mathcal{M} \models A$, for every $\mathcal{L}$-formula A. Let $\mathsf{x}_1, \ldots, \mathsf{x}_n, \mathsf{y}_1, \ldots, \mathsf{y}_n \in \mathsf{M}_1$ and let $\mathsf{X}_m, \ldots, \mathsf{X}_m, \mathsf{Y}_1, \ldots, \mathsf{Y}_m \in \mathsf{M}_2$ (we will use abbreviations $\vec{\mathsf{x}}, \vec{\mathsf{X}}$ etc.). Note that $=_1^{\mathcal{M}}$ and $=_2^{\mathcal{M}}$ are congruence relations on M_1 and M_2, respectively. That is

(i) $=_1^{\mathcal{M}}$ and $=_2^{\mathcal{M}}$ are equivalence relations (axioms **E1′** to **E3″**).
(ii) For every n-ary number function symbol f of $\mathcal{L}$:
$(\mathsf{x}_1, \mathsf{y}_1) \in =^{\mathcal{M}}, \ldots, (\mathsf{x}_n, \mathsf{y}_n) \in =^{\mathcal{M}}, (\mathsf{X}_1, \mathsf{Y}_1) \in =^{\mathcal{M}}, \ldots, (\mathsf{X}_m, \mathsf{Y}_m) \in =^{\mathcal{M}}$
implies $\big(f^{\mathcal{M}}(\vec{\mathsf{x}}, \vec{\mathsf{X}}), f^{\mathcal{M}}(\vec{\mathsf{y}}, \vec{\mathsf{Y}})\big) \in =^{\mathcal{M}}$ (axiom **E4′**).
(iii) Analogously for set function symbols (axiom **E4″**).
(iv) For every n-ary relation symbol R of $\mathcal{L}$:
$(\mathsf{x}_1, \mathsf{y}_1) \in =^{\mathcal{M}}, \ldots, (\mathsf{x}_n, \mathsf{y}_n) \in =^{\mathcal{M}}, (\mathsf{X}_1, \mathsf{Y}_1) \in =^{\mathcal{M}}, \ldots, (\mathsf{X}_m, \mathsf{Y}_m) \in =^{\mathcal{M}}$ and
$(\vec{\mathsf{x}}, \vec{\mathsf{X}}) \in R^{\mathcal{M}}$ implies $(\vec{\mathsf{y}}, \vec{\mathsf{Y}}) \in R^{\mathcal{M}}$ (axiom **E5**).

For $\mathsf{x} \in \mathsf{M}_1$ let $[\mathsf{x}]$ be the equivalence class of x w.r.t. $=^{\mathcal{M}}$. Accordingly, let $[\mathsf{X}]$ be the equivalence class of $\mathsf{X} \in \mathsf{M}_2$ w.r.t. $=^{\mathcal{M}}$. Let $\mathsf{M}_1' = \{[\mathsf{x}] \mid \mathsf{x} \in \mathsf{M}_1\}$ and $\mathsf{M}_2' = \{[\mathsf{X}] \mid \mathsf{X} \in \mathsf{M}_2\}$. Let $\mathcal{M}'$ be a structure with universe $(\mathsf{M}_1', \mathsf{M}_2')$ and

$$f^{\mathcal{M}'}([\vec{\mathsf{x}}], [\vec{\mathsf{X}}]) = [f^{\mathcal{M}}(\vec{\mathsf{x}}, \vec{\mathsf{X}})] \tag{3.1}$$

$$F^{\mathcal{M}'}([\vec{\mathsf{x}}], [\vec{\mathsf{X}}]) = [F^{\mathcal{M}}(\vec{\mathsf{x}}, \vec{\mathsf{X}})] \tag{3.2}$$

$$R^{\mathcal{M}'} = \big\{\big([\vec{\mathsf{x}}], [\vec{\mathsf{X}}]\big) \mid (\vec{\mathsf{x}}, \vec{\mathsf{X}}) \in R^{\mathcal{M}}\big\} \tag{3.3}$$

for all function and relation symbols f, F, R of $\mathcal{L}$. Note that by the above definition, $=_1^{\mathcal{M}'}$ is the relation $\big\{\big([\mathsf{x}_1], [\mathsf{y}_1]\big) \mid (\mathsf{x}_1, \mathsf{x}_2) \in =_1^{\mathcal{M}}\big\}$ (accordingly for $=_2$). Hence $=_1^{\mathcal{M}'}$ ($=_2^{\mathcal{M}'}$) is the equality relation over M_1' (M_2') and hence $\mathcal{M}'$ is a proper structure. For each assignment σ for $\mathcal{M}$ let σ' be an assignment for $\mathcal{M}'$ with

$$\sigma'(a) = [\sigma(a)] \tag{3.4}$$

for some variable a. Next we show by structural induction on terms that for any $\mathcal{L}$-term t (T) and for every assignment σ:

$$t^{\mathcal{M}'}[\sigma'] = \big[t^{\mathcal{M}}[\sigma]\big] \tag{3.5}$$

If t is a variable a, then by definition $a^{\mathcal{M}'}[\sigma'] = \sigma'(a) = [\sigma(a)] = \big[t^{\mathcal{M}}[\sigma]\big]$. If t is of the form $f(t_1, \ldots, t_n, T_1, \ldots, T_m)$, then

$$t^{\mathcal{M}'}[\sigma'] = f^{\mathcal{M}'}(t_1^{\mathcal{M}'}[\sigma'], \ldots, t_n^{\mathcal{M}'}[\sigma'], T_1^{\mathcal{M}'}[\sigma'], \ldots, T_m^{\mathcal{M}'}[\sigma'])$$
$$\overset{\text{Induction Hypothesis}}{=}$$
$$f^{\mathcal{M}'}\big(\big[t_1^{\mathcal{M}}[\sigma]\big], \ldots, \big[t_n^{\mathcal{M}}[\sigma]\big], \big[T_1^{\mathcal{M}}[\sigma]\big], \ldots, \big[T_m^{\mathcal{M}}[\sigma]\big]\big)$$

By (3.1) this is equal to

$$\big[f^{\mathcal{M}}\big(t_1^{\mathcal{M}}[\sigma], \ldots, t_n^{\mathcal{M}}[\sigma], T_1^{\mathcal{M}}[\sigma], \ldots, T_m^{\mathcal{M}}[\sigma]\big)\big] = \big[t^{\mathcal{M}}[\sigma]\big].$$

The same holds for set terms T. Now we are ready to show by structural induction on formulas A that for all σ

$$\mathcal{M}[\sigma] \models A \quad \Leftrightarrow \quad \mathcal{M}'[\sigma'] \models A.$$

If A is of the form $R(t_1, \ldots, t_n, T_1, \ldots, T_m)$, then

$$\begin{aligned}\mathcal{M}'[\sigma'] \models R(\ldots) &\Leftrightarrow \big(t_1^{\mathcal{M}'}[\sigma'], \ldots, T_m^{\mathcal{M}'}[\sigma']\big) \in R^{\mathcal{M}'} \\ &\overset{\text{by (3.5)}}{\Leftrightarrow} \big(\big[t_1^{\mathcal{M}}[\sigma]\big], \ldots, \big[T_m^{\mathcal{M}}[\sigma]\big]\big) \in R^{\mathcal{M}'} \\ &\overset{\text{by (3.3)}}{\Leftrightarrow} \big(t_1^{\mathcal{M}}[\sigma], \ldots, T_m^{\mathcal{M}}[\sigma]\big) \in R^{\mathcal{M}} \Leftrightarrow \mathcal{M}[\sigma] \models A.\end{aligned}$$

If A is of the form $B \wedge C$, $B \vee C$ or $\neg B$, then we can just apply the induction hypothesis and are done. If A is of the form $\forall x B$, then

$$\mathcal{M}'[\sigma'] \models \forall x B \Leftrightarrow \mathcal{M}'[\sigma'([\mathsf{x}]/x)] \models B \text{ for all } [\mathsf{x}] \in \mathsf{M}_1'$$

By applying the induction hypothesis and (3.4) we conclude that this is equivalent to

$$\mathcal{M}[\sigma(\mathsf{x}/x)] \models B \text{ for all } \mathsf{x} \in \mathsf{M}_1 \Leftrightarrow \mathcal{M}[\sigma] \models \forall x B.$$

The cases $\forall X B$, $\exists x B$, $\exists X B$ are similar. □

Notation We write $\Phi_{\varepsilon_{\mathcal{L}}}$ as an abbreviation for $\Phi \cup \varepsilon_{\mathcal{L}}$.

Lemma 3.22. *Let Φ be a set of formulas and A be a formula. Then $\Phi \models A$ iff A holds in all weak models of $\Phi_{\varepsilon_{\mathcal{L}}}$* [i].

[i] Note that this is a "weak" version of logical consequence.

Proof. First, assume $\Phi \models A$. Let $\mathcal{M}$ be a weak model of $\Phi_{\varepsilon_\mathcal{L}}$. We have to show that $\mathcal{M} \models A$. By lemma 3.21 there is a proper structure $\mathcal{M}'$ s.t. $\mathcal{M}'$ satisfies the same formulas as $\mathcal{M}$. Hence $\mathcal{M}' \models \Phi$ and $\mathcal{M}' \models A$ (because $\Phi \models A$) and therefore also $\mathcal{M} \models A$.

Now assume that A holds in all weak models of $\Phi_{\varepsilon_\mathcal{L}}$. Let $\mathcal{M}$ be a model of Φ (we have to show that $\mathcal{M} \models A$). Since A holds in all weak models of $\Phi_{\varepsilon_\mathcal{L}}$ and $\mathcal{M}$ *is* a (weak) model of $\Phi_{\varepsilon_\mathcal{L}}$ [i], we have $\mathcal{M} \models A$. □

We can now reformulate and prove derivational completeness and anchored completeness with equality.

Theorem 3.23 (Derivational Soundness and Completeness of LK^2 with Equality)**.** *Let Φ be a set of $\mathcal{L}$-formulas. $\Phi \models (\Gamma \vdash \Delta)^\bigstar$ iff $\Gamma \vdash \Delta$ has an LK^2-$\Phi_{\varepsilon_\mathcal{L}}$ proof.*

Proof. The soundness direction is straightforward. Assume that $\Gamma \vdash \Delta$ has an LK^2-$\Phi_{\varepsilon_\mathcal{L}}$ proof. By the soundness theorem 3.10 we have $\Phi_{\varepsilon_\mathcal{L}} \models (\Gamma \vdash \Delta)^\bigstar$. For every proper structure $\mathcal{M}$ we have $\mathcal{M} \models \Phi \Leftrightarrow \mathcal{M} \models \Phi_{\varepsilon_\mathcal{L}}$. Therefore $\Phi \models (\Gamma \vdash \Delta)^\bigstar$.

Now to completeness. By lemma 3.22 $(\Gamma \vdash \Delta)^\bigstar$ holds in all weak models of $\Phi_{\varepsilon_\mathcal{L}}$. If we treat $=_1$ and $=_2$ as ordinary relation symbols (or replace them with other symbols), we can use the previous completeness theorem 3.13 to conclude that $\Gamma \vdash \Delta$ has an LK^2-$\Phi_{\varepsilon_\mathcal{L}}$ proof. □

Theorem 3.24 (Anchored Completeness of LK^2 with Equality)**.** *Let Φ be a set of $\mathcal{L}$-formulas closed under substitution of terms for free variables. If $\Phi \models (\Gamma \vdash \Delta)^\bigstar$, then $\Gamma \vdash \Delta$ has an anchored LK^2-$\Phi_{\varepsilon_\mathcal{L}}$ proof.*

Proof. By lemma 3.22 $(\Gamma \vdash \Delta)^\bigstar$ holds in all weak models of $\Phi_{\varepsilon_\mathcal{L}}$. Since $\varepsilon_\mathcal{L}$ is closed under substitution of terms for free variables, we can again treat $=_1$ and $=_2$ as ordinary relation symbols (or replace them with other symbols) to apply the previous anchored completeness theorem 3.16. Therefore $\Gamma \vdash \Delta$ has an anchored LK^2-$\Phi_{\varepsilon_\mathcal{L}}$ proof. □

Anchored proofs are interesting because they share the so-called subformula property.

Definition 3.25 (Subformula)**.** *The set sub(A) of subformulas of a formula*

[i] Note that all proper structures are models of $\varepsilon_\mathcal{L}$.

A is defined inductively by

$$\begin{aligned} sub(A) &= \{A\} \textit{ for atomic } A \\ sub(A_1 \vee A_2) &= sub(A_1) \cup sub(A_2) \cup \{A_1 \vee A_2\} \textit{ (analogously for } \wedge) \\ sub(\neg A) &= sub(A) \cup \{\neg A\} \\ sub(QxA(x)) &= \bigcup \{sub(A(t)) \mid t \textit{ a term}\} \cup \{QxA\} \textit{ for all quantifiers } Q \end{aligned}$$

we call B a subformula of A if $B \in sub(A)$.

Note that formulas of the form $\forall x A(x)$ or $\exists x A(x)$ have infinitely many subformulas.

Lemma 3.26 (Subformula Property of LK^2-Φ)**.** *Let Φ be a set of formulas, closed under substitution of terms for free variables, and let Π be an anchored LK^2-Φ proof of a sequent $\Gamma \vdash \Delta$. Then every formula in every sequent of Π is a subformula of a formula in $\Gamma \vdash \Delta$ or of a formula in Φ.*

Proof. The proof is by induction on the number of sequents in Π. The base case where Π consists of just a nonlogical axiom is obvious. Then we need to examine all rules of LK^2 and check that every formula in the top sequent is a subformula of a formula in the end sequent or of a formula in Φ. This is straight-forward for the structural, propositional and quantifier rules [i]. For the cut rule, we use the fact that every cut formula is in Φ. □

3.3.5 The Compactness Theorem

The well-known compactness theorem is an immediate consequence of the completeness theorem 3.24. Note that there are several forms of the compactness theorem (cf. [9], for example). Here, we only use one form that is useful in section 4.6.3.

Theorem 3.27 (Compactness)**.** *If a formula A is a logical consequence of a set Φ, then A is a logical consequence of some finite subset of Φ.*

Proof. Immediately from the completeness theorem 3.24 and the fact that LK^2 proofs are finite objects. □

[i] For the quantifier rules, note that the formulas $A(b)$ and $A(t)$ are subformulas of $\forall x A(x)$ and $\exists x A(x)$ (accordingly for the "set rules").

4 The Theory $\mathbf{V^1}$

In this section we present the two-sorted theory $\mathbf{V^1}$ which is part of the hierarchy $\mathbf{V^0} \subset \mathbf{V^1} \subseteq \mathbf{V^2} \subseteq \ldots$, where for $i \geq 1$, $\mathbf{V^i}$ characterises the $(i-1)$-th level of the polynomial hierarchy. Thus $\mathbf{V^1}$ characterises P in the sense that the *provably total* functions of $\mathbf{V^1}$ are exactly the polynomial time computable functions. Before we give the definition of $\mathbf{V^1}$, we introduce the notion of a bounded formula and a useful syntactical hierarchy of formulas.

Definition 4.1 (Bounded Formula)**.** *Given a number variable x and a set variable X, let t be a number term not involving x and X. Then $\exists x \leq tA$ stands for $\exists x(x \leq t \wedge A)$, $\forall x \leq tA$ stands for $\forall x(x \leq t \rightarrow A)$, $\exists X \leq tA$ stands for $\exists X(|X| \leq t \wedge A)$ and $\forall X \leq tA$ stands for $\forall X(|X| \leq t \rightarrow A)$. Quantifiers in this form are called bounded and a formula is called bounded if all its quantifiers are bounded.*

Notation $\exists \vec{x} \leq \vec{t}A$ stands for $\exists x_1 \leq t_1 \ldots \exists x_n \leq t_n A$ for some $n \geq 0$, where no x_i occurs in any t_j. Accordingly for $\forall \vec{x} \leq \vec{t}$, $\exists \vec{X} \leq \vec{t}$, $\forall \vec{X} \leq \vec{t}$.

Definition 4.2 (Σ^B_i-, Π^B_i- and Σ^1_1-formulas)**.** *$\Sigma^B_0 = \Pi^B_0$ is the set of $\mathcal{L}^2_\mathcal{A}$-formulas where all quantifiers are bounded number quantifiers (with possibly free set variables)* [i]*. Σ^B_{i+1} (resp. Π^B_{i+1}) is the set of all $\mathcal{L}^2_\mathcal{A}$-formulas of the form $\exists \vec{X} \leq \vec{t}A(\vec{X})$ (resp. $\forall \vec{X} \leq \vec{t}A(\vec{X})$), where $A(\vec{X})$ is a Σ^B_i-formula (resp. a Π^B_i-formula) and the number terms in $\vec{t}$ are over $\mathcal{L}^2_\mathcal{A}$ and do not involve any variable in $\vec{X}$. Σ^1_1 is the set of $\mathcal{L}^2_\mathcal{A}$-formulas of the form $\exists \vec{X} A(\vec{X})$, where A is a Σ^B_0-formula. For a language $\mathcal{L} \supseteq \mathcal{L}^2_\mathcal{A}$ the classes $\Sigma^B_i(\mathcal{L})$ and $\Pi^B_i(\mathcal{L})$ are defined as above with the exception that the underlying language is $\mathcal{L}$ instead of $\mathcal{L}^2_\mathcal{A}$, provided that the terms in $\vec{t}$ are still over $\mathcal{L}^2_\mathcal{A}$.*

Note that for Σ^B_i and Π^B_i, no number quantifiers are in front of set quantifiers. For example, the formula $\forall x \leq t_1 \exists X \leq t_2 A$ is not in Σ^B_1.

When it is clear from the context (or unimportant), we do not mention the underlying language $\mathcal{L}$ explicitly. Note that $\Sigma^B_0 \subseteq \Sigma^B_1 \subseteq \Sigma^B_2 \subseteq \ldots$ and $\Pi^B_0 \subseteq \Pi^B_1 \subseteq \Pi^B_2 \subseteq \ldots$ and for $i \geq 0$ we have $\Sigma^B_i \subseteq \Pi^B_{i+1}$ and $\Pi^B_i \subseteq \Sigma^B_{i+1}$.

4.1 Definition of $\mathbf{V^1}$

Notation We write $X(s)$ instead of $s \in X$.

[i] Σ^B_0 roughly corresponds to one-sorted Δ_0.

The theory $\mathbf{V^1}$ is axiomatised by the following axioms and a comprehension axiom scheme (defined later).

B1. $x + 1 \neq 0$
B2. $x + 1 = y + 1 \rightarrow x = y$
B3. $x + 0 = x$
B4. $x + (y + 1) = (x + y) + 1$
B5. $x \times 0 = 0$
B6. $x \times (y + 1) = (x \times y) + x$
B7. $(x \leq y \wedge y \leq x) \rightarrow x = y$
B8. $x \leq x + y$
B9. $0 \leq x$
B10. $x \leq y \vee y \leq x$
B11. $x \leq y \leftrightarrow x < y + 1$
B12. $x \neq 0 \rightarrow \exists y \leq x(y + 1 = x)$
L1. $X(y) \rightarrow y < |X|$
L2. $y + 1 = |X| \rightarrow X(y)$
SE. $\Big(|X| = |Y| \wedge \forall i < |X|\big(X(i) \leftrightarrow Y(i)\big)\Big) \rightarrow X = Y$

B1 to **B8** are the same as in $\mathbf{I\Delta_0}$ and **B9**-**B12** are theorems of $\mathbf{I\Delta_0}$. **L1** and **L2** define the relation $||$ as the least upper bound relation and **SE** (set equality) states that if two sets have the same elements, then they are equal. It is worth noting that the other direction of the axiom **SE** is valid. That is

$$\mathcal{M} \models X = Y \rightarrow \Big(|X| = |Y| \wedge \forall i < |X|\big(X(i) \leftrightarrow Y(i)\big)\Big)$$

for every $\mathcal{L}$-structure $\mathcal{M}$ with $\mathcal{L} \supseteq \mathcal{L}^2_{\mathcal{A}}$. The reason is that $=$ (actually $=_2$) is always interpreted as the true equality relation.

Definition 4.3 (Φ-**COMP**)**.** *Let Φ be a set of formulas. Then Φ-**COMP** is the set of all formulas of the form*

$$\exists X \leq y \forall z < y\big(X(z) \leftrightarrow A(z)\big),$$

where $A(z) \in \Phi$ and X does not occur free in $A(z)$.

Note that A above may have free number and set variables other than z. The reason why X must not occur free in $A(z)$ is that otherwise the (unsatisfiable) axiom

$$\exists X \leq y \forall z < y\big(X(z) \leftrightarrow \neg X(z)\big)$$

would make our theory inconsistent. Intuitively, the comprehension axiom states that for each $(1,0)$-ary, Φ-definable relation R, defined by A, there exists a set $S = \{x \mid R(x)\}$.

Definition 4.4 ($\mathbf{V^1}$)**.** $\mathbf{V^1}$ *is the theory axiomatised by the axioms* **B1**-**B12**, $\mathbf{L1}, \mathbf{L2}, \mathbf{SE}$ *and* Σ_1^B-**COMP**.

Definition 4.5 (Unique Existence $\exists!$)**.** *The notation* $\exists! X A(X)$ *stands for* $\exists X(A(X) \wedge \forall Y(A(Y) \rightarrow X = Y))$ *(similarly for number variables).*

It is worth noting that in every theory $\mathbf{T}$ containing $\mathbf{V^1}$, it follows from the extensionality axiom **SE** that comprehension is unique (in the sense of definition 4.5). That is, if $\mathbf{T}$ proves a formula $\exists X \leq y \forall z < y\big(X(z) \leftrightarrow A(z)\big)$, then $\mathbf{T}$ also proves the formula $\exists! X \leq y \forall z < y\big(X(z) \leftrightarrow A(z)\big)$.

The theory $\mathbf{V^0}$ has the same axioms as $\mathbf{V^1}$ with the exception that the comprehension axiom is restricted to Σ_0^B-formulas.

Definition 4.6 ($\mathbf{V^0}$)**.** $\mathbf{V^0}$ *is the theory axiomatised by the axioms* **B1**-**B12**, $\mathbf{L1}, \mathbf{L2}, \mathbf{SE}$ *and* Σ_0^B-**COMP**.

Of course $\mathbf{V^0} \subseteq \mathbf{V^1}$ [i]. In section 4.8 we will define a theory $\widetilde{\mathbf{V^1}}$ as an extension of $\mathbf{V^0}$ and it is therefore useful to strengthen a few results and prove them for $\mathbf{V^0}$ instead of $\mathbf{V^1}$. To begin with, we state a few theorems of $\mathbf{V^0}$ (and $\mathbf{V^1}$, of course) that we use in later proofs. They are actually theorems of $\mathbf{I\Delta_0}$.

Lemma 4.7 (Theorems of $\mathbf{V^0}$)**.**

(1) $\mathbf{V^0} \models x \leq 0 \rightarrow x = 0$
(2) $\mathbf{V^0} \models \neg x < 0$
(3) $\mathbf{V^0} \models x < x + 1$
(4) $\mathbf{V^0} \models 0 < x + 1$
(5) $\mathbf{V^0} \models x \leq x$

Proof. (1) Follows from **B7** and **B9**.
(2) Let $\mathcal{M}[\sigma] \models \mathbf{V^0}$ and assume to the contrary that $\mathcal{M}[\sigma] \models x \leq 0 \wedge x \neq 0$. Then it follows from (1) that $\mathcal{M}[\sigma] \models x = 0$, a contradiction. Hence $\mathcal{M}[\sigma] \models \neg x < 0$.
(3) By **B11** we have $\mathcal{M}[\sigma] \models x \leq x \rightarrow x < x + 1$. Then it follows from **B8** and **B3** that $\mathcal{M}[\sigma] \models x < x + 1$.
(4) We have to show $\mathcal{M}[\sigma] \models 0 \leq x + 1 \wedge 0 \neq x + 1$. The RHS of the

[i] Actually, $\mathbf{V^0} \subsetneq \mathbf{V^1}$.

conjunction is **B1**[i] and the LHS follows from **B9**.
(5) cf. (3) □

4.2 Induction in $\mathbf{V^1}$

$\mathbf{V^1}$ does not contain (explicitly) the induction axiom scheme. But its axioms provide enough strength to prove the induction axioms. In the following we will show that $\mathbf{V^0} \models \Sigma^B_0$-**IND** and $\mathbf{V^1} \models \Sigma^B_1$-**IND**. We restate the definition of number induction (cf. definition 2.12).

Definition 4.8 (Number Induction Axiom Scheme)**.** *Let Φ be a set of two-sorted formulas. Then Φ-***IND** *is the set of formulas of the form*

$$\Big(A(0) \wedge \forall x\big(A(x) \rightarrow A(x+1)\big)\Big) \rightarrow \forall x A(x)$$

where $A \in \Phi$.

We first show that $\mathbf{V^0}$ proves the following formula.

Definition 4.9 (X-**MIN**)**.**

$$X\text{-}\mathbf{MIN} \equiv 0 < |X| \rightarrow \exists x < |X|\big(X(x) \wedge \forall y < x \neg X(y)\big) \tag{4.1}$$

Intuitively, X-**MIN** states that every nonempty set X has a smallest element x.

Lemma 4.10. $\mathbf{V^0} \models X$-**MIN**.

Proof. We prove this lemma with semantical arguments. Let $\mathcal{M}$ be *any* model of $\mathbf{V^0}$, $(\mathsf{M}_1, \mathsf{M}_2)$ its universe and σ an arbitrary (but fixed) assignment. Let $A(z)$ be the Σ^B_0-formula $\forall y \leq z \neg X(y)$. Then the following formula is a logical consequence of Σ^B_0-**COMP**

$$\mathcal{M}[\sigma] \models \exists Y \leq |X| \forall z < |X|\big(Y(z) \leftrightarrow \forall y \leq z \neg X(y)\big). \tag{4.2}$$

Intuitively, (4.2) states that for every set X there exists a set Y that consists of the numbers smaller than every element of X. For an arbitrary set $\mathsf{X} \in \mathsf{M}_2$, let $\mathsf{Y} \in \mathsf{M}_2$ be the set that satisfies the existential quantifier in (4.2). We will show that $|Y|$ (respectively $|Y|^{\mathcal{M}}[\sigma(\mathsf{Y}/Y)]$) is just the witness for (4.1), i.e. the smallest element of X, assuming $0 < |X|$. Formally, we need to show:
(i) $\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models X(|Y|)$,

[i]Note that we can assume the symmetry of $=$ because of the requirement that $=^{\mathcal{M}}$ is always the true equality relation.

(ii) $\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \forall y < |Y| \neg X(y)$
while assuming

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models 0 < |X|. \tag{4.3}$$

We do a case analysis on Y. First, suppose that Y is empty, i.e.

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \forall y \neg Y(y). \tag{4.4}$$

By the contraposition of **L2** we have

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \forall y\big(\neg Y(y) \to y + 1 \neq |Y|\big) \tag{4.5}$$

From (4.4) and (4.5) we conclude that

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \forall y\big(y + 1 \neq |Y|\big). \tag{4.6}$$

From the contraposition of **B12** it follows that

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \big(\neg\exists y \leq |Y|(y + 1 = |Y|)\big) \to |Y| = 0 \tag{4.7}$$

and as a logical consequence thereof

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \big(\neg\exists y(y + 1 = |Y|)\big) \to |Y| = 0. \tag{4.8}$$

and

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \big(\forall y(y + 1 \neq |Y|)\big) \to |Y| = 0 \tag{4.9}$$

From (4.6) and (4.9) we conclude that

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models |Y| = 0. \tag{4.10}$$

By lemma 4.7 ($\neg x < 0$) and (4.10) the condition (ii) trivially holds[i]. By (4.3) and (4.2) we obtain

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models Y(0) \leftrightarrow \forall y \leq 0 \neg X(y) \tag{4.11}$$

and with lemma 4.7 (1)

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models Y(0) \leftrightarrow \neg X(0) \tag{4.12}$$

Since Y is empty by (4.4) we have

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models X(0) \tag{4.13}$$

[i]Recall that $=^{\mathcal{M}}$ is always the true equality relation.

which proves (i). Now suppose that Y is *not* empty, i.e.

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models Y(y) \tag{4.14}$$

Then by **L1**

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models y < |Y| \tag{4.15}$$

and by lemma 4.7 (2)

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models |Y| \neq 0 \tag{4.16}$$

since otherwise we had $\mathcal{M}[\dots] \models y < 0$. By **B12** and the above we obtain

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)(\mathsf{z}/z)] \models z + 1 = |Y| \tag{4.17}$$

for some $\mathsf{z} \in \mathsf{M}_1$ and by **L2**

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)(\mathsf{z}/z)] \models Y(z) \tag{4.18}$$

By the contraposition of **L1** we obtain (recall that $<$ is an abbreviation)

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)(\mathsf{z}/z)] \models \neg(z+1 \leq |Y| \wedge z+1 \neq |Y|) \rightarrow \neg Y(z+1) \tag{4.19}$$

and it follows from (4.17) that

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)(\mathsf{z}/z)] \models \neg Y(z+1) \tag{4.20}$$

By (4.2) and (4.18) we obtain

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)(\mathsf{z}/z)] \models \forall y \leq z \neg X(y) \tag{4.21}$$

and with **B11** and (4.17)

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models \forall y < |Y| \neg X(y) \tag{4.22}$$

which proves (ii). From (4.20) and the contraposition of (4.2) we conclude[i]

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{z}/z)] \models \exists y \leq z+1 \; X(y) \tag{4.23}$$

and hence

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{z}/z)(\mathsf{y}/y)] \models y \leq z+1 \wedge X(y) \quad \text{for some } \mathsf{y} \in \mathsf{M}_1 \tag{4.24}$$

It turns out that $\mathcal{M}[\sigma(\mathsf{z}/z)(\mathsf{y}/y)] \models y = z+1$ because otherwise

$$\mathcal{M}[\sigma(\mathsf{z}/z)(\mathsf{y}/y)] \models y < z+1 \quad \text{and} \quad \mathcal{M}[\sigma(\mathsf{z}/z)(\mathsf{y}/y)] \models y \leq z$$

by **B11**, which contradicts (4.21). Therefore we have proved (i):

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{Y}/Y)] \models X(|Y|) \tag{4.25}$$

□

[i]Note that we make use of lemma 2.8.

Convention From now on, we will not always explicitly mention the universe $(\mathsf{M}_1, \mathsf{M}_2)$ of a model $\mathcal{M}$. When we write x (or another letter in the same font), we mean an element of M_1 and when we write X we mean an element of M_2.

Now consider the following weak form of induction.

Definition 4.11. X-**IND** $\equiv \Big(X(0) \wedge \forall y < z\big(X(y) \to X(y+1)\big)\Big) \to X(z)$

Lemma 4.12. $\mathbf{V^0} \models X$-**IND**.

Proof. We give a proof by contradiction. Assume $\mathbf{V^0} \not\models X$-**IND**. Then $\mathbf{V^0} \not\models \forall z X$-**IND** (otherwise $\mathbf{V^0} \models X$-**IND**) and there exists a model $\mathcal{M}$ of $\mathbf{V^0}$ with $\mathcal{M} \not\models \forall z X$-**IND**. Therefore

$$\mathcal{M}[\sigma] \models \neg \forall z X\text{-}\mathbf{IND} \quad \text{for some assignment } \sigma, \tag{4.26}$$

and hence

$$\mathcal{M}[\sigma] \models \exists z \neg X\text{-}\mathbf{IND} \tag{4.27}$$

Thus we obtain (using De Morgan's Laws) for some z

$$\mathcal{M}[\sigma(\mathsf{z}/z)] \models X(0) \wedge \neg X(z) \wedge \forall y < z\big(X(y) \to X(y+1)\big). \tag{4.28}$$

By Σ^B_0-**COMP** we have for some Y

$$\mathcal{M}[\sigma(\mathsf{z}/z)(\mathsf{Y}/Y)] \models |Y| \leq z+1 \wedge \Big(\forall y < z+1\big(Y(y) \leftrightarrow \neg X(y)\big)\Big) \tag{4.29}$$

By Lemma 4.7 (4) ($x < x+1$) and the fact that $\mathcal{M}[\sigma(\mathsf{z}/z) \models \neg X(z)$ (by (4.28)) we obtain

$$\mathcal{M}[\sigma(\mathsf{z}/z)(\mathsf{Y}/Y)] \models Y(z) \tag{4.30}$$

Now we need to establish $0 < |Y|$ in order to apply Y-**MIN**. By **L1** we obtain

$$\mathcal{M}[\sigma(\mathsf{z}/z)(\mathsf{Y}/Y)] \models z < |Y| \tag{4.31}$$

and by lemma 4.7 (2) (since otherwise "$z < 0$")

$$\mathcal{M}[\sigma(\mathsf{z}/z)(\mathsf{Y}/Y)] \models 0 \neq |Y| \tag{4.32}$$

From **B9** and the above we obtain

$$\mathcal{M}[\sigma(\mathsf{Y}/Y)] \models 0 < |Y| \tag{4.33}$$

By Y-**MIN** and the above we have for some $\mathsf{y_0} \in \mathsf{M}_1$ (the *least element* of Y)

$$\mathcal{M}[\sigma(\mathsf{y_0}/y_0)(\mathsf{Y}/Y)] \models y_0 < |Y| \wedge Y(y_0) \wedge \big(\forall x < y_0 \neg Y(x)\big) \tag{4.34}$$

Then $\mathcal{M}[\sigma(\mathsf{y}_0/y_0)] \models y_0 \neq 0$ since otherwise $\mathcal{M}[\sigma] \models \neg X(0)$ by (4.29) and lemma 4.7 (4) $(0 < x + 1)$ which contradicts (4.28). Then by **B12**

$$\mathcal{M}[\sigma(\mathsf{y}_0/y_0)(\mathsf{x}_0/x_0)] \models x_0 \leq y_0 \wedge x_0 + 1 = y_0 \quad \text{for some } \mathsf{x}_0 \in \mathsf{M}_1 \tag{4.35}$$

Then from the above, lemma 4.7 (3) $(x < x + 1)$ and (4.34) we obtain

$$\mathcal{M}[\sigma(\mathsf{Y}/Y)(\mathsf{x}_0/x_0)] \models \neg Y(x_0) \tag{4.36}$$

By (4.29) we obtain $\mathcal{M}[\sigma(\mathsf{x}_0/x_0)] \models X(x_0)$. But since $\mathcal{M}[\sigma(\mathsf{Y}/Y)(\mathsf{x}_0/x_0)] \models Y(x_0 + 1)$ (recall $\mathsf{x}_0 + 1 = \mathsf{y}_0$) we obtain again by (4.29) $\mathcal{M}[\sigma(\mathsf{x}_0/x_0)] \models \neg X(x_0 + 1)$ which contradicts our assumption (4.28). Therefore we have proved $\mathbf{V^0} \models X$-**IND**.

□

Now it is easy, using the previous lemma, to show that $\mathbf{V^0} \models \Sigma^B_0$-**IND** and $\mathbf{V^1} \models \Sigma^B_1$-**IND**.

Theorem 4.13. *Let* $\mathbf{T} \supseteq \mathbf{V^0}$ *and assume that* $\mathbf{T} \models \Phi$-**COMP**, *for some set of formulas* Φ. *Then* $\mathbf{T} \models \Phi$-**IND**.

Proof. Let $A(x) \in \Phi$. We need to show that

$$\mathcal{T} \models \Big(A(0) \wedge \forall y\big(A(y) \rightarrow A(y+1)\big)\Big) \rightarrow A(x) \tag{4.37}$$

Let $\mathcal{M}$ be a model of $\mathcal{T}$ and let σ be an arbitrary assignment. Assume

$$\mathcal{M}[\sigma] \models A(0) \wedge \forall y\big(A(y) \rightarrow A(y+1)\big) \tag{4.38}$$

By Φ-**COMP** we have for some X

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| \leq z + 1 \ \wedge \ \forall y < z + 1\big(X(y) \leftrightarrow A(y)\big) \tag{4.39}$$

We conclude from (4.39), (4.38) and lemma 4.7 (4) $(0 < x + 1)$ that

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models X(0) \tag{4.40}$$

By the right conjunct of (4.39), the right conjunct of (4.38) and lemma 4.7 (3) $(x < x + 1)$ we obtain

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models \forall y < z\big(X(y) \rightarrow X(y+1)\big) \tag{4.41}$$

Taken together we obtain

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models X(0) \wedge \forall y < z\big(X(y) \rightarrow X(y+1)\big) \tag{4.42}$$

Since $\mathcal{M}[\sigma(\mathsf{X}/X)] \models X$-**IND** (lemma 4.12) we get

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models X(z) \tag{4.43}$$

and with lemma 4.7 (3) ($x < x + 1$) and (4.39) we finally obtain

$$\mathcal{M}[\sigma] \models A(z) \tag{4.44}$$

which completes our proof. □

Corollary 4.14. $\mathbf{V^0} \models \Sigma_0^B$-**IND** *and* $\mathbf{V^1} \models \Sigma_1^B$-**IND**.

We now list a few theorems of $\mathbf{V^0}$ [i] that we will use later. We prove only some of them. Proof sketches for the others can be found in [6].

Lemma 4.15 (More Theorems of $\mathbf{V^0}$)**.**

(1) $\mathbf{V^0} \models (x + y) + z = x + (y + z)$ *(Associativity of +)*
(2) $\mathbf{V^0} \models (x \times y) \times z = x \times (y \times z)$ *(Associativity of ×)*
(3) $\mathbf{V^0} \models x + y = y + x$ *(Commutativity of +)*
(4) $\mathbf{V^0} \models x \times y = y \times x$ *(Commutativity of ×)*
(5) $\mathbf{V^0} \models x \leq 0 \rightarrow x = 0$
(6) $\mathbf{V^0} \models x \leq y \rightarrow \exists z(x + z = y)$
(7) $\mathbf{V^0} \models x \leq y \leftrightarrow x + z \leq y + z$
(8) $\mathbf{V^0} \models x \leq y \rightarrow x \times z \leq y \times z$
(9) $\mathbf{V^0} \models x < y \leftrightarrow x + 1 \leq y$
(10) $\mathbf{V^0} \models x \leq y \wedge y \leq z \rightarrow x \leq z$ *(Transitivity of ≤)*
(11) $\mathbf{V^0} \models x \leq y + 1 \leftrightarrow (x \leq y \vee x = y + 1)$
(14) $\mathbf{V^0} \models 0 \leq x$
(12) $\mathbf{V^0} \models x \leq y \wedge z \leq v \rightarrow x + z \leq y + v$
(13) $\mathbf{V^0} \models x \leq y \wedge z \leq v \rightarrow x \times z \leq y \times v$
(15) $\mathbf{V^0} \models x < y \rightarrow x < y + z$
(16) $\mathbf{V^0} \models x + 1 \leq y \rightarrow x < y$
(17) $\mathbf{V^0} \models x < y \wedge y < z \rightarrow x < z$ *(Transitivity of <)*

Proof. (13): By theorem 4.13 we can use the Σ_0^B-**IND** axiom scheme on the variable y. Let $\mathcal{M}$ be a model of $\mathbf{V^0}$ and σ an arbitrary assignment. We first need to show that $\mathcal{M}[\sigma] \models x \leq 0 \wedge z \leq v \rightarrow x + z \leq 0 + v$. Assume

[i] They are also theorems of $\mathbf{I\Delta_0}$.

$\mathcal{M}[\sigma] \models x \leq 0 \wedge z \leq v$. By sublemma (5) we have $\mathcal{M}[\sigma] \models x = 0$. Hence it suffices to show that $\mathcal{M}[\sigma] \models 0 + z \leq 0 + v$, which follows from the axiom **B3**, sublemma (3) and the fact that $\mathcal{M}[\sigma] \models z \leq v$. For the induction step we assume that

$$\mathcal{M}[\sigma] \models x \leq y \wedge z \leq v \rightarrow x + z \leq y + v \tag{4.45}$$

We need to establish $\mathcal{M}[\sigma] \models x \leq y+1 \wedge z \leq v \rightarrow x+z \leq (y+1)+v$. Assume that $\mathcal{M}[\sigma]$ satisfies the premise of this implication. Then, by sublemma (11), either $\mathcal{M}[\sigma] \models x \leq y$ or $\mathcal{M}[\sigma] \models x = y + 1$. In the former case we have $\mathcal{M}[\sigma] \models x + z \leq y + v$ by (4.45). Since $\mathcal{M}[\sigma] \models y + v \leq (y + 1) + v$ (by sublemmas (3), (1) and axiom **B8**) we can apply sublemma (10) to obtain $\mathcal{M}[\sigma] \models x + z \leq (y + 1) + v$. In the case of $\mathcal{M}[\sigma] \models x = y + 1$, we need to show $\mathcal{M}[\sigma] \models (y + 1) + z \leq (y + 1) + v$. This follows from sublemmas (7) and (3).

(14): The proof of (14) is also by induction on y and is very similar to the proof of (13). For the base case we need sublemma (4) instead of (3) and axiom **B5** instead of **B3**. For the induction step we assume that

$$\mathcal{M}[\sigma] \models x \leq y \wedge z \leq v \rightarrow x \times z \leq y \times v \tag{4.46}$$

We need to establish $\mathcal{M}[\sigma] \models x \leq y+1 \wedge z \leq v \rightarrow x \times z \leq (y+1) \times v$. Assume that $\mathcal{M}[\sigma]$ satisfies the premise of this implication. Then, by sublemma (11), either $\mathcal{M}[\sigma] \models x \leq y$ or $\mathcal{M}[\sigma] \models x = y + 1$. In the former case we have $\mathcal{M}[\sigma] \models x \times z \leq y \times v$ by (4.46). By sublemma (4) and **B6** we have $\mathcal{M}[\sigma] \models (y+1) \times v = (y \times v) + y$ and hence $\mathcal{M}[\sigma] \models y \times v \leq (y+1) \times v$ by **B8**. Then we can apply sublemma (10) to obtain $\mathcal{M}[\sigma] \models x \times z \leq (y + 1) \times v$. In the case of $\mathcal{M}[\sigma] \models x = y+1$, we need to show $\mathcal{M}[\sigma] \models (y+1) \times z \leq (y+1) \times v$. This follows from sublemmas (8) and (4).

(15): Let $\mathcal{M} \models \mathbf{V^0}$ and σ arbitrary. Assume $\mathcal{M}[\sigma] \models x < y$. We have $\mathcal{M}[\sigma] \models x \leq y + z$ by (13) and (12). I.e we have to show $\mathcal{M}[\sigma] \models x \neq y + z$. Assume to the contrary that $\mathcal{M}[\sigma] \models x = y + z$. By (9), $\mathcal{M}[\sigma] \models x + 1 \leq y$ and by (6) and (3) $\mathcal{M}[\sigma] \models \exists w(x + 1 + w = y)$. Let w be such a w. Then $\mathcal{M}[\sigma(\mathsf{w}/w)] \models x = x + 1 + w + z$, which contradicts our assumption.

(16): Let $\mathcal{M} \models \mathbf{V^0}$ and σ arbitrary. Assume $\mathcal{M}[\sigma] \models x+1 \leq y$. Then, by **B11**, $\mathcal{M}[\sigma] \models x+1 \leq y+1 \wedge x+1 \neq y+1$. By (7), $\mathcal{M}[\sigma] \models x \leq y \wedge x+1 \neq y+1$ and by **B2**, $\mathcal{M}[\sigma] \models x \leq y \wedge x \neq y$.

(17): Follows from (10). □

Using the above results, we can show that $\mathbf{V^0}$ and $\mathbf{V^1}$ prove so-called *number minimisation schemes*.

Definition 4.16 (Φ-**MIN**). *For a set Φ of formulas, the set Φ-**MIN** consists of all the formulas of the form*

$$A(y) \to \exists x \leq y\big(A(x) \wedge \neg\exists z < xA(z)\big)$$

for $A(x) \in \Phi$.

Lemma 4.17. *Let $\mathbf{T} \supseteq \mathbf{V^0}$ and assume that $\mathbf{T} \models \Phi$-**COMP**, for some set of formulas Φ. Then $\mathbf{T} \models \Phi$-**MIN**.*

Proof. Let $A(x) \in \Phi$. We have to show that

$$\mathbf{T} \models A(y) \to \exists x \leq y\big(A(x) \wedge \neg\exists z < xA(z)\big) \tag{4.47}$$

Let $\mathcal{M}$ be a model of $\mathbf{T}$ and let σ be an arbitrary assignment. Assume

$$\mathcal{M}[\sigma] \models A(y) \tag{4.48}$$

We have to show that the RHS of (4.47) holds in $\mathcal{M}[\sigma]$. By $\mathbf{T} \models \Phi$-**COMP**, we have for some X

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| \leq y + 1 \wedge \forall x < y + 1\big(X(x) \leftrightarrow A(x)\big) \tag{4.49}$$

By $\mathbf{T} \models X$-**MIN**, we have

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models 0 < |X| \to \exists x < |X|\big(X(x) \wedge \forall y < x\neg X(y)\big) \tag{4.50}$$

It follows from (4.48), (4.49), lemma 4.7 (3)[i] and the axioms of $\mathbf{V^1}$ that $\mathcal{M}[\sigma(\mathsf{X}/X)] \models 0 < |X|$. Hence the RHS of (4.50) holds in $\mathcal{M}[\sigma(\mathsf{X}/X)]$. Let x be this "smallest" element that satisfies the quantifier $\exists x < |X|$ in the RHS of (4.50). Then we have

$$\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{x}/x)] \models x < |X| \wedge X(x) \wedge \forall y < x\neg X(y) \tag{4.51}$$

We now show that x satisfies the existential quantifier in (4.47). For simplicity, we argue "in" $\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{x}/x)]$. Since $\mathsf{x} < |X|$ and $|X| \leq y + 1$, it follows from lemma 4.15 (10),(17) that $\mathsf{x} < y+1$ and from (9),(7) that $\mathsf{x} \leq y$. Hence, by (4.51) we have $X(\mathsf{x})$ and by (4.49) $A(\mathsf{x})$. And again from (4.51) and (4.49) it follows that $\neg\exists z < \mathsf{x}A(z)$ holds in (4.47). □

[i] $x < x + 1$

4.3 Extensions of Theories

In this section we define what it means for a function or a predicate to be *definable* in a theory. Note that we only talk about functions and relations in the natural numbers and in finite sets of natural numbers. For convenience, we use the same symbol for a relation (function) in the real world and as a relation (function) symbol of our logical language.

In the following definitions, we assume that $\underline{\mathbb{N}_2}'$ is an expansion of the standard model $\underline{\mathbb{N}_2}$ where the respective relation and function symbols R, f, F get their intended interpretations, i.e. $R^{\underline{\mathbb{N}_2}'} = R, f^{\underline{\mathbb{N}_2}'} = f, F^{\underline{\mathbb{N}_2}'} = F$, and the extra symbols of $\mathcal{L} \setminus \mathcal{L}^2_{\mathcal{A}}$ get their intended interpretations. For the notation $\exists!$ see definition 4.5.

Definition 4.18 (Definable Relation)**.** *Let $\mathcal{L} \supseteq \mathcal{L}^2_{\mathcal{A}}$ be a language and let Φ be a set of $\mathcal{L}$-formulas. Let $R(\vec{x}, \vec{X}) \subseteq \mathbb{N}^m \times \mathcal{P}_{\mathbf{f}}(\mathbb{N})^n$ be a (real world) (m, n)-ary two-sorted relation and assume that the symbol R is not in $\mathcal{L}$. We call R Φ-definable if there is a formula $A(\vec{x}, \vec{X}) \in \Phi$*[i] *such that*

$$\underline{\mathbb{N}_2}' \models R(\vec{x}, \vec{X}) \leftrightarrow A(\vec{x}, \vec{X}).$$

We then call $R(\vec{x}, \vec{X}) \leftrightarrow A(\vec{x}, \vec{X})$ the[ii] defining axiom *for R.*

Definition 4.19 (Definable Function)**.** *Let $\mathbf{T}$ be a theory over some language $\mathcal{L} \supseteq \mathcal{L}^2_{\mathcal{A}}$ and Φ as above. Let $f(\vec{x}, \vec{X}) : \mathbb{N}^m \times \mathcal{P}_{\mathbf{f}}(\mathbb{N})^n \to \mathbb{N}$ be a (real world) (m, n)-ary number function and let $F(\vec{x}, \vec{X}) : \mathbb{N}^m \times \mathcal{P}_{\mathbf{f}}(\mathbb{N})^n \to \mathcal{P}_{\mathbf{f}}(\mathbb{N})$ be a set function, respectively. Assume that the symbols f and F are not in $\mathcal{L}$. We call f Φ-definable in $\mathbf{T}$ if its graph (relation) is Φ-definable in $\mathbf{T}$ and $\mathbf{T} \models \forall\vec{x}\forall\vec{X}\exists! y A(\vec{x}, y, \vec{X})$. That is if there is a formula $A(\vec{x}, y, \vec{X}) \in \Phi$ s.t.*

$$\underline{\mathbb{N}_2}' \models y = f(\vec{x}, \vec{X}) \leftrightarrow A(\vec{x}, y, \vec{X}).$$

We then call $y = f(\vec{x}, \vec{X}) \leftrightarrow A(\vec{x}, y, \vec{X})$ the defining axiom *for f.*

We call F Φ-definable in $\mathbf{T}$ if its graph (relation) is Φ-definable in $\mathbf{T}$ and $\mathbf{T} \models \forall\vec{x}\forall\vec{X}\exists! Y\, A(\vec{x}, \vec{X}, Y)$. That is if there is a formula $A(\vec{x}, \vec{X}, Y) \in \Phi$ s.t.

$$\underline{\mathbb{N}_2}' \models Y = F(\vec{x}, \vec{X}) \leftrightarrow A(\vec{x}, \vec{X}, Y).$$

We then call $Y = F(\vec{x}, \vec{X}) \leftrightarrow A(\vec{x}, \vec{X}, Y)$ the defining axiom *for F.*

[i] with all free variables indicated

[ii] Even if there is more than one, we will assume that one specific defining axiom has been chosen and speak of *the* defining axiom.

Note that for relations, the notion of definability is independent of a theory and depends only on the standard model (or an expansion thereof) and the underlying language. For set functions we introduce another notion of definability called *bit-definability*. A set function is called Φ-bit-definable if (roughly) its bit graph relation[i] is Φ-definable.

Definition 4.20 (Bit-definable Function)**.** *Let Φ be a set of $\mathcal{L} \supseteq \mathcal{L}^2_\mathcal{A}$-formulas. A set function $F(\vec{x}, \vec{X})$ (where F is not in $\mathcal{L}$) is called Φ-bit-definable if there is a formula $A(i, \vec{x}, \vec{X}) \in \Phi$ and an $\mathcal{L}^2_\mathcal{A}$-term $t(\vec{x}, \vec{X})$ s.t.*

$$\underline{\mathbb{N}_2}' \models F(\vec{x}, \vec{X})(i) \leftrightarrow i < t(\vec{x}, \vec{X}) \wedge A(i, \vec{x}, \vec{X}).$$

We then call $F(\vec{x}, \vec{X})(i) \leftrightarrow i < t(\vec{x}, \vec{X}) \wedge A(i, \vec{x}, \vec{X})$ the bit-defining axiom for F.

The bit-defining axiom of a (total) set function F can easily be obtained from its defining axiom $Y = F(\vec{x}, \vec{X}) \leftrightarrow A_F(\vec{x}, \vec{X}, Y)$ like this:

$$F(\vec{x}, \vec{X})(i) \leftrightarrow i < |Y| \wedge Y(i) \wedge A_F(\vec{x}, \vec{X}, Y)$$

Definition 4.21 (Conservative Extension)**.** *Let $\mathbf{T}_1, \mathbf{T}_2$ be two theories over the languages $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively, and $\mathbf{T}_2 \supseteq \mathbf{T}_1$. We call $\mathbf{T}_2$ a conservative extension of $\mathbf{T}_1$ if every $\mathcal{L}_1$-formula in $\mathbf{T}_2$ is also in $\mathbf{T}_1$.*

Bit-definability of set functions is a weaker concept than definability in a theory. Not all set functions that are Φ-bit-definable, for some Φ, are Φ-definable in e.g. $\mathbf{V^1}$. However, in sections 4.4 and 4.6.3 we show that bit-definability is useful. Also in section 4.6.3 we will show that adding definable predicates and functions to an existing theory results in a conservative extension of that theory.

We want to fix a set Φ so that the class of definable functions only depends on the proving power of the underlying theory. For this purpose we define here what we mean by a *provably total function*.

Definition 4.22 (Provably Total Function)**.** *A (number or set) function is called provably total in a theory $\mathbf{T}$ iff it is Σ^1_1-definable in $\mathbf{T}$.*

We now show that in extensions of $\mathbf{V^0}$, the provably total functions are closed under composition. In order to do this, we need the two auxiliary lemmas 4.23 and 4.24.

Lemma 4.23. $\mathbf{V^0} \models \forall x \exists X \; x = |X|$.

[i] Definition 4.34

Proof. Let $\mathcal{M}$ be a model of $\mathbf{V^0}$ and let σ be an arbitrary assignment. By Σ_0^B-**COMP** we have $\mathcal{M}[\sigma] \models \exists X \leq x \forall y < x\big(X(y) \leftrightarrow 0 = 0\big)$ and hence

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| \leq x \wedge \forall y < x\ X(y), \text{ for some } \mathsf{X}. \tag{4.52}$$

We have to show that $\mathcal{M}[\sigma(\mathsf{X}/X)] \models x = |X|$. Assume to the contrary that

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models x \neq |X|. \tag{4.53}$$

Then, by (4.52), we have $\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| < x \wedge \forall y < x\ X(y)$ and as a consequence $\mathcal{M}[\sigma(\mathsf{X}/X)] \models X(|X|)$. It follows from axiom **L1** that $\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| < |X|$, which is impossible because $=^{\mathcal{M}}$ is the true equality relation. Therefore we have obtained a contradiction and it follows that our assumption (4.53) was wrong. Hence $\mathcal{M}[\sigma(\mathsf{X}/X)] \models x = |X|$ and we are done. □

The following lemma is useful with regard to the formula replacement lemma 2.8.

Lemma 4.24 (Existential Quantifier Lemma)**.** *The formula*

$$\exists \vec{x} \exists \vec{X} A(\vec{x}, \vec{X}) \otimes \exists \vec{y} \exists \vec{Y} B(\vec{y}, \vec{Y}) \leftrightarrow \exists \vec{x} \exists \vec{y} \exists \vec{X} \exists \vec{Y} \big(A(\vec{x}, \vec{X}) \otimes B(\vec{y}, \vec{Y})\big)$$

where $\otimes$ is either $\wedge$ or $\vee$, is valid, provided that the formulas on the RHS and the LHS have the same free variables.

Proof. We proceed by induction on the number n of existential quantifiers in $\exists \vec{x} \exists \vec{X}$ and $\exists \vec{y} \exists \vec{Y}$. The base case $n = 0$ holds trivially. For the induction step we only show one case. The other cases are proved analogously. Consider a formula $\exists x C(x) \otimes D$. It is easy to verify that this formula is provably equivalent to $\exists x\big(C(x) \otimes D\big)$. Then we apply the induction hypothesis to $C(x) \otimes D$ and are done. Note that we can always rename bound variables in order to avoid name clashes. □

In the proof of the lower bound of $\mathbf{V^1}$ (theorem 4.79), we will use the fact that the provably total functions of $\mathbf{V^1}$ are closed under function composition.

Lemma 4.25. *Let $\mathbf{T} \supseteq \mathbf{V^0}$ be a theory over a language $\mathcal{L} \supseteq \mathcal{L}^2_{\mathcal{A}}$. Then the provably total functions of $\mathbf{T}$ are closed under function composition.*

Proof. Let $\vec{x}, \vec{X}$ stand for $x_1, \ldots, x_k, X_1, \ldots, X_l$. Suppose that the functions $h(x_1, \ldots, x_n, X_1, \ldots, X_m)$, $H(x_1, \ldots, x_n, X_1, \ldots, X_m)$, $g_i(\vec{x}, \vec{X})$, for $1 \leq i \leq$

n, $G_j(\vec{x}, \vec{X})$, for $1 \leq j \leq m$, are Σ^1_1-definable in $\mathbf{T}$. . We have to show that the functions

$$f(\vec{x}, \vec{X}) = h\big(g_1(\vec{x}, \vec{X}), \ldots, g_n(\vec{x}, \vec{X}), G_1(\vec{x}, \vec{X}), \ldots, G_m(\vec{x}, \vec{X})\big), \tag{4.54}$$

$$F(\vec{x}, \vec{X}) = H\big(g_1(\vec{x}, \vec{X}), \ldots, g_n(\vec{x}, \vec{X}), G_1(\vec{x}, \vec{X}), \ldots, G_m(\vec{x}, \vec{X})\big) \tag{4.55}$$

are also Σ^1_1-definable in $\mathbf{T}$. Let

$$A_{g_i}(\vec{x}, y, \vec{X}), A_{G_j}(\vec{x}, \vec{X}, Y), A_h(\vec{z}, y, \vec{Z}), A_H(\vec{z}, \vec{Z}, Y) \tag{4.56}$$

be the RHS of the Σ^1_1-defining axioms for the above functions (cf. definition 4.19). Then f has the following defining axiom

$$\begin{aligned} y = f(\vec{x}, \vec{X}) \leftrightarrow \exists z_1 \ldots \exists z_n \exists Z_1 \ldots \exists Z_m \\ \big(A_{g_1}(\vec{x}, z_1, \vec{X}) \wedge \ldots A_{g_n}(\vec{x}, z_n, \vec{X}) \wedge \\ A_{G_1}(\vec{x}, \vec{X}, Z_1) \wedge \cdots \wedge A_{G_m}(\vec{x}, \vec{X}, Z_m) \wedge \\ A_h(\vec{z}, y, \vec{Z})\big) \end{aligned} \tag{4.57}$$

where $\vec{z} = z_1, \ldots, z_n$ and $\vec{Z} = Z_1, \ldots, Z_m$. Note that (4.57) is not a Σ^1_1-formula. However, according to lemma 4.24, it is equivalent to a Σ^1_1-formula because the existential quantifiers of the formulas (4.56) can be put in front, and, by lemma 4.23, the n existential number quantifiers can be replaced by existential set quantifiers. Let $A(\vec{x}, y, \vec{X})$ be the RHS of the axiom (4.57). We have to show that $\mathbf{T} \models \forall \vec{x} \forall X \exists! y A(\vec{x}, y, \vec{X})$. By assumption, we have

$$\begin{aligned} &\mathbf{T} \models \forall \vec{x} \forall \vec{X} \exists! y A_{g_i}(\vec{x}, y, \vec{X}), \\ &\mathbf{T} \models \forall \vec{x} \forall \vec{X} \exists! Y A_{G_j}(\vec{x}, \vec{X}, Y), \\ &\mathbf{T} \models \forall \vec{x} \forall \vec{X} \exists! y A_h(\vec{x}, y, \vec{X}), \end{aligned}$$

for all g_i, G_j. As a consequence, the quantifiers $\exists z_i, \exists Z_j$ in (4.57) are uniquely satisfied. Since the y in $A_h(\vec{z}, y, \vec{Z})$ is also unique, it follows that $\mathbf{T} \models \forall \vec{x} \forall X \exists! y A(\vec{x}, y, \vec{X})$. The proof for F is analogous. □

.

4.4 Complexity Theory

In the one-sorted context of bounded arithmetic, elements of complexity classes are considered subsets (i.e. relations) of $\mathbb{N}$. For example, P is the set of all relations $R(x_1, \ldots, x_n)$, $n \geq 1$, over $\mathbb{N}$ such that some polytime Turing machine, given input $x_1, \ldots, x_n$ (in binary notation, separated by blanks), decides whether $R(x_1, \ldots, x_n)$ holds. Contrary to the one-sorted context, in the two-sorted context the relations $R(x_1, \ldots, x_n, X_1, \ldots, X_m)$ have arguments of both sorts. Now, numbers are presented in *unary* notation and sets in *binary* notation using the following encoding. Let $S \subseteq \mathbb{N}$. If we write $S(i)$ for $i \in S$ and 1 if $S(i)$ holds (0 otherwise), then we can define an encoding $w(S)$ as follows:

$$w(S) = S(n)S(n-1)\ldots S(1)S(0),$$

where n is the largest number in the set S. We further let $w(\emptyset)$ be the empty string. For example $w\big(\{2,3,6,7\}\big) = 11001100$. Note that the mapping w is injective but not surjective (since all encodings $w(S)$ begin with 1). Therefore we just drop the first 1 and (re-)define $w(S) = S(n-1)\ldots S(0)$ to obtain a bijection.

For example, two-sorted P (polynomial time) is the set of all relations $R(\vec{x}, \vec{X})$ where some Turing machine, given input $x_1, \ldots, x_m$ in unary notation (separated by blanks) and input $w(X_1)\ldots w(X_n)$ (separated by blanks), decides whether $R(\vec{x}, \vec{X})$ holds or not. The two-sorted polynomial hierarchy PH is defined accordingly. Note that a numerical relation is in two-sorted P iff it is computed in time $2^{O(n)}$ on some deterministic Turing machine (because we need 2^n steps to read the input). Note that this encoding of sets naturally leads to the term of "bit-definability" since $A(i)$ implies that the i-th bit of the string representation of A holds. We now introduce the (small) complexity class AC^0 (see also [1]).

Definition 4.26 (AC^0)**.** *A relation $R(\vec{x}, \vec{X})$ is in AC^0 iff some alternating Turing machine accepts R in time $\mathcal{O}(\log n)$ with a constant number of alternations.*

One way of relating logic to complexity classes are so-called representation theorems. We show that a relation is in a complexity class iff it is definable by a certain type of formula. The following theorem (see [6] for details) connects AC^0 and the language $\mathcal{L}^2_{\mathcal{A}}$.

Theorem 4.27 (Σ^B_0 Representation Theorem)**.** *A relation $R(\vec{x}, \vec{X})$ is in AC^0 iff it is Σ^B_0-definable.*

For the sake of completeness, we state representation theorems for the class P and the polynomial hierarchy PH (although we will not make use of these theorems). Details are in [6].

Theorem 4.28 (Σ_i^B Representation Theorem)**.** *For $i \geq 1$, a relation $R(\vec{x}, \vec{X})$ is in the i-th level Σ_i^{P} of PH iff it is Σ_i^B-definable, e.g. R is in NP iff it is Σ_1^B-definable.*

Theorem 4.29 (Σ_1^1 Representation Theorem)**.** *A relation $R(\vec{x}, \vec{X})$ is recursively enumerable iff it is Σ_1^1-definable.*

Grädel (cf. for example [5]) proved a representation theorem for the class P.

Definition 4.30 (Σ_1^B-HORN-formula)**.** *A Σ_1^B-HORN-formula is an $\mathcal{L}_{\mathcal{A}}^2$-formula of the form*

$$\exists Y_1 \ldots \exists Y_n \forall y_1 \leq t_1 \ldots \forall y_m \leq t_m B \tag{4.58}$$

where $n, m \geq 0$ and B is a quantifier-free formula in conjunctive normal form and each clause contains at most one positive occurrence of a literal of the form $Z_i(t)$. Additionally, no terms of the form $|Z_i|$ occur in (4.58)*.* (4.58) *may contain free number and set variables (even in the form $|X|$) and clauses of B may contain any number of positive or negative literals of the form $X(t)$.*

Theorem 4.31 (Σ_1^B-HORN Representation Theorem)**.** *A relation $R(\vec{x}, \vec{X})$ is in P iff it is Σ_1^B-HORN-definable.*

4.4.1 Two-sorted Functions

So far, complexity classes are defined in terms of (two-sorted) *relations*. Now we associate to each complexity class C a class of functions FC (number and set functions). For example FP is the class of polynomial time computable functions. We now define what it means for a function to be polynomially bounded.

Definition 4.32 (Polynomially Bounded Function)**.** *A number function f or a set function F is polynomially bounded if there exists a polynomial $p(\vec{x}, \vec{y})$ such that $f(\vec{x}, \vec{Y}) \leq p(\vec{x}, |\vec{Y}|)$ or $|F(\vec{x}, \vec{Y})| \leq p(\vec{x}, |\vec{Y}|)$ for all $\vec{x} \in \mathbb{N}^k$, for some k, and all $\vec{Y} \in \mathcal{P}_{\mathbf{f}}(\mathbb{N})^l$, for some l* [i].

[i] Recall that $\mathcal{P}_{\mathbf{f}}(\mathbb{N})$ denotes the set of all finite subsets of $\mathbb{N}$.

Recall that $|X|$ denotes the function "one plus the largest element of X, or 0 if X is empty", i.e. $|X|$ is equal to the length of the binary encoding $w(X)$.

Definition 4.33 (Graph of a Function)**.** *Given a number function $f(\vec{x}, \vec{X})$, its graph $G_f(y, \vec{x}, \vec{X})$ is the relation $\{(y, \vec{x}, \vec{X}) \mid y = f(\vec{x}, \vec{X})\}$. Analogously, for a set function $F(\vec{x}, \vec{X})$, its graph is the relation $\{(\vec{x}, \vec{X}, Y) \mid Y = F(\vec{x}, \vec{X})\}$.*

Definition 4.34 (Bit Graph of a Function)**.** *Given a set function $F(\vec{x}, \vec{X})$, its bit graph $B_F(i, \vec{x}, \vec{X})$ is the relation $\{(i, \vec{x}, \vec{X}) \mid i \in F(\vec{x}, \vec{X})\}$.*

Definition 4.35 (Function Class)**.** *Let C be a two-sorted complexity class (of relations). Then the corresponding class FC of functions consists of all polynomially bounded number functions whose graphs are in C, together with all polynomially bounded set functions whose* bit *graphs are in C.*

For example, the set functions in FAC^0 are those polynomially bounded functions whose bit graphs are in AC^0. In [11, 12, 10] and others, these functions are called *rudimentary*. The following corollary is an immediate consequence of definition 4.35 and the Σ_0^B representation theorem 4.27.

Corollary 4.36. *A set function is in FAC^0 iff it is polynomially bounded and its bit graph is Σ_0^B-definable. A number function is in FAC^0 iff it is polynomially bounded and its graph is Σ_0^B-definable.*

The next corollary follows immediately from the above.

Corollary 4.37. *A set function is in FAC^0 iff it is Σ_0^B-bit-definable.*

Our goal is to show that the provably total functions of $\mathbf{V^1}$ are exactly the functions in FP. In [4], Cobham first introduced a machine-independent characterisation of FP and we will use this characterisation for proving the latter fact about $\mathbf{V^1}$. First, we define the function $\mathrm{chop}(x, X)$ as the function that returns all elements $y \in X$ that are strictly smaller than x. I.e. $\mathrm{chop}(x, X)$ returns the initial segment of length x of the binary string $w(X)$. chop has the Σ_0^B-bit-defining axiom

$$\mathrm{chop}(x, X)(z) \leftrightarrow z < x \wedge X(z) \tag{4.59}$$

and is therefore in FAC^0 by corollary 4.37.

Notation We often write $X^{<t}$ instead of $\mathrm{chop}(t, X)$ to make things more readable.

Definition 4.38 (Bounded Recursion on Notation)**.** *A set function $F(y, \vec{x}, \vec{X})$ is defined by bounded recursion on notation from set functions $G(\vec{x}, \vec{X})$ and $H(y, \vec{x}, \vec{X}, Z)$ iff*

$$F(0, \vec{x}, \vec{X}) = G(\vec{x}, \vec{X}) \tag{4.60}$$

$$F(y+1, \vec{x}, \vec{X}) = H\big(y, \vec{x}, \vec{X}, F(y, \vec{x}, \vec{X})\big)^{<t(y,\vec{x},\vec{X})} \tag{4.61}$$

for some polynomial t in $y, \vec{x}, |\vec{X}|$ (i.e. t is an $\mathcal{L}^2_{\mathcal{A}}$-term).

Now we state a two-sorted version of Cobham's theorem. For details, see [4] and [6].

Theorem 4.39 (Cobham's Characterisation of FP)**.** *A set function is in FP iff it can be obtained from FAC^0 set functions by finitely many applications of composition and bounded recursion on notation.*

We could also define the notion of bounded recursion on notation for number functions. However, the following lemma shows that this is not explicitly necessary.

Lemma 4.40. *A number function $f(\vec{x}, \vec{X})$ is in FP iff there exists a set function $F(\vec{x}, \vec{X})$ in FP s.t. $f(\vec{x}, \vec{X}) = |F(\vec{x}, \vec{X})|$.*

Proof. Given a number function $f(\vec{x}, \vec{X})$ in FP we can easily define a set function $F(\vec{x}, \vec{X})$ as

$$F(\vec{x}, \vec{X}) = \{z \mid z < f(\vec{x}, \vec{X})\}$$

with $|F(\vec{x}, \vec{X})| = f(\vec{x}, \vec{X})$. The converse direction of the lemma is obvious. □

4.5 Parikh's Theorem

Parikh's theorem is useful in proving the lower bound of $\mathbf{V^1}$ (section 4.7). For the proof of Parikh's theorem, as well as for the proof of the witnessing theorem 4.95, it is important that the proofs are in so-called free variable normal form.

Definition 4.41 (Free Variable Normal Form (FVNF))**.** *Let* Π *be an* LK^2-Φ *proof of the sequent* $\Gamma \vdash \Delta$*. We call the free variables in* $\Gamma \vdash \Delta$ *parameter variables of* Π*. We say that* Π *is in free variable normal form if* 1*. no free variable is eliminated (in the sense that it occurs in the top sequent but not in the bottom sequent) in* Π *by any rule except the* $\forall$*-right and* $\exists$*-left rules where, in addition, no eigenvariable is a parameter variable, and* 2*. every nonparameter free variable in* Π *is used exactly once as an eigenvariable.*

Lemma 4.42 (FVNF Lemma)**.** *Let* $\mathcal{L}$ *be a language containing at least one number constant symbol and one set constant symbol. Every* LK^2-Φ *proof can be transformed into an* LK^2-Φ *proof (with the same end sequent) in free variable normal form.*

Proof. Note that the only rules, other than the ones mentioned above, that can eliminate a free variable are the $\forall$-left and $\exists$-right rules and the cut rule. Also note that Π is a tree. In this context, when we say a free variable b occurs "above" a sequent, we mean that b occurs somewhere between this sequent and a leaf of the tree, i.e. an axiom. We can transform Π into free variable normal form by the following procedure: Select an upper-most rule in Π which eliminates a free variable (abort the procedure if there is none). If the rule is $\forall$-right or $\exists$-left and the eliminated eigenvariable b (M) occurs somewhere in Π other than above this rule, then replace b (M) by a new variable b' (M') (which does not occur in Π) in every sequent above this rule. If the rule is $\forall$-left, $\exists$-right or cut, then replace every variable that is eliminated by the rule by the same constant symbol c (C) in every sequent above this rule. Repeat the procedure as long as necessary. □

Note that the language $\mathcal{L}^2_{\mathcal{A}}$ does not contain a constant symbol for sets. Therefore, in order to put an LK^2-Φ proof into free variable normal form it might be necessary to extend the underlying theory by allowing the set constant symbol $\varnothing$ and adding the axiom $|\varnothing| = 0$ to Φ. The following lemma makes clear that this results in a conservative extension (cf. definition 4.21) of the theory.

Lemma 4.43. *Let* $\mathbf{T} \supseteq \mathbf{V^0}$ *and let* $\mathcal{L}$ *be the language of* $\mathbf{T}$*. The theory* $\mathbf{T}'$ *obtained by adding the axiom* $|\varnothing| = 0$ *to* $\mathbf{T}$ *is a conservative extension of* $\mathbf{T}$*.*

Proof. Let $\mathcal{M}$ be a model of $\mathbf{T}$ and σ an arbitrary assignment. We show that there exists a set X with $|\mathsf{X}| = 0$ in the universe of $\mathcal{M}$ and that this X is unique. By Σ_0^B-**COMP**, there exists an X s.t. $\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| \leq 0$. By lemma 4.15 (5) we have $\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| = 0$. From lemma 4.7 (2) and the axiom **SE** it follows that X is unique. We can conclude from this that the models of $\mathbf{T}$ and $\mathbf{T}'$ differ only in the sense that models of $\mathbf{T}'$ map the constant symbol $\varnothing$ to the empty set. Hence every $\mathcal{L}$-formula in $\mathbf{T}'$ is also in $\mathbf{T}$. □

Corollary 4.44. *Let* $\mathbf{T} \supseteq \mathbf{V^0}$. *Then* $\mathbf{T} \models \exists! X\big(|X| = 0\big) \wedge \exists! X\, \forall x \neg X(x)$.

Proof. By the proof of lemma 4.43 above we have $\mathbf{T} \models \exists! X\ |X| = 0$ and the rest follows from the axiom **L1** and lemma 4.7 (2)[i]. □

$\mathbf{V^1}$ (as well as $\mathbf{V^0}$) is a so-called *polynomially bounded theory*, a concept we now define. We say that a number term $t(\vec{x}, \vec{X})$ is a bounding term for a function f or F in a theory $\mathbf{T}$ if $\mathbf{T} \models \forall \vec{x} \forall \vec{X}\ f(\vec{x}, \vec{X}) \leq t(\vec{x}, \vec{X})$ or $\mathbf{T} \models \forall \vec{x} \forall \vec{X}\ |F(\vec{x}, \vec{X})| \leq t(\vec{x}, \vec{X})$, respectively. We call f or F *polynomially bounded* in $\mathbf{T}$ if f (F) has an $\mathcal{L}^2_{\mathcal{A}}$-bounding term in $\mathbf{T}$.

Definition 4.45 (Polynomially Bounded Theory)**.** *A theory* $\mathbf{T}$ *over* $\mathcal{L}$ *is called polynomially bounded if (1)* $\mathbf{T} \supseteq \mathbf{V^0}$*, (2) it can be axiomatised by a set of bounded formulas, (3) every function* f *or* F *of* $\mathcal{L}$ *is polynomially bounded in* $\mathbf{T}$.

$\mathbf{V^0}$ and $\mathbf{V^1}$ are polynomially bounded theories since all their axioms are bounded formulas and, because they are $\mathcal{L}^2_{\mathcal{A}}$-theories, the functions $0, 1, +, \times, ||$ all have (trivial) $\mathcal{L}^2_{\mathcal{A}}$-bounding terms.

Lemma 4.46 (Monotonicity of $\mathcal{L}^2_{\mathcal{A}}$-terms)**.** *Let* $t(x_1, \ldots, x_n)$ *be an* $\mathcal{L}^2_{\mathcal{A}}$*-number term. Then*

$$\mathbf{V^0} \models x_1 \leq y_1 \wedge \cdots \wedge x_n \leq y_n \rightarrow t(x_1, \ldots, x_n) \leq t(y_1, \ldots, y_n)$$

Proof. The proof is by structural induction on the term $t(\vec{x})$. For convenience, we write $\vec{x} \leq \vec{y}$ for $x_1 \leq y_1 \wedge \cdots \wedge x_n \leq y_n$. Let $\mathcal{M}$ be a model of $\mathbf{V^0}$. In the base case, $t(\vec{x}) \equiv x_i$ for some i and $\mathcal{M} \models \vec{x} \leq \vec{y} \rightarrow x_i \leq y_i$ holds trivially. The cases $t(\vec{x}) \equiv 0$ and $t(\vec{x}) \equiv 1$ follow from lemma 4.7 (5). If $t(\vec{x})$ has the form $t_1(\vec{x}) + t_2(\vec{x})$, then, by induction hypothesis,

$$\begin{aligned} &\mathbf{V^0} \models \vec{x} \leq \vec{y} \rightarrow t_1(\vec{x}) \leq t_1(\vec{y}), \\ &\mathbf{V^0} \models \vec{x} \leq \vec{y} \rightarrow t_2(\vec{x}) \leq t_2(\vec{y}) \end{aligned} \tag{4.62}$$

[i] $\neg x < 0$

We have to show that for an arbitrary σ, $\mathcal{M}[\sigma] \models \vec{x} \leq \vec{y} \rightarrow t_1(\vec{x}) + t_2(\vec{x}) \leq t_1(\vec{y}) + t_2(\vec{y})$. Assume $\mathcal{M}[\sigma] \models \vec{x} \leq \vec{y}$. Then, by (4.62), $\mathcal{M}[\sigma] \models t_1(\vec{x}) \leq t_1(\vec{y})$ and $\mathcal{M}[\sigma] \models t_2(\vec{x}) \leq t_2(\vec{y})$ and it follows from lemma 4.15 (13) that $\mathcal{M}[\sigma] \models t_1(\vec{x}) + t_2(\vec{x}) \leq t_1(\vec{y}) + t_2(\vec{y})$. The case $t(\vec{x}) \equiv t_1(\vec{x}) \times t_2(\vec{x})$ is proved analogously using lemma 4.15 (14). The last case is $t(\vec{x}) \equiv |X|$ (note that set variables are the only $\mathcal{L}^2_\mathcal{A}$ set terms). Then the lemma follows trivially from lemma 4.7 (5). □

Note that $\mathcal{L}^2_\mathcal{A}$-number terms represent polynomials. The next lemma shows that in a theory with only polynomially bounded functions, all terms are polynomially bounded. We need this lemma for the proof of Parikh's theorem.

Lemma 4.47. *Let $\mathbf{T} \supseteq \mathbf{V^0}$ be a theory and $\mathcal{L} \supseteq \mathcal{L}^2_\mathcal{A}$ be the language of $\mathbf{T}$. If all functions of $\mathcal{L}$ are polynomially bounded in $\mathbf{T}$, then for each $\mathcal{L}$-number term $s(\vec{x}, \vec{X})$ and $\mathcal{L}$-set term $S(\vec{x}, \vec{X})$, there is an $\mathcal{L}^2_\mathcal{A}$-number term $t(\vec{x}, \vec{X})$ s.t.*

$$\mathbf{T} \models s(\vec{x}, \vec{X}) \leq t(\vec{x}, \vec{X}), \text{ and}$$
$$\mathbf{T} \models |S(\vec{x}, \vec{X})| \leq t(\vec{x}, \vec{X}), \text{ respectively,}$$

and all terms involved contain only the variables $\vec{x}$ and $\vec{X}$.

Proof. The proof is by structural induction on s and S. If s is a variable x, then $\mathbf{T} \models x \leq x$ by lemma 4.7 (5). If S is a variable X, then also $\mathbf{T} \models |X| \leq |X|$ (note that $|X|$ is an $\mathcal{L}^2_\mathcal{A}$-number term). If s is of the form

$$f\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, t_n(\vec{x_n}, \vec{X_n}), T_1(\vec{y_1}, \vec{Y_1}), \ldots, T_m(\vec{y_m}, \vec{Y_m})\big)$$

then, by the induction hypothesis, we have for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$

$$\begin{aligned} &\mathbf{T} \models t_i(\vec{x_i}, \vec{X_i}) \leq \overline{t_i}(\vec{x_i}, \vec{X_i}) \text{ for some } \mathcal{L}^2_\mathcal{A}\text{-number term } \overline{t_i}(\vec{x_i}, \vec{X_i}), \\ &\mathbf{T} \models |T_j(\vec{y_j}, \vec{Y_j})| \leq \overline{r_j}(\vec{y_j}, \vec{Y_j}) \text{ for some } \mathcal{L}^2_\mathcal{A}\text{-number term } \overline{r_j}(\vec{y_j}, \vec{Y_j}) \end{aligned} \tag{4.63}$$

Since f is polynomially bounded in $\mathbf{T}$, it has an $\mathcal{L}^2_\mathcal{A}$-bounding term $t(\vec{x}, \vec{X})$ in $\mathbf{T}$. Hence we have[i]

$$\mathbf{T} \models f\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, T_m(\vec{y_m}, \vec{Y_m})\big) \leq t\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, T_m(\vec{y_m}, \vec{Y_m})\big) \tag{4.64}$$

[i] We use $t_1, \ldots, T_m$ as an abbreviation for $t_1, \ldots, t_n, T_1, \ldots, T_m$.

Since $t(\vec{x}, \vec{X})$ is an $\mathcal{L}^2_{\mathcal{A}}$-number term, the only set terms in $t(\vec{x}, \vec{X})$ are set variables. Furthermore, a set variable X_i of $\vec{X}$ can only occur in a subterm $|X_i|$. Therefore $t(\vec{x}, \vec{X})$ can be rewritten as $t(\vec{x}, |\vec{X}|)$ and (4.64) as

$$\mathbf{T} \models f\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, T_m(\vec{y_m}, \vec{Y_m})\big) \leq t\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, |T_m(\vec{y_m}, \vec{Y_m})|\big) \tag{4.65}$$

By the sublemma below and (4.63) we have

$$\mathbf{T} \models t\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, |T_m(\vec{y_m}, \vec{Y_m})|\big) \leq t\big(\overline{t_1}(\vec{x_1}, \vec{X_1}), \ldots, \overline{r_m}(\vec{y_m}, \vec{Y_m})\big) \tag{4.66}$$

It then follows from transitivity (lemma 4.15 (10)) that

$$\mathbf{T} \models f\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, T_m(\vec{y_m}, \vec{Y_m})\big) \leq t\big(\overline{t_1}(\vec{x_1}, \vec{X_1}), \ldots, \overline{r_m}(\vec{y_m}, \vec{Y_m})\big) \tag{4.67}$$

which is what we want since the RHS of (4.67) is an $\mathcal{L}^2_{\mathcal{A}}$-number term.

The case where S is of the form $F\big(t_1(\vec{x_1}, \vec{X_1}), \ldots, T_m(\vec{y_m}, \vec{Y_m})\big)$ is proved in the same way replacing $f(\ldots)$ with $|F(\ldots)|$.

Sublemma. *Given an $\mathcal{L}^2_{\mathcal{A}}$-number term $t(\vec{x})$ and $\mathcal{L}$-number terms $\vec{s} = s_1, \ldots, s_n, \vec{r} = r_1, \ldots, r_n$ with $\mathbf{T} \models s_i \leq t_i$ for all $i = 1, \ldots, n$. Then $\mathbf{T} \models t(s_1, \ldots, s_n) \leq t(r_1, \ldots, r_n)$.*
Proof. Follows immediately from the monotonicity of $\mathcal{L}^2_{\mathcal{A}}$-terms (lemma 4.46). □

Now we are ready to prove the following special case of Parikh's theorem from which the general form (theorem 4.49) will follow.

Lemma 4.48 (Parikh's Theorem, Special Case)**.** *Let $\mathbf{T}$ be a polynomially bounded theory and $A(y, \vec{x}, \vec{X})$ be a bounded formula with all free variables indicated. Assume*

$$\mathbf{T} \models \exists y A(y, \vec{x}, \vec{X}). \tag{4.68}$$

Then there exists an $\mathcal{L}^2_{\mathcal{A}}$-term $t(\vec{x}, \vec{X})$ with no variables other than $\vec{x}, \vec{X}$ s.t.

$$\mathbf{T} \models \exists y \leq t(\vec{x}, \vec{X}) A(y, \vec{x}, \vec{X}).$$

Proof. Let Φ be the set of all axioms of $\mathbf{T}$, closed under substitution of terms for free variables. Note that $\Phi \subseteq \mathbf{T}$ and Φ is also an axiomatisation of $\mathbf{T}$. From the anchored completeness theorem 3.24, we conclude that the sequent $\vdash \exists y A(y, \vec{a}, \vec{M})$ [i] has an anchored LK^2-Φ proof Π [ii]. Π features the subformula

[i] Here, the bound variables have been replaced by free variables according to our variable convention on page 21.

[ii] Note that $\exists y A(y, \vec{a}, \vec{M})$ is a logical consequence of (4.68) and theories are closed under logical consequence.

property (lemma 3.26) and by lemma 4.41 we can assume that Π is in free variable normal form by adding the axiom $|\varnothing| = 0$ to Φ if the language of $\mathbf{T}$ does not contain a set constant symbol. Since $\mathbf{T}$ is polynomially bounded, all axioms in Φ are bounded formulas. Hence, by the subformula property, every formula in every sequent in Π is either bounded or is equal (syntactically) to $\exists y A(y, \vec{a}, \vec{M})$. Additionally, $\exists y A(y, \vec{a}, \vec{M})$ cannot occur in an antecedent. The reason is the following: If some antecedent contained this formula, then it would have been eliminated somewhere along Π by either the cut or the $\neg$-right rule. The former case is not possible since cuts are restricted to bounded formulas (i.e. Φ) and in the latter case, a formula $\neg\exists y A(y, \vec{a}, \vec{M})$ would occur in Π which contradicts the subformula property. We will now convert Π to an LK^2-Φ proof $\hat{\Pi}$ of $\exists y \leq t'(\vec{a}, \vec{M}) A(y, \vec{a}, \vec{M})$ for some number term $t'(\vec{a}, \vec{M})$ with no variables other than $\vec{a}, \vec{M}$. Then $\mathbf{T} \models \exists y \leq t'(\vec{a}, \vec{M}) A(y, \vec{x}, \vec{X})$ follows from the completeness theorem. It then follows from lemma 4.47 and transitivity (lemma 4.15 (10)) that there exits an $\mathcal{L}^2_{\mathcal{A}}$-term $t(\vec{a}, \vec{M})$ s.t. $\mathbf{T} \models \exists y \leq t(\vec{a}, \vec{M}) A(y, \vec{x}, \vec{X})$. However, the case where the extra constant symbol $\varnothing$ was added needs special care. In this case $t(\vec{a}, \vec{M}, |\varnothing|)$ may contain $|\varnothing|$ [i]. But since then Φ contains $|\varnothing| = 0$, we can conclude that $\Phi \models \exists y \leq t(\vec{a}, \vec{M}, 0) A(y, \vec{x}, \vec{X})$ and hence $\mathbf{T} \models \exists y \leq t(\vec{a}, \vec{M}, 0) A(y, \vec{x}, \vec{X})$ since $\mathbf{T} \cup \{|\varnothing| = 0\}$ is a conservative extension of $\mathbf{T}$ (lemma 4.43).

The procedure that converts Π is defined inductively on the depth of a sequent S in Π. It replaces every sequent S in Π by a suitable sequent $\hat{\mathsf{S}}$, sometimes adding a short derivation. We give an exact definition of the procedure and make clear that the following **claim** holds:

For every sequent S of Π: If S does not contain $\exists y A(y, \vec{a}, \vec{M})$, then $\hat{\mathsf{S}} = \mathsf{S}$. Otherwise $\hat{\mathsf{S}}$ is the same as S with the exception that all occurrences of $\exists y A(y, \vec{a}, \vec{M})$ are replaced by one single occurrence of $\exists y \leq t A(y, \vec{a}, \vec{M})$, for some number term t that does only contain variables which occur free in S.

Note that the cases where S does not contain $\exists y A(y, \vec{a}, \vec{M})$ are trivial (just let $\hat{\mathsf{S}} = \mathsf{S}$). For convenience, we treat all cedents as multisets and ignore the order of the formulas. It will be clear that this is not a restriction since we only ignore finitely many applications of the exchange rules of LK^2.

Base case: If S is an axiom, then it does not contain $\exists y A(y, \vec{a}, \vec{M})$ (since all non-logical axioms are bounded).

Case I: S is obtained by the inference

$$\frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \exists y A(y, \vec{a}, \vec{M})}\,(\textbf{weakening-right})$$

[i] Note that $\varnothing$ cannot occur in $\exists y A(y, \vec{x}, \vec{X})$ because $\mathbf{T}$ does not contain $\varnothing$.

where S is the bottom sequent. Let S_1 denote the top sequent of this inference. If Δ does not contain $\exists y A(y, \vec{a}, \vec{M})$, then simply let $\hat{\mathsf{S}} = \Gamma \vdash \Delta, \exists y \leq 0 A(y, \vec{a}, \vec{M})$. If Δ contains $\exists y A(y, \vec{a}, \vec{M})$, then apply the induction hypothesis and let $\hat{\mathsf{S}} = \hat{\mathsf{S}_1}$.

Case II: S is obtained using the $\exists$**-right** rule on the formula $A(s, \vec{a}, \vec{M})$ for some number term s. Hence S is the bottom sequent in the inference

$$\frac{\Gamma \ \vdash \ \Delta, A(s, \vec{a}, \vec{M})}{\Gamma \ \vdash \ \Delta, \exists y A(y, \vec{a}, \vec{M})}$$

We distinguish two cases: Either Δ contains $\exists y A(y, \vec{a}, \vec{M})$ or not. If Δ does not contain $\exists y A(y, \vec{a}, \vec{M})$, then it follows from lemma 4.7 (5) that there is an LK^2-Φ proof of $\vdash s \leq s$. We can replace S with the derivation

$$\dfrac{\dfrac{\dfrac{}{\vdash \ s \leq s}}{\Gamma \ \vdash \ \Delta, s \leq s} \qquad \Gamma \ \vdash \ \Delta, A(s, \vec{a}, \vec{M})}{\Gamma \ \vdash \ \Delta, s \leq s \wedge A(s, \vec{a}, \vec{M})}}{\Gamma \ \vdash \ \Delta, \exists y \leq s A(y, \vec{a}, \vec{M})}$$

where $\hat{\mathsf{S}}$ is the bottom sequent. Note that if s contains free variables, then they still occur in the original sequent S since Π is in free variable normal form and therefore no free variable is eliminated by the $\exists$-right rule.

If Δ contains one or more occurrences of $\exists y A(y, \vec{a}, \vec{M})$, then, by the induction hypothesis, the (modified) sequent $\Gamma \vdash \Delta, A(s, \vec{a}, \vec{M})$ has the form

$$\Gamma \vdash \Delta', \exists y \leq t A(y, \vec{a}, \vec{M}), A(s, \vec{a}, \vec{M}) \tag{4.69}$$

As in the above case, we can derive a sequent

$$\Gamma \vdash \Delta', \exists y \leq t A(y, \vec{a}, \vec{M}), \exists y \leq s A(y, \vec{a}, \vec{M}) \tag{4.70}$$

From the axiom **B8** and lemma 4.15 (3) it follows that $\mathbf{T} \models s \leq s + t$ and $\mathbf{T} \models t \leq s + t$. Then, by transitivity of $\leq$ (lemma 4.15 (10)), we have

$$\begin{aligned} \mathbf{T} &\models \exists y \leq s A(y, \vec{a}, \vec{M}) \rightarrow \exists y \leq (s + t) A(y, \vec{a}, \vec{M}), \\ \mathbf{T} &\models \exists y \leq t A(y, \vec{a}, \vec{M}) \rightarrow \exists y \leq (s + t) A(y, \vec{a}, \vec{M}) \end{aligned}$$

and by completeness there are LK^2-Φ proofs of the sequents

$$\exists y \leq s A(y, \vec{a}, \vec{M}) \vdash \exists y \leq (s + t) A(y, \vec{a}, \vec{M}), \tag{4.71}$$

$$\exists y \leq t A(y, \vec{a}, \vec{M}) \vdash \exists y \leq (s + t) A(y, \vec{a}, \vec{M}) \tag{4.72}$$

Using the weakening rule on (4.70), (4.71) and (4.72) allows us to apply the cut rule twice to obtain the sequent $\hat{\mathsf{S}}$

$$\Gamma \vdash \Delta', \exists y \leq (s+t) A(y, \vec{a}, \vec{M}) \tag{4.73}$$

Again, the variables of s occur in the original sequent (4.69). The same holds for t (by induction hypothesis).

If S is obtained using the $\exists$-right rule on a formula $B(s')$, different from $A(s, \vec{a}, \vec{M})$, then S is the bottom sequent of the inference

$$\frac{\Gamma \ \vdash \ \Delta, s' \leq rB(s')}{\Gamma \ \vdash \ \Delta, \exists y \leq rB(y)}$$

In this case no problems arise. We can just apply the induction hypothesis to the top sequent and then apply the $\exists$-right rule.

Case II′: S is obtained using the **set $\exists$-right** rule on a formula $B(S)$. Then S is the bottom sequent of the inference

$$\frac{\Gamma \ \vdash \ \Delta, S \leq rB(S)}{\Gamma \ \vdash \ \Delta, \exists Y \leq rB(Y)}$$

This case poses no problems and we can just apply the induction hypothesis to the top sequent and then apply the set $\exists$-right rule (as above).

Case III: S is obtained by the **$\exists$-left** rule. In this case, the $\exists$ quantifier introduced is bounded because $\exists y A(y, \vec{a}, \vec{M})$ is the only unbounded formula in Π and it never occurs in an antecedent (cf. discussion above). Thus S is the bottom sequent of an inference

$$\frac{b \leq r \wedge B(b), \Gamma \ \vdash \ \Delta}{\exists x \leq rB(x), \Gamma \ \vdash \ \Delta} \ (\exists\text{-left})$$

If Δ does not contain $\exists y A(y, \vec{a}, \vec{M})$, then nothing needs to be done, i.e. let $\hat{\mathsf{S}} = \mathsf{S}$. The case where Δ contains $\exists y A(y, \vec{a}, \vec{M})$, however, requires special care. By the induction hypothesis, the top sequent was converted to

$$b \leq r \wedge B(b), \Gamma \ \vdash \ \Delta', \exists y \leq s(b) A(y, \vec{a}, \vec{M}) \tag{4.74}$$

But here, the eigenvariable b possibly occurs in the introduced term $s(b)$. Therefore the restriction of the $\exists$-left rule might be violated. In order to apply the $\exists$-left rule to (4.74) we need to replace $s(b)$ by a term that does not contain b and whose variables occur free in S. Here we use the fact that the functions of $\mathbf{T}$ are polynomially bounded. By lemma 4.47, there are $\mathcal{L}^2_{\mathcal{A}}$-terms r' and $s'(b)$ (with the same variables as r and $s(b)$, respectively) s.t. $\mathbf{T} \models r \leq r'$ and $\mathbf{T} \models s(b) \leq s'(b)$. Note that all variables except b in $s(b)$

occur free in S by induction hypothesis and that r (and thus r') does not contain b because of the eigenvariable restriction. We now show that $s'(r')$ is the term we are looking for. We first establish

$$\mathbf{T} \models b \leq r \to s(b) \leq s'(r'). \tag{4.75}$$

Let $\mathcal{M}$ be a model of $\mathbf{T}$ and σ an arbitrary assignment. Assume $\mathcal{M}[\sigma] \models b \leq r$. Then $\mathcal{M}[\sigma] \models b \leq r'$ (lemma 4.15 (10)). Lemma 4.46 implies $\mathcal{M}[\sigma] \models s'(b) \leq s'(r')$ and it follows again from lemma 4.15 (10) that $\mathcal{M}[\sigma] \models s(b) \leq s'(r')$. Therefore (4.75) holds. It is easy to check (using again lemma 4.15 (10)) that the following holds

$$\mathbf{T} \models b \leq r \wedge \exists y \leq s(b) A(y, \vec{a}, \vec{M}) \to \exists y \leq s'(r') A(y, \vec{a}, \vec{M})$$

and therefore (by the completeness theorem) the following sequent has an LK^2-Φ proof

$$b \leq r, \exists y \leq s(b) A(y, \vec{a}, \vec{M}) \ \vdash \ \exists y \leq s'(r') A(y, \vec{a}, \vec{M}). \tag{4.76}$$

With structural rules and $\wedge$-left we obtain

$$b \leq r \wedge B(b), \Gamma, \exists y \leq s(b) A(y, \vec{a}, \vec{M}) \ \vdash \ \Delta', \exists y \leq s'(r') A(y, \vec{a}, \vec{M}). \tag{4.77}$$

Now we can apply the cut rule[i] with cut formula $\exists y \leq s(b) A(y, \vec{a}, \vec{M})$ on (4.74) and (4.77) and obtain the sequent

$$b \leq r \wedge B(b), \Gamma \ \vdash \ \Delta', \exists y \leq s'(r') A(y, \vec{a}, \vec{M}) \tag{4.78}$$

Since $s'(r')$ does not contain b we can now apply the $\exists$-left rule to obtain the sequent $\hat{\mathsf{S}}$

$$\exists x \leq r B(x), \Gamma \ \vdash \ \Delta', \exists y \leq s'(r') A(y, \vec{a}, \vec{M})$$

Case III′: S is obtained by the **set $\exists$-left** rule. In this case, the $\exists$ quantifier introduced is again bounded and S is the bottom sequent of an inference

$$\frac{M \leq r \wedge B(M), \Gamma \ \vdash \ \Delta}{\exists X \leq r B(X), \Gamma \ \vdash \ \Delta}$$

We proceed exactly as in case III with the eigenvariable b replaced by M.

Case IV: S is obtained by the **$\forall$-right** rule. In this case, the $\forall$ quantifier introduced is bounded (see case III). Thus S is the bottom sequent of an inference

[i] Note that this application of the cut rule needs not be anchored.

$$\frac{\Gamma \vdash \Delta, b \leq r \to B(b)}{\Gamma \vdash \Delta, \forall x \leq r B(x)}\,(\forall\text{-right})$$

If Δ does not contain $\exists y A(y, \vec{a}, \vec{M})$, then let $\hat{\mathsf{S}} = \mathsf{S}$. The case where Δ contains $\exists y A(y, \vec{a}, \vec{M})$ requires again special care because the eigenvariable restriction might be violated. In this case, by the induction hypothesis, the top sequent was converted to

$$\Gamma \vdash \Delta', \exists y \leq s(b) A(y, \vec{a}, \vec{M}), b \leq r \to B(b) \tag{4.79}$$

In order to apply the $\forall$-right rule to (4.79) we need to replace $s(b)$ by a term that does not contain b and whose variables occur free in S. We proceed as in case III to obtain an LK^2-Φ proof of the sequent (cf. (4.76))

$$b \leq r, \exists y \leq s(b) A(y, \vec{a}, \vec{M}) \vdash \exists y \leq s'(r') A(y, \vec{a}, \vec{M}). \tag{4.80}$$

Using the $\neg$-right rule together with weakenings and an application of the $\vee$-right rule (recall that $\neg b \leq r \vee B(b) \equiv b \leq r \to B(b)$) we obtain

$$\Gamma, \exists y \leq s(b) A(y, \vec{a}, \vec{M}) \vdash \Delta', \exists y \leq s'(r') A(y, \vec{a}, \vec{M}), b \leq r \to B(b) \tag{4.81}$$

Now we can apply the cut rule with cut formula $\exists y \leq s(b) A(y, \vec{a}, \vec{M})$ on (4.81) and (4.79) and obtain the sequent

$$\Gamma \vdash \Delta', \exists y \leq s'(r') A(y, \vec{a}, \vec{M}), b \leq r \to B(b) \tag{4.82}$$

Since $s'(r')$ does not contain b we can now apply the $\forall$-right rule to obtain the sequent $\hat{\mathsf{S}}$

$$\Gamma \vdash \Delta', \exists y \leq s'(r') A(y, \vec{a}, \vec{M}), \forall x \leq r B(x).$$

Case IV′: S is obtained by the **set $\forall$-right** rule. In this case, S is the bottom sequent of an inference

$$\frac{\Gamma \vdash \Delta, M \leq r \to B(M)}{\Gamma \vdash \Delta, \forall X \leq r B(X)}\,(\forall\text{-right})$$

We proceed exactly as in case IV with the eigenvariable b replaced by M.

Case V: S is obtained by a rule with two premises, i.e. **$\wedge$-right**, **$\vee$-left** or **cut**. Note that in all cases, the formula $\exists y A(y, \vec{a}, \vec{M})$ can only occur in the context Δ (cf. discussion above). Here, the following problem arises: The contexts (i.e. Δ) of the two premises might have been converted to $\Delta', \exists y \leq t_1 A(y, \vec{a}, \vec{M})$ and $\Delta', \exists y \leq t_2 A(y, \vec{a}, \vec{M})$ with $t_1 \not\equiv t_2$. In this case we can proceed as in case II to convert the two premises such that Δ has the form $\Delta', \exists y \leq (t_1 + t_2) A(y, \vec{a}, \vec{M})$.

Case VI: S is obtained by the inference

$$\frac{\Gamma \vdash \Delta, \exists y A(y, \vec{a}, \vec{M}), \exists y A(y, \vec{a}, \vec{M})}{\Gamma \vdash \Delta, \exists y A(y, \vec{a}, \vec{M})}\ (\textbf{contraction-right})$$

Let S_1 denote the top sequent of this inference. In this case we apply the induction hypothesis and set $\hat{\mathsf{S}} = \hat{\mathsf{S}_1}$.

Case VII: In all remaining cases (especially the $\forall$**-left** rules) we can apply the induction hypothesis to the top sequent and then apply the corresponding rule again. □

Theorem 4.49 (Parikh's Theorem)**.** *Let* $\mathbf{T}$ *be a polynomially bounded theory and* $A(\vec{x}, \vec{y}, \vec{X}, \vec{Y})$ *be a bounded formula with all free variables indicated. Assume*

$$\mathbf{T} \models \exists \vec{y} \exists \vec{Y}\, A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}). \tag{4.83}$$

Then there exists an $\mathcal{L}^2_{\mathcal{A}}$*-number term* $t(\vec{x}, \vec{X})$ *with no variables other than* $\vec{x}, \vec{X}$ *s.t.*

$$\mathbf{T} \models \exists \vec{y} \leq t(\vec{x}, \vec{X})\ \exists \vec{Y} \leq t(\vec{x}, \vec{X})\ A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}). \tag{4.84}$$

Proof of Parikh's Theorem from Lemma 4.48. Let $\mathcal{M}$ be a model of $\mathbf{T}$ and σ an arbitrary assignment. By the assumption (4.83) we have

$$\mathcal{M}[\sigma] \models \exists \vec{y} \exists \vec{Y} A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}). \tag{4.85}$$

From axiom **B8** and lemma 4.15 it follows that for all $i = 1, \ldots, n$ and for all $j = 1, \ldots, m$

$$\begin{aligned} \mathcal{M}[\sigma] &\models y_i \leq y_1 + \cdots + y_n + |Y_1| + \cdots + |Y_m| \\ \mathcal{M}[\sigma] &\models |Y_j| \leq y_1 + \cdots + y_n + |Y_1| + \cdots + |Y_m| \end{aligned} \tag{4.86}$$

and therefore[i]

$$\mathcal{M}[\sigma] \models \exists z \exists \vec{y} \leq z \exists \vec{Y} \leq z A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}). \tag{4.87}$$

Since $\exists \vec{y} \leq z \exists \vec{Y} \leq z A(\vec{x}, \vec{y}, \vec{X}, \vec{Y})$ is a bounded formula we can now apply lemma 4.48 and obtain

$$\mathcal{M}[\sigma] \models \exists z \leq t(\vec{x}, \vec{X}) \exists \vec{y} \leq z \exists \vec{Y} \leq z A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}). \tag{4.88}$$

for some $\mathcal{L}^2_{\mathcal{A}}$-number term $t(\vec{x}, \vec{X})$ with all variables indicated. By transitivity (lemma 4.15 (10)) we obtain

$$\mathcal{M}[\sigma] \models \exists \vec{y} \leq t(\vec{x}, \vec{X}) \exists \vec{Y} \leq t(\vec{x}, \vec{X}) A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}). \tag{4.89}$$

Therefore $\mathbf{T} \models \exists \vec{y} \leq t(\vec{x}, \vec{X}) \exists \vec{Y} \leq t(\vec{x}, \vec{X}) A(\vec{x}, \vec{y}, \vec{X}, \vec{Y})$ and we are done. □

[i] by mapping "z" to $(y_1 + \cdots + y_n + |Y_1| + \cdots + |Y_m|)^{\mathcal{M}}[\sigma]$

4.6 Properties of $\mathbf{V^1}$

In this section we present some definitions and results that are needed to prove the lower and the upper bound of $\mathbf{V^1}$ in the following sections.

4.6.1 Set Arrays and Sequence Numbers

For the prove of the next lemma we need the notion of *maximum nesting depth*.

Definition 4.50 (Maximum Nesting Depth $\sharp$)**.** *Let $\mathcal{F}$ be a set of function symbols. The maximum nesting depth $\sharp_{\mathcal{F}} t$ of a term t with respect to $\mathcal{F}$ is defined as follows*

$$\sharp_{\mathcal{F}} x = 0,\ \sharp_{\mathcal{F}} X = 0,$$
$$\sharp_{\mathcal{F}} f(t_1, \ldots, T_n) = \begin{cases} \max(\sharp_{\mathcal{F}} t_1, \ldots, \sharp_{\mathcal{F}} T_n) & \text{if } f \notin \mathcal{F} \\ \max(\sharp_{\mathcal{F}} t_1, \ldots, \sharp_{\mathcal{F}} T_n) + 1 & \text{if } f \in \mathcal{F}, \end{cases}$$
$$\sharp_{\mathcal{F}} F(t_1, \ldots, T_n) = \begin{cases} \max(\sharp_{\mathcal{F}} t_1, \ldots, \sharp_{\mathcal{F}} T_n) & \text{if } F \notin \mathcal{F} \\ \max(\sharp_{\mathcal{F}} t_1, \ldots, \sharp_{\mathcal{F}} T_n) + 1 & \text{if } F \in \mathcal{F}. \end{cases}$$

The maximum nesting depth $\sharp_{\mathcal{F}} A$ of a formula A with respect to $\mathcal{F}$ is defined as follows

$$\sharp_{\mathcal{F}} P(t_1, \ldots, T_n) = \max(\sharp_{\mathcal{F}} t_1, \ldots, \sharp_{\mathcal{F}} T_n),$$
$$\sharp_{\mathcal{F}} (A \otimes B) = \max(\sharp_{\mathcal{F}} A, \sharp_{\mathcal{F}} B),$$
$$\sharp_{\mathcal{F}} \exists x A = \sharp_{\mathcal{F}} \exists X A = \sharp_{\mathcal{F}} \forall x A = \sharp_{\mathcal{F}} \forall X A = \sharp_{\mathcal{F}} A$$

The next lemma shows that we can add a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-bit-defining axiom to a polynomially bounded theory, and then, for every $\Sigma^B_0(\mathcal{L})$-formula over the resulting language $\mathcal{L}$, the resulting theory contains a provably equivalent $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula. It will later follow from lemma 4.72 that the resulting theory is even a conservative extension.

Remark 4.51. Note that theorem 4.76 in section 4.6.3 only guarantees that the equivalent formula is Σ^B_1.

Lemma 4.52 (Σ^B_0-Transformation Lemma)**.** *Let $\mathbf{T}$ be a polynomially bounded theory and let $\mathcal{L}$ be its language. Assume that $\mathcal{L}$ has the same predicate symbols as $\mathcal{L}^2_{\mathcal{A}}$. Further, assume that for every number function f in $\mathcal{L}$, $\mathbf{T}$ contains a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-defining axiom for f, and for every set function F in $\mathcal{L}$, $\mathbf{T}$ contains a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-bit-defining axiom for F. Then for every $\Sigma^B_0(\mathcal{L})$-formula A^+ there is a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula A s.t.*

$$\mathbf{T} \models A^+ \leftrightarrow A \tag{4.90}$$

Proof. The proof is by induction on the maximum nesting depth $\sharp_{\mathcal{L}\setminus\mathcal{L}^2_{\mathcal{A}}} A^+$ of A^+ w.r.t to function symbols in $\mathcal{L}\setminus\mathcal{L}^2_{\mathcal{A}}$. If $\sharp_{\mathcal{L}\setminus\mathcal{L}^2_{\mathcal{A}}} A^+ = 0$, then A^+ is already a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula and there is nothing to prove.

For the induction step it is enough to consider atomic formulas. By assumption, we have to consider the predicate symbols of $\mathcal{L}^2_{\mathcal{A}}$: $\in, \leq, =$. We show the case of $\in$. The other cases are similar. We first assume that A^+ has the form $F(\vec{t},\vec{T})(s)$. By assumption, $\mathbf{T}$ contains a bit-defining axiom

$$F(\vec{x},\vec{X})(i) \leftrightarrow i < r(\vec{x},\vec{X}) \wedge A_F(i,\vec{x},\vec{X})$$

for F, where $r(\vec{x},\vec{X})$ is an $\mathcal{L}^2_{\mathcal{A}}$-term and $A_F(i,\vec{X},\vec{X})$ is a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula. Hence

$$\mathbf{T} \models A^+ \leftrightarrow s < r(\vec{t},\vec{T}) \wedge A_F(s,\vec{t},\vec{T}) \tag{4.91}$$

Note that in the RHS of (4.91) the terms $\vec{T}$ can only occur in the form $|\vec{T}|$. Hence, when we consider only the function symbols not in $\mathcal{L}^2_{\mathcal{A}}$, every atomic subformula of the RHS is of the form $B(\vec{s'})$ where $B(\vec{x})$ is an atomic $\mathcal{L}^2_{\mathcal{A}}$-formula. Let n be the arity of $\vec{s'}$. Hence, for all $i = 1,\ldots,n$, the term s'_i is either of the form $f(\vec{t},\vec{T})$ or $|G(\vec{t},\vec{T})|$ for $f, G \in \mathcal{L}\setminus\mathcal{L}^2_{\mathcal{A}}$. In both cases it follows either from the defining axiom of f or from lemma 4.53 below and from the fact that $\mathbf{T}$ is polynomially bounded (cf. definition 4.45) that there exists a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula $C_i(z,\vec{x},\vec{X})$ and an $\mathcal{L}^2_{\mathcal{A}}$-term $r'_i(\vec{x},\vec{X})$ s.t.

$$\mathbf{T} \models z = s'_i \leftrightarrow z \leq r'_i(\vec{t},\vec{T}) \wedge C_i(z,\vec{t},\vec{T})$$

Therefore

$$\mathbf{T} \models B(\vec{s'}) \leftrightarrow \exists z_1 \leq r'_1(\vec{t},\vec{T}) \ldots \exists z_n \leq r'_n(\vec{t},\vec{T}) B(z_1,\ldots,z_n) \wedge \bigwedge_{i=1,\ldots,n} C_i(z_i,\vec{t},\vec{T}) \tag{4.92}$$

Note that the maximum nesting depth of the RHS of (4.92) is strictly smaller than of A^+. Therefore we can apply the induction hypothesis to obtain a (provably) equivalent $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula for every atomic subformula of the RHS of (4.91). It is easy to check that this gives us a $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula that is provably equivalent to A^+.

Now assume that A^+ has the form $X(s)$. When we consider only the function symbols not in $\mathcal{L}^2_{\mathcal{A}}$, then, as above, we can write A^+ as $A^+(\vec{s'})$ where every term s'_i is either of the form $f(\vec{t},\vec{T})$ or $|G(\vec{t},\vec{T})|$, for $f, G \in \mathcal{L}\setminus\mathcal{L}^2_{\mathcal{A}}$. By the same argument as above we have

$$\mathbf{T} \models z = s'_i \leftrightarrow z \leq r_i(\vec{t},\vec{T}) \wedge C_i(z,\vec{t},\vec{T})$$

for some $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formula $C_i(z,\vec{x},\vec{X})$ and an $\mathcal{L}^2_{\mathcal{A}}$-term $r_i(\vec{x},\vec{X})$. Then $\mathbf{T}$ proves the corresponding formula of the form (4.92) and we can apply the induction hypothesis. □

Lemma 4.53 (Auxiliary Lemma for Lemma 4.52)**.** *Assume that a theory* $\mathbf{T} \supseteq \mathbf{V^0}$ *contains the* $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$*-bit-defining axiom*

$$F(\vec{x}, \vec{X})(i) \leftrightarrow i < t(\vec{x}, \vec{X}) \wedge A(i, \vec{x}, \vec{X}) \tag{4.93}$$

for a set function F*. Then there exists a* $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$*-formula* $B(z, \vec{x}, \vec{x})$ *s.t.*

$$\mathbf{T} \models z = |F(\vec{x}, \vec{X})| \leftrightarrow B(z, \vec{x}, \vec{X})$$

Proof. It follows from the axioms of $\mathbf{V^0}$ and corollary 4.44 (which shows that the empty set is unique) that

$$\mathbf{T} \models z = |F(\vec{x}, \vec{X})| \leftrightarrow \exists x < z\Big(\big(x + 1 = z \wedge F(\vec{x}, \vec{X})(x)\big) \vee z = 0\Big)$$

By replacing $F(\vec{x}, \vec{X})(x)$ with the RHS of (4.93) we obtain what we want. □

Note that it follows from lemma 4.52 that if we extend $\mathbf{V^1}$ by Σ^B_0-defining axioms and Σ^B_0-bit-defining axioms, then the resulting theory $\mathbf{V^1}(\mathcal{L})$ also proves $\Sigma^B_1(\mathcal{L})$-**COMP** and $\Sigma^B_1(\mathcal{L})$-**IND**, where $\mathcal{L}$ is the resulting language. This is due to the fact that by lemma 4.52, all $\Sigma^B_0(\mathcal{L})$-subformulas of $\Sigma^B_1(\mathcal{L})$-**COMP** and $\Sigma^B_1(\mathcal{L})$-**IND** have provably equivalent $\Sigma^B_0(\mathcal{L}^2_{\mathcal{A}})$-formulas in $\mathbf{V^1}(\mathcal{L})$.

It will be useful to be able to encode multiple numbers into one number. In order to do this we need a pairing function. Since this function needs not be bijective (only injective) we can use a function that is simpler than the well-known pairing function of Cantor.

Definition 4.54 (Pairing Function)**.** *For number terms* s, t *we define*

$$\langle s, t\rangle \equiv_{def} (s + t) \times (s + t + 1) + \underline{2} \times t$$

Of course we can encode arbitrary n-tuples by defining $\langle s_1, \ldots, s_n\rangle \equiv_{def} \langle\langle s_1, \ldots, s_{n-1}\rangle, s_n\rangle$. The pairing function is injective (in $\mathbf{V^0}$ and also in $\mathbf{I\Delta_0}$), but the proof thereof is cumbersome and we omit it.

Lemma 4.55. $\mathbf{V^0}$ *proves that the pairing function is injective, i.e.*[i]

$$\mathbf{V^0} \models \langle x_1, y_1\rangle = \langle x_2, y_2\rangle \leftrightarrow (x_1 = x_2 \wedge y_1 = y_2)$$

Corollary 4.56.

$$\mathbf{V^0} \models \langle x_1, \ldots, x_n\rangle = \langle y_1, \ldots, y_n\rangle \leftrightarrow (x_1 = y_1 \wedge \cdots \wedge x_n = y_n)$$

[i]Note that the direction $\leftarrow$ trivially holds because of the requirement that $=$ is interpreted as the equality relation.

Proof. By induction on n. By lemma 4.55 we have

$$\begin{aligned}\mathbf{V^0} \models \langle\langle x_1, \ldots, x_{n-1}\rangle, x_n\rangle = \langle\langle y_1, \ldots, y_{n-1}\rangle, y_n\rangle \leftrightarrow \\ (\langle x_1, \ldots, x_{n-1}\rangle = \langle y_1, \ldots, y_{n-1}\rangle \wedge x_n = y_n)\end{aligned}$$

By the induction hypothesis we have

$$\begin{aligned}\mathbf{V^0} \models \langle\langle x_1, \ldots, x_{n-1}\rangle, x_n\rangle = \langle\langle y_1, \ldots, y_{n-1}\rangle, y_n\rangle \leftrightarrow \\ (x_1 = y_1 \wedge \cdots \wedge x_{n-1} = y_{n-1} \wedge x_n = y_n)\end{aligned}$$

and are done. □

Lemma 4.57 (Strict Monotonicity of $\langle x, y\rangle$)**.**

$$\mathbf{V^0} \models (x_1 < y_1 \wedge \cdots \wedge x_n < y_n) \rightarrow \langle x_1, \ldots, x_n\rangle < \langle y_1, \ldots, y_n\rangle$$

Proof. It follows from the monotonicity of $\mathcal{L}^2_{\mathcal{A}}$-terms (lemma 4.46) that

$$\mathbf{V^0} \models (x_1 \leq y_1 \wedge \cdots \wedge x_n \leq y_n) \rightarrow \langle x_1, \ldots, x_n\rangle \leq \langle y_1, \ldots, y_n\rangle$$

Then the lemma follows from corollary 4.56 (contraposition of direction $\rightarrow$). □

We now show that $\mathbf{V^0}$ contains the comprehension axioms (definition 4.3) also for more than one variable, a result we will use later.

Definition 4.58 (Φ-**MULTICOMP**)**.** *Let Φ be a set of formulas. Then* Φ-**MULTICOMP** *is the set of all formulas of the form*

$$\exists X \leq \langle y_1, \ldots, y_n\rangle \forall z_1 < y_1 \ldots \forall z_n < y_n \big(X(\langle z_1, \ldots, z_n\rangle) \leftrightarrow A(z_1, \ldots, z_n)\big)$$

where $n \geq 2$, $A(z_1, \ldots, z_n) \in \Phi$ and X does not occur free in $A(z_1, \ldots, z_n)$.

Lemma 4.59 (Multiple Comprehension)**.** *Let $\mathbf{T} \supseteq \mathbf{V^0}$ be a theory, $\mathcal{L}$ its language and assume $\mathbf{T} \models \Sigma^B_0(\mathcal{L})$-**COMP**. Then $\mathbf{T} \models \Sigma^B_0(\mathcal{L})$-**MULTICOMP**.*

Proof. We show the case $n = 2$ and it will be clear that the same proof works for all $n \geq 2$. I.e. we have to show

$$\mathbf{T} \models \exists X \leq \langle y_1, y_2\rangle \forall z_1 < y_1 \forall z_2 < y_2 \big(X(\langle z_1, z_2\rangle) \leftrightarrow A(z_1, z_2)\big). \tag{4.94}$$

By Σ^B_0-**COMP** we have

$$\begin{aligned}\mathbf{T} \models \exists X \leq \langle y_1, y_2\rangle \forall z < \langle y_1, y_2\rangle \\ \Big(X(z) \leftrightarrow \exists v_1 < z \exists v_2 < z \big(z = \langle v_1, v_2\rangle \wedge A(v_1, v_2)\big)\Big)\end{aligned} \tag{4.95}$$

Let $\mathcal{M}$ be a model of $\mathbf{T}$ and let σ be an arbitrary assignment. By (4.95) we have for some X

$$\begin{aligned}\mathcal{M}[\sigma(\mathsf{X}/X)] \models |X| \leq \langle y_1, y_2\rangle \wedge \forall z < \langle y_1, y_2\rangle \\ \Big(X(z) \leftrightarrow \exists v_1 < z \exists v_2 < z\big(z = \langle v_1, v_2\rangle \wedge A(v_1, v_2)\big)\Big)\end{aligned} \tag{4.96}$$

We want to show that

$$\mathcal{M}[\sigma(\mathsf{X}/X)] \models \forall z_1 < y_1 \forall z_2 < y_2\big(X(\langle z_1, z_2\rangle) \leftrightarrow A(z_1, z_2)\big) \tag{4.97}$$

i.e. for arbitrary $\mathsf{z_1}, \mathsf{z_2}$

$$\begin{aligned}&\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{z_1}/z_1)(\mathsf{z_2}/z_2)] \models \\ &z_1 < y_1 \rightarrow \Big(z_2 < y_2 \rightarrow \underbrace{\big(X(\langle z_1, z_2\rangle) \leftrightarrow A(z_1, z_2)\big)}_{(*)}\Big)\end{aligned} \tag{4.98}$$

So assume $\mathcal{M}[\sigma(\mathsf{X}/X)(\mathsf{z_1}/z_1)(\mathsf{z_2}/z_2) \models z_1 < y_1 \wedge z_2 < y_2$. We have to show that $(*)$ holds in this $\mathcal{M}[\sigma(\dots)]$. By lemma 4.57, $\mathcal{M}[\sigma(\dots)] \models \langle z_1, z_2\rangle < \langle y_1, y_2\rangle$ and therefore by (4.96)

$$\begin{aligned}&\mathcal{M}[\sigma(\dots)] \models\ X(\langle z_1, z_2\rangle) \leftrightarrow \\ &\exists v_1 < \langle z_1, z_2\rangle \exists v_2 < \langle z_1, z_2\rangle\big(\langle z_1, z_2\rangle = \langle v_1, v_2\rangle \wedge A(v_1, v_2)\big)\Big)\end{aligned} \tag{4.99}$$

Because $=^{\mathcal{M}}$ is the true equality relation it follows that $(*)$ holds in $\mathcal{M}[\sigma(\dots)]$. Therefore (4.97) holds and (4.94) as well. □

Since the above pairing function (definition 4.54) encodes a finite set of numbers (actually a finite sequence of numbers) as one number, a set X can be viewed as an "array" of sets of natural numbers. And therefore the encoding $w(X)$ encodes an array of binary strings (cf. section 4.4). Now we can use the above concepts to define a function $\mathrm{row}(i, X)$ that returns row i of the "array" X. Intuitively, $\mathrm{row}(i, X)$ contains a number z iff the pair $\langle i, z\rangle$ is in the set X.

Definition 4.60. *The function* $\mathrm{row}(i, X)$ *has the* Σ_0^B*-bit-defining axiom*

$$\mathrm{row}(i, X)(z) \leftrightarrow z < |X| \wedge X(\langle i, z\rangle)$$

We often write $X[i]$ *instead of* $\mathrm{row}(i, X)$. $\mathbf{T}(\mathrm{row})$ *is the extension of* $\mathbf{T}$ *obtained by adding the above axiom to* $\mathbf{T}$.

Corollary 4.61. row *is in* FAC^0.

Proof. By corollary 4.36. □

Notation Let $\mathbf{T}$ be a theory and let $\mathcal{F}$ be a list of functions that are definable in $\mathbf{T}$ or bit-definable. Then $\mathbf{T}(\mathcal{F})$ denotes the theory that is obtained by adding to $\mathbf{T}$ the defining axioms or the bit-defining axioms for the function symbols in $\mathcal{F}$.

The following lemma shows that in $\mathbf{V^0}(\mathrm{row})$, for any n sets $X_1, \ldots, X_n$ there exists a set Y (an "array") that encodes $X_1, \ldots, X_n$.

Lemma 4.62.

$$\mathbf{V^0}(\mathrm{row}) \models \exists Y \leq \langle \underline{n}, |X_1| + \cdots + |X_n| \rangle \ \big(X_1 = Y[0] \wedge \cdots \wedge X_n = Y[\underline{n-1}]\big)$$

Proof. By lemma 4.52 and lemma 4.59,

$$\mathbf{V^0}(\mathrm{row}) \models \Sigma^B_1(\mathrm{row})\text{-}\mathbf{MULTICOMP}$$

and hence

$$\mathbf{V^0}(\mathrm{row}) \models \exists Y \leq \langle \underline{n}, |X_1| + \cdots + |X_n| \rangle \ \forall x < \underline{n} \forall y < |X_1| + \cdots + |X_n| \\ \Big(Y(\langle x, y \rangle) \leftrightarrow \big(x = 0 \wedge X_1(y)\big) \vee \cdots \vee \big(x = \underline{n-1} \wedge X_n(y)\big)\Big)$$

Then the lemma follows from definition 4.60, the fact that $\mathbf{V^0}$ proves the formula $x < y \rightarrow x < y + z$ (lemma 4.15) and the axiom **SE**. □

With the help of the row function, multiple existential set quantifiers can be collapsed into a single one, as the next lemma shows.

Definition 4.63 (single Σ^B_1-formula)**.** *A Σ^B_1-formula of the form $\exists X \leq tA(X)$, where $A(X) \in \Sigma^B_0$, is called a* single Σ^B_1*-formula.*

Lemma 4.64. *Let $\mathbf{T} \supseteq \mathbf{V^0}(\mathrm{row})$ be a polynomially bounded theory over a language $\mathcal{L}$. Then for every $\Sigma^B_1(\mathcal{L})$-formula A there is a single $\Sigma^B_1(\mathcal{L})$-formula A' s.t. $\mathbf{T} \models A \leftrightarrow A'$.*

Proof. By assumption, A has the form

$$\exists X_1 \leq t_1 \ldots \exists X_n \leq t_n B(X_1, \ldots, X_n)$$

where $B(X_1, \ldots, X_n) \in \Sigma^B_0(\mathcal{L})$. Consider the single $\Sigma^B_1(\mathcal{L})$-formula

$$A' \equiv \exists Z \leq \langle \underline{n}, t_1 + \cdots + t_n \rangle \\ \big(|Z[0]| \leq t_1 \wedge \cdots \wedge |Z[\underline{n-1}]| \leq t_n \wedge B(Z[0], \ldots, Z[\underline{n-1}])\big)$$

We have to show that $\mathbf{T} \models A \leftrightarrow A'$. The direction $A' \rightarrow A$ is logically valid and the direction $A \rightarrow A'$ follows from lemma 4.62 and the monotonicity of $\mathcal{L}^2_{\mathcal{A}}$-terms (lemma 4.46). □

We use a similar idea to let a set X encode a sequence $x_0, \ldots, x_n$ of numbers[i] and we define the function seq to extract an x_i. Intuitively, $\mathrm{seq}(i, X)$ returns the smallest number of $X[i]$, or $|X|$ if $X[i]$ is empty.

Definition 4.65. *The function* $\mathrm{seq}(i, X)$ *has the defining axiom*

$$y = \mathrm{seq}(i, X) \leftrightarrow \big(y < |X| \wedge X(\langle i, y\rangle) \wedge \forall z < y \neg X(\langle i, z\rangle)\big) \vee \big(\forall z < |X| \neg X(\langle i, z\rangle) \wedge y = |X|\big)$$

We often write $X^{[i]}$ *instead of* $\mathrm{seq}(i, X)$.

Since seq has a Σ^B_0-defining axiom, it follows from lemma 4.36 that $\mathrm{seq} \in \mathsf{FAC}^0$. Later, it follows from lemma 4.78 that the function seq is Σ^B_0-definable in $\mathbf{V^0}$.

4.6.2 The Replacement Scheme

In section 4.8 we need the fact that any formula of the form

$$\forall x \leq t_1 \exists X \leq t_2 A(x, X) \tag{4.100}$$

where $A(x, X)$ is Σ^B_0, is provably equivalent in $\mathbf{V^0}$ to a Σ^B_1-formula. Informally, the idea is to use the function row introduced in section 4.6.1 to encode the finitely many values of X for every $x \leq t_1$ by a set array Z. To characterise formulas of the form (4.100) we introduce the new formula classes ${}_e\Sigma^B_i$ and ${}_e\Pi^B_i$ (the e stands for "extended").

Definition 4.66 (${}_e\Sigma^B_i$ and ${}_e\Pi^B_i$)**.** *Let* $\mathcal{L} \supseteq \mathcal{L}^2_{\mathcal{A}}$. *Then*

$${}_e\Sigma^B_0(\mathcal{L}) = {}_e\Pi^B_0(\mathcal{L}) = \Sigma^B_0(\mathcal{L}).$$

For $i \geq 0$, ${}_e\Sigma^B_{i+1}(\mathcal{L})$ *is* ${}_e\Pi^B_i(\mathcal{L})$, *closed under* $\vee, \wedge, \forall x \leq t, \exists x \leq t$ *and* $\exists X \leq t$. ${}_e\Pi^B_{i+1}(\mathcal{L})$ *is* ${}_e\Sigma^B_i(\mathcal{L})$, *closed under* $\vee, \wedge, \forall x \leq t, \exists x \leq t$ *and* $\forall X \leq t$.

Definition 4.67 (Replacement Scheme)**.** *Let* Φ *be a set of* $\mathcal{L}$*-formulas. The replacement scheme* Φ-**REPL** *is the set of all* $\mathcal{L} \cup \{\mathrm{row}\}$*-formulas of the form*

$$\forall x \leq b \exists X \leq c\ A(x, X) \ \rightarrow\ \exists Z \leq \langle b, c\rangle \forall x \leq b\big(\big|Z[x]\big| \leq c \wedge A(x, Z[x])\big) \tag{4.101}$$

where $A(x, X) \in \Phi$.

We first show that the right to left direction of (4.101) is a valid formula.

[i] Note that it is not sufficient to "sort" X since X is not a multiset.

Lemma 4.68. *Let B_L be the LHS and B_R be the RHS of* (4.101)*. Then the formula $B_R \to B_L$ is valid.*

Proof. Note that the above statement is stronger than $B_R \Rightarrow B_L$ (cf. definition 2.5 of logical consequence). We have to show that for every structure $\mathcal{M}$ and for every assignment σ, if $\mathcal{M}[\sigma] \models B_R$, then $\mathcal{M}[\sigma] \models B_L$. So assume $\mathcal{M}[\sigma] \models B_R$. Then there is a Z s.t. for all x we have

$$\mathcal{M}[\sigma(\mathsf{Z}/Z)(\mathsf{x}/x)] \models x \leq b \to \big|Z[x]\big| \leq c \wedge A(x, Z[x])$$

Let Z' denote the element $\mathcal{M}^{|Z[x]|}[\sigma]$. Then it is obvious that Z' satisfies the existential set quantifier in B_L. □

We want to show that $\mathbf{V^1}$ and the theory $\widetilde{\mathbf{V^1}}$ that we will introduce in section 4.8 prove Σ^B_1-**REPL**.

Lemma 4.69. *Let $\mathbf{T} \supseteq \mathbf{V^0}$ be a polynomially bounded theory over a language $\mathcal{L}$. Assume that $\mathbf{T} \models \Sigma^B_{i+1}(\mathcal{L})$-**IND**, for some $i \geq 0$. Then $\mathbf{T}$(row) $\models \Sigma^B_{i+1}(\mathcal{L})$-**REPL**.*

Proof sketch. We first prove that $\mathbf{T}$ proves $\Pi^B_i(\mathcal{L})$-**REPL** and we first show the case where $i = 0$. Let $A(x, X)$ be a $\Pi^B_0(\mathcal{L})$-formula, i.e. a $\Sigma^B_0(\mathcal{L})$-formula. Let $\mathcal{M} \models \mathbf{T}$(row) and let σ be an arbitrary assignment. Assume that the LHS of the replacement scheme (4.101) holds in $\mathcal{M}[\sigma]$, i.e.

$$\mathcal{M}[\sigma] \models \forall x \leq b \exists X \leq c\ A(x, X). \tag{4.102}$$

Then we have to show that

$$\mathcal{M}[\sigma] \models \exists Z \leq \langle b, c\rangle \forall x \leq b\big(\big|Z[x]\big| \leq c \wedge A(x, Z[x])\big) \tag{4.103}$$

(4.103) is a $\Sigma^B_1(\mathcal{L})$-formula. Therefore we can apply $\Sigma^B_1(\mathcal{L})$-**IND**. We show

$$\mathcal{M}[\sigma] \models z \leq b \to \exists Z \leq \langle z, c\rangle \forall x \leq z\big(\big|Z[x]\big| \leq c \wedge A(x, Z[x])\big) \tag{4.104}$$

Let $B(z)$ be the formula in (4.104). $\mathcal{M}[\sigma] \models B(0)$ follows from (4.102) (for "$x = 0$"). Then $\mathcal{M}[\sigma] \models B(z + 1)$ follows from the induction hypothesis $\mathcal{M}[\sigma] \models B(z)$, (4.102) and Σ^B_0-**COMP**.

For $i \geq 0$, let $A(x, X)$ be a $\Pi^B_i(\mathcal{L})$-formula. Then (4.103) is not a $\Sigma^B_{i+1}(\mathcal{L})$-formula. But, according to lemma 4.23, it is equivalent to

$$\exists Z \leq \langle b, c\rangle \forall Y \leq b\big(\big|Z[|Y|]\big| \leq c \wedge A(|Y|, Z[|Y|])\big) \tag{4.105}$$

which is equivalent to a $\Sigma^B_{i+1}(\mathcal{L})$-formula and hence we can proceed by induction as in the case of $i = 0$. Hence $\mathbf{T}$(row) $\models \Pi^B_i(\mathcal{L})$-**REPL**. It then follows that $\mathbf{T}$(row) also proves $\Sigma^B_{i+}(\mathcal{L})$-**REPL**. For details, see [6]. □

Now we can prove the lemma that allows us to eliminate formulas of the form (4.100). Recall (definition 4.63) that a single Σ_1^B-formula is a Σ_1^B-formula with one single (bounded) existential set quantifier.

Lemma 4.70. *Let* $\mathbf{T}$ *be a polynomially bounded theory over a language* $\mathcal{L}$ *with* $\mathbf{T}(\mathrm{row}) \models \Sigma_0^B(\mathcal{L})$-**REPL**. *Then for every* ${}_e\Sigma_1^B(\mathcal{L})$*-formula* A *there is a single* $\Sigma_1^B(\mathcal{L})$*-formula* A' *s.t.* $\mathbf{T} \models A \leftrightarrow A'$.

Proof. We prove the lemma by structural induction (cf. definition 4.66) on the ${}_e\Sigma_1^B(\mathcal{L})$-formula A. The base case is trivial because then A is a $\Sigma_0^B(\mathcal{L})$-formula and obviously $\mathbf{T} \models A \leftrightarrow \exists X \leq tA$, for some X not occurring in A. For the induction step, the only interesting case is $A \equiv \forall x \leq tB(x)$, where $B(x)$ is a ${}_e\Sigma_1^B(\mathcal{L})$-formula. By the induction hypothesis, $B(x)$ is equivalent in $\mathbf{T}$ to a single $\Sigma_1^B(\mathcal{L})$-formula $\exists X \leq t'B'(x, X)$, where $B'(x, X) \in \Sigma_0^B(\mathcal{L})$, hence

$$\mathbf{T} \models A \leftrightarrow \forall x \leq t \exists X \leq t'B'(x, X). \tag{4.106}$$

From $\mathbf{T}(\mathrm{row}) \models \Sigma_0^B(\mathcal{L})$-**REPL** and lemma 4.68 it follows that the RHS of (4.106) is equivalent in $\mathbf{T}(\mathrm{row})$ to a single $\Sigma_1^B(\mathcal{L} \cup \{\mathrm{row}\})$-formula A'. And by the Σ_0^B-Transformation lemma 4.52, there exists a provably equivalent (in $\mathbf{T}(\mathrm{row})$) formula A'' without occurrences of the row function (i.e. we eliminate row in every atomic subformula of A'). It follows that $\mathbf{T}(\mathrm{row}) \models A \leftrightarrow A''$ and since $\mathbf{T}(\mathrm{row})$ is a conservative extension of $\mathbf{T}$ by lemma 4.72 below[i], we also have $\mathbf{T} \models A \leftrightarrow A''$.

Then the other cases are easily proved with the help of lemma 4.64 which allows us to collapse several bounded set quantifiers into a single one. □

Corollary 4.71. *For every* ${}_e\Sigma_1^B$*-formula* A*, there is a single* Σ_1^B*-formula* A' *s.t.* $\mathbf{V^1} \models A \leftrightarrow A'$.

4.6.3 Conservative Extensions and Transformations

The following lemma shows that adding definable predicates and functions to an existing theory results in a conservative extension of that theory.

Lemma 4.72 (Extension by Definition)**.** *Assume that* $\mathbf{T}_1$ *results from* $\mathbf{T}_2$ *by adding to* $\mathbf{T}_1$ *the defining axioms of* Φ*-definable predicates and functions, for some* Φ*. Then* $\mathbf{T}_2$ *is a conservative extension of* $\mathbf{T}_1$.

Proof. Let A be a formula in the language of $\mathbf{T}_1$ and assume that $\mathbf{T}_2 \models A$. Let $\mathcal{M}$ be a model of $\mathbf{T}_1$. We extend $\mathcal{M}$ by interpreting the "new" predicate

[i]Note that row $\in \mathsf{FAC}^0$ by corollary 4.61. It will follow from lemma 4.78 that all FAC^0 functions are Σ_0^B-definable in $\mathbf{V^0}$.

and function symbols such that the corresponding defining axioms hold in $\mathcal{M}$. This interpretation is uniquely determined by the defining axioms and the totality condition (in the case of functions). Let $\mathcal{M}'$ be the obtained model. Then $\mathcal{M}'$ is a model of $\mathbf{T}_2$, hence $\mathcal{M}' \models A$. Since $\mathcal{M}$ is identical to $\mathcal{M}'$ with the exception of the new symbols, we also have $\mathcal{M} \models A$. Because $\mathcal{M}$ is an arbitrary model of $\mathbf{T}_1$, we have $\mathbf{T}_1 \models A$. □

It is important to note, however, that extending a theory $\mathbf{T}$ by adding a bit-definable set function together with its bit-defining axiom does not in general lead to a conservative extension of $\mathbf{T}$. However, the following lemma holds. Recall definition 4.32 about polynomially bounded functions.

Definition 4.73 (Σ^B_0-Closure)**.** *Let Φ be a set of formulas over $\mathcal{L} \supseteq \mathcal{L}^2_\mathcal{A}$. Then $\Sigma^B_0(\Phi)$ is the closure of Φ under the operations $\neg, \vee, \wedge$ and bounded number quantification. That is, if $A, B \in \Sigma^B_0(\Phi)$ and t is an $\mathcal{L}^2_\mathcal{A}$-term not containing x, then the following formulas are also in $\Sigma^B_0(\Phi)$: $\neg A, (A \vee B), (A \wedge B), \forall x \leq tA, \exists x \leq tA$.*

Note that $\Sigma^B_0(\Sigma^B_0) = \Sigma^B_0$ but $\Sigma^B_0(\Sigma^B_1) \neq \Sigma^B_1$. We will use the next lemma when we prove in lemma 4.78 that the FAC^0 functions are Σ^B_0-definable in $\mathbf{V^0}$ (i.e. the lower bound of $\mathbf{V^0}$).

Lemma 4.74. *Let $\mathbf{T} \supseteq \mathbf{V^0}$ be a theory over $\mathcal{L}$, and Φ a set of $\Sigma^B_0(\mathcal{L})$-formulas. Suppose $\mathbf{T} \models \Phi$-**COMP**. Then any polynomially bounded number function whose graph is Φ-definable is $\Sigma^B_0(\Phi)$-definable in $\mathbf{T}$. And any polynomially bounded set function that is Φ-bit-definable[i] is $\Sigma^B_0(\Phi)$-definable in $\mathbf{T}$.*

Proof. Let F be a polynomially bounded set function that is Φ-bit-definable. Then there are an $\mathcal{L}^2_\mathcal{A}$-term $t(\vec{x}, \vec{X})$ and a formula $A \in \Phi$ s.t. F has the following bit-defining axiom

$$F(\vec{x}, \vec{X})(i) \leftrightarrow i < t(\vec{x}, \vec{X}) \wedge A(i, \vec{x}, \vec{X})$$

The graph of F can be $\Sigma^B_0(\Phi)$-defined from its bit graph as

$$Y = F(\vec{x}, \vec{X}) \;\leftrightarrow\; |Y| \leq t(\vec{x}, \vec{X}) \;\wedge\; \forall i < t(\vec{x}, \vec{X})\big(Y(i) \leftrightarrow A(i, \vec{x}, \vec{X})\big) \tag{4.107}$$

Let G_F stand for the RHS of (4.107). Since $\mathbf{T} \models \Phi$-**COMP** we can use comprehension for the formula $A(i, \vec{x}, \vec{X})$ to obtain

$$\mathbf{T} \models \exists Y \leq t(\vec{x}, \vec{X}) \forall i < t(\vec{x}, \vec{X})\big(Y(i) \leftrightarrow A(i, \vec{x}, \vec{X})\big)$$

[i] i.e. the graph relation of F is Φ-definable.

which is equal to[i] $\mathbf{T} \models \exists Y G_F$. Because comprehension is unique by the axiom **SE** (see page 34) it follows that $\mathbf{T} \models \exists! Y G_F$. Since theories are closed under universal quantification we obtain $\mathbf{T} \models \forall \vec{x} \forall \vec{X} \exists! Y G_F$. Therefore F is $\Sigma_0^B(\Phi)$-definable in $\mathbf{T}$ (because $G_F \in \Sigma_0^B(\Phi)$).

Next, let f be a polynomially bounded number function whose graph is Φ-definable. Then the graph of f has a defining axiom which we can strengthen because f is bounded by a polynomial. Let the $\mathcal{L}_{\mathcal{A}}^2$-term $t(\vec{x}, \vec{X})$ be this bounding polynomial. The defining axiom of f's graph is then

$$y = f(\vec{x}, \vec{X}) \;\leftrightarrow\; y < t(\vec{x}, \vec{X}) \wedge A(y, \vec{x}, \vec{X})$$

for a formula $A \in \Phi$[ii]. We need to show that $\mathbf{T}$ proves the uniqueness of y. Since $\mathbf{T} \models \Phi$-**MIN** (theorem 4.17) we have

$$\mathbf{T} \models A(y, \vec{x}, \vec{X}) \rightarrow \Big(\exists z \leq y \big(A(z, \vec{x}, \vec{X}) \wedge \neg \exists z' < z A(z', \vec{x}, \vec{X}) \big) \Big)$$

Therefore, there is a least element y s.t. $A(y, \vec{x}, \vec{X})$ holds and the axioms **B1**-**B12** quarantee that this y is unique. Therefore we can define f in $\mathbf{T}$ by the following defining axiom

$$y = f(\vec{x}, \vec{X}) \;\leftrightarrow\; \forall z < y \neg A(z, \vec{x}, \vec{X}) \wedge y < t(\vec{x}, \vec{X}) \rightarrow A(y, \vec{x}, \vec{X})$$

and it follows from the above discussion that

$$\mathbf{T} \models \forall \vec{x} \forall \vec{X} \exists! y \forall z < y \neg A(z, \vec{x}, \vec{X}) \wedge y < t(\vec{x}, \vec{X}) \rightarrow A(y, \vec{x}, \vec{X})$$

□

The next corollary follows from lemma 4.74, lemma 4.72 and the fact that $\Sigma_0^B\big(\Sigma_0^B(\mathcal{L})\big) = \Sigma_0^B(\mathcal{L})$.

Corollary 4.75. *Let* $\mathbf{T}$ *be a theory over a language* $\mathcal{L}$ *and assume* $\mathbf{T} \models \Sigma_0^B(\mathcal{L})$-**COMP**. *The theory resulting from* $\mathbf{T}$ *by adding to* $\mathbf{T}$ *the* $\Sigma_0^B(\mathcal{L})$*-defining axioms or the* $\Sigma_0^B(\mathcal{L})$*-bit-defining axioms for a collection of number or set functions is a conservative extension of* $\mathbf{T}$.

In contrast to lemma 4.52, we now state a transformation theorem for Σ_1^1-definable functions.

[i] recall definition 4.1 of $\exists X \leq t$

[ii] We use the trivial fact that if f is bounded by some polynomial p, then the function value of f is sharply bounded (i.e. $<$) by some polynomial, namely $p + 1$. See definition 4.32.

Theorem 4.76 (Σ_1^B-Transformation)**.** *Let* $\mathbf{T}$ *be a polynomially bounded theory and* $\mathcal{L} \supseteq \mathcal{L}_{\mathcal{A}}^2 \cup \{\mathrm{row}\}$ *its language. Assume that* $\mathbf{T} \models \Sigma_0^B(\mathcal{L})$-**REPL**. *Let* $\mathbf{T}'$ *be an extension of* $\mathbf{T}$*, obtained by adding to* $\mathbf{T}$ *a* $\Sigma_1^1(\mathcal{L})$*-definable function (in* $\mathbf{T}$*) and its defining axiom. Let* $\mathcal{L}'$ *be the resulting language of* $\mathbf{T}'$*. Then*

(a) $\mathbf{T}'$ *is a conservative extension of* $\mathbf{T}$*.*

(b) $\mathbf{T}'$ *is polynomially bounded.*

(c) For every $\Sigma_0^B(\mathcal{L}')$*-formula* A^+*, there is a* $\Sigma_1^B(\mathcal{L})$*-formula* A *s.t.* $\mathbf{T}' \models A^+ \leftrightarrow A$*.*

(d) For every $\Sigma_1^B(\mathcal{L}')$*-formula* A^+*, there is a* $\Sigma_1^B(\mathcal{L})$*-formula* A *s.t.* $\mathbf{T}' \models A^+ \leftrightarrow A$*.*

(e) $\mathbf{T}' \models \Sigma_1^B(\mathcal{L}')$-**REPL**.

Proof. (a) is an immediate consequence of theorem 4.72.

(b): It follows from Parikh's theorem that the Σ_1^1-definable functions of $\mathbf{T}$ are polynomially bounded. Hence $\mathbf{T}'$ is polynomially bounded as well.

(c): We show the case where the $\Sigma_1^1(\mathcal{L})$-definable function is a set function. Let $F(\vec{x}, \vec{X})$ be this function. It follows from Parikh's theorem that F is $\Sigma_1^B(\mathcal{L})$-definable in $\mathbf{T}$ (cf. *(b)*). Thus there is a $\Sigma_1^B(\mathcal{L})$-formula $A_F(\vec{x}, \vec{X}, Y)$ s.t.

$$\underline{\mathbb{N}_2}' \models Y = F(\vec{x}, \vec{X}) \leftrightarrow A_F(\vec{x}, \vec{X}, Y), \tag{4.108}$$

$$\mathbf{T}' \models Y = F(\vec{x}, \vec{X}) \leftrightarrow A_F(\vec{x}, \vec{X}, Y), \tag{4.109}$$

$$\mathbf{T} \models \forall \vec{x} \forall \vec{X} \exists! Y \leq t(\vec{x}, \vec{X})\ A_F(\vec{x}, \vec{X}, Y) \tag{4.110}$$

for some $\mathcal{L}_{\mathcal{A}}^2$-term $t(\vec{x}, \vec{X})$[i]. We show that there is a ${}_e\Sigma_1^B(\mathcal{L})$-formula A' s.t. $\mathbf{T}' \models A^+ \leftrightarrow A'$. Then by lemma 4.70 there is a single $\Sigma_1^B(\mathcal{L})$-formula A s.t. $\mathbf{T} \models A' \leftrightarrow A$ and hence $\mathbf{T}' \models A^+ \leftrightarrow A$. We first show that it is sufficient to consider the case where A^+ is an atomic formula.

claim: Assume that for every atomic subformula B^+ of A^+ there is a ${}_e\Sigma_1^B(\mathcal{L})$-formula B s.t. $\mathbf{T}' \models B^+ \leftrightarrow B$. Then there is a ${}_e\Sigma_1^B(\mathcal{L})$-formula A s.t. $\mathbf{T}' \models A^+ \leftrightarrow A$.

The claim is easily proved by structural induction on A^+. The base case is trivial and the induction step follows immediately from the definition 4.66 of ${}_e\Sigma_1^B$.[ii]

Now we show *(c)* by induction on the maximum nesting depth $\sharp_F A^+$ of F in A^+ (cf. definition 4.50 on p. 62). In the base case F does not occur

[i] It follows also from Parikh's theorem that $\exists! Y$ can be bounded in (4.110).

[ii] Note that the case $A^+ \equiv \forall \vec{X} \leq \vec{t} A'$ does not occur since $A^+ \in {}_e\Sigma_1^B(\mathcal{L}')$.

in A^+ and there is nothing to prove. Assume that $F(\vec{s}, \vec{S})$ occurs in A^+. We can write A^+ as $A^+\big(F(\vec{s}, \vec{S})\big)$. From (4.109) and (4.110) it follows that

$$\mathbf{T}' \models A^+\big(F(\vec{s}, \vec{S})\big) \leftrightarrow \exists Y \leq t(\vec{s}, \vec{S})\big(A_F(\vec{s}, \vec{S}, Y) \wedge A^+(Y)\big) \tag{4.111}$$

where $A_F(\vec{s}, \vec{S}, Y) \in \Sigma^B_1(\mathcal{L}')$ and $A^+ \in \Sigma^B_0(\mathcal{L}')$. It follows from lemma 4.24 that the RHS of (4.111) is equivalent in $\mathbf{T}'$ to a $\Sigma^B_1(\mathcal{L}')$-formula. Note that in the RHS of (4.111) we have reduced the maximum nesting depth of F. Therefore we can apply the induction hypothesis to all atomic subformulas and are done.

For number functions the proof is analogously.

(d): Follows from *(c)* because a $\Sigma^B_1(\mathcal{L}')$-formula has the form $\exists \vec{X} \leq tB(\vec{X})$, where $B(\vec{X})$ is $\Sigma^B_0(\mathcal{L}')$. Then we can apply *(c)* to $B(\vec{X})$ and obtain an equivalent $\Sigma^B_1(\mathcal{L})$-formula $\exists \vec{Y} \leq sC(\vec{Y})$. It is immediate that $\mathbf{T}'$ proves the $\Sigma^B_1(\mathcal{L})$-formula

$$\exists \vec{X} \leq tB(\vec{X}) \leftrightarrow \exists \vec{X} \leq t \exists \vec{Y} \leq sC(\vec{Y}).$$

(e): Follows from the proof of lemma 4.69 (case $i = 0$) and (d). □

Corollary 4.77. *Let $\mathbf{T}_0$ be a polynomially bounded theory and $\mathcal{L} \supseteq \mathcal{L}^2_{\mathcal{A}} \cup \{\mathrm{row}\}$ its language. Assume that $\mathbf{T}_0 \models \Sigma^B_0(\mathcal{L})$-**REPL**. Let $\mathbf{T}_0 \subset \mathbf{T}_1 \subset \mathbf{T}_2 \subset \ldots$ be a sequence of extensions of $\mathbf{T}_0$ where $\mathcal{L}_i$ is the language of $\mathbf{T}_i$ and $\mathbf{T}_{i+1}$ is obtained from $\mathbf{T}_i$ by adding to $\mathbf{T}_i$ the defining axiom of a function that is $\Sigma^1_1(\mathcal{L}_i)$-definable in $\mathbf{T}_i$. Let*

$$\mathbf{T} = \bigcup_{i \geq 0} \mathbf{T}_i$$

*and let $\mathcal{L}$ be the resulting language of $\mathbf{T}$. Then $\mathbf{T}$ is polynomially bounded and a conservative extension of $\mathbf{T}_0$, $\mathbf{T} \models \Sigma^B_1(\mathcal{L})$-**REPL** and each function in $\mathcal{L}$ is $\Sigma^1_1(\mathcal{L}_0)$-definable in $\mathbf{T}_0$. Furthermore, for every $\Sigma^B_1(\mathcal{L})$-formula A^+ there is a $\Sigma^B_1(\mathcal{L}_0)$-formula A s.t. $\mathbf{T} \models A^+ \leftrightarrow A$.*

Proof. We first show by induction on i that

(a) $\mathbf{T}_i$ is polynomially bounded,

(b) $\mathbf{T}_i$ is a conservative extension of $\mathbf{T}_0$,

(c) $\mathbf{T}_i \models \Sigma^B_1(\mathcal{L}_i)$-**REPL**,

(d) for every $\Sigma^B_1(\mathcal{L}_i)$-formula A^+ there is a $\Sigma^B_1(\mathcal{L}_0)$-formula A s.t. $\mathbf{T}_i \models A^+ \leftrightarrow A$.

The base case $i = 0$ holds by assumption. The induction step follows immediately from theorem 4.76. By theorem 4.76 it is also clear that $\mathbf{T}$ is polynomially bounded, $\mathbf{T} \models \Sigma_1^B(\mathcal{L})$-**REPL** and that $\mathbf{T}$ proves the equivalence of every $\Sigma_1^B(\mathcal{L})$-formula with some $\Sigma_1^B(\mathcal{L}_0)$-formula. It is a consequence of the compactness theorem 3.27 that $\mathbf{T}$ is a conservative extension of $\mathbf{T}_0$. □

Note that by lemma 4.69, $\mathbf{V^1} \models \Sigma_1^B$-**REPL**. We will make use of corollary 4.77 when we prove the witnessing theorem in section 4.9 by setting $\mathbf{T}_0 = \mathbf{V^1}$ and $\mathbf{T}_1 = \mathbf{V^1}(\mathrm{row})$.

4.7 The Lower Bound of $\mathbf{V^1}$

In this section we show that each polytime function is provably total, i.e. Σ^1_1-definable, in $\mathbf{V^1}$. This result can be proved using Turing machine computations encoded as formulas. However, we will use Cobham's characterisation of FP instead (theorem 4.39). We will also use Parikh's theorem 4.49.

The following lemma will serve as the base case in the inductive proof of theorem 4.79. It characterises the lower bound of $\mathbf{V^0}$.

Lemma 4.78. *Every function in* FAC^0 *is* Σ^B_0*-definable in* $\mathbf{V^0}$.

Proof. By corollaries 4.36 and 4.37, a number function is in FAC^0 iff it is polynomially bounded and its graph is Σ^B_0-definable, and a set function is in FAC^0 iff it is Σ^B_0-bit-definable instead. Because the Σ^B_0-closure (cf. definition 4.73) of Σ^B_0 is Σ^B_0, it follows from lemma 4.74 that every FAC^0 function is Σ^B_0-definable in $\mathbf{V^0}$. □

Theorem 4.79 (Lower Bound of $\mathbf{V^1}$)**.** *A (number or set) function in* FP *is* Σ^1_1*-definable in* $\mathbf{V^1}$.

Proof. We show that the set functions in FP are Σ^1_1-definable in $\mathbf{V^1}$. It then follows from lemma 4.40 that the number functions in FP are also Σ^1_1-definable in $\mathbf{V^1}$.

The proof is by induction on the number of applications of composition and bounded recursion on notation needed to define a function F from functions in FAC^0. The base case follows from lemma 4.78. By lemma 4.25, the provably total functions of $\mathbf{V^1}$ are closed under composition. Therefore it is sufficient for the induction step to show the case of bounded recursion on notation. Suppose that $G(\vec{x}, \vec{X})$ and $H(y, \vec{x}, \vec{X}, Z)$ are Σ^1_1-definable in $\mathbf{V^1}$, and $F(y, \vec{x}, \vec{X})$ is defined from G and H by bounded recursion on notation, i.e.

$$F(0, \vec{x}, \vec{X}) = G(\vec{x}, \vec{X}) \tag{4.112}$$

$$F(y+1, \vec{x}, \vec{X}) = H\big(y, \vec{x}, \vec{X}, F(y, \vec{x}, \vec{X})\big)^{<t(y,\vec{x},\vec{X})} \tag{4.113}$$

for some polynomial $t(y, \vec{x}, \vec{X})$[i]. Our goal is to find a Σ^1_1-defining axiom for the function F. Note that the RHS of (4.113) is a function composition. Therefore we can define the function $H^<$ as

$$H^<(y, \vec{x}, \vec{X}, Z) = \mathrm{chop}\big(t(y, \vec{x}, \vec{X}), H(y, \vec{x}, \vec{X}, Z)\big) \tag{4.114}$$

[i]Note that this polynomial corresponds to an $\mathcal{L}^2_{\mathcal{A}}$-number term.

The function chop (cf. page 50) is in FP by lemma 4.78 and it follows from Cobham's theorem 4.39 that $H^<$ is in FP. Therefore we replace (4.113) by

$$F(y+1,\vec{x},\vec{X}) = H^<\big(y,\vec{x},\vec{X},F(y,\vec{x},\vec{X})\big) \tag{4.115}$$

It follows from lemma 4.25 that $H^<$ is also Σ^1_1-definable in $\mathbf{V^1}$. Every function value $F(y,\vec{x},\vec{X})$ is *finitely generated* by y recursive steps. Therefore we try to encode the values $F(0,\vec{x},\vec{X}),\dots,F(y,\vec{x},\vec{X})$ as a set array (cf. definition 4.60). A first guess for the defining axiom of F is

$$\begin{aligned} Y = F(y,\vec{x},\vec{X}) \leftrightarrow \exists W\Big(& W[0] = G(\vec{x},\vec{X}) \wedge \\ & \forall z < y\big(W[z+1] = H^<(z,\vec{x},\vec{X},W[z])\big) \wedge \\ & Y = W[y]\Big) \end{aligned} \tag{4.116}$$

We should replace the equations involving G and $H^<$ in (4.116) with the defining axioms of G and $H^<$, respectively. By the induction hypothesis, G and $H^<$ have Σ^1_1-defining axioms. Therefore (cf. definition 4.19), G has a defining axiom of the form

$$Y = G(\vec{x},\vec{X}) \leftrightarrow \exists\vec{U}\,A_G(\vec{x},\vec{X},Y,\vec{U}) \tag{4.117}$$

where $A_G(\vec{x},\vec{X},Y,\vec{U}) \in \Sigma^B_0$, and $H^<$ has a defining axiom of the form

$$Y = H^<(y,\vec{x},\vec{X},Z) \leftrightarrow \exists\vec{V}\,A_{H^<}(y,\vec{x},\vec{X},Z,Y,\vec{V}) \tag{4.118}$$

where $A_{H^<}(y,\vec{x},\vec{X},Z,Y,\vec{V}) \in \Sigma^B_0$. Further we have

$$\mathbf{V^1} \models \exists!Y\exists\vec{U}\,A_G(\vec{x},\vec{X},Y,\vec{U}), \tag{4.119}$$

$$\mathbf{V^1} \models \exists!Y\exists\vec{V}\,A_{H^<}(y,\vec{x},\vec{X},Z,Y,\vec{V}). \tag{4.120}$$

and by Parikh's theorem

$$\mathbf{V^1} \models \exists!Y \le t_1(\vec{x},\vec{X})\ \exists\vec{U} \le t_1(\vec{x},\vec{X})A_G(\vec{x},\vec{X},Y,\vec{U}), \tag{4.121}$$

$$\mathbf{V^1} \models \exists!Y \le t_2(y,\vec{x},\vec{X},Z)\ \exists\vec{V} \le t_2(y,\vec{x},\vec{X},Z)A_{H^<}(y,\vec{x},\vec{X},Z,Y,\vec{V}) \tag{4.122}$$

for some $\mathcal{L}^2_{\mathcal{A}}$-terms $t_1(\dots)$ and $t_2(\dots)$. We can even simplify (4.122) using the fact that each function value of $H^<$ is, according to (4.114), bounded by the $\mathcal{L}^2_{\mathcal{A}}$-number term $t(y,\vec{x},\vec{X})$[i]:

$$\mathbf{V^1} \models \exists!Y \le t(y,\vec{x},\vec{X})\ \exists\vec{V} \le t_2(y,\vec{x},\vec{X},Z)A_{H^<}(y,\vec{x},\vec{X},Z,Y,\vec{V}) \tag{4.123}$$

[i]It is easy to check that $\mathbf{V^1}(\mathrm{chop}) \models X^{<t} \le t$.

Now we can introduce the defining axioms for G and $H^<$ in (4.116) and obtain

$$\begin{aligned} Y =& F(y,\vec{x},\vec{X}) \;\leftrightarrow\; \exists W\Big(\exists \vec{U} \le t_1(\vec{x},\vec{X}) A_G(\vec{x},\vec{X},W[0],\vec{U}) \wedge \\ & \forall z < y\Big(\exists \vec{V} \le t_2(y,\vec{x},\vec{X},W[z]) A_{H^<}(z,\vec{x},\vec{X},W[z],W[z+1],\vec{V})\Big) \wedge \\ & Y = W[y]\Big) \end{aligned} \tag{4.124}$$

The problem with (4.124) is that it is not a Σ^1_1-formula because of the quantifiers $\exists \vec{V}$ which occur after $\forall z < y$. Let k be the number of existential quantifiers in $\exists \vec{V}$ and let $\vec{V}[z]$ stand for $V_1[z],\ldots,V_k[z]$. Then it is easy to verify (using lemma 4.62) that the following formula is equivalent to (4.124) in the (extended) standard model $\underline{\mathbb{N}_2}'$ (cf. definition 4.19).

$$\begin{aligned} Y =& F(y,\vec{x},\vec{X}) \;\leftrightarrow\; \exists W \exists \vec{U} \le t_1(\vec{x},\vec{X}) \\ & \exists \vec{V} \le \Big\langle \underline{k}, t_2\big(y,\vec{x},\vec{X},t_1(\vec{x},\vec{X})\big) + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big) + \ldots \\ & + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big)\Big\rangle \\ & \Big(A_G(\vec{x},\vec{X},W[0],\vec{U}) \wedge \\ & \forall z < y\Big(A_{H^<}(z,\vec{x},\vec{X},W[z],W[z+1],\vec{V}[z])\Big) \wedge \\ & Y = W[y]\Big) \end{aligned} \tag{4.125}$$

In the above formula, $t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big) + \cdots + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big)$ denotes a $k-1$ times addition. It remains to show the existence and uniqueness in $\mathbf{V^1}$ of Y in the RHS of (4.125). We prove this in the extension $\mathbf{V^1}$(row) of $\mathbf{V^1}$, obtained by adding the Σ^B_0-bit-defining axiom of the function row to $\mathbf{V^1}$. I.e. we show

$$\begin{aligned} \mathbf{V^1}(\text{row}) \models \exists! Y \exists W \exists \vec{U} \le t_1(\vec{x},\vec{X}) \exists \vec{V} \le \Big\langle \ldots \Big\rangle \\ \Big(A_G(\vec{x},\vec{X},W[0],\vec{U}) \wedge \\ \forall z < y\Big(A_{H^<}(z,\vec{x},\vec{X},W[z],W[z+1],\vec{V}[z])\Big) \wedge \\ Y = W[y]\Big) \end{aligned} \tag{4.126}$$

It then follows from lemma 4.52 that the $\Sigma_0^B(\mathcal{L}^2_{\mathcal{A}} \cup \{\text{row}\})$-formulas $Y = W[y]$, $A_G(\dots)$ and $A_{H^<}(\dots)$ can be replaced by equivalent (in $\mathbf{V^1}(\text{row})$) $\Sigma_0^B(\mathcal{L}^2_{\mathcal{A}})$-formulas and since $\mathbf{V^1}(\text{row})$ is a conservative extension of $\mathbf{V^1}$ it follows that F is also Σ^1_1-definable in $\mathbf{V^1}$. The fact that $\mathbf{V^1}(\text{row})$ is a conservative extension of $\mathbf{V^1}$ follows from corollary 4.61 ($\text{row} \in \mathsf{FAC}^0$), lemma 4.78 and lemma 4.72. We modify the RHS of (4.125) to obtain a Σ^B_1-formula.

$$\begin{aligned}
&\exists W \leq \Big\langle y+1, t_1(\vec{x},\vec{X}) + t(y,\vec{x},\vec{X})\Big\rangle \\
&\exists \vec{U} \leq t_1(\vec{x},\vec{X}) \\
&\exists \vec{V} \leq \Big\langle \underline{k}, t_2\big(y,\vec{x},\vec{X},t_1(\vec{x},\vec{X})\big) + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big) + \dots \\
&+ t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big)\Big\rangle \\
&\Big(A_G(\vec{x},\vec{X},W[0],\vec{U}) \wedge \\
&\forall z < y\Big(A_{H^<}(z,\vec{x},\vec{X},W[z],W[z+1],\vec{V}[z])\Big)\Big)
\end{aligned} \tag{4.127}$$

Note that we dropped the last conjunct of (4.125) for the moment. Let $B(y)$ denote the formula (4.127). We want to show that $\mathbf{V^1}(\text{row})$ proves $B(y)$. Since $B(y)$ is a Σ^B_1-formula we can apply the number induction axiom (cf. theorem 4.13) on y. Let $\mathcal{M}$ be a model of $\mathbf{V^1}(\text{row})$ and let σ be an arbitrary assignment. We first have to show that $\mathcal{M}[\sigma] \models B(0)$. It suffices to show[i]

$$\mathcal{M}[\sigma] \models \exists W \leq \big\langle 1, t_1(\vec{x},\vec{X}) + t(0,\vec{x},\vec{X})\big\rangle \exists \vec{U} \leq t_1(\vec{x},\vec{X}) A_G(\vec{x},\vec{X},W[0],\vec{U}) \tag{4.128}$$

Let W'_0 be the set that satisfies the quantifier $\exists! Y$ in (4.121). By lemma 4.62 there exists a W_0 s.t.

$$\mathcal{M}[\sigma(\mathsf{W}'_0/W'_0)(\mathsf{W}_0/W_0)] \models |W_0| \leq \big\langle 1, |W'_0|\big\rangle \wedge W'_0 = W_0[0] \tag{4.129}$$

and hence, by the monotonicity of $\mathcal{L}^2_{\mathcal{A}}$-terms (lemma 4.46) and transitivity,

$$\mathcal{M}[\sigma(\mathsf{W}'_0/W'_0)(\mathsf{W}_0/W_0)] \models |W_0| \leq \big\langle 1, t_1(\vec{x},\vec{X})\big\rangle \wedge W'_0 = W_0[0].$$

Then it is obvious that this W_0 also satisfies (4.128).

For the induction step we need to show that $\mathcal{M}[\sigma] \models \forall y\big(B(y) \to B(y+1)\big)$. And since σ is arbitrary it is enough to show $\mathcal{M}[\sigma] \models B(y) \to B(y+1)$.

[i] cf. lemma 4.7 (2).

Assume $\mathcal{M}[\sigma] \models B(y)$, hence we have

$$\begin{aligned}\mathcal{M}[\sigma(\mathsf{W}/W)] \models |W| \leq \Big\langle & y+1, t_1(\vec{x},\vec{X}) + t(y,\vec{x},\vec{X})\Big\rangle \\ & \exists \vec{U} \leq t_1(\vec{x},\vec{X}) \\ & \exists \vec{V} \leq \Big\langle \underline{k}, t_2\big(y,\vec{x},\vec{X},t_1(\vec{x},\vec{X})\big) + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big) + \ldots \\ & + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big)\Big\rangle \\ & \Big(A_G(\vec{x},\vec{X},W[0],\vec{U}) \wedge \\ & \forall z < y\Big(A_{H^<}(z,\vec{x},\vec{X},W[z],W[z+1],\vec{V}[z])\Big)\Big)\end{aligned} \tag{4.130}$$

for some W. By (4.123), we further have

$$\begin{aligned}\mathcal{M}[\sigma(\mathsf{W}/W)(\mathsf{W}'/W')] \models |W'| & \leq t(y+1,\vec{x},\vec{X}) \\ \exists \vec{V} & \leq t_2\big(y+1,\vec{x},\vec{X},W[y]\big) \\ A_{H^<}(y+1,\vec{x},\vec{X},W[y],W',\vec{V}) & \end{aligned} \tag{4.131}$$

for some W'. It follows from (4.121) that

$$\mathcal{M}[\sigma(\mathsf{W}/W)(\mathsf{W}'/W')] \models y = 0 \rightarrow W[y] \leq t_1(\vec{x},\vec{X})$$

and it follows from (4.123) that

$$\mathcal{M}[\sigma(\mathsf{W}/W)(\mathsf{W}'/W')] \models y \neq 0 \rightarrow W[y] \leq t(y,\vec{x},\vec{X}).$$

In any case, by the monotonicity of $\mathcal{L}^2_{\mathcal{A}}$-terms (lemma 4.46) and transitivity, we have

$$\mathcal{M}[\sigma(\mathsf{W}/W)(\mathsf{W}'/W')] \models W[y] \leq t_1(\vec{x},\vec{X}) + t(y,\vec{x},\vec{X}). \tag{4.132}$$

and thus we can adjust (4.131) and obtain

$$\begin{aligned}\mathcal{M}[\sigma(\mathsf{W}/W)(\mathsf{W}'/W')] \models |W'| & \leq t(y+1,\vec{x},\vec{X}) \\ \exists \vec{V} & \leq t_2\big(y+1,\vec{x},\vec{X},t_1(\vec{x},\vec{X}) + t(y,\vec{x},\vec{X})\big) \\ A_{H^<}(y+1,\vec{x},\vec{X},W[y],W',\vec{V}) & \end{aligned} \tag{4.133}$$

Informally, we now have a set (array) W that encodes the function values $F(i,\vec{x},\vec{X})$, for all $i \leq y$, and a set W' that contains the value $F(y+1,\vec{x},\vec{X})$.

Our objective is to concatenate them to obtain a set W'' that satisfies the quantifier $\exists W$ in $B(y+1)$. Note that we cannot apply lemma 4.62 because there is not necessarily a numeral $\underline{n}$ such that $\mathcal{M}[\sigma] \models y+1 = \underline{n}$. The reason is that $\mathbf{V^1}$ has models other than the standard model $\underline{\mathbb{N}_2}$ [i]. The concatenation is done by multiple comprehension (lemma 4.59):

$$\begin{aligned}
&\mathcal{M}[\sigma(\mathsf{W}/W)(\mathsf{W}'/W')] \models \\
&\exists W'' \leq \Big\langle y+1+1, t_1(\vec{x},\vec{X}) + t(y+1,\vec{x},\vec{X}) \Big\rangle \\
&\forall z_1 < y+1+1 \forall z_2 < t_1(\vec{x},\vec{X}) + t(y+1,\vec{x},\vec{X}) \\
&\Big(W''(\langle z_1, z_2\rangle) \leftrightarrow \big(z_1 < y+1 \wedge W(\langle z_1, z_2\rangle)\big) \vee \big(z_1 = y+1 \wedge W'(z_2)\big)\Big)
\end{aligned} \tag{4.134}$$

Hence $\mathbf{V^1}(\text{row}) \models B(y)$. It follows that the "existence part" of (4.126) holds since, informally, $W[y]$ satisfies the quantifier $\exists Y$.

Now to uniqueness. Recall definition 4.5 of $\exists!$. It is enough to show

$$\begin{aligned}
\mathbf{V^1}(\text{row}) \models\ & \\
& \exists \vec{U} \leq t_1(\vec{x},\vec{X}) \\
& \exists \vec{V} \leq \Big\langle \underline{k}, t_2\big(y,\vec{x},\vec{X},t_1(\vec{x},\vec{X})\big) + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big) + \ldots \\
& + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big)\Big\rangle \\
& \Big(A_G(\vec{x},\vec{X},W_1[0],\vec{U}) \wedge \\
& \forall z < y \Big(A_{H^<}(z,\vec{x},\vec{X},W_1[z],W_1[z+1],\vec{V}[z])\Big) \wedge Y = W_1[y]\Big) \wedge \\
& \exists \vec{U} \leq t_1(\vec{x},\vec{X}) \\
& \exists \vec{V} \leq \Big\langle \underline{k}, t_2\big(y,\vec{x},\vec{X},t_1(\vec{x},\vec{X})\big) + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big) + \ldots \\
& + t_2\big(y,\vec{x},\vec{X},t(y,\vec{x},\vec{X})\big)\Big\rangle \\
& \Big(A_G(\vec{x},\vec{X},W_2[0],\vec{U}) \wedge \\
& \forall z < y \Big(A_{H^<}(z,\vec{x},\vec{X},W_2[z],W_2[z+1],\vec{V}[z])\Big) \wedge Y = W_2[y]\Big) \\
& \rightarrow \big(i \leq y \rightarrow W_1[i] = W_2[i]\big)
\end{aligned} \tag{4.135}$$

[i] Note that **PA** has non-standard models (cf. for example [9]) and since every model of (two-sorted) **PA** is also a model of $\mathbf{V^1}$, $\mathbf{V^1}$ also has non-standard models.

Let $C(i)$ denote the RHS of (4.135). Since $C(i)$ is a Σ_1^B-formula we can again apply the number induction axiom. Let $\mathcal{M} \models \mathbf{V^1}(\mathrm{row})$ and let σ be an arbitrary assignment. $\mathcal{M}[\sigma] \models C(0)$ follows immediately from the uniqueness of Y in (4.119). For the induction step, assume $\mathcal{M}[\sigma] \models C(i)$. We show $\mathcal{M}[\sigma] \models C(i+1)$. If $\mathcal{M}[\sigma] \not\models i+1 \leq y$, then $\mathcal{M}[\sigma] \models C(i+1)$ holds trivially. So assume $\mathcal{M}[\sigma] \models i+1 \leq y$. Then, by lemma 4.15 (16) [i], $\mathcal{M}[\sigma] \models i < y$. Then $\mathcal{M}[\sigma] \models W_1[i+1] = W_2[i+1]$ follows from the induction hypothesis (cf. (4.135)) and the uniqueness of Y in (4.120). $\square$

[i] $x+1 \leq y \to x < y$

4.8 An Alternative Axiomatisation of $\mathbf{V^1}$

We show that $\mathbf{V^1}$ can be axiomatised by Σ^B_0-**COMP** and Σ^B_1-**IND** instead of Σ^B_1-**COMP**. The idea is then to replace the axiom scheme Σ^B_1-**IND** with a new LK^2 induction rule. The reason is that the formulas of Σ^B_1-**COMP** are in general not equivalent to Σ^B_1-formulas and, in the proof of the witnessing theorem for $\mathbf{V^1}$, we want to avoid that non-Σ^1_1-formulas occur in LK^2 proofs.

Definition 4.80 ($\widetilde{\mathbf{V^1}}$)**.** $\widetilde{\mathbf{V^1}}$ *is the theory axiomatised by the axioms* **B1**-**B12**, **L1**, **L2**, **SE**, Σ^B_0-**COMP** *and* Σ^B_1-**IND**.

Note that $\widetilde{\mathbf{V^1}}$ is also an extension of $\mathbf{V^0}$. Our objective is to prove that $\mathbf{V^1} = \widetilde{\mathbf{V^1}}$ (theorem 4.85 below). To this end, we will need to prove that $\widetilde{\mathbf{V^1}}$ contains the Σ^B_1-**COMP** axioms. This proof is based on the fact that formulas of the form $\forall x \leq t_1 \exists X \leq t_2 A(x, X)$, where $A(x, X)$ is Σ^B_0, are provably equivalent in $\mathbf{V^0}$ to a Σ^B_1-formula. This was proved in section 4.6.2. For the proof of theorem 4.85 below we need a number function numones(y, X) that returns the number of elements of a set X that are strictly smaller than y. numones has the following Σ^B_1-defining axiom

$$\begin{aligned} \text{numones}(y, X) = z \;\leftrightarrow\; & z \leq y \wedge \\ \exists Z \leq 1 + \langle y, y\rangle \Big(& Z^{[0]} = 0 \wedge Z^{[y]} = z \wedge \\ \forall u < y \Big(& \big(X(u) \rightarrow Z^{[u+1]} = Z^{[u]} + 1\big) \wedge \\ & \big(\neg X(u) \rightarrow Z^{[u+1]} = Z^{[u]}\big) \Big) \Big) \end{aligned} \tag{4.136}$$

The proofs of the next three lemmas are left as an exercise.

Lemma 4.81. numones *is* Σ^B_1*-definable in* $\widetilde{\mathbf{V^1}}$*, i.e.* (4.136) *is a defining axiom of* numones *in* $\widetilde{\mathbf{V^1}}$.

Lemma 4.82.

$$\begin{aligned} \widetilde{\mathbf{V^1}}(\text{numones}) \models \exists x < y \Big(& X(x) \wedge \neg Y(x) \wedge \forall u < y \Big(u \neq x \rightarrow \big(X(u) \leftrightarrow Y(u)\big) \Big) \Big) \\ & \rightarrow \text{numones}(y, X) = \text{numones}(y, Y) + 1 \end{aligned}$$

Intuitively, the following axiom scheme Φ-**MAX** states that for every formula $A(x) \in \Phi$, if $A(0)$, then there exists a maximum number $x \leq y$ that satisfies $A(x)$.

Definition 4.83 (Φ-**MAX**)**.** Φ-**MAX** *is the set of formulas of the form*

$$A(0) \to \exists x \leq y\Big(A(x) \wedge \neg\exists z \leq y\big(x < z \wedge A(z)\big)\Big), \tag{4.137}$$

for all $A(x) \in \Phi$.

Lemma 4.84. $\widetilde{\mathbf{V^1}} \models \Sigma^B_1$-**MAX**.

Theorem 4.85. $\mathbf{V^1} = \widetilde{\mathbf{V^1}}$

Proof. The direction $\widetilde{\mathbf{V^1}} \subseteq \mathbf{V^1}$ follows from theorem 4.13.

For the other direction, it is sufficient to show that $\widetilde{\mathbf{V^1}}$ proves the Σ^B_1-**COMP** axiom scheme:

$$\exists X \leq y \forall z < y\big(X(z) \leftrightarrow A(z)\big) \tag{4.138}$$

where $A(z) \in \Sigma^B_1$. Unfortunately, (4.138) is not a Σ^B_1-formula and is not (in general) equivalent to one[i]. Therefore we cannot apply Σ^B_1-**IND** directly on (4.138). Consider the formula

$$\forall z < y\big(Y(z) \to A(z)\big) \tag{4.139}$$

(4.139) is equivalent to a Σ^B_1-formula in $\widetilde{\mathbf{V^1}}$ by lemma 4.70 [ii]. Let $B(y, Y)$ denote this Σ^B_1-formula. Now consider the formula

$$C(w, y) \equiv \exists Y \leq y\Big(B(y, Y) \wedge w = \text{numones}(y, Y)\Big) \tag{4.140}$$

Now we need the Σ^B_1-**MAX** scheme. By lemma 4.81, numones is Σ^1_1-definable in $\widetilde{\mathbf{V^1}}$. It then follows from corollary 4.77 that $\widetilde{\mathbf{V^1}}$(numones) is a conservative extension of $\widetilde{\mathbf{V^1}}$ and that every Σ^B_1(numones)-formula is provably equivalent in $\widetilde{\mathbf{V^1}}$(numones) to some $\Sigma^B_1(\mathcal{L}^2_{\mathcal{A}})$-formula. Since $\widetilde{\mathbf{V^1}} \models \Sigma^B_1$-**MAX** (lemma 4.84), it follows that $\widetilde{\mathbf{V^1}}$(numones) $\models \Sigma^B_1$(numones)-**MAX**.

Now we apply Σ^B_1(numones)-**MAX** to $C(w, y)$. Let $\mathcal{M} \models \widetilde{\mathbf{V^1}}$(numones) and let σ be arbitrary. Then we have

$$\mathcal{M}[\sigma] \models C(0, y) \to \exists x \leq y\Big(C(x, y) \wedge \neg\exists z \leq y(x < z \wedge C(z, y))\Big) \tag{4.141}$$

We first show that $\mathcal{M} \models C(0, y)$, i.e.

$$\mathcal{M}[\sigma] \models \exists Y \leq y\Big(B(y, Y) \wedge 0 = \text{numones}(y, Y)\Big) \tag{4.142}$$

[i] With regard to lemma 4.70, the problem is the negation of $\neg X(z)$ in the direction $\leftarrow$.

[ii] Note that $\neg Y(z) \vee \exists \vec{X} \leq \vec{t} A(z) \leftrightarrow \exists \vec{X} \leq t\big(\neg Y(z) \vee A(z)\big)$ is valid.

Lemma 4.44 shows that there is a unique empty set and it is clear that the empty set satisfies the quantifier $\exists Y$ in (4.142). Hence we have $\mathcal{M} \models C(0, y)$ and by (4.141)

$$\mathcal{M}[\sigma] \models \exists x \leq y \Big(C(x, y) \wedge \neg \exists z \leq y (x < z \wedge C(z, y)) \Big) \tag{4.143}$$

i.e. there exists a "maximum" $\mathsf{w_0}$ s.t.

$$\mathcal{M}[\sigma(\mathsf{w_0}/w_0)] \models w_0 \leq y \wedge \Big(C(w_0, y) \wedge \neg \exists z \leq y (w_0 < z \wedge C(z, y)) \Big) \tag{4.144}$$

Hence $\mathcal{M}[\sigma(\mathsf{w_0}/w_0)] \models C(w_0, y)$, i.e.

$$\mathcal{M}[\sigma(\mathsf{w_0}/w_0)(\mathsf{Y}/Y)] \models Y \leq y \wedge \big(B(y, Y) \wedge w_0 = \text{numones}(y, Y) \big) \tag{4.145}$$

for some Y. It follows from lemma 4.82 that Y satisfies the quantifier $\exists X$ in (4.138).

□

4.9 The Upper Bound of $\mathbf{V^1}$

Now we show that $\mathbf{V^1}$ is not too strong, i.e. that all provably total functions in $\mathbf{V^1}$ are in FP. This will follow from the so-called witnessing theorem for $\mathbf{V^1}$. For its proof we use the theory $\widetilde{\mathbf{V^1}}$ for which we showed the identity to $\mathbf{V^1}$ in section 4.8.

4.9.1 The System LK^2-$\widetilde{\mathbf{V^1}}$

Instead of using the axioms of $\widetilde{\mathbf{V^1}}$ as non-logical axioms in the system LK^2, we replace the Σ^B_1-**IND** axiom scheme by a new inference rule.

We assume that in LK^2, terms do not contain bound variables $x, y, z, \ldots$ and $X, Y, Z, \ldots$, bound variables do not occur free in formulas and free variables do not occur bound (cf. variable convention on page 21).

Definition 4.86 (Φ-Ind rule)**.** *The Φ-Ind rule consists of inferences of the form*

$$(\Phi\text{-}Ind)\ \frac{\Gamma, A(b) \vdash A(b+1), \Delta}{\Gamma, A(0) \vdash A(t), \Delta}$$

where Φ is a set of formulas and the eigenvariable b does not occur in the bottom sequent.

Definition 4.87 (LK^2-$\widetilde{\mathbf{V^1}}$)**.** *The rules of the sequent system LK^2-$\widetilde{\mathbf{V^1}}$ consist of the rules of LK^2 (section 3.3.1) and the Σ^B_1-Ind rule. The non-logical axioms of LK^2-$\widetilde{\mathbf{V^1}}$ consist of the sets* **B1**-**B12**, **L1**, **L2**, **SE**, Σ^B_0-**COMP** *and $\varepsilon_{\mathcal{L}}$, all closed under substitution of terms for free variables*[i]*.*

By an LK^2-$\Psi + \Phi$-Ind proof we mean an LK^2-Ψ proof where additionally the Φ-Ind rule is admitted.

Theorem 4.88 (Soundness of LK^2-$\Psi + \Phi$-Ind)**.** *Let Φ, Ψ be sets of formulas. If a sequent $\Gamma \vdash \Delta$ has an LK^2-$\Psi + \Phi$-Ind proof, then $(\Gamma \vdash \Delta)^\star$ is a theorem of the theory axiomatised by $\Psi \cup \Phi$-**IND**.*

Proof. From our definition of a theory (definition 2.9) it follows that it suffices to check that $\Psi \cup \Phi$-**IND** $\models (\Gamma \vdash \Delta)^\star$. This is proved by induction on the number of sequents in a proof. The derivational soundness theorem 3.10 provides already all the cases except, of course, the case of the Φ-Ind rule. Thus it remains to show that if the top sequent of this rule is a logical

[i] Substitution of terms for free variables is necessary in order to prove anchored completeness.

consequence of $\Psi \cup \Phi$-**IND**, then so is the bottom sequent[i]. Let $\mathcal{M}$ be a model of $\Psi \cup \Phi$-**IND** and assume that $\mathcal{M}$ satisfies the top sequent, i.e.

$$\mathcal{M} \models \bigwedge_{\Gamma} \wedge A(b) \to A(b+1) \vee \bigvee_{\Delta} \tag{4.146}$$

Note that

$$\bigwedge_{\Gamma} \to \big(A(b) \to A(b+1)\big) \vee \bigvee_{\Delta} \tag{4.147}$$

is provably equivalent[ii] to the formula in (4.146). And because the eigenvariable b does not occur in $\Gamma \cup \Delta$, we have

$$\mathcal{M} \models \bigwedge_{\Gamma} \to \forall b\big(A(b) \to A(b+1)\big) \vee \bigvee_{\Delta} \tag{4.148}$$

We have to show that $\mathcal{M}$ satisfies the bottom sequent, i.e.

$$\mathcal{M}[\sigma] \models \bigwedge_{\Gamma} \wedge A(0) \to A(t) \vee \bigvee_{\Delta} \tag{4.149}$$

for an arbitrary assignment σ. So assume

$$\mathcal{M}[\sigma] \models \bigwedge_{\Gamma} \wedge A(0) \tag{4.150}$$

By (4.148), we have $\mathcal{M}[\sigma] \models \forall b\big(A(b) \to A(b+1)\big) \vee \bigvee_{\Delta}$. If $\mathcal{M}[\sigma] \models \bigvee_{\Delta}$, then (4.149) is obviously satisfied. If $\mathcal{M}[\sigma] \models \forall b\big(A(b) \to A(b+1)\big)$, then, by $\mathcal{M} \models \Phi$-**IND**, we have $\mathcal{M}[\sigma] \models \forall x A(x)$ and hence $\mathcal{M}[\sigma] \models A(t)$ and (4.149) follows. □

Corollary 4.89 (Soundness of LK^2-$\widetilde{\mathbf{V^1}}$)**.** *If a sequent $\Gamma \vdash \Delta$ has an LK^2-$\widetilde{\mathbf{V^1}}$ proof, then $(\Gamma \vdash \Delta)^\star$ is a theorem of $\widetilde{\mathbf{V^1}}$.*

We generalise definition 3.15 of an *anchored proof.*

Definition 4.90 (Anchored LK^2-$\Psi + \Phi$-Ind Proof)**.** *An LK^2-$\Psi + \Phi$-Ind proof is called anchored provided that every cut formula is either in Ψ or is one of the formulas $A(0), A(t)$ in an instance of the Φ-Ind rule.*

[i] Note that in this case, the bottom sequent is *not* a logical consequence of the top sequent, as it is the case with all other rules.

[ii] i.e. the corresponding formula of the form $A \leftrightarrow B$ is valid.

Theorem 4.91 (Anchored Completeness of LK^2-$\Psi_{\varepsilon_\mathcal{L}}$ + Φ-Ind)**.** *Let Φ, Ψ be sets of $\mathcal{L}$-formulas and let Φ', Ψ' be their closures under substitution of terms for free variables, respectively. If $(\Gamma \vdash \Delta)^\star$ is a logical consequence of $\Psi_{\varepsilon_\mathcal{L}} \cup \Phi$-**IND**, then $\Gamma \vdash \Delta$ has an anchored LK^2-$\Psi'_{\varepsilon_\mathcal{L}}$ + Φ'-Ind proof.*

Proof. As in the proof of lemma 4.91, we slightly modify the proof of the completeness lemma 3.12. Here, a sequent is called *active* if it is a leaf of the proof tree and is not directly derivable from a logical axiom of LK^2 or from a non-logical axiom of $\Psi_{\varepsilon_\mathcal{L}}$ using only weakening and exchange rules. Recall that in our enumeration of (A_i, t_j, T_k), the terms t_j and T_k do not contain bound variables (variable convention on page 21). We begin with an end sequent $\Gamma \vdash \Delta$ and work upwards by subsequently modifying the proof Π.

Loop: Let (A_i, t_j, T_k) be the next tuple in the enumeration.

1. *Step:* If $A_i \in \Psi'_{\varepsilon_\mathcal{L}}$, then every active sequent $\Pi \vdash \Omega$ is replaced with the derivation

$$\dfrac{\dfrac{}{\Pi \vdash \Omega, A_i}\,(\text{weakening}) \qquad \Pi, A_i \vdash \Omega}{\Pi \vdash \Omega}\,(\text{cut})$$

2. *Step:* If $A_i \in \Phi'$ and A_i contains at least one free occurrence of some free variable c, then let b be a fresh free number variable not yet used in Π, and let $A_i(b)$ be the result of substituting b for c in $A_i(c)$. Then every active sequent $\Pi \vdash \Omega$ is replaced with the derivation

$$\dfrac{\Pi \vdash \Omega, A_i(0) \qquad \dfrac{\dfrac{\Pi, A_i(b) \vdash A_i(b+1), \Omega}{\Pi, A_i(0) \vdash A_i(t_j), \Omega} \qquad \dfrac{A_i(t_j), \Pi \vdash \Omega}{A_i(t_j), \Pi, A_i(0) \vdash \Omega}}{\Pi, A_i(0) \vdash \Omega}}{\Pi \vdash \Omega}$$

   where the top-left inference is by the Φ-Ind rule and the other inferences are by the weakening, exchange and the cut rule (note that the cut formulas are in Φ).

   (For the other steps, proceed as in steps 2 to 4 of the completeness lemma 3.12.)

End of the loop.

If the above procedure terminates, then Π is an anchored LK^2-$\Psi'_{\varepsilon_\mathcal{L}}$ + Φ'-Ind proof of $\Gamma \vdash \Delta$. Again, we have to show that if it does not terminate, then $\Psi_{\varepsilon_\mathcal{L}} \cup \Phi$-**IND** $\not\models (\Gamma \vdash \Delta)^\bigstar$.

So assume that the above procedure runs forever and let Π denote the resulting infinite tree. Now we modify Π by "fading out" all the intermediate rule applications introduced at step 2. I.e. we replace each subtree of Π of the form

$$\dfrac{\Pi \vdash \Omega, A_i(0) \qquad \dfrac{\dfrac{\Pi, A_i(b) \vdash A_i(b+1), \Omega}{\Pi, A_i(0) \vdash A_i(t_j), \Omega} \qquad \dfrac{A_i(t_j), \Pi \vdash \Omega}{A_i(t_j), \Pi, A_i(0) \vdash \Omega}}{\Pi, A_i(0) \vdash \Omega}}{\Pi \vdash \Omega}$$

with

$$\dfrac{\Pi \vdash \Omega, A_i(0) \qquad \Pi, A_i(b) \vdash A_i(b+1), \Omega \qquad A_i(t_j), \Pi \vdash \Omega}{\Pi \vdash \Omega}$$

Note that the resulting tree (we still call it Π) is no longer an LK^2-$\Psi'_{\varepsilon_\mathcal{L}}$+$\Phi'$-Ind proof. But that's not a problem since Π only serves the purpose of constructing a structure.

Π still has an infinite path π (by König's Lemma 3.11) and we use π to define a structure $\mathcal{M}$ and an assignment σ s.t. $\mathcal{M} \models \Psi_{\varepsilon_\mathcal{L}} \cup \Phi$-**IND** and $\mathcal{M}[\sigma] \not\models (\Gamma \vdash \Delta)^\bigstar$. Let σ and $\mathcal{M}$ be as in the completeness lemma 3.12. Hence, for any predicate symbol P, $(\vec{t}, \vec{T}) \in P^\mathcal{M}$ holds iff the formula $P(\vec{t}, \vec{T})$ occurs in the antecedent of some sequent in π. We claim again that every formula A occurring in an antecedent along π is true in $\mathcal{M}[\sigma]$ and that every formula A occurring in a succedent along π is false in $\mathcal{M}[\sigma]$. The argument is again the same as in the proof of the completeness lemma 3.12. Therefore $\mathcal{M}[\sigma] \not\models (\Gamma \vdash \Delta)^\bigstar$.

It remains to show that $\mathcal{M} \models \Psi_{\varepsilon_\mathcal{L}} \cup \Phi$-**IND**. Every formula $A_i \in \Psi'_{\varepsilon_\mathcal{L}}$ occurs in an antecedent of π (cf. step 1). Therefore, by the above claim, we have $\mathcal{M}[\sigma] \models A_i$ and hence $\mathcal{M}[\sigma] \models \Psi'_{\varepsilon_\mathcal{L}}$. But since $\Psi'_{\varepsilon_\mathcal{L}}$ is closed under substitution of terms for free variables, we have $\mathcal{M} \models \Psi'_{\varepsilon_\mathcal{L}}$ and hence $\mathcal{M} \models \Psi_{\varepsilon_\mathcal{L}}$.

To show that $\mathcal{M} \models \Phi$-**IND** we consider an arbitrary formula $A_i \in \Phi'$. We assume that Φ' is nonempty, otherwise what we claim is trivially true. With respect to step 2 (recall that we have modified Π and hence π!), we distinguish three cases, whereas the third case excludes the first and the second.

If π contains a sequent of the form $\Pi \vdash \Omega, A_i(0)$, then, by the above claim, $\mathcal{M}[\sigma] \not\models A_i(0)$. If π contains a sequent of the form $\Pi, A_i(b) \vdash A_i(b+1), \Omega$, then $\mathcal{M}[\sigma] \models A_i(b)$ and $\mathcal{M}[\sigma] \not\models A_i(b+1)$ and hence $\mathcal{M}[\sigma] \not\models A_i(b) \rightarrow$

$A_i(b+1)$. As a consequence, $\mathcal{M}[\sigma] \not\models \forall x(A_i(x) \to A_i(x+1))$. In both cases we have $\mathcal{M}[\sigma] \models A_i$-**IND**.

If π does neither contain $\Pi \vdash \Omega, A_i(0)$ nor $\Pi, A_i(b) \vdash A_i(b+1), \Omega$, then, for every number term t_j, π contains $A_i(t_j), \Pi \vdash \Omega$ (note that every tuple (A_i, t_j, T_k) occurs infinitely often in our enumeration). It follows that, for every t_j, the formula $A_i(t_j)$ occurs in an antecedent of π. Therefore, by the above claim, $\mathcal{M}[\sigma] \models A_i(t_j)$ for every number term t_j. Since the "number" universe of $\mathcal{M}$ consists precisely of all number terms and σ maps terms to themselves we have $\mathcal{M}[\sigma] \models \forall x A_i(x)$ and hence $\mathcal{M}[\sigma] \models A_i$-**IND**.

Since the choice of $A_i \in \Phi'$ was arbitrary, it follows that $\mathcal{M}[\sigma] \models \Phi'$-**IND** and (since Φ'-**IND** is closed under substitution of terms for free variables) $\mathcal{M} \models \Phi'$-**IND**. Therefore $\mathcal{M} \models \Phi$-**IND** and we are done. □

The following corollary is immediate from the above and theorem 4.85.

Corollary 4.92. *Every theorem of* $\mathbf{V^1}$ *has an anchored* LK^2-$\widetilde{\mathbf{V^1}}$ *proof.*

The advantage of LK^2-$\Psi'_{\varepsilon_\mathcal{L}} + \Phi'$-Ind proofs is that they contain only cut formulas of $\Phi' \cup \Psi_{\varepsilon_\mathcal{L}}$ and not of Φ'-**IND**. Therefore, for LK^2-$\widetilde{\mathbf{V^1}}$, the cut formulas are in Σ^B_1. We now restate the subformula property (lemma 3.26) for the case of LK^2-$\Psi + \Phi$-Ind proofs.

Lemma 4.93 (Subformula Property of LK^2-$\Psi + \Phi$-Ind)**.** *Let* Ψ *and* Φ *be sets of formulas, closed under substitution of terms for free variables, and let* Π *be an anchored* LK^2-$\Psi + \Phi$-*Ind proof of a sequent* $\Gamma \vdash \Delta$*. Then every formula in every sequent of* Π *is a subformula of a formula in* $\Gamma \vdash \Delta$ *or of a formula in* $\Psi \cup \Phi$*.*

Proof. The proof is the same as for lemma 3.26. The only new case is the Φ-Ind rule, where the formulas $A(b), A(b+1), A(0), A(t)$ are all in Φ. □

Obviously, the subformula property holds also for LK^2-$\widetilde{\mathbf{V^1}}$, as a special case of LK^2-$\Psi + \Phi$-Ind. Recall (page 67) that $\mathbf{T}(\mathcal{F})$ is the theory $\mathbf{T}$ together with the defining axioms or bit-defining axioms for functions $\mathcal{F}$, that are definable in $\mathbf{T}$ or bit-definable. Note that by theorem 4.79, all functions in FP are definable in $\mathbf{V^1}$. The witnessing theorem 4.95 follows from the following special case.

Lemma 4.94 (Witnessing Lemma)**.** *Let* $A(\vec{x}, \vec{X}, Y)$ *be a* Σ^B_0*-formula and assume that* $\mathbf{V^1} \models \exists Y A(\vec{x}, \vec{X}, Y)$*. Then there exists a set function* $F \in \mathsf{FP}$ *such that*

$$\mathbf{V^1}(F) \models A\big(\vec{x}, \vec{X}, F(\vec{x}, \vec{X})\big).$$

Proof. We assume that $\mathbf{V^1}$ contains a Σ^1_1-formula $\exists Y A(\vec{x}, \vec{X}, Y)$. By the above anchored completeness theorem (corollary 4.92) there is an anchored LK^2-$\widetilde{\mathbf{V^1}}$ proof Π of the sequent $\vdash \exists Y A(\vec{x}, \vec{X}, Y)$. We assume that Π is in free variable normal form (cf. page 52) by adding the axiom $|\varnothing| = 0$ to $\mathbf{V^1}$ (and $\widetilde{\mathbf{V^1}}$, respectively). Additionally, free variables can be eliminated by the Σ^B_1-Ind rule. Recall that in LK^2-$\widetilde{\mathbf{V^1}}$, the only nonlogical axioms are Σ^B_0-**COMP** and the axioms **B1** to **SE** on page 35, i.e. they are either Σ^B_1- or Σ^B_0-formulas. From the subformula property (lemma 3.26), we can conclude that all formulas in Π are either Σ^1_1- or Σ^B_0-formulas. Note that $\Sigma^B_1 \subseteq \Sigma^1_1$ and that the Σ^1_1-formulas have at most one existential set quantifier in front. Therefore, every sequent in Π is of the form

$$\exists X_1 A_1(X_1), \ldots, \exists X_m A_m(X_m), \Gamma \vdash \Delta, \exists Y_1 B_1(Y_1), \ldots, \exists Y_n B_n(Y_n) \quad (4.151)$$

where $m, n \geq 0$ and all A_i and B_j as well as Γ, Δ are Σ^B_0-formulas. As in the proof of Parikh's theorem 4.49 we treat the sequents as multisets and ignore applications of the exchange rules. We proceed by induction on the depth in the prooftree Π of a sequent S of the form (4.151) and show that there exists a finite set of polytime functions $\mathcal{F} = \{\ldots, F_1, \ldots, F_n, \ldots\}$ s.t.

$$\begin{aligned} \mathbf{V^1}(\mathcal{F}) \models \Big(& A_1(N_1), \ldots, A_m(N_m), \Gamma \vdash \\ & \Delta, B_1\big(F_1(\vec{a}, \vec{M}, \vec{N})\big), \ldots, B_n\big(F_n(\vec{a}, \vec{M}, \vec{N})\big) \Big)^\star \end{aligned} \quad (4.152)$$

where $\vec{a}, \vec{M}$ are exactly the free variables of the sequent (4.151) and $\vec{N}$ are new distinct free variables replacing the bound variables $X_1, \ldots, X_m$ (note that $\vec{a}$ and $\vec{M}$ may be different for different sequents.). In the following, we write $\hat{\mathsf{S}}$ for the corresponding sequent of the form (4.152).

From this we can conclude that for the end sequent $\vdash \exists Y A(\vec{x}, \vec{X}.Y)$ of Π there is a set $\mathcal{F}$ of polytime set functions and an $F \in \mathcal{F}$ s.t. $\mathbf{V^1}(\mathcal{F}) \models A(\vec{a}, \vec{M}, F(\vec{a}, \vec{M}))$. By theorem 4.79 (lower bound of $\mathbf{V^1}$), all functions in $\mathcal{F}$ are Σ^1_1-definable in $\mathbf{V^1}$ and it follows from corollary 4.77 that $\mathbf{V^1}(\mathcal{F})$ is a conservative extension of $\mathbf{V^1}(F)$. Therefore $\mathbf{V^1}(F) \models A(\vec{a}, \vec{M}, F(\vec{a}, \vec{M}))$. Because $\mathbf{V^1}(F)$ is also a conservative extension of $\mathbf{V^1}(F) \setminus \{|\varnothing| = 0\}$ (lemma 4.43) and since we will never use the constant $\varnothing$ for the defining axioms of F, we have $\mathbf{V^1}(F) \setminus \{|\varnothing| = 0\} \models A(\vec{a}, \vec{M}, F(\vec{a}, \vec{M}))$ and are done.

In the following inductive proof we know that no free variable (and hence no parameter of previously introduced witnessing functions) is eliminated by any rule except $\exists$-left, set $\exists$-left, $\forall$-right[i] and Σ^B_1-Ind. This follows from the

[i]Note that the set $\forall$-rules do not occur because all formulas in Π are Σ^1_1 or Σ^B_0.

fact that Π is in free variable normal form. We will treat these four cases accordingly.

Base case: S is an axiom of LK^2-$\widetilde{\mathbf{V^1}}$. If S contains only Σ_0^B-formulas, then what we claim is trivially satisfied. Otherwise, if S is a logical axiom, then it has the form $\exists Y B(\vec{a}, \vec{M}, Y) \vdash \exists Y B(\vec{a}, \vec{M}, Y)$ with all free variables indicated. We need to find a polytime witnessing function F s.t.

$$\mathbf{V^1}(F) \models B(\vec{a}, \vec{M}, N) \to B(\vec{a}, \vec{M}, F(\vec{a}, \vec{M}, N)). \tag{4.153}$$

The Σ_0^B-bit-defining axiom $F(\vec{a}, \vec{M}, N)(z) \leftrightarrow N(z)$ defines such a witnessing function. If S is the Σ_0^B-**COMP** axiom, then it has the form

$$\vdash \exists X \leq b \forall z < b\big(X(z) \leftrightarrow B(z, b, \vec{a}, \vec{M})\big) \tag{4.154}$$

with all free variables indicated, and we have to find a polytime witnessing function F s.t.

$$\mathbf{V^1}(F) \models |F(b, \vec{a}, \vec{M})| \leq b \wedge \forall z < b\big(F(b, \vec{a}, \vec{M})(z) \leftrightarrow B(z, b, \vec{a}, \vec{M})\big) \tag{4.155}$$

The Σ_0^B-bit-defining axiom

$$F(b, \vec{a}, \vec{M})(z) \leftrightarrow z < b \wedge B(z, b, \vec{a}, \vec{M}) \tag{4.156}$$

defines such a function F.

Case II: S is obtained by an application of the Σ_1^B-Ind rule. This is the most interesting and most difficult case. Hence S is the bottom sequent of an inference

$$(\Sigma_1^B\text{-Ind})\ \frac{\Pi, \exists X \leq r(b) B(b, X) \ \vdash\ \exists X \leq r(b+1) B(b+1, X), \Omega}{\Pi, \exists X \leq r(0) B(0, X) \ \vdash\ \exists X \leq r(t) B(t, X), \Omega} \tag{4.157}$$

where the eigenvariable b does not occur in the bottom sequent S and $B(x, X)$ is a Σ_0^B-formula. Note that if the induction formula is in Σ_0^B, then we can just apply the induction hypothesis and are done.

In the following, for a set Π in a sequent S, let $\hat{\Pi}$ denote the conversion of Π in $\hat{\mathsf{S}}$ according to (4.152). By induction hypothesis there is a collection $\mathcal{F}$ of polytime functions including a function F s.t.

$$\begin{aligned} \mathbf{V^1}(\mathcal{F}) \models \hat{\Pi}, |N| \leq r(b) \wedge B(b, N) \vdash \\ |F(\vec{a}, b, \vec{M}, N)| \leq r(b+1) \wedge B\big(b+1, F(\vec{a}, b, \vec{M}, N)\big), \hat{\Omega} \end{aligned} \tag{4.158}$$

Note that b (the eigenvariable) and N do not occur in $\hat{\Pi}$ and can only occur in $\hat{\Omega}$ as arguments to witnessing functions. We need to find a new witnessing

function F' for the formula $\exists X \leq r(t)B(t,X)$ in S. Since F is in FP, the following function F', defined by bounded recursion on notation, is also in FP by Cobham's theorem 4.39 (below, we do not write the irrelevant parameters):

$$\begin{aligned} F'(0,X) &= X \\ F'(y+1,X) &= \mathrm{chop}\Big(r(y+1),\ F\big(y,F'(y,X)\big)\Big) \end{aligned} \tag{4.159}$$

where r is the polynomial represented by the $\mathcal{L}^2_{\mathcal{A}}$-term r in (4.157).

Let $F_1(b,N_1),\ldots,F_m(b,N_m)$ (we omit the irrelevant parameters $\vec{a},\vec{M}$) be the witnessing functions of $\hat{\Omega}$ in (4.158). For each $i=1,\ldots,m$, we define the function composition $F'_i(b,N_i,) = F_i\big(b,F'(b,N_i)\big)$. Note that all F'_i are in FP according to Cobham's theorem 4.39. If we replace N with $F'(b,N)$ in (4.158) we obtain

$$\begin{aligned} &\hat{\Pi}, |F'(b,N)| \leq r(b) \wedge B(b,F'(b,N)) \vdash \\ &|F(b,F'(b,N))| \leq r(b+1) \wedge B\big(b+1,F(b,F'(b,N))\big), \hat{\Omega}' \end{aligned} \tag{4.160}$$

where, additionally, $\hat{\Omega}'$ is $\hat{\Omega}$ with $F'_i(b,N_i)$ replacing $F_i(b,N_i)$, for all i (recall that all N_i are distinct). Let $\mathcal{F}' = \mathcal{F} \cup \{F',F'_1,\ldots,F'_m\}$. (4.160) is also a theorem of $\mathbf{V^1}(\mathcal{F})'$ because N (and every N_i) is implicitly universally quantified in (4.158). Now consider the sequent

$$\begin{aligned} &\hat{\Pi}, |F'(b,N)| \leq r(b) \wedge B\big(b,F'(b,N)\big) \vdash \\ &|F'(b+1,N)| \leq r(b+1) \wedge B\big(b+1,F'(b+1,N)\big), \hat{\Omega}' \end{aligned} \tag{4.161}$$

where $F(b,F'(b,N))$ was replaced by $F'(b+1,N)$. It follows from the definition (4.159) of F' that (4.161) is also a theorem of $\mathbf{V^1}(\mathcal{F}')$. Fortunately, the sequent (4.161) has the form

$$\hat{\Pi}, C(b,N) \vdash C(b+1,N), \hat{\Omega}' \tag{4.162}$$

where

$$C(b,N) \equiv |F'(b,N)| \leq r(b) \wedge B\big(b,F'(b,N)\big)$$

and $C(b,N)$ is a $\Sigma^B_0(\mathcal{F}')$-formula. Unfortunately, b occurs as a parameter (and nowhere else) to the witnessing functions F'_i (in $\hat{\Omega}'$). But b must not occur in the desired sequent $\hat{\mathsf{S}}$ [i]. Therefore we try to remove b from $\hat{\Omega}'$ and introduce the number function $h(X)$ with the following defining axiom

$$h(X) = y \leftrightarrow y \leq t \wedge \big(\neg C(y+1,X) \vee y = t\big) \wedge \forall x \leq y C(x,X) \tag{4.163}$$

[i] Note that the fact that b occurs in $\hat{\Omega}'$ prohibits us from applying the $\Sigma^B_1(\mathcal{F}')$-IND rule on (4.162)

where the term t comes from (4.157). Intuitively, $h(X)$ is the smallest $y \leq t$ such that $\neg C(y{+}1, X)$, or t if there exists no $y \leq t$ that satisfies $\neg C(y{+}1, X)$. It will soon be clear what we need this function for. For each $i = 1, \ldots, m$ we define

$$F_i''(X) = F_i'\big(h(X), X\big)$$

F_i'' is in FP according to Cobham's theorem 4.39. Let $\hat{\Omega}''$ be like $\hat{\Omega}'$ with the exception that each occurrence of $F_i'(b, N_i)$ is replaced with $F_i''(N_i)$. We define

$$F''(X) = F'(t, X) \tag{4.164}$$

It follows again from theorem 4.39 that F'' is in FP. Now we define the sequent $\hat{\mathsf{S}}$ (recall (4.152)) as

$$\begin{aligned} \hat{\mathsf{S}} \ = \hat{\Pi}, |N| \leq r(0) \wedge B(0, N) \vdash \\ |F''(N)| \leq r(t) \wedge B\big(t, F''(N)\big), \hat{\Omega}'' \end{aligned} \tag{4.165}$$

Let $\mathcal{F}'' = \mathcal{F}' \cup \{h, F_1'', \ldots, F_m'', F''\}$. We finally have to show that $\hat{\mathsf{S}}$ is a theorem of $\mathbf{V^1}(\mathcal{F}'')$. It follows from the definition of F' ($F'(0, N) = N$) that (4.165) is equivalent in $\mathbf{V^1}(\mathcal{F}'')$ to

$$\begin{aligned} \hat{\Pi}, |F'(0, N)| \leq r(0) \wedge B\big(0, F'(0, N)\big) \vdash \\ |F''(N)| \leq r(t) \wedge B\big(t, F''(N)\big), \hat{\Omega}'' \end{aligned} \tag{4.166}$$

Then it follows from the definition (4.164) of F'' that (4.166) is equivalent in $\mathbf{V^1}(\mathcal{F}'')$ to

$$\begin{aligned} \hat{\Pi}, |F'(0, N)| \leq r(0) \wedge B\big(0, F'(0, N)\big) \vdash \\ |F'(t, N)| \leq r(t) \wedge B\big(t, F'(t, N)\big), \hat{\Omega}'' \end{aligned} \tag{4.167}$$

which is identical to

$$\hat{\Pi}, C(0, N) \vdash C(t, N), \hat{\Omega}''. \tag{4.168}$$

Hence it suffices to show that $\mathbf{V^1}(\mathcal{F}'')$ proves (4.168). Since b is implicitly universally quantified in (4.162), it follows that[i]

$$\mathbf{V^1}(\mathcal{F}'') \ \models \ \hat{\Pi}, C\big(h(N), N\big) \vdash C\big(h(N) + 1, N\big), \hat{\Omega}''. \tag{4.169}$$

Now, by the definition of the function h, we have

$$\mathbf{V^1}(\mathcal{F}'') \models C(0, N) \vdash C\big(h(N), N\big), \ \text{and} \tag{4.170}$$

$$\mathbf{V^1}(\mathcal{F}'') \models C\big(h(N) + 1, N\big) \vdash C(t, N). \tag{4.171}$$

[i] If in doubt about $\hat{\Omega}''$, check the definition of $\hat{\Omega}''$.

The reason is the following (we argue[i] "in" the theory $\mathbf{V^1}(\mathcal{F}'')$): By the last conjunct of (4.163), $C\big(h(N), N\big)$ holds. Therefore $\mathbf{V^1}(\mathcal{F}'')$ proves (4.170). Now consider the middle conjunct of (4.163). If the first disjunct holds, then $\neg C\big(h(N)+1, N\big)$. If the second disjunct holds, then $h(N) = t$ and $C(t, N)$ holds by the last conjunct. In any case, $\mathbf{V^1}(\mathcal{F}'')$ proves (4.171).

Now we can use the rules of LK^2 (since they are sound w.r.t. to the axioms of $\mathbf{V^1}(\mathcal{F}'')$). First we apply the cut rule (after the obvious applications of the weakening rules) to (4.169) and (4.170) and obtain the sequent

$$\hat{\Pi}, C(0, N) \vdash C\big(h(N)+1, N\big), \hat{\Omega}'' \tag{4.172}$$

Then we do the same with (4.172) and (4.171) and obtain (4.168). Hence $\mathbf{V^1}(\mathcal{F}'') \models \hat{\mathsf{S}}$.

Case III: S is obtained by an inference of the form

$$\frac{\Pi \ \vdash \ \Omega, B(T)}{\Pi \ \vdash \ \Omega, \exists X B(X)} \text{ (set } \exists\text{-right)}$$

Let S_1 denote the top sequent. Note that since $\mathbf{V^1}$ is an $\mathcal{L}^2_{\mathcal{A}}$-theory, T can either be a free variable or the constant symbol $\varnothing$ (introduced when putting $\mathsf{\Pi}$ in free variable normal form). If T is a free variable M, then M must occur in Π or Ω (because $\mathsf{\Pi}$ is in FVNF). We need to find a witnessing function F s.t. M and the witnessing term $F(\vec{a}, \vec{M}, \vec{N}, M)$ are always mapped to equal elements in the universe. The Σ^B_0-bit-defining axiom

$$F(\vec{a}, \vec{M}, \vec{N}, M)(z) \leftrightarrow z < |M| \wedge M(z) \tag{4.173}$$

defines such a function F. Then the induction hypothesis $\mathbf{V^1}(\mathcal{F}) \models \hat{\mathsf{S}_1}$ implies $\mathbf{V^1}(\mathcal{F} \cup F) \models \hat{\mathsf{S}}$. Note that $F \in \mathsf{FP}$ by corollary 4.37. If T is $\varnothing$, then let F be the function defined by the Σ^B_0-bit-defining axiom

$$F(\vec{a}, \vec{M}, \vec{N})(z) \leftrightarrow z < 0 \tag{4.174}$$

It follows from lemma 4.7 (2)[ii] and the contrapositions of the axioms **L2** and **B12** that $\mathbf{V^1}(\mathcal{F} \cup F) \models |F(\vec{a}, \vec{M}, \vec{N})| = 0$, and then from axiom **SE** that $\mathbf{V^1} \models F(\vec{a}, \vec{M}, \vec{N}) = \varnothing$ (note that we added $|\varnothing| = 0$ to $\mathbf{V^1}$). Then the induction hypothesis $\mathbf{V^1}(\mathcal{F}) \models \hat{\mathsf{S}_1}$ implies $\mathbf{V^1}(\mathcal{F} \cup F) \models \hat{\mathsf{S}}$.

Case IV: S is obtained by an inference of the form

$$\frac{\Pi, B(M) \ \vdash \ \Omega}{\Pi, \exists X B(X) \ \vdash \ \Omega} \text{ (set } \exists\text{-left)}$$

[i] Note that we implicitly use the number axioms of $\mathbf{V^1}$ and some its theorems (lemma 4.15).

[ii] $\mathbf{V^1} \models \neg z < 0$

where the eigenvariable M does not occur in S. Let S_1 denote the top sequent. Note that $\hat{\mathsf{S}}$ contains a new free variable N for the quantifier variable X. For S we can use the same witnessing functions as for S_1 with $F_i(\vec{a}, N, \vec{M}, \vec{N})$ replacing $F_i(\vec{a}, M, \vec{M}, \vec{N})$ for each witnessing function F_i of S_1. Then we can apply the induction hypothesis and are done.

Case V: S is obtained by an inference of the form

$$\frac{\Pi, b \leq t \wedge B(b) \ \vdash \ \Omega}{\Pi, \exists x \leq t B(x) \ \vdash \ \Omega} \, (\exists\text{-left})$$

where the eigenvariable b does not occur in S. We cannot use the same witnessing functions as for the top sequent because the argument b is eliminated. Therefore, the argument b of the witnessing functions of the top sequent needs to be replaced by something. We define the function $g(\vec{a}, \vec{M})$ as the minimum $b \leq t$ that satisfies $B(b)$. g has the following Σ_0^B-defining axiom

$$y = g(\vec{a}, \vec{M}) \leftrightarrow y \leq t \wedge B(y) \wedge \forall x < y \neg B(x) \tag{4.175}$$

Since g is obviously polynomially bounded (by t), it follows from corollary 4.36 that g is in FP (even in FAC^0). For each witnessing function $F_i(\vec{a}, b, \vec{M}, \vec{N})$ of the top sequent we can now define a witnessing function $F_i'(\vec{a}, \vec{M}, \vec{N})$ for $\hat{\mathsf{S}}$ by function composition as follows

$$F_i'(\vec{a}, \vec{M}, \vec{N}) = F_i\big(\vec{a}, g(\vec{a}, \vec{M}), \vec{M}, \vec{N}\big)$$

By Cobham's theorem 4.39, F_i' is in FP.

Case V': The case where S is obtained by the $\forall$-right rule is proved analogously.

Case VI: S is obtained by an inference of the form

$$\frac{\Pi \ \vdash \ \Omega, A \qquad \Pi \ \vdash \ \Omega, B}{\Pi \ \vdash \ \Omega, A \wedge B} \, (\wedge\text{-right})$$

Let S_1 and S_2 denote the two top sequents, respectively. The problem with this rule is that the witnessing functions of S_1 and S_2 are not necessarily the same and it is not immediately clear which functions we have to chose. We illustrate this by the example where Ω consists of the single formula $\exists X C(X)$. Then, by the induction hypothesis, we have

$$\mathbf{V^1}(\mathcal{F}_1) \models \bigwedge \hat{\Pi} \rightarrow A \vee C\big(F_1(\dots)\big), \tag{4.176}$$

$$\mathbf{V^1}(\mathcal{F}_2) \models \bigwedge \hat{\Pi} \rightarrow B \vee C\big(F_2(\dots)\big) \tag{4.177}$$

for some witnessing functions $F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$. We need a new witnessing function F s.t.

$$\mathbf{V^1}(\mathcal{F}_1 \cup \mathcal{F}_2 \cup \{F\}) \models \bigwedge \hat{\Pi} \to (A \wedge B) \vee C\big(F(\dots)\big)$$

Note that we cannot simply take $F = F_1$ because we do not know which disjunct in the RHS of (4.176) is true. Informally, if $C\big(F_1(\dots)\big)$ is true, then we can take F_1, otherwise we take F_2[i]. Hence F can be defined as

$$F(\dots)(z) \leftrightarrow \Big(C\big(F_1(\dots)\big) \wedge F_1(\dots)(z)\Big) \vee \Big(\neg C\big(F_1(\dots)\big) \wedge F_2(\dots)(z)\Big) \tag{4.178}$$

In general, Ω has the form (cf. (4.151))

$$\Omega \;\equiv\; \Omega', \exists Y_1 C_1(Y_1), \dots, \exists Y_n C_n(Y_n),$$

for $n \geq 0$, where all C_i, all formulas in Ω' as well as A and B are Σ_0^B-formulas. For $i = 1, \dots, n$, let $F_i^1(\dots)$ and $F_i^2(\dots)$ be the witnessing functions in $\hat{\mathsf{S}}_1$ and $\hat{\mathsf{S}}_2$, respectively. Then we define the new witnessing functions F_i as in (4.178), with F_i^1 replacing F_1 and F_i^2 replacing F_2.

Case VI': The case where S is obtained by the $\vee$-left rule is proved analogously.

Case VII: S is obtained by an inference of the form

$$\frac{\Pi \;\vdash\; \Omega, A, A}{\Pi \;\vdash\; \Omega, A}\,(\text{contraction-right})$$

If A is a Σ_0^B-formula, then we can just apply the induction hypothesis and are done. If A has the form $\exists X B(X)$, then, by induction hypothesis, there are witnessing functions $F_1 \in \mathcal{F}$ and $F_2 \in \mathcal{F}$ s.t.

$$\mathbf{V^1}(\mathcal{F}) \models \bigwedge \hat{\Pi} \to B\big(F_1(\vec{a}, \vec{M})\big) \vee B\big(F_2(\vec{a}, \vec{M})\big) \vee \bigvee \hat{\Omega} \tag{4.179}$$

We need a new witnessing function F s.t.

$$\mathbf{V^1}(\mathcal{F} \cup \{F\}) \models \bigwedge \hat{\Pi} \to B\big(F(\vec{a}, \vec{M})\big) \vee \bigvee \hat{\Omega}$$

As in case VI, we cannot simply take $F = F_1$ or $F = F_2$ because we do not know which of the disjuncts in the RHS of (4.179) is true. An appropriate function F is defined by (replace $\dots$ with $\vec{a}, \vec{M}$)

$$F(\dots)(z) \leftrightarrow \Big(B\big(F_1(\dots)\big) \wedge F_1(\dots)(z)\Big) \vee \Big(\neg B\big(F_1(\dots)\big) \wedge F_2(\dots)(z)\Big)$$

[i]Note that it doesn't matter if $C\big(F_2(\dots)\big)$ is false, because then $A \wedge B$ is true.

Case VII': If S is obtained from the contraction-left rule and the formula A is of the form $\exists X B(X)$, then we just replace every witnessing function $F_i(\ldots, N_1, N_2)$ with $F_i(\ldots, N, N)$, where N is the new fresh variable introduced for X in S and N_1, N_1 are the (distinct) free variables for X in the top sequent.

Case VIII: S is obtained by an inference of the form

$$\frac{\Pi \vdash \Omega, A \qquad \Pi, A \vdash \Omega}{\Pi \vdash \Omega}\,(\text{Cut})$$

Let $\mathsf{S_1}$ and $\mathsf{S_2}$ denote the two top sequents, respectively. Assume first that the cut formula A is a Σ_0^B-formula. Then we proceed similarly as in case VI. Let $F_1^1, \ldots, F_n^1$ be the winteesing functions for Ω in $\hat{\mathsf{S}_1}$ and $F_1^2, \ldots, F_n^2$ in $\hat{\mathsf{S}_2}$, respectively. Then we define new witnessing function $F_1, \ldots, F_n$ for $\hat{\mathsf{S}}$ by

$$F_i(\ldots)(z) \leftrightarrow \Big(\big(\neg A \wedge F_i^1(\ldots)(z)\big) \vee \big(A \wedge F_i^2(\ldots)(z)\big)\Big) \tag{4.180}$$

If A is not a Σ_0^B-formula, then A has the form $\exists X B(X)$ and $B(X)$ is a Σ_0^B-formula. Let $G(\vec{a}, \vec{M})$ be the witnessing function for $B(X)$ in $\hat{\mathsf{S}_1}$ and let N be the fresh introduced for X in $\hat{\mathsf{S}_2}$. Let $F_1^1(\vec{a}, \vec{M}), \ldots, F_n^1(\vec{a}, \vec{M})$ be the witnessing functions for Ω in $\hat{\mathsf{S}_1}$ and let $F_1^2(\vec{a}, \vec{M}, N), \ldots, F_n^2(\vec{a}, \vec{M}, N)$ be the witnessing functions for Ω in $\hat{\mathsf{S}_2}$. The new witnessing functions $F_1, \ldots, F_n$ for Ω in $\hat{\mathsf{S}}$ are defined by

$$\begin{aligned} F_i(\vec{a}, \vec{M})(z) \leftrightarrow &\Big(\neg B\big(G(\vec{a}, \vec{M})\big) \wedge F_i^1(\vec{a}, \vec{M})(z)\Big) \vee \\ &\Big(B\big(G(\vec{a}, \vec{M})\big) \wedge F_i^2(\vec{a}, \vec{M}, G(\vec{a}, \vec{M}))(z)\Big) \end{aligned}$$

Case IX: The case where S is obtained by a weakening rule is easy. In the case of weakening-left, nothing needs to be done. For weakening-right we can introduce an *arbitrary* witnessing function in FP for the case that the introduced formula has the form $\exists X B(X)$.

Case X: There is nothing to do for the exchange, $\neg$-introduction, $\exists$-right, $\forall$-left, $\vee$-right and $\wedge$-left rules (see also the remark about the subformula property at the beginning of the proof). Note that $\exists$-right and $\forall$-left do not eliminate free variables. □

The following general witnessing theorem follows from the above special case.

Theorem 4.95 (Witnessing Theorem)**.** *Let $A(\vec{x}, \vec{y}, \vec{X}, \vec{Y})$ be a Σ_0^B-formula and suppose that*

$$\mathbf{V^1} \models \exists \vec{y} \exists \vec{Y} A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}).$$

Then there exist functions $f_1, \dots, f_m, F_1, \dots, F_n \in \mathsf{FP}$ *such that*

$$\begin{aligned}&\mathbf{V^1}(f_1, \dots, f_m, F_1, \dots, F_n) \models \\ &A\big(\vec{x}, f_1(\vec{x}, \vec{X}), \dots, f_m(\vec{x}, \vec{X}), \vec{X}, F_1(\vec{x}, \vec{X}), \dots, F_n(\vec{x}, \vec{X})\big).\end{aligned}$$

Proof from Lemma 4.94. Assume $\mathbf{V^1} \models \exists \vec{y} \exists \vec{Y} A(\vec{x}, \vec{y}, \vec{X}, \vec{Y})$. Then we also have

$$\mathbf{V^1}(\mathrm{row}) \models \exists \vec{y} \exists \vec{Y} A(\vec{x}, \vec{y}, \vec{X}, \vec{Y}). \tag{4.181}$$

Let m and n be the "arities" of $\vec{y}$ and $\vec{Y}$, respectively. We claim that

$$\begin{aligned}\mathbf{V^1}(\mathrm{row}) \models \exists Z \big(&\big|Z[0]\big| = y_1 \wedge \dots \wedge \big|Z[\underline{m-1}]\big| = y_m \wedge \\ &Z[\underline{m}] = Y_1 \wedge \dots \wedge Z[\underline{m+n-1}] = Y_n\big)\end{aligned} \tag{4.182}$$

The proof of this claim is analogous to the proof of lemma 4.62. Then it follows from (4.181) and (4.182) that

$$\mathbf{V^1}(\mathrm{row}) \models \exists Z \underbrace{A\big(\vec{x}, \big|Z[0]\big|, \dots, \big|Z[\underline{m-1}]\big|, \vec{X}, Z[\underline{m}], \dots, Z[\underline{m+n-1}]\big)}_{B(\vec{x}, \vec{X}, Z)} \tag{4.183}$$

where $B(\vec{x}, \vec{X}, Z)$ is a $\Sigma_0^B(\mathcal{L}^2_{\mathcal{A}} \cup \{\mathrm{row}\})$-formula. By lemma 4.52, there is a $\Sigma_0^B(\mathcal{L}^2_{\mathcal{A}})$-formula $B'(\vec{x}, \vec{X}, Z)$ s.t.

$$\mathbf{V^1}(\mathrm{row}) \models B(\vec{x}, \vec{X}, Z) \leftrightarrow B'(\vec{x}, \vec{X}, Z) \tag{4.184}$$

Since $\mathbf{V^1}(\mathrm{row})$ is a conservative extension of $\mathbf{V^1}$ (by corollary 4.77), it follows from (4.184) and (4.183) that

$$\mathbf{V^1} \models \exists Z B'(\vec{x}, \vec{X}, Z) \tag{4.185}$$

Now we can apply the witnessing lemma 4.94. I.e. there exists a set function $F \in \mathsf{FP}$ s.t.

$$\mathbf{V^1}(F) \models B'\big(\vec{x}, \vec{X}, F(\vec{x}, \vec{X})\big) \tag{4.186}$$

Therefore, by (4.184),

$$\mathbf{V^1}(\mathrm{row}, F) \models B\big(\vec{x}, \vec{X}, F(\vec{x}, \vec{X})\big) \tag{4.187}$$

and by (4.183)

$$\begin{aligned}\mathbf{V^1}(\mathrm{row}, F) \models A\big(&\vec{x}, \big|F(\vec{x}, \vec{X})[0]\big|, \dots, \big|F(\vec{x}, \vec{X})[\underline{m-1}]\big|, \\ &\vec{X}, F(\vec{x}, \vec{X})[\underline{m}], \dots, F(\vec{x}, \vec{X})[\underline{m+n-1}]\big)\end{aligned} \tag{4.188}$$

Now let

$$f_1(\vec{x}, \vec{X}) = \big|F(\vec{x}, \vec{X})[0]\big|, \ldots, f_m(\vec{x}, \vec{X}) = \big|F(\vec{x}, \vec{X})[m-1]\big|,$$
$$F_1(\vec{x}, \vec{X}) = F(\vec{x}, \vec{X})[m], \ldots, F_n(\vec{x}, \vec{X}) = F(\vec{x}, \vec{X})[m+n-1]$$

Let $\mathcal{F}$ denote $\{f_1, \ldots, f_m, F_1, \ldots, F_n\}$. Then we have

$$\mathbf{V^1}(\mathcal{F} \cup \{\mathrm{row}, F\}) \models A\big(\vec{x}, f_1, \ldots, f_m, \vec{X}, F_1, \ldots, F_n\big)$$

where f_i stands for $f_i(\vec{x}, \vec{X})$ (analogously for F_i). $\mathbf{V^1}(\mathcal{F} \cup \{\mathrm{row}, F\})$ is a conservative extension of $\mathbf{V^1}(\mathcal{F})$ by corollary 4.77. Therefore

$$\mathbf{V^1}(\mathcal{F}) \models A\big(\vec{x}, f_1, \ldots, f_m, \vec{X}, F_1, \ldots, F_n\big)$$

and we are done. □

The following corollary follows immediately from the above witnessing theorem.

Corollary 4.96. *Every number or set function that is Σ^1_1-definable in $\mathbf{V^1}$ is in* FP*.*

And together with the lower bound of $\mathbf{V^1}$ (theorem 4.79) we obtain our main result.

Corollary 4.97 ($\mathbf{V^1}$ characterises FP)**.** *The Σ^1_1-definable functions of $\mathbf{V^1}$ are exactly the polynomial time computable functions* FP*.*

And we can strengthen the above using Parikh's theorem 4.49.

Corollary 4.98. *The Σ^B_1-definable functions of $\mathbf{V^1}$ are exactly the polynomial time computable functions* FP*.*

References

[1] Bovet, D. P., and Crescenzi, P. *Introduction to the theory of complexity*. Prentice Hall International (UK) Ltd., Hertfordshire, UK, UK, 1994.

[2] Buss, S. *Bounded Arithmetic*. Bibliopolis, 1986.

[3] Buss, S. *An Introduction to Proof Theory*. Elsevier Science B.V., 1998.

[4] Cobham, A. The intrinsic computational difficulty of functions. A. Y. Bar-Hillel, North-Holland, Ed., Logic, Methodology and Philosophy of Science, proceedings of the second International Congress, held in Jerusalem, 1964.

[5] Cook, S., and Kolokolova, A. A second-order system for polynomial-time reasoning based on graedel's theorem. *Electronic Colloquium on Computational Complexity (ECCC) 8*, 24 (2001).

[6] Cook, S., and Nguyen, P. Foundations of proof complexity: Bounded arithmetic and propositional translations. 2006.

[7] Gentzen, G. Untersuchungen über das logische Schliessen. In *Logik-Texte: Kommentierte Auswahl zur Geschichte der Modernen Logik (vierte Auflage)*, K. Berka and L. Kreiser, Eds. Akademie-Verlag, Berlin, 1986, pp. 206–262.

[8] Parikh, R. J. Existence and feasibility in arithmetic. *Journal of Symbolic Logic 36* (1971), 494–508.

[9] van Dalen, D. *Logic and Structure*, 4th ed. Springer-Verlag Berlin Heidelberg, 2004.

[10] Wrathall, C. Rudimentary predicates and relative computation. *SIAM J. Comput. 7*, 2 (1978), 194–209.

[11] Zambella, D. Notes on polynomially bounded arithmetic. *The Journal of Symbolic Logic 61*, 3 (1996), 942–966.

[12] Zambella, D. End extensions of models of linearly bounded arithmetic. *Annals of Pure and Applied Logic 88*, 2-3 (1997), 263–277.