FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Provability Interpretations of a Many-Sorted Polymodal Logic

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Computational Intelligence

eingereicht von

Gerald Berger
Matrikelnummer 0958520

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung: A.o.Univ.Prof. Dr. Hans Tompits
Prof. Lev D. Beklemishev (Steklov Mathematical Institute, Moskau)

Wien, 24.02.2015

(Unterschrift Verfasser) (Unterschrift Betreuung)

Erklärung zur Verfassung der Arbeit

Gerald Berger
Ottakringer Straße 48/20-21, 1170 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit—einschließlich Tabellen, Karten und Abbildungen—, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

(Ort, Datum) (Unterschrift Verfasser)

Acknowledgments

First and foremost I would like to express my sincerest gratitude to my supervisors. I thank Hans Tompits for agreeing to supervise me on this thesis. His comments and suggestions greatly improved the presentation and the content of this work. Apart from that, the support I received from him during my studies goes far beyond the scope of this thesis and I thank him for that. When one is interested in provability logics it is inevitable to hit upon the work of S.N. Artemov and his former students, among them Lev D. Beklemishev. It was him who formulated the problems this thesis is concerned with. Moreover, his work on provability logics and his comments on my work were valuable sources of progress. I thank him for agreeing to supervise me on this topic despite the fact that we, unfortunately, have never met in person. I am glad that I could exchange thoughts with one of the leading researchers in provability logics.

Apart from the people who provided professional support for my thesis, I want to thank those who were important for the success of my studies. I thank my family for supporting my studies in all imaginable aspects. Furthermore, many thanks go to my friends who accompanied me during my studies. Concerning this thesis, I especially want to mention Benjamin Kiesl who showed great interest in the progress of my work and with whom I attended many lectures on logics and mathematics.

Abstract

Provability logics constitute a well-studied branch of nonclassical logics and find interpretations in systems formalizing elementary number theory. The polymodal provability logic GLP, due to G.K. Japaridze, received considerable interest in the literature. GLP is arithmetically complete for an arithmetical interpretation which is closely related to the partial uniform reflection principles in formal arithmetic. Furthermore, the closed fragment of GLP allows to develop an ordinal notation system up to ε_0. Based on these observations, L.D. Beklemishev provided an alternative proof of G. Gentzen's consistency proof for Peano Arithmetic (PA) by transfinite induction up to ε_0. This ordinal analysis is carried out in the framework of graded provability algebras, which enable one to capture proof-theoretic information of the theory under consideration. The graded provability algebra of a theory can—from a logical point of view—be considered as a many-sorted variant of GLP.

In this thesis, we investigate this many-sorted variant of GLP which assigns sorts $\alpha \leq \omega$ to propositional variables. Thereby, propositional variables of sort $n < \omega$ are arithmetically interpreted as Π_{n+1}-sentences. In response to a question posed by Beklemishev, we show in the first part of this thesis that the resulting many-sorted modal logic is arithmetically complete with respect to a class of arithmetical interpretations which satisfies the aforementioned restriction.

Since these proof-theoretic applications can already be carried out in a positive fragment of GLP, we follow in the second part of this thesis recent trends concerning the investigation of such positive fragments of GLP. In the style of a work due to Beklemishev, we define a many-sorted positive reflection calculus where we, from the point of view of arithmetic, interpret the modal diamonds as different forms of reflection in formal arithmetic. Thereby, the restriction to the positive fragment allows for a richer arithmetical interpretation of propositional variables: these are not interpreted as single arithmetical sentences but as primitive recursive numerations of possibly infinite arithmetical theories. There, variables of sort $n < \omega$ are interpreted as Π_{n+1}-axiomatized extensions of PA, while variables of sort ω are interpreted as arbitrary extensions thereof. This interpretation enables us to introduce an additional modal operator $\langle\omega\rangle$ which is interpreted as the full uniform reflection schema in arithmetic that knows no finite, yet a recursive axiomatization. We prove that our reflection calculus is arithmetically complete with respect to this interpretation.

Kurzfassung

Beweisbarkeitslogiken stellen einen wohlstudierten Zweig nichtklassischer Logiken dar und finden Interpretationen in Systemen, welche die elementare Zahlentheorie formalisieren. Die polymodale Beweisbarkeitslogik GLP von G.K. Japaridze erfuhr reges Interesse in der Literatur. GLP ist vollständig bezüglich einer arithmetischen Interpretation, welche eng mit den partiellen uniformen Reflexionsprinzipien der formalen Arithmetik zusammenhängt. Weiters erlaubt das geschlossene Fragment von GLP die Entwicklung eines Systems zur Notation von Ordinalzahlen bis ε_0. Basierend auf diesen Beobachtungen lieferte L.D. Beklemishev einen zum Gentzen'schen alternativen Beweis zur Konsistenz der Peano Arithmetik (PA) per transfiniter Induktion bis ε_0. Diese beweistheoretische Analyse wird im Rahmen von sortierten Beweisbarkeitsalgebren durchgeführt, welche einem erlauben beweistheoretische Informationen der betrachteten Theorie zu erfassen. Die sortierte Beweisbarkeitsalgebra einer Theorie kann—von einem logischen Standpunkt betrachtet—als mehrsortige Variante von GLP aufgefasst werden.

In dieser Arbeit untersuchen wir diese mehrsortige Variante von GLP, welche aussagenlogischen Variablen Sorten $\alpha \leq \omega$ zuweist. Dabei wird jede aussagenlogische Variable der Sorte $n < \omega$ nur durch Π_{n+1}-Sätze interpretiert. In Beantwortung einer von Beklemishev gestellten Frage zeigen wir im ersten Teil dieser Arbeit, dass diese mehrsortige Logik arithmetisch vollständig bezüglich einer geeigneten arithmetischen Interpretation ist, welche der zuvor Einschränkung bezüglich der Sorten genügt.

Da die beweistheoretischen Anwendungen von GLP bereits in dem positiven Fragment derselben erfolgen können, folgen wir im zweiten Teil dieser Arbeit jüngsten Untersuchungen positiver Fragmente von GLP. In Anlehnung an eine Arbeit von Beklemishev definieren wir einen mehrsortigen positiven Reflexionskalkül, wobei wir die modalen Diamanten als verschiedene Formen der Reflexion in der formalen Arithmetik auffassen. Dabei erlaubt uns die Beschränkung auf das positive Fragment eine reichhaltigere Interpretation der aussagenlogischen Variablen: Diese werden nicht als arithmetische Sätze, sondern als primitiv rekursive Aufzählungen von möglicherweise unendlichen arithmetischen Theorien interpretiert. Hierbei werden Variablen der Sorte $n < \omega$ durch Π_{n+1}-axiomatisierbare Erweiterungen von PA instanziert, während jene der Sorte ω durch beliebige Erweiterungen derselben interpretiert werden. Diese Interpretation gestattet die Einführung eines modalen Operators $\langle\omega\rangle$, welcher als das Schema der vollen uniformen Reflexion der Arithmetik interpretiert wird, für welches es keine endliche, jedoch eine rekursive Axiomatisierung gibt. Wir zeigen, dass unser Reflexionskalkül arithmetisch vollständig bezüglich dieser Interpretation ist.

Contents

CHAPTER 1

Introduction

In the course of proving his celebrated incompleteness theorems, Gödel, in his seminal paper of 1931 [19], demonstrated how a sufficiently strong theory can encode properties talking about itself—most importantly, he showed how such a theory can talk about its own theorems in terms of formal provability. Later on in 1933 [20], Gödel chose the language of propositional modal logic to provide an adequate semantics for intuitionistic logic—according to Brouwer, intuitionistic truth means provability. Gödel briefly mentions that provability in formal systems may be viewed as a modal operator.

It was then Löb [30] who discovered a principle of provability which, together with some elegant principles already known, constitutes the axiomatic basis of the well-known provability logic GL. In particular, Löb's principle allows one to establish (a formalized version of) the second incompleteness theorem by purely modal reasoning. A landmark result by Solovay [39] reveals that this logic is adequate for provability in Peano Arithmetic (PA) with the interpretation of the modality $\Box$ as the standard Gödelian provability predicate.

Provability logics have since then been vividly studied and ties have been established between mathematical logic and the more isolated field of nonclassical logics. Most importantly for us, provability logics have found a manifold of interpretations in arithmetical theories. In this context, provability logics are mostly modal logics which axiomatize properties of certain provability predicates of arithmetical theories.

One of the logics which received considerable interest in the literature is the polymodal provability logic GLP due to Japaridze [25]. GLP is formulated over a modal language with modalities $[n]$ for every natural number n. Japaridze showed that GLP is arithmetically complete for sound extensions of Peano Arithmetic when $[n]$ is interpreted as being "provable under n nested applications of the ω-rule" (see Section 3.3). Later, Ignatiev [24] simplified the work of Japaridze and showed that GLP is complete for an even broader class of arithmetical interpretations. Most importantly, for a sound extension T of PA, it turns out that GLP is arithmetically complete for T for the interpretation of $[n]$ as (formalized) provability in the theory $T + \mathsf{Th}_{\Pi_n}(\mathbb{N})$, where $\mathsf{Th}_{\Pi_n}(\mathbb{N})$ denotes the set of all true Π_n-sentences. The arithmetical interpretation of the dual operator $\langle n\rangle := \neg[n]\neg$ is called *n-consistency* and is equivalent to the uniform reflection principle for Π_{n+1}-formulas over T (see Section 2.5). Ignatiev further showed that GLP exhibits Craig interpolation, has a fixed-point property, and that there is a universal model for the closed fragment of GLP (i.e., the set of theorems with no variables) based on the ordinal ε_0.

Beklemishev [2] brought the study of GLP into mainstream proof theory concerning ordinal analysis. The background can be roughly sketched as follows (cf. also Rathjen [32] some background on ordinal analysis). Ever since the work of Gentzen [17, 18], it has been a principal aim of proof theory to assign ordinals to theories which should somehow measure the "proof-theoretic strength" of the theory under consideration. Gentzen showed that Peano Arithmetic is consistent by transfinite induction up to ε_0 and furthermore that transfinite induction up to every ordinal less than ε_0 is provable in PA. Hence, it seems natural to define the *proof-theoretic ordinal* of a theory T as the least ordinal such that the theory plus transfinite induction up to this ordinal proves the consistency of T. Similarly, we could define the proof-theoretic ordinal to be the supremum of the order types which T is able to prove to be well-founded (recall that ε_0 is the supremum of $\omega, \omega^\omega, \omega^{\omega^\omega}, \ldots$). The problem with these definitions is that the notions crucially depend on the representation of the ordinal notation system within T. To wit, a well-known example due to Kreisel shows that one can define a primitive recursive well-ordering of order type ω such that a comparatively weak theory can prove the consistency of a strong theory by transfinite induction on this ordering [27]. The tentative conclusion we can draw from the possibility of such pathological orderings is that in order to gain "natural" representations of ordinals, the notion of proof-theoretic ordinal should disregard many syntactical details of the theory under consideration.

Beklemishev [2] proposes an approach to the ordinal analysis of PA which addresses these issues. A proof-theoretic analysis of PA based on the notion of *graded provability algebras* is suggested which allows one to capture enough syntactic information in order to canonically recover an ordinal notation system up to ε_0. This permits one to obtain Gentzen's results in a rather abstract fashion. Furthermore, based on these notions, Beklemishev [4] provided a combinatorial statement undecided by PA.

The notion of graded provability algebra proposed by Beklemishev bears the structure of a Lindenbaum algebra of an arithmetical theory T enriched by additional operators $\langle n\rangle_T$ for every natural number n, denoting n-consistency. Given such a structure $\mathcal{M}_T$, one can associate a *stratification* $P_0 \subset P_1 \subset \cdots \subseteq \mathcal{M}_T$ with it, where the sets P_n correspond to Π_{n+1}-sentences. An algebra with such a stratification can be regarded as a many-sorted algebra, where the operator $\langle n\rangle_T$ maps elements from $\mathcal{M}_T$ to P_n. It is the modal logic of such many-sorted algebras we are going to investigate in this thesis.

More precisely, we consider the modal logic which contains variables of sort n for every $n < \omega$. The arithmetical interpretation of a variable of sort n ranges over Π_{n+1}-sentences. In addition to the postulates of GLP, our logic will contain the axiom of Σ_{n+1}-*completeness*, that is,

$$\neg p \to [n]\neg p,$$

where p is a variable of sort n. The notion of sort can be naturally extended to capture all polymodal formulas (i.e., terms in the language of many-sorted algebras). Substitution in the logic under consideration is then restricted to respect the sorts.

Since the proof-theoretic applications of GLP mentioned before can already take place in a positive fragment of the same, positive fragments of provability logics have received interest recently. Dashkov [13] showed that the positive fragment of GLP can be axiomatized by a positive calculus which is decidable in polynomial time. Furthermore, he mentions that an arithmetical interpretation of the positive fragment of GLP can be richer than the standard one: propositional variables can be interpreted as possibly infinite arithmetical theories rather than single sentences. The arithmetical interpretation of modal formulas is then not restricted to adhere to finitely axiomatizable concepts only. In particular, the arithmetical interpretation of positive formulas can result in theories of unbounded arithmetical complexity. Hence, this interpretation allows to introduce a modality $\langle\omega\rangle$ which is interpreted as the full uniform reflection principle in arithmetic. Beklemishev [7] showed that a suitable *reflection calculus* capturing these notions is arithmetically sound and complete with respect to the aforementioned arithmetical interpretation, where propositional variables are formally interpreted as primitive recursive numerations of theories extending PA. We follow these lines and investigate the positive fragment of our many-sorted variant by defining a suitable many-sorted version of Beklemishev's reflection calculus and show that it is arithmetically complete with respect to an arithmetical interpretation which treats variables of sort $n < \omega$ as Π_{n+1}-axiomatized extensions of PA. As in the work of Beklemishev, our calculus will also contain a modality $\langle\omega\rangle$ which, from an arithmetical point of view, corresponds to the full uniform reflection principle in arithmetic.

1.1 Structure of the Thesis

After this introductory chapter, we continue in Chapter 2 with an exposition of some background knowledge. We collect some basic facts on formal arithmetic in Section 2.2 and continue to review the famous limitative results of reasonably strong arithmetical theories in Sections 2.3 and 2.4. Section 2.5 contains a brief treatment of the reflection principles in arithmetic. In Section 2.6, we introduce the basic notions of provability logics and recite Solovay's famous theorems.

Chapter 3 is devoted to the study of our many-sorted variant of GLP, denoted by GLP^*. It is well-known that GLP is not sound and complete for any class of Kripke frames [24]. Therefore, Ignatiev [24] identified a logic weaker than GLP which is already sound and complete for a decent class of Kripke frames and allows one to prove properties about GLP by reducing GLP to that logic. Beklemishev [5] isolates an even more convenient subsystem of GLP, denoted by J, which he uses to simplify the arithmetical completeness theorem for GLP [6]. We will analogously define a logic J^* which is weaker than GLP^* and prove in Section 3.4 that J^* is complete with respect to a nice class of Kripke models. In Section 3.5, we show that GLP^* is arithmetically complete with respect to the broad class of arithmetical interpretations identified by Ignatiev which in addition respect our conditions on the sorts of variables. Afterwards, we will discuss some corollaries and extensions of this theorem.

In Chapter 4, we introduce many-sorted variants of the reflection calculi studied by Beklemishev [7] and Dashkov [13]. Here, the arithmetical interpretation of propositional variables of sort n is given by primitive recursive enumerations of Π_{n+1}-axiomatized extensions of PA, while the variables of sort ω can be assigned arbitrary extensions of PA. Additionally, we introduce a modality $\langle\omega\rangle$ which, as in the work of Beklemishev [7], receives the full uniform reflection principle as its arithmetical semantics. For $n < \omega$, the modal operator $\langle n\rangle$ is interpreted as the uniform reflection schemata restricted to Π_{n+1} formulas, that is, n-consistency. The arithmetical interpretation of our many-sorted positive calculus is subject of Section 4.2. Kripke semantics is subsequently treated in Section 4.3. Following Dashkov [13], we show in Section 4.4 that our reflection calculus axiomatizes the positive fragment of $\mathsf{GLP}^*_{\omega+1}$ which is the logic GLP^* enriched by a modality $\langle\omega\rangle$ and suitable axioms. In the style of Beklemishev [7], we continue in Section 4.5 to prove that this resulting system is arithmetically complete with respect to the aforementioned interpretation.

CHAPTER 2

Preliminaries

We assume familiarity with the basics of classical first-order logic as well as the treatment of metamathematics as a branch of number theory. In the sequel, we will briefly define the basic logical concepts in order to fix notation and terminology. We will then introduce the arithmetical theories of our interest in Section 2.2 and summarize some well-known properties about them. For more details on the contents of the first two sections, we refer the reader to standard textbooks on mathematical logic, e.g., Shoenfield [35], Boolos et al. [12], as well as to text books on formal arithmetic like Hájek and Pudlák [21]. Some famous limitative results concerning the metamathematics of arithmetic are repeated in Sections 2.3 and 2.4. We continue with a brief exposition of the so called reflection principles in arithmetic in Section 2.5, which we will encounter later in Chapter 4. At the end of this chapter, in Section 2.6, we briefly discuss existing provability logics which are relevant for us.

2.1 Basics

A *first-order language* $\mathcal{L}$ consists of *logical*, *nonlogical*, and *auxiliary symbols*. The nonlogical symbols are specified by pairwise disjoint sets of *predicate* (*relation*), *function*, and *constant symbols*, where each predicate and function symbol has an associated positive *arity*. Throughout the text, we assume that every first-order language $\mathcal{L}$ implicitly contains a binary predicate symbol $=$ called *equality* (which we write in infix notation for convenience). The logical symbols of $\mathcal{L}$ consist of equality, $\forall$, $\rightarrow$, $\neg$, as well as a countably infinite supply of (*individual*) *variables*. Furthermore, the symbols (,), and , are called *auxiliary symbols*. We assume that every language has the same logical and auxiliary symbols. Hence, a first-order language is determined by the choice of its nonlogical symbols. When exhibiting syntactic objects, we agree to let $x, y, z, \ldots$ (possibly with subscripts) be metavariables for individual variables. We say that a language $\mathcal{L}'$ *extends* a language $\mathcal{L}$ (symbolically $\mathcal{L} \subseteq \mathcal{L}'$) if the nonlogical symbols of $\mathcal{L}$ are contained in $\mathcal{L}'$. If $\mathcal{L} \subseteq \mathcal{L}'$ then $\mathcal{L}'$ is an *extension* of $\mathcal{L}$.

Given a language $\mathcal{L}$, the set of $\mathcal{L}$-*terms* is defined inductively in the usual manner: (i) individual variables and constant symbols are $\mathcal{L}$-terms, (ii) if f is a function symbol of arity n and $t_1, \ldots, t_n$ are $\mathcal{L}$-terms then $f(t_1, \ldots, t_n)$ is an $\mathcal{L}$-term. If $t_1, \ldots, t_n$ are $\mathcal{L}$-terms and R is a predicate symbol of arity n then $R(t_1, \ldots, t_n)$ is an *atomic* $\mathcal{L}$-*formula*. The set of $\mathcal{L}$-*formulas* (*formulas over* $\mathcal{L}$) is defined inductively:

(i) Every atomic $\mathcal{L}$-formula is an $\mathcal{L}$-formula.

(ii) If φ and ψ are $\mathcal{L}$-formulas then $\neg\varphi$ and $(\varphi \to \psi)$ are $\mathcal{L}$-formulas.

(iii) If x is an individual variable and φ an $\mathcal{L}$-formula then $\forall x\, \varphi$ is an $\mathcal{L}$-formula.

We introduce the usual abbreviations for *logical connectives* different from $\to$ and $\neg$. We set $\varphi \vee \psi := \neg\varphi \to \psi$, $\varphi \wedge \psi := \neg(\neg\varphi \vee \neg\psi)$, $\varphi \leftrightarrow \psi := (\varphi \to \psi) \wedge (\psi \to \varphi)$, and $\exists x\, \varphi := \neg\forall x\, \neg\varphi$. As we did here in these definitions, we omit parentheses whenever possible and assign $\neg$, $\forall x$ the highest and $\to$, $\leftrightarrow$ the least binding priority. Furthermore, we sometimes also introduce additional parentheses for the sake of readability (possibly parentheses of a different style). We usually omit "$\mathcal{L}$" in the terms "$\mathcal{L}$-term" and "$\mathcal{L}$-formula" whenever $\mathcal{L}$ is clear from context. Adhering to standard mathematical notation, we write $t_1 \neq t_2$ instead of $\neg t_1 = t_2$.

The notion of an expression (i.e., a term or formula) *occurring* in another expression is defined the usual way. In particular, the notions of *free* and *bound occurrence* of a variable in an $\mathcal{L}$-formula are defined as usual. A variable x in a formula is called *free* if x has a free occurrence in φ. An expression is called *closed* if no variable has a free occurrence in it. A closed $\mathcal{L}$-formula φ is also called $\mathcal{L}$-*sentence*. If a formula φ is of the form $\mathsf{Q}x_1\mathsf{Q}x_2 \cdots \mathsf{Q}x_n\, \psi$ ($\mathsf{Q} \in \{\forall, \exists\}$), we abbreviate the sequence $\mathsf{Q}x_1\mathsf{Q}x_2 \cdots \mathsf{Q}x_n$ of quantifiers in φ by $\mathsf{Q}x_1, x_2, \ldots, x_n$. When we denote a formula φ by $\varphi(x_1, \ldots, x_n)$, we indicate that all free variables are among $x_1, \ldots, x_n$. We may also abbreviate a list $x_1, \ldots, x_n$ of variables by $\vec{x}$. The notion of *substitution* of a term t for all free occurrences of a variable x in a formula φ is defined as usual. We indicate substitution by $\varphi(x_1/t_1, \ldots, x_n/t_n)$ and omit the $x_1, \ldots, x_n$ whenever they are clear from context.

Turning to semantics, an $\mathcal{L}$-*structure* is a pair $\mathfrak{A} = \langle M, \cdot^{\mathfrak{A}} \rangle$, where M is a non-empty set (called *universe of* $\mathfrak{A}$) and $\cdot^{\mathfrak{A}}$ is a function which assigns

(i) to every n-ary ($n \geq 0$) predicate symbol R from $\mathcal{L}$ a relation $R^{\mathfrak{A}} \subseteq M^n$;

(ii) to every n-ary ($n \geq 0$) function symbol f from $\mathcal{L}$ a total function $f^{\mathfrak{A}}\colon M^n \to M$;

(iii) to every constant symbol c from $\mathcal{L}$ an element $c^{\mathfrak{A}} \in M$.

An $\mathfrak{A}$-*valuation* is a function v which assigns every variable an element from M. For two $\mathfrak{A}$-valuations v and v' and all variables x, we define

$$v' \sim_x v \iff_{df} v'(y) = v(y) \text{ for all variables } y \neq x.$$

We define the *value* $[\![t]\!]_{\mathfrak{A},v} \in M$ *with respect to* $\mathfrak{A}$ *and* v for all $\mathcal{L}$-terms t, as well as the *truth value* $[\![\varphi]\!]_{\mathfrak{A},v} \in \{0,1\}$ *with respect to* $\mathfrak{A}$ *and* v for all $\mathcal{L}$-formulas recursively:

(i) $[\![x]\!]_{\mathfrak{A},v} = v(x)$, for every variable x;

(ii) $[\![c]\!]_{\mathfrak{A},v} = c^{\mathfrak{A}}$, for every constant symbol c from $\mathcal{L}$;

(iii) $[\![f(t_1,\dots,t_n)]\!]_{\mathfrak{A},v} = f^{\mathfrak{A}}([\![t_1]\!]_{\mathfrak{A},v},\dots,[\![t_n]\!]_{\mathfrak{A},v})$, for every n-ary function symbol f from $\mathcal{L}$;

(iv) $[\![R(t_1,\dots,t_n)]\!]_{\mathfrak{A},v} = R^{\mathfrak{A}}([\![t_1]\!]_{\mathfrak{A},v},\dots,[\![t_n]\!]_{\mathfrak{A},v})$, for every n-ary predicate symbol R from $\mathcal{L}$;

(v) $[\![\neg\varphi]\!]_{\mathfrak{A},v} = 1 - [\![\varphi]\!]_{\mathfrak{A},v}$; $\quad [\![\varphi \to \psi]\!]_{\mathfrak{A},v} = 1$, if $[\![\varphi]\!]_{\mathfrak{A},v} \leq [\![\psi]\!]_{\mathfrak{A},v}$ and 0 otherwise;

(vi) $[\![\forall x\, \varphi]\!]_{\mathfrak{A},v} = \inf\{[\![\varphi]\!]_{\mathfrak{A},v'} \mid v' \sim_x v\}$.

For a formula φ, we write $\mathfrak{A}, v \models \varphi$ whenever $[\![\varphi]\!]_{\mathfrak{A},v} = 1$. $\mathfrak{A}$ is a *model* of φ, if $\mathfrak{A}, v \models \varphi$ for all $\mathfrak{A}$-valuations v. We abbreviate this fact by $\mathfrak{A} \models \varphi$. Likewise, $\mathfrak{A}$ is a model of a set of formulas T, if $\mathfrak{A} \models \varphi$ for every $\varphi \in T$. The *theory of* $\mathfrak{A}$ is the set $\mathsf{Th}(\mathfrak{A}) := \{\varphi \mid \mathfrak{A} \models \varphi,\ \varphi \text{ is an } \mathcal{L}\text{-sentence}\}$.

An $\mathcal{L}$-*theory* (simply *theory* if $\mathcal{L}$ is clear from context) is just a set of $\mathcal{L}$-sentences. For an $\mathcal{L}$-theory T, $\mathcal{L}$ is called the *language of* T. For a formula φ, we write $T \models \varphi$ and say that T (*logically*) *entails* φ, if every model of T is also a model of φ. The notion of $\mathcal{L}$-*proof in* T is defined as usual. For a theory T, we write $T \vdash \varphi$ if φ is provable in T. In this case, we say that φ is a *theorem* of T. For a set of $\mathcal{L}$-formulas Γ, we write $T \vdash \Gamma$ if $T \vdash \gamma$ for all $\gamma \in \Gamma$. For the notion of theoremhood, we assume a set of *logical axioms* present which together with the standard *logical rules of inference* suffices that our notion of theoremhood is strongly sound and complete with respect to our notion of entailment, i.e., for every φ, $T \vdash \varphi$ iff $T \models \varphi$. Sentences from T are called (*nonlogical*) *axioms of* T. It is clear that in order to specify a theory we only need to specify the nonlogical symbols of its language plus its nonlogical axioms, i.e., the logical machinery necessary to derive all and only the sentences which are entailed by the theory are assumed to be implicitly given by the notion of proof.

Two theories are (*deductively*) *equivalent* if they have the same language and prove the same theorems. Given two theories T, S in the same language, we write $T + S$ to denote the theory $T \cup S$. For a sentence φ in the language of T, we also write $T + \varphi$ instead of $T + \{\varphi\}$. A theory T is *axiomatizable* if there is a decidable set of sentences whose theorems coincide with the theorems of T. A theory S is an *extension* of T if the language of S extends the language of T and every theorem of T is also a theorem of S. S is a *finite extension* of T if there are sentences $\varphi_1,\dots,\varphi_n$ such that $T + \{\varphi_1,\dots,\varphi_n\}$ and S are deductively equivalent.

During our discussion, we often implicitly make use of the following concepts, which formally capture common mathematical practice (cf. Shoenfield [35], Kunen [29]).

Definition 2.1.1. Let $\mathcal{L} \subseteq \mathcal{L}'$ and Γ a set of sentences over $\mathcal{L}$. If $P \in \mathcal{L}' \setminus \mathcal{L}$ is an n-ary predicate symbol, we mean by a *definition of* P *over* $\mathcal{L}$ *and* Γ a sentence

$$\forall \vec{x}\, (P(\vec{x}) \leftrightarrow \varphi(\vec{x})),$$

where $\varphi(\vec{x})$ is a formula over $\mathcal{L}$. Similarly, for an n-ary function symbol $f \in \mathcal{L}' \setminus \mathcal{L}$, we say that $\forall \vec{x}\, \varphi(\vec{x}, f(\vec{x}))$ is a *definition of* f *over* $\mathcal{L}$ *and* Γ if $\varphi(\vec{x}, y)$ is a formula over $\mathcal{L}$ and additionally $\Gamma \vdash \forall \vec{x}\, \exists! y\, \varphi(\vec{x}, y)$.

A set of sentences $\Gamma' \supseteq \Gamma$ over $\mathcal{L}'$ is an *extension by definitions of* Γ if every sentence from $\Gamma' \setminus \Gamma$ is a definition of some symbol in $\mathcal{L}' \setminus \mathcal{L}$ over $\mathcal{L}$ and Γ. $\dashv$

Theorem 2.1.2. *Let* $\Gamma' \supseteq \Gamma$ *be an extension by definitions of* Γ*, where* Γ *is a set of sentences over* $\mathcal{L}$ *and* Γ' *is a set of sentences over* $\mathcal{L}'$*, respectively.*

(i) *If* φ *is a sentence over* $\mathcal{L}$ *then* $\Gamma \vdash \varphi$ *iff* $\Gamma' \vdash \varphi$.

(ii) *If* $\varphi(\vec{x})$ *is a formula over* $\mathcal{L}'$ *then there is a* $\psi(\vec{x})$ *over* $\mathcal{L}$ *with exactly the same free variables as* $\varphi(\vec{x})$ *such that* $\Gamma' \vdash \forall \vec{x}\, (\varphi(\vec{x}) \leftrightarrow \psi(\vec{x}))$.

2.2 Formal Arithmetic

The *language of arithmetic*, $\mathcal{L}_0$, is the first-order language with (besides equality) the binary predicate symbol $\leq$, the binary function symbols $+$, $\cdot$, the unary function symbol s (*successor*), and the constant symbol 0. Henceforth, unless stated otherwise, all formulas we consider will be formulas from a language extending $\mathcal{L}_0$. For terms t_1, t_2 in the language of arithmetic, we introduce the abbreviations $t_1 < t_2 :\equiv t_1 \leq t_2 \wedge t_1 \neq t_2$, $t_1 > t_2 :\equiv \neg t_1 \leq t_2$, and $t_1 \geq t_2 :\equiv \neg t_1 < t_2$. Furthermore, we recursively define

$$\overline{0} :\equiv 0 \quad \text{and} \quad \overline{n+1} :\equiv s(\overline{n}).$$

For $n \in \omega$,[1] the term $\overline{n}$ is called *numeral* and is a natural representation of n in the language of arithmetic.

The *standard model of arithmetic* $\mathbb{N}$ has as its universe $\omega = \{0, 1, 2, \ldots\}$ and assigns the previously mentioned nonlogical symbols their usual meaning. In particular, the denotation of s is the successor function $\lambda x.x + 1$. We call the theory of the structure $\mathbb{N}$ (i.e., the set of all sentences true in the standard model of arithmetic) *true arithmetic*. In the following, by a *true sentence* we mean a sentence true in the standard model of arithmetic.

Let t be a term which has no occurrence of x and φ a formula. We introduce the abbreviations

$$\forall x \leq t\, \varphi :\equiv \forall x\, (x \leq t \rightarrow \varphi),$$
$$\exists x \leq t\, \varphi :\equiv \exists x\, (x \leq t \wedge \varphi),$$

and similarly for the symbols $<$, $>$, and $\geq$. Occurrences of quantifiers of form $\forall x \leq t$ and $\exists x \leq t$ are called *bounded*. A formula is called *bounded* if every occurrence of a quantifier in it is bounded. Obviously, a quantifier occurrence of the form $\forall x < t$ ($\exists x < t$, respectively) can be rewritten into a logically equivalent bounded occurrence of the

[1] We consider ω to be the set of natural numbers and we often regard each natural number $n \in \omega$ to be a set consisting of all and only its predecessors.

respective form. Hence, we also call such occurrences bounded. Notice that the notion of bounded formula heavily depends on the choice of our language. In our discussion, this language will always be clear from context.

Arithmetical Theories

An *arithmetical theory* (henceforth simply theory) is just a theory whose formulas are in a language extending $\mathcal{L}_0$. Most importantly, we will confine ourselves to *Peano Arithmetic* (defined below), though many results of the present text extends to much weaker theories.

Let T be a theory. We say that a formula $\varphi(x_1, \ldots, x_k)$ having exactly k variables free *defines* a relation $R \subseteq \omega^k$ *in* T, if

$$(n_1, \ldots, n_k) \in R \iff T \vdash \varphi(\overline{n}_1, \ldots, \overline{n}_k),$$
$$(n_1, \ldots, n_k) \notin R \iff T \vdash \neg\varphi(\overline{n}_1, \ldots, \overline{n}_k),$$

for all $(n_1, \ldots, n_k) \in \omega^k$. R is *definable in* T, if there exists such a formula defining R. Similarly, a function $f\colon \omega^k \to \omega$ is definable in T, if its graph $f \subseteq \omega^{k+1}$ is. We say that a function $f\colon \omega^k \to \omega$ is *represented by* $\varphi(x_1, \ldots, x_{k+1})$ *in* T, if whenever $f(n_1, \ldots, n_k) = m$ then

$$T \vdash \forall y\,(\varphi(\overline{n}_1, \ldots, \overline{n}_k, y) \leftrightarrow y = \overline{m}).$$

Such an f is *representable in* T if there is such a $\varphi(x_1, \ldots, x_{k+1})$ having exactly $k+1$ variables free. A relation (function, respectively) is *arithmetically definable* if it is definable in true arithmetic.

Remark. Note that in the case of true arithmetic, the notions of definability and representability of a function coincide.

Definition 2.2.1. We define the classes of Σ_n and Π_n-*formulas* for all $n \in \omega$ inductively:

(i) The classes of Σ_0 and Π_0-formulas are the class of all bounded $\mathcal{L}_0$-formulas (i.e., bounded formulas in the language of arithmetic). This class is commonly called the class of Δ_0-*formulas.*

(ii) The class of Σ_{n+1}-formulas are all formulas of the form $\exists x\, \varphi(x, \vec{y})$, where $\varphi(x, \vec{y})$ is a Π_n-formula.

(iii) Similarly, the class of Π_{n+1}-formulas are all formulas of the form $\forall x\, \varphi(x, \vec{y})$, where $\varphi(x, \vec{y})$ is a Σ_n-formula.

We say that a relation $R \subseteq \omega^n$ is *in* Σ_n (Π_n, respectively), if it is arithmetically definable by a Σ_n-formula (Π_n-formula, respectively). A relation is *in* Δ_n iff it is both in Σ_n and Π_n. If a relation is in Σ_n (Π_n, Δ_n, respectively), we also say that it is a Σ_n-relation (Π_n, Δ_n-relation, respectively). The same terminology is used for functions. A Σ_n-*sentence* (Π_n-, Δ_n-*sentence*, respectively) is just a Σ_n-formula (Π_n-, Δ_n-formula, respectively) with no free occurrences of variables. $\dashv$

The classes of formulas defined above form the *arithmetical hierarchy*. A formula belonging to one of the classes Γ is said to be *of arithmetical complexity* Γ.

A formula is in *prenex form*, if it is of the form

$$\mathsf{Q}_1 x_1 \mathsf{Q}_2 x_2 \cdots \mathsf{Q}_n x_n\, \varphi(x_1, x_2, \ldots, x_n, \vec{y}), \qquad \mathsf{Q}_i \in \{\forall, \exists\}\ (i = 1, 2, \ldots, n),$$

where $\varphi(x_1, x_2, \ldots, x_n, \vec{y})$ has no quantifier occurrence. It is well-known that for every formula, there exists a formula in prenex form logically equivalent to it. Since every formula is equivalent to one in prenex form, it immediately follows that every formula is logically equivalent to some Σ_n-formula and some Π_k-formula for some $n, k \geq 0$ (possibly quantifying over dummy variables).

Proposition 2.2.2. *For all* $n \geq 0$*,*

(i) Σ_n *and* Π_n*-relations are closed under unions and intersections;*

(ii) *the complement of a* Σ_n*-relation* (Π_n*-relation*) *is a* Π_n*-relation* (Σ_n*-relation*)*;* Δ_n*-relations are thus closed under complements;*

(iii) *if* $n > 0$*,* Σ_n*-relations are closed under existential projections and* Π_n*-relations are closed under universal projections, respectively.*

A theory T is *sound* if all its theorems are true in $\mathbb{N}$. Similarly, for a class of sentences Γ, we say that T is Γ-*sound* if all sentences from Γ are true whenever they are theorems of T.

The theory Q (called *minimal arithmetic*) is axiomatized by the following axioms (the free variables are supposed to be bound by universal quantifiers):

$$\begin{aligned} s(x) &\neq 0, \\ s(x) = s(y) &\rightarrow x = y, \\ x \neq 0 &\rightarrow \exists y\, x = s(y), \\ x + 0 &= x, \\ x + s(y) &= s(x + y), \\ x \cdot 0 &= 0, \\ x \cdot s(y) &= (x \cdot y) + x, \\ x \leq y &\leftrightarrow \exists z\, z + x = y. \end{aligned}$$

Peano Arithmetic (PA) is the theory obtained from Q by adding induction axioms

(Ind) $$\varphi(0, \vec{y}) \wedge \forall x\, (\varphi(x, \vec{y}) \rightarrow \varphi(s(x), \vec{y})) \rightarrow \forall x\, \varphi(x, \vec{y}),$$

for all formulas $\varphi(x, \vec{y})$.

Peano Arithmetic allows one to establish major theorems of number theory. In particular, PA allows to develop syntactical and metamathematical notions of itself.

Example 2.2.3. Let us prove for purposes of illustration that $\mathsf{PA} \vdash \forall x, y\,(x+y = y+x)$. To this end, we first prove some auxiliary results. Let $\varphi(x) \coloneqq 0 + x = x + 0$. We prove $\mathsf{PA} \vdash \forall x\, \varphi(x)$. An instance of the induction axiom is then

$$\varphi(0) \wedge \forall x\,(\varphi(x) \to \varphi(s(x))) \to \forall x\, \varphi(x).$$

Hence, it suffices to prove $\varphi(0)$ and $\forall x\,(\varphi(x) \to \varphi(s(x)))$. The statement $0 + 0 = 0 + 0$ is valid by pure logic, hence provable in PA. We reason in PA and suppose $0 + x = x + 0$. We infer

$$0 + s(x) = s(0 + x) = s(x + 0) = s(x) = s(x) + 0,$$

which proves the first claim. We now show $\mathsf{PA} \vdash \forall x, y\; s(x) + y = x + s(y)$ by induction on y in PA (i.e., taking $s(x) + y = x + s(y)$ as the induction formula). We reason in PA as follows. If $y = 0$, then $s(x) + 0 = 0 + s(x) = s(0 + x) = s(x + 0) = x + s(0)$. Suppose now $s(x) + y = x + s(y)$. We have

$$s(x) + s(y) = s(s(x) + y) = s(x + s(y)) = x + s(s(y)),$$

which proves the second auxiliary claim. Now $\forall x, y\; x + y = y + x$ is proved by induction on x, taking $x + y = y + x$ as induction formula. The case $x = 0$ has already been established, so suppose $x + y = y + x$; we prove $s(x) + y = y + s(x)$. Observe

$$s(x) + y = x + s(y) = s(x + y) = s(y + x) = y + s(x),$$

so we are finished. $\dashv$

Definition 2.2.4. Let T be a theory and $\varphi(\vec{x})$ be a formula having exactly the variables among $\vec{x}$ free. We say that $\varphi(\vec{x})$ is Σ_n *in* T if there is a Σ_n-formula $\psi(\vec{x})$ with the same free variables as $\varphi(\vec{x})$ such that $T \vdash \forall \vec{x}\,(\varphi(\vec{x}) \leftrightarrow \psi(\vec{x}))$ (and similarly for the the classes Π_n). We say that $\varphi(\vec{x})$ is Δ_n *in* T if it is both Σ_n and Π_n in T. $\dashv$

When T is clear from context, we often omit the phrase "in T" and identify the corresponding classes of formulas modulo provable equivalence in T. In this context when we talk about Σ_n-formulas (Π_n-formula, respectively), we mean the class of formulas whose every formula is provably equivalent to a Σ_n-formula (Π_n-formula, respectively). By a Δ_n*-formula* we then mean a formula which is Δ_n in T. We have the following well-known closure properties [21, 11].

Proposition 2.2.5. *Let T be an extension of* PA.

(i) *For $n \geq 0$, the class of formulas which are Σ_n (Π_n, respectively) in T is closed under conjunction, disjunction, bounded universal, and bounded existential projection.*

(ii) *For $n \geq 1$, the class of formulas which are Σ_n in T is closed under existential projection and the class of formulas which are Π_n in T is closed under universal projection, respectively.*

We now recall some basic facts about the concepts we have introduced so far [12, 21].

Theorem 2.2.6. *A Σ_1-sentence φ is true iff $\mathsf{Q} \vdash \varphi$.*

Note that it follows from the first incompleteness theorems (Theorems 2.4.4 and 2.4.5) that no consistent axiomatizable extension of Q proves all true Π_1-sentences (see Boolos et al. [12] as well as Hájek and Pudlák [21]).

Theorem 2.2.7. *A function is recursive iff it is in Δ_1. Furthermore, a set is recursively enumerable iff it is in Σ_1.*

Theorem 2.2.8. *Every function in Δ_1 is representable in Q.*

Proof. Let f be a function (say, one-place) in Δ_1 which is defined by the Σ_1-formula $\varphi(x, y)$. Let

$$\psi(x, y) :\equiv \varphi(x, y) \wedge \forall z < y\, \neg\varphi(x, z).$$

We prove that $\psi(x, y)$ represents f in Q. Assume $f(n) = m$. It is sufficient to prove that $\mathsf{Q} \vdash \psi(\overline{n}, \overline{m})$ and $\mathsf{Q} \vdash y \neq \overline{m} \to \neg\psi(\overline{n}, y)$. The first claim is easily seen as $\mathsf{Q} \vdash \varphi(\overline{n}, \overline{m})$ and also $\mathsf{Q} \vdash \forall z < \overline{m}\, \neg\varphi(\overline{n}, z)$ (since the corresponding sentences are true). It remains to show

$$\mathsf{Q} \vdash y \neq \overline{m} \to \neg\psi(\overline{n}, y),$$

that is,

$$\mathsf{Q} \vdash y \neq \overline{m} \to \neg\varphi(\overline{n}, y) \vee \exists z < y\, \varphi(\overline{n}, z).$$

We reason in Q as follows. Suppose that $y \neq \overline{m}$. It is well-known that, provably in Q, either $y = \overline{m}$, $y > \overline{m}$, or $y < \overline{m}$. Therefore, either $y > \overline{m}$ or $y < \overline{m}$. Suppose first that $y < \overline{m}$. As we argued before, we have $\forall z < \overline{m}\, \neg\varphi(\overline{n}, z)$. Therefore also $\neg\varphi(\overline{n}, y)$. Suppose now $y > \overline{m}$. We have $\varphi(\overline{n}, \overline{m})$, therefore $\varphi(\overline{n}, z)$ for some $z < y$ as desired. This completes the proof. ☐

Remark. When we reason inside an arithmetical theory, we often ease notation and use conventional mathematical notation instead if no confusion arises.

The following concepts will be useful (see Hájek and Pudlák [21]).

Definition 2.2.9. We say that a formula $\varphi(\vec{x}, y)$ *defines a total function in T*, if $T \vdash \forall\vec{x}\, \exists! y\, \varphi(\vec{x}, y)$. Suppose that $\varphi(\vec{x}, y)$ is a formula which arithmetically defines a function $f \colon \omega^k \to \omega$. Then f is called *provably total in T*, if $\varphi(\vec{x}, y)$ defines a total function in T. Likewise, f is a *T-provably total Σ_n-function*, if $\varphi(\vec{x}, y)$ is Σ_n in T and provably total in T. ⊣

The following fact is well-known (cf. Hájek and Pudlák [21]):

Theorem 2.2.10. *Every primitive recursive function is a PA-provably total Σ_1-function.*

It follows that we can safely introduce a function symbol for every primitive recursive function into the language of arithmetic (cf. Definition 2.1.1 and Theorem 2.1.2). It will be convenient to do so in the following.

2.3 Arithmetization of Metamathematics

Arithmetical theories formalize certain portions of elementary number theory, i.e., sentences in the language of arithmetic are intended to express certain properties about numbers. In this section we are interested in the development of the syntax of PA within PA itself.[2] It was a fundamental insight of Gödel [19] that this is possible and that the machinery provided by the arithmetization of metamathematics allows one to construct true statements which are undecidable in PA.

The way we interpret statements which are proved by PA is determined by truth in the standard model $\mathbb{N}$, i.e., this model defines the standard meaning of the statements proved by PA. Hence, given our standard model $\mathbb{N}$, all objects PA can talk about are numbers—it is the *intended range* of the quantified variables which determines the meaning of the sentences inferred in PA. It should thus be clear that developing the syntax of PA in PA will be different from the development of elementary number theory in PA (cf. Boolos [11]). This issue is addressed by the notion of *Gödel numbering*: objects of syntax are assigned natural numbers in such a way that certain statements of our informal metatheory can be expressed in the language of arithmetic. Furthermore, the goal of this assignment is that certain true statements concerning the syntax of PA become provable in PA itself. Hence, syntax is crafted into a branch of number theory.

We do not develop a particular Gödel numbering here and prove that this assignment of numbers to expressions has all the desirable properties a decent Gödel numbering has. We refer the reader interested in such an elaboration to (among many sources) Boolos [11] and Hájek and Pudlák [21]. We assume a standard global assignment $\ulcorner \cdot \urcorner$ of expressions (terms, formulas, etc.) to natural numbers. Given any expression τ, we call $\ulcorner \tau \urcorner$ the *code* or *Gödel number of* τ. Note that $\ulcorner \tau \urcorner$, being a natural number, "lives" in our informal metatheory and has a natural representation in $\mathcal{L}_0$ through the term $\overline{\ulcorner \tau \urcorner}$. However, when presenting formulas in the arithmetical language, we usually write $\ulcorner \tau \urcorner$ instead of $\overline{\ulcorner \tau \urcorner}$.

Such a coding is assumed to allow us to arithmetically define many elementary notions of our metatheory. Most importantly, our Gödel numbering is assumed to allow for a one-to-one encoding of finite sequences of natural numbers. Among them are the following [3, 11, 21, 16]:

- $\mathsf{Seq}(x)$: "x is the code of a sequence";
- $\mathsf{Formula}(x)$: "x is the code of a formula";
- $\mathsf{LogAx}(x)$: "x is the code of a logical axiom";

[2] One does not need the entire strength of PA to develop certain metamathematical notions. However, we will restrict our discussion to PA. Interested readers may consult Hájek and Pudlák [21] for an extensive treatment of the development of metamathematics in first-order arithmetic.

- $\mathsf{MP}(x, y, z)$: "z follows from an application of modus ponens from x and y";
- $\mathsf{Gen}(x, y)$: "y follows from an application of generalization from x";
- $\mathsf{Ax}_{\mathsf{PA}}(x)$: "$x$ is the code of a nonlogical axiom of PA".

Many basic properties of these notions are then also verifiable in (extensions of) PA [11, 21, 16]. Furthermore, the functions and predicates assumed for manipulating sequences can be defined to be primitive recursive and we introduce function and predicate symbols defining them. With this machinery at hand we can already define the most important notion of our metatheory—the notion of proof:

$$\begin{aligned}\mathsf{Prf}_{\mathsf{PA}}(y, x) :=\ & \mathsf{Seq}(y) \wedge \mathsf{lh}(y) > 0 \wedge \mathsf{end}(y) = x \wedge \\ & \forall i < \mathsf{lh}(y)(\\ & \quad \mathsf{LogAx}(y_i) \vee \\ & \quad \mathsf{Ax}_{\mathsf{PA}}(y_i) \vee \\ & \quad \exists j, k < i\, \mathsf{MP}(y_j, y_k, y_i) \vee \\ & \quad \exists j < i\, \mathsf{Gen}(y_j, y_i)).\end{aligned}$$

(Here, end is a definition of the primitive recursive function *end* which assigns to every finite sequence its last element. Furthermore, y_i denotes the i-th element of the sequence coded by y.) The predicate $\mathsf{Prf}_{\mathsf{PA}}$ arithmetically defines the set of (codes of) provable theorems of PA. Its definition formalizes the (informal) definition of Hilbert's notion of proof.

We are now able to formally express provability in PA:

$$\mathsf{Prv}_{\mathsf{PA}}(x) := \exists y\, \mathsf{Prf}_{\mathsf{PA}}(y, x).$$

$\mathsf{Prf}_{\mathsf{PA}}$ can be defined to be Δ_1 in PA. In particular, in the wording of Gödel, $\mathsf{Prf}_{\mathsf{PA}}$ is *entscheidungsdefinit*, i.e.,

$$\begin{aligned}\mathbb{N} \models \mathsf{Prf}_{\mathsf{PA}}(\overline{n}, \overline{m}) &\implies \mathsf{PA} \vdash \mathsf{Prf}_{\mathsf{PA}}(\overline{n}, \overline{m}), \\ \mathbb{N} \models \neg\mathsf{Prf}_{\mathsf{PA}}(\overline{n}, \overline{m}) &\implies \mathsf{PA} \vdash \neg\mathsf{Prf}_{\mathsf{PA}}(\overline{n}, \overline{m}).\end{aligned}$$

$\mathsf{Prv}_{\mathsf{PA}}$ is then Σ_1 in PA. Let us for convenience introduce some additional notation as also done in Beklemishev [3]. If T is an axiomatizable extension of PA, we introduce formulas $\mathsf{Prf}_T(y, x)$ and $\mathsf{Prv}_T(x)$ as above, where (for constructing these) we assume an appropriate bounded formula $\mathsf{Ax}_T(x)$ which arithmetically defines the axioms of T. We abbreviate $\mathsf{Prv}_T(x)$ by $\Box_T(x)$ and often write $\Box_T\varphi$ instead of $\Box_T(\ulcorner\varphi\urcorner)$ if no confusion arises. The formula $\Box_T(x)$ is called (*standard*) *provability predicate for* T and arithmetically defines the set of Gödel numbers of all provable theorems of T.

2.4 Limitative Results

Let $sub_{\vec{x}}(a, b_1, \ldots, b_n)$ (where $\vec{x} = x_1, \ldots, x_n$) be the primitive recursive function whose value at $a, b_1, \ldots, b_n$ is the Gödel number of the result of respectively substituting the numerals $\overline{b}_1, \ldots, \overline{b}_n$ for the variables $x_1, \ldots, x_n$ in the formula with Gödel number a (see also Boolos [11]). Let $\mathsf{sub}_{\vec{x}}$ be a function symbol for a definition of that function. Following Beklemishev [3] and Smoryński [36, 38], for any formula $\varphi(x_1, \ldots, x_n)$, we abbreviate by $\ulcorner\varphi(\dot{y}_1, \ldots, \dot{y}_n)\urcorner$ the term

$$\mathsf{sub}_{\vec{x}}(\ulcorner\varphi(x_1, \ldots, x_n)\urcorner, y_1, \ldots, y_n).$$

Given a provability predicate $\Box_T(x)$ of T, we also write $\Box_T\varphi(\dot{x}_1, \ldots, \dot{x}_n)$ instead of $\Box_T(\ulcorner\varphi(\dot{x}_1, \ldots, \dot{x}_n)\urcorner)$. Furthermore, we often consider primitive recursive families of formulas φ_n which depend on a parameter $n \in \omega$. In this case, $\ulcorner\varphi_x\urcorner$ denotes a term for the primitive recursive function whose value at any given $n \in \omega$ is the Gödel number of φ_n. It will be convenient to introduce additional notational conventions for these cases. We assume that our first-order language contains variables $\alpha, \beta, \ldots$ of a second sort whose values range over the codes of formulas. Every formula which contains occurrences of such variables can be naturally translated into a formula in the original one-sorted first-order language of the corresponding theory. We also make use of variables $\alpha, \beta, \ldots$ when they are not necessary from a formal point of view, but increase readability.

The following statements are central for the derivation of limitative results [3, 11].

Lemma 2.4.1 (Generalized diagonal lemma). *For any formula $\varphi(x, x_1, \ldots, x_n)$ there exists a formula $\psi(x_1, \ldots, x_n)$, having exactly the variables of φ except x free, such that*

$$\mathsf{PA} \vdash \psi(x_1, \ldots, x_n) \leftrightarrow \varphi(\ulcorner\psi(x_1, \ldots, x_n)\urcorner, x_1, \ldots, x_n).$$

Proof. Let k be the Gödel number of

$$\varphi(\mathsf{sub}_x(x, x), x_1, \ldots, x_n)$$

and $\psi(x_1, \ldots, x_n)$ the formula

$$\varphi(\mathsf{sub}_x(\overline{k}, \overline{k}), x_1, \ldots, x_n).$$

It suffices to show that

$$\mathsf{PA} \vdash \mathsf{sub}_x(\overline{k}, \overline{k}) = \ulcorner\psi(x_1, \ldots, x_n)\urcorner.$$

The formula $\varphi(\mathsf{sub}_x(x, x), x_1, \ldots, x_n)$ has Gödel number k. Hence,

$$sub_x(k, k) = \ulcorner\varphi(\mathsf{sub}_x(\overline{k}, \overline{k}), x_1, \ldots, x_n)\urcorner = \ulcorner\psi(x_1, \ldots, x_n)\urcorner.$$

Therefore, $\mathsf{sub}_x(\overline{k}, \overline{k}) = \ulcorner\psi(x_1, \ldots, x_n)\urcorner$ is true and hence provable in PA. □

Corollary 2.4.2 (Diagonal lemma). *For any formula $\varphi(x)$ there exists a sentence ψ such that*

$$\mathsf{PA} \vdash \psi \leftrightarrow \varphi(\ulcorner\psi\urcorner).$$

Corollary 2.4.3. *For any formula $\varphi(v, x_1, \ldots, x_n)$ there is a formula $\psi(x_1, \ldots, x_n)$ such that*

$$\mathsf{PA} \vdash \psi(x_1, \ldots, x_n) \leftrightarrow \varphi(\ulcorner \psi(\dot{x}_1, \ldots, \dot{x}_n)\urcorner).$$

Proof. Apply Lemma 2.4.1 on $\varphi(v/\mathsf{sub}_{\vec{x}}(x, x_1, \ldots, x_n))$ (for $\vec{x} = x_1, \ldots, x_n$). □

Let T be an extension of PA. We say that φ is a *Gödel sentence for* T if $T \vdash \varphi \leftrightarrow \neg\Box_T\varphi$. By the previous results, it is clear that every such T has a Gödel sentence. Call T *ω-consistent* if there is no formula $\varphi(x)$ such that $T \vdash \exists x\, \varphi(x)$ but $T \vdash \neg\varphi(\overline{n})$ for every $n \in \omega$. T is *ω-inconsistent* if it is not ω-consistent. A little thought shows that every ω-consistent theory is also consistent. The converse is not true as we will see now. Call a sentence φ *undecidable in* T if neither $T \vdash \varphi$ nor $T \vdash \neg\varphi$. We are now able to derive Gödel's first incompleteness theorem. Let us briefly reconsider Gödel's results [19] which are among the most celebrated ones in mathematical logic.

Theorem 2.4.4 (Gödel's first incompleteness theorem). *Let T be an ω-consistent axiomatizable extension of* PA*. Then a Gödel sentence for T is undecidable in T.*

Proof. Let φ_G be such that $T \vdash \varphi_G \leftrightarrow \neg\Box_T\varphi_G$. Suppose $T \vdash \varphi_G$. Then $\mathsf{Prf}_T(\overline{n}, \ulcorner\varphi_G\urcorner)$ is true for some $n \in \omega$ and so is $\Box_T\varphi_G$, whence $T \vdash \Box_T\varphi_G$ follows. But then $T \vdash \neg\varphi_G$ and T is inconsistent, a contradiction. Suppose now that $T \vdash \neg\varphi_G$. Then $T \vdash \Box_T\varphi_G$. But $\neg\mathsf{Prf}_T(\overline{n}, \ulcorner\varphi_G\urcorner)$ is true for every $n \in \omega$ and so provable in T. Hence, T is ω-inconsistent, a contradiction. □

Let us also briefly recite the strengthened version obtained by Rosser in 1936 [33].

Theorem 2.4.5. *Let T be a consistent axiomatizable extension of* PA*. Then there is a sentence which is undecidable in T.*

Proof. Let φ_R be a sentence such that

$$T \vdash \varphi_R \leftrightarrow \exists y\, (\mathsf{Prf}_T(y, \ulcorner\neg\varphi_R\urcorner) \wedge \forall z < y\, \neg\mathsf{Prf}_T(z, \ulcorner\varphi_R\urcorner)).$$

We show that φ_R is undecidable in T. Suppose that $T \vdash \varphi_R$. Then for some $m \in \omega$, $\mathsf{Prf}_T(\overline{m}, \ulcorner\varphi_R\urcorner)$ is true and hence provable in T. Since $T \vdash \varphi_R$, we also have

$$T \vdash \exists y\, (\mathsf{Prf}_T(y, \ulcorner\neg\varphi_R\urcorner) \wedge \forall z < y\, \neg\mathsf{Prf}_T(z, \ulcorner\varphi_R\urcorner)).$$

Consider such a y and reason in T. We either have $y > \overline{m}$, $y < \overline{m}$, or $y = \overline{m}$. Clearly, $y > \overline{m}$ is impossible, therefore $y \leq \overline{m}$. So

$$T \vdash \exists y \leq \overline{m}\, \mathsf{Prf}_T(y, \ulcorner\neg\varphi_R\urcorner).$$

Since $T \nvdash \neg\varphi_R$, the formula $\forall y \leq \overline{m}\, \neg\mathsf{Prf}_T(y, \ulcorner\neg\varphi_R\urcorner)$ is true and so provable in T. Therefore, we arrive at contradiction to the consistency of T. Suppose now that $T \vdash \neg\varphi_R$. Then, for some $m \in \omega$ the formula

$$\mathsf{Prf}_T(\overline{m}, \ulcorner\neg\varphi_R\urcorner) \wedge \forall z < \overline{m}\, \neg\mathsf{Prf}_T(z, \ulcorner\varphi_R\urcorner)$$

is true and so provable in T. But then also $T \vdash \varphi_R$, a contradiction to the consistency of T. $\square$

Note that the assumption of ω-consistency is only needed for proving that T does not prove the negation of the Gödel sentence of interest. The theory $T + \neg\varphi$ for a φ such that $T \vdash \varphi \leftrightarrow \Box_T \varphi$ is an example of a consistent but ω-inconsistent theory. Observe that in proving the previous theorem we used the fact that

$$T \vdash \varphi \implies T \vdash \Box_T \varphi.$$

This constitutes one of the well-known *Löb's conditions*[3] [11, 21, 38, 3].

Theorem 2.4.6 (Löb's conditions). *Let T be an axiomatizable extension of* PA*. For all sentences φ, ψ,*

(L1) *if $T \vdash \varphi$ then $\mathsf{PA} \vdash \Box_T \varphi$;*

(L2) $\mathsf{PA} \vdash \Box_T(\varphi \to \psi) \to (\Box_T \varphi \to \Box_T \psi)$*;*

(L3) $\mathsf{PA} \vdash \Box_T \varphi \to \Box_T \Box_T \varphi$*.*

Furthermore, for any $\varphi(x), \psi(x)$,

(L4) *if $T \vdash \varphi(x)$ then $\mathsf{PA} \vdash \Box_T \varphi(\dot{x})$;*

(L5) $\mathsf{PA} \vdash \Box_T(\varphi(\dot{x}) \to \psi(\dot{x})) \to (\Box_T \varphi(\dot{x}) \to \Box_T \psi(\dot{x}))$*;*

(L6) $\mathsf{PA} \vdash \Box_T \varphi(\dot{x}) \to \Box_T \Box_T \varphi(\dot{x})$*.*

Notice that by $T \vdash \forall x\, \varphi(x) \to \varphi(x)$ we have $\mathsf{PA} \vdash \Box_T(\forall x\, \varphi(x) \to \varphi(\dot{x}))$, whence $\mathsf{PA} \vdash \Box_T \forall x\, \varphi(x) \to \Box_T \varphi(\dot{x})$ follows. Hence,

$$\mathsf{PA} \vdash \Box_T \forall x\, \varphi(x) \to \forall x\, \Box_T \varphi(\dot{x}).$$

Items (L3) and (L6) are special instances of the more general principle called *provable Σ_1-completeness* [21, 3, 38]:

Proposition 2.4.7. *Let T be an axiomatizable extension of* PA*.*

(i) *For every Σ_1-sentence φ,* $\mathsf{PA} \vdash \varphi \to \Box_T \varphi$*.*

(ii) *For every Σ_1-formula $\varphi(x_1, \dots, x_n)$ having exactly the variables $x_1, \dots, x_n$ free,* $\mathsf{PA} \vdash \varphi(x_1, \dots, x_n) \to \Box_T \varphi(\dot{x}_1, \dots, \dot{x}_n)$*.*

Let T be an axiomatizable extension of PA with the formula $\mathsf{Ax}_T(\alpha)$ arithmetically defining the axioms of T. For any finite extension of T of the form $T + \varphi$, we assume that $\mathsf{Ax}_{T+\varphi}(\alpha)$ is naturally given by

$$\mathsf{Ax}_T(\alpha) \lor \alpha = \ulcorner \varphi \urcorner.$$

[3] Sufficient properties of $\Box_T$ to derive the second incompleteness theorem were first offered by Hilbert and Bernays (see Hilbert and Bernays [23]). However, the present formulation is due to Löb [30] and is more convenient for many purposes (cf. also Smoryński [38] for a discussion).

We then have the following formalization of the deduction theorem (see Feferman [16]).

Proposition 2.4.8. *Let T be an axiomatizable extension of* PA *and φ be a sentence. Then for all ψ,*

$$\mathsf{PA} \vdash \Box_{T+\varphi}\,\psi \leftrightarrow \Box_T(\varphi \to \psi).$$

Similarly, for all $\psi(x_1, \ldots, x_n)$,

$$\mathsf{PA} \vdash \Box_{T+\varphi}\,\psi(\dot{x}_1, \ldots, \dot{x}_n) \leftrightarrow \Box_T(\varphi \to \psi(\dot{x}_1, \ldots, \dot{x}_n)).$$

Now let $\bot$ abbreviate a statement in the language of arithmetic which is contradictory (e.g., $0 \neq 0$). We abbreviate by $\mathsf{Con}(T)$ (called *consistency assertion for T*) the sentence $\neg\Box_T\bot$. Löb's conditions permit us to easily prove Gödel's second incompleteness theorem [19].

Theorem 2.4.9 (Gödel's second incompleteness theorem). *Let T be an axiomatizable extension of* PA*. Then, $T \vdash \mathsf{Con}(T)$ iff T is inconsistent.*

Proof. If T is inconsistent then certainly $T \vdash \mathsf{Con}(T)$. We show the other direction by proving that $T \vdash \mathsf{Con}(T) \to \varphi_G$, where φ_G is a Gödel sentence for T. So let φ_G be such that $T \vdash \varphi_G \leftrightarrow \neg\Box_T\varphi_G$. Then $T \vdash \varphi_G \to (\Box_T\varphi_G \to \bot)$ and so $T \vdash \Box_T\varphi_G \to \Box_T(\Box_T\varphi_G \to \bot)$. We know that

$$T \vdash \Box_T(\Box_T\varphi_G \to \bot) \to (\Box_T\Box_T\varphi_G \to \Box_T\bot)$$

and so $T \vdash \Box_T\varphi_G \to (\Box_T\Box_T\varphi_G \to \Box_T\bot)$. Since $T \vdash \Box_T\varphi_G \to \Box_T\Box_T\varphi_G$, we obtain $T \vdash \Box_T\varphi_G \to \Box_T\bot$ and thus, using its contrapositive form, $T \vdash \mathsf{Con}(T) \to \varphi_G$ as desired. $\Box$

The incompleteness theorems are proved considering fixed points of $\neg\Box_T(\alpha)$. These fixed points which can be understood to express a sentences own unprovability turn out to be undecidable. In 1952, Henkin [22] asked the question whether fixed points of $\Box_T(\alpha)$ are always provable. A well-known theorem of Löb [30] answers this question. The following theorem is a formalized version of Löb's theorem.

Theorem 2.4.10. *Let T be an axiomatizable extension of* PA*. For any sentence φ,*

$$\mathsf{PA} \vdash \Box_T(\Box_T\varphi \to \varphi) \to \Box_T\varphi.$$

Proof. We follow the proof of Smoryński [37]. Let ψ be such that

$$\mathsf{PA} \vdash \psi \leftrightarrow \Box_T(\psi \to \varphi), \tag{2.1}$$

which exists by the diagonal lemma (Corollary 2.4.2). It follows that

$$\mathsf{PA} \vdash \Box_T\psi \leftrightarrow \Box_T\Box_T(\psi \to \varphi),$$

whence by $\mathsf{PA} \vdash \Box_T(\psi \to \varphi) \to \Box_T\Box_T(\psi \to \varphi)$ we obtain

$$\mathsf{PA} \vdash \psi \to \Box_T\psi.$$

Therefore also $\mathsf{PA} \vdash \psi \to \Box_T\varphi$. The tautology $\varphi \to (\psi \to \varphi)$ allows us to infer

$$\mathsf{PA} \vdash \Box_T\varphi \to \Box_T(\psi \to \varphi),$$

whence it follows that

$$\mathsf{PA} \vdash \Box_T\varphi \to \psi.$$

Hence, $\mathsf{PA} \vdash \psi \leftrightarrow \Box_T\varphi$. Note that substitution of provably equivalent sentences in the scope of $\Box_T$ is legitimate by (L1) and (L2) (a proof of this is similar to the proof of Proposition 3.2.4). Thus we obtain

$$\mathsf{PA} \vdash \Box_T(\Box_T\varphi \to \varphi) \leftrightarrow \Box_T\varphi,$$

by performing a substitution of $\Box_T\varphi$ for ψ in (2.1). □

From that we easily obtain Löb's theorem:

Corollary 2.4.11 (Löb's theorem). *Let T be an axiomatizable extension of* PA. *For any sentence φ, $T \vdash \Box_T\varphi \to \varphi$ iff $T \vdash \varphi$.*

Proof. The direction from right to left is immediate. For the other direction, suppose $T \vdash \Box_T\varphi \to \varphi$. Then $T \vdash \Box_T(\Box_T\varphi \to \varphi)$. We invoke Theorem 2.4.10 and obtain $T \vdash \Box_T(\Box_T\varphi \to \varphi) \to \Box_T\varphi$ and so $T \vdash \Box_T\varphi$. Hence, $T \vdash \varphi$ as required. □

Therefore, Löb's theorem settles Henkin's question. As a concluding remark of this section, note that (as often remarked by Kreisel) the second incompleteness theorem easily follows from Löb's theorem when we take $\bot$ for φ (cf. Smoryński [37]). Conversely, we may prove Löb's theorem using the second incompleteness theorem as follows [36]. Suppose that $T \nvdash \varphi$. Then $T + \neg\varphi$ is consistent. By the second incompleteness theorem, we have that $T + \neg\varphi \nvdash \mathsf{Con}(T + \neg\varphi)$. Therefore, $T + \neg\varphi \nvdash \neg\Box_T(\neg\varphi \to \bot)$ and so $T + \neg\varphi \nvdash \neg\Box_T\varphi$, whence $T \nvdash \Box_T\varphi \to \varphi$ follows.

2.5 Reflection Principles

Given a theory T, the *reflection principles over T* are certain schemata of formulas expressing the soundness of T [3, 28]. We have to rely on schemata since, as a well-known result of Tarski [40] shows us, there is no truth definition for T inside T [36]. Therefore, no formula in the language of arithmetic exists which asserts that everything provably in T is true. We will encounter reflection principles in Chapter 4, where modalities of a positive polymodal calculus receive an arithmetical interpretation via reflection principles. In this section, we introduce reflection principles and summarize properties about them which are relevant for us. The material contained in this section is mainly taken from the references Beklemishev [3], Kreisel and Lévy [28], and Smoryński [36], where the interested reader may also find many additional results on reflection principles.

Let T be an axiomatizable extension of PA. The two forms of reflection we are interested in are the following schemata:

- *Local reflection schema for* T, $\mathsf{Rfn}(T)$:

$$\Box_T \varphi \to \varphi,$$

  for all sentences φ.

- *Uniform reflection schema for* T, $\mathsf{RFN}(T)$:

$$\forall x_1, \dots, x_n \, (\Box_T \varphi(\dot{x}_1, \dots, \dot{x}_n) \to \varphi(x_1, \dots, x_n)),$$

  for all formulas $\varphi(x_1, \dots, x_n)$.

Note that Gödel's first incompleteness theorem (Theorem 2.4.4) tells us that if T is consistent, there is an instance of $\mathsf{Rfn}(T)$ which is not provable in T. For there is a sentence φ_G such that $T \vdash \varphi_G \leftrightarrow \neg\Box_T \varphi_G$, so in particular $T \vdash \neg\Box_T \varphi_G \to \varphi_G$, whence the provability of $\Box_T \varphi_G \to \varphi_G$ would imply $T \vdash \varphi_G$, contradicting Gödel's first incompleteness theorem. Furthermore, Löb's theorem (Corollary 2.4.11) tells us that no nontrivial instance of $\mathsf{Rfn}(T)$ is provable provided T is consistent.

Clearly, $\mathsf{PA} + \mathsf{RFN}(T) \vdash \mathsf{Rfn}(T)$ and $\mathsf{PA} + \mathsf{Rfn}(T) \vdash \mathsf{Con}(T)$. Therefore, $\mathsf{RFN}(T)$ is stronger than $\mathsf{Rfn}(T)$ and certainly, T does neither prove $\mathsf{RFN}(T)$ nor $\mathsf{Rfn}(T)$. Let Γ be a class of formulas. We denote by $\mathsf{Rfn}_\Gamma(T)$ ($\mathsf{RFN}_\Gamma(T)$, respectively) the local reflection principle for T (uniform reflection principle for T, respectively) instantiated over formulas from Γ. (The classes of formulas we are interested in are the Σ_n and Π_n classes; the corresponding reflection principle are called *partial reflection principles*.) Furthermore, we may restrict the uniform reflection schema for T to range over formulas with one free variable, since the general case is reducible to this one by a coding of finite sequences (cf. Kreisel and Lévy [28]).

Theorem 2.5.1. *Let T be an axiomatizable extension of* PA*. Over* PA*, the following are deductively equivalent:*

(i) $\mathsf{Con}(T)$*;*

(ii) $\mathsf{Rfn}_{\Pi_1}(T)$*;*

(iii) $\mathsf{RFN}_{\Pi_1}(T)$*.*

Proof. The directions from (ii) to (i) and (iii) to (ii) are clear from our previous discussion. Let $\varphi(x)$ be a Π_1-formula. We prove that $\mathsf{PA} + \mathsf{Con}(T) \vdash \forall x \, (\Box_T \varphi(\dot{x}) \to \varphi(x))$. Indeed, by provable Σ_1-completeness,

$$\mathsf{PA} \vdash \neg\varphi(x) \to \Box_T \neg\varphi(\dot{x}),$$

but also

$$\mathsf{PA} + \mathsf{Con}(T) \vdash \Box_T \neg\varphi(\dot{x}) \to \neg\Box_T\varphi(\dot{x}),$$

whence propositional logic gives us $\mathsf{PA} + \mathsf{Con}(T) \vdash \Box_T\varphi(\dot{x}) \to \varphi(x)$ as required. □

Corollary 2.5.2. *Let T be an axiomatizable extension of* PA *and suppose $T \vdash \varphi$, where φ is a Π_1-sentence. Then* $\mathsf{PA} + \mathsf{Con}(T) \vdash \varphi$.

Proof. By $T \vdash \varphi$ we have $\mathsf{PA} \vdash \Box_T\varphi$, whence $\mathsf{PA} + \mathsf{Con}(T) \vdash \Box_T\varphi \to \varphi$ gives us $\mathsf{PA} + \mathsf{Con}(T) \vdash \varphi$ as desired. □

Corollary 2.5.2 has an interesting interpretation concerning the philosophy of mathematics and Hilbert's program.[4] If we consider a part of Hilbert's program which asks for conservation: whenever a statement about *real objects* (i.e., the objects having an intuitive meaning, see Kleene [26]) is provable by means of *ideal objects* (those objects opposed to the real ones), then it should also be provable by referring to real objects only. However, by the previous corollary, we can make the following observation. Assume that T is a comparatively strong theory which contains portions of ideal mathematics, while suppose we declare PA to be a system formalizing real mathematics. If φ is a real universal statement and φ is provable in T, then the assumption of T being consistent establishes the provability of φ in real mathematics. Hence, in a certain sense, this reduces Hilbert's conservation program to the consistency program.

Theorem 2.5.3. *Let T be an axiomatizable extension of* PA*. For $n \geq 1$, the schemata* $\mathsf{RFN}_{\Sigma_n}(T)$ *and* $\mathsf{RFN}_{\Pi_{n+1}}(T)$ *are deductively equivalent over* PA*.*

Proof. Let $\forall y\, \varphi(y, x)$ be a Π_{n+1}-formula, where $\varphi(y, x)$ is a Σ_n-formula. Then,

$$\mathsf{PA} + \mathsf{RFN}_{\Sigma_n}(T) \vdash \Box_T \forall y\, \varphi(y, \dot{x}) \to \forall y\, \Box_T\varphi(\dot{y}, \dot{x}).$$

But also

$$\mathsf{PA} + \mathsf{RFN}_{\Sigma_n}(T) \vdash \forall y\, \Box_T\varphi(\dot{y}, \dot{x}) \to \forall y\, \varphi(y, x),$$

whence $\mathsf{PA} + \mathsf{RFN}_{\Sigma_n}(T) \vdash \Box_T\forall y\, \varphi(y, \dot{x}) \to \forall y\, \varphi(y, x)$ follows. □

Although we do not have a truth definition for all formulas in the language of arithmetic, we have the following result (cf. Hájek and Pudlák [21]).

Theorem 2.5.4. *For each $n \geq 0$ there is a Π_n-formula* $\mathsf{True}_{\Pi_n}(x)$ *such that for every Π_n-formula $\varphi(x_1, \ldots, x_n)$,*

$$\mathsf{PA} \vdash \varphi(x_1, \ldots, x_n) \leftrightarrow \mathsf{True}_{\Pi_n}(\ulcorner \varphi(\dot{x}_1, \ldots, \dot{x}_n) \urcorner).$$

An analogous statement holds for the classes Σ_n, for $n \geq 0$.

[4] Private communications with Matthias Baaz. See also Zach [41] for an excellent overview on Hilbert's program.

Due to these partial truth definitions, the partial uniform reflection principles are subject to finite axiomatization.

Lemma 2.5.5. *Let T be an axiomatizable extension of* PA*. For each $n \geq 0$, the schema* $\mathsf{RFN}_{\Pi_n}(T)$ *is deductively equivalent over* PA *to the instance*

$$\forall x\, (\Box_T \mathsf{True}_{\Pi_n}(\dot{x}) \to \mathsf{True}_{\Pi_n}(x)). \tag{2.2}$$

An analogous statement holds for the classes Σ_n, for $n \geq 0$.

Proof. By the previous theorem we easily infer

$$\mathsf{PA} \vdash \forall x_1, \dots, x_n\, \Box_T(\varphi(\dot{x}_1, \dots, \dot{x}_n) \leftrightarrow \mathsf{True}_{\Pi_n}(\ulcorner \varphi(\dot{x}_1, \dots, \dot{x}_n)\urcorner)).$$

Let $\varphi(x)$ be a Π_n-formula. We have

$$\mathsf{PA} \vdash \Box_T \varphi(\dot{x}) \to \Box_T \mathsf{True}_{\Pi_n}(\ulcorner \varphi(\dot{x})\urcorner),$$

whence it follows by (2.2) that

$$\mathsf{PA} \vdash \Box_T \varphi(\dot{x}) \to \mathsf{True}_{\Pi_n}(\ulcorner \varphi(\dot{x})\urcorner).$$

Hence, by Theorem 2.5.4 we conclude $\mathsf{PA} \vdash \Box_T \varphi(\dot{x}) \to \varphi(x)$. □

Corollary 2.5.6. *For $n \geq 0$, the schemata* $\mathsf{RFN}_{\Sigma_n}(T)$ *and* $\mathsf{RFN}_{\Pi_n}(T)$ *are finitely axiomatizable over* PA*.*

Let Γ be a class of formulas. We say that an extension S of T is *of complexity* Γ if there is a theory S' which is deductively equivalent to S such that all sentences from $S' \setminus T$ are from Γ.

Theorem 2.5.7. *Let T be an axiomatizable extension of* PA*.*

(i) $\mathsf{Rfn}_{\Pi_n}(T)$ *is not contained in any consistent finite extension of T of complexity Σ_n.*

(ii) $\mathsf{RFN}_{\Pi_n}(T)$ *is not contained in any consistent extension of T of complexity Σ_n.*

Dual statements respectively hold for $\mathsf{Rfn}_{\Sigma_n}(T)$ *and* $\mathsf{RFN}_{\Sigma_n}(T)$*.*

Proof. For (i), suppose that $T + \varphi \vdash \mathsf{Rfn}_{\Pi_n}(T)$ for some Σ_n-sentence φ. Then,

$$T + \varphi \vdash \Box_T \neg\varphi \to \neg\varphi,$$

whence it follows that

$$T \vdash \varphi \to (\Box_T \neg\varphi \to \neg\varphi).$$

By pure logic, we have $T \vdash \neg\varphi \to (\Box_T \neg\varphi \to \neg\varphi)$ and thus $T \vdash \Box_T \neg\varphi \to \neg\varphi$, whence the formalized version of Löb's theorem (Theorem 2.4.10) gives us $T \vdash \neg\varphi$, i.e., $T + \varphi$ is inconsistent.

For (ii), suppose that U is an extension of T of complexity Σ_n and assume that $U \vdash \mathsf{RFN}_{\Pi_n}(T)$. Since $\mathsf{RFN}_{\Pi_n}(T)$ is finitely axiomatizable over T, we have $U_0 \vdash \mathsf{RFN}_{\Pi_n}(T)$ for some finite $U_0 \subseteq U$. But then also $U_0 \vdash \mathsf{Rfn}_{\Pi_n}(T)$, whence by (i) it follows that U_0 is inconsistent and so U is inconsistent too. □

Remark. It can even be shown that $\mathsf{Rfn}_{\Pi_n}(T)$ is not contained in any consistent axiomatizable extension of T by Σ_n-sentences [3] (dually for $\mathsf{Rfn}_{\Sigma_n}(T)$).

Corollary 2.5.8. *Let T be an axiomatizable extension of* PA*.*

(i) $\mathsf{Rfn}(T)$ *is not contained in any consistent finite extension of T.*

(ii) $\mathsf{RFN}(T)$ *is not contained in any consistent extension of T of bounded arithmetical complexity.*

For more results on reflection principles, we refer the reader to Kreisel and Lévy [28], Smoryński [36], Beklemishev [3], as well as Artemov and Beklemishev [1].

Let us now turn to notions of consistency and provability which we will encounter several times in this thesis. For $n \geq 1$, let $\mathsf{Th}_{\Pi_n}(\mathbb{N})$ denote the set of all true Π_n-sentences. We say that a theory T is *n-consistent* if $T + \mathsf{Th}_{\Pi_n}(\mathbb{N})$ is consistent [3]. So T is n-consistent if there is no true Π_n-sentence whose negation is provable in T. Formally, this is expressible by

$$\mathsf{Con}_n(T) :\equiv \forall \alpha\, (\mathsf{True}_{\Pi_n}(\alpha) \to \neg \Box_T \neg \mathsf{True}_{\Pi_n}(\dot{\alpha})).$$

Note that $\mathsf{Con}_n(T)$ is a Π_{n+1}-sentence. Dually to n-consistency, we say that φ is *n-provable in T* if $T + \neg\varphi$ is not n-consistent, that is, iff φ is provable in $T + \mathsf{Th}_{\Pi_n}(\mathbb{N})$. We use the abbreviation

$$[n]_T \varphi :\equiv \neg \mathsf{Con}_n(T + \neg\varphi)$$

to formally express the notion of n-provability in T. Furthermore, we stipulate that $[0]_T$ and $\mathsf{Con}_0(T)$ correspond to $\Box_T$ and $\mathsf{Con}(T)$, respectively. The formula $[n]_T(x)$ which obeys the above definition is then a Σ_{n+1}-formula. Following the conventions of modal languages, we abbreviate $\mathsf{Con}_n(T + \varphi)$ by $\langle n \rangle_T \varphi$. If T is axiomatizable then clearly

$$\mathsf{PA} \vdash \langle n \rangle_T \varphi \leftrightarrow \neg [n]_T \neg \varphi,$$

for every sentence φ. Note that from

$$\mathsf{PA} \vdash [n]_T \varphi \leftrightarrow \exists \alpha\, (\mathsf{True}_{\Pi_n}(\alpha) \wedge \Box_{T+\neg\varphi} \neg \mathsf{True}_{\Pi_n}(\dot{\alpha}))$$

we easily obtain that

$$\mathsf{PA} \vdash [n]_T \varphi \leftrightarrow \exists \alpha\, (\mathsf{True}_{\Pi_n}(\alpha) \wedge \Box_T (\mathsf{True}_{\Pi_n}(\dot{\alpha}) \to \varphi)),$$

by an application of the formalized deduction theorem (Proposition 2.4.8).

We use the same notational conventions for $[n]_T$ as for $\Box_T$. The following conditions are natural liftings of Löb's conditions (Theorem 2.4.6) to the more general case of n-provability (cf. Beklemishev [3]).

Theorem 2.5.9. *Let T be an axiomatizable extension of* PA. *For all $n \geq 0$ and all sentences φ, ψ,*

(i) *if $T \vdash \varphi$ then $\mathsf{PA} \vdash [n]_T\varphi$;*

(ii) $\mathsf{PA} \vdash [n]_T(\varphi \to \psi) \to ([n]_T\varphi \to [n]_T\psi)$*;*

(iii) $\mathsf{PA} \vdash [n]_T\varphi \to [n]_T[n]_T\varphi$.

Similar statements hold for formulas with free variables.

Note that a consequence of these conditions is that a generalization of the formalized version of Löb's theorem is provable, i.e., for all $n \geq 0$,

$$\mathsf{PA} \vdash [n]_T([n]_T\varphi \to \varphi) \to [n]_T\varphi.$$

The last of these conditions follows from the more general property of *provable* Σ_{n+1}-*completeness.*

Proposition 2.5.10. *Let T be an axiomatizable extension of* PA. *For every Σ_{n+1}-sentence φ, $\mathsf{PA} \vdash \varphi \to [n]_T\varphi$. Furthermore, for every Σ_{n+1}-formula $\varphi(x_1, \ldots, x_k)$ which has exactly the variables $x_1, \ldots, x_k$ free, we have*

$$\mathsf{PA} \vdash \varphi(x_1, \ldots, x_k) \to [n]_T\varphi(\dot{x}_1, \ldots, \dot{x}_k).$$

Lemma 2.5.11. *For $n \geq 0$, the schema $\mathsf{RFN}_{\Pi_{n+1}}(T)$ is equivalent to $\mathsf{Con}_n(T)$ over* PA.

Proof. The case of $n = 0$ is just Theorem 2.5.1. For $n > 0$ it is sufficient to show that $\mathsf{Con}_n(T)$ is equivalent to $\mathsf{RFN}_{\Sigma_n}(T)$ over PA by virtue of Theorem 2.5.3. We note that $\neg\mathsf{True}_{\Pi_n}(x)$ is a Σ_n-formula, so

$$\mathsf{PA} + \mathsf{RFN}_{\Sigma_n}(T) \vdash \forall\alpha\,(\Box_T\neg\mathsf{True}_{\Pi_n}(\dot{\alpha}) \to \neg\mathsf{True}_{\Pi_n}(\alpha)),$$

i.e., $\mathsf{PA} + \mathsf{RFN}_{\Sigma_n}(T) \vdash \mathsf{Con}_n(T)$.

Conversely, let $\varphi(x)$ be a Σ_n-formula. We know that

$$\begin{aligned} \mathsf{PA} \vdash \varphi(x) &\leftrightarrow \mathsf{True}_{\Sigma_n}(\ulcorner\varphi(\dot{x})\urcorner) \\ &\leftrightarrow \neg\mathsf{True}_{\Pi_n}(\ulcorner\neg\varphi(\dot{x})\urcorner). \end{aligned}$$

Furthermore, $\mathsf{PA} + \mathsf{Con}_n(T) \vdash \Box_T\neg\mathsf{True}_{\Pi_n}(\ulcorner\neg\varphi(\dot{x})\urcorner) \to \neg\mathsf{True}_{\Pi_n}(\ulcorner\neg\varphi(\dot{x})\urcorner)$, whence it follows that $\mathsf{PA} + \mathsf{Con}_n(T) \vdash \Box_T\varphi(\dot{x}) \to \varphi(x)$. □

2.6 Provability Logics

According to Artemov and Beklemishev [1], the origins of provability logics may be traced back to a paper by Gödel [20], where he attempted to formalize the notion of provability for the intuitionistic propositional calculus in order to cope with the Brouwer's interpretation of intuitionistic truth as provability. Gödel's approach consists of an embedding of intuitionistic propositional logic into the modal system S4. The logic S4 is formulated over a modal language which contains a countably infinite supply of propositional variables, the usual propositional connectives (including the constant $\bot$ denoting falsity), and the modal operator $\Box$. The intended meaning of $\Box\varphi$ is then "φ is provable". The dual operator to $\Box$ is denoted by $\Diamond$ and is defined by $\Diamond := \neg\Box\neg$. S4 is axiomatized by the following axiom schemes and rules of inference:

(i) all propositional tautologies;

(ii) $\Box\varphi \to \varphi$;

(iii) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$;

(iv) $\Box\varphi \to \Box\Box\varphi$;

(v) if φ then infer $\Box\varphi$ (*necessitation*);

(vi) if $\varphi \to \psi$ and φ then infer ψ (*modus ponens*).

However, this system can be easily seen to be inadequate for the interpretation of $\Box\varphi$ as formal provability in theories extending PA. For by $\Box\bot \to \bot$ (which is equivalent to $\neg\Box\bot$), we would obtain that the theory under consideration proves its own consistency, contradicting the incompleteness theorems. Therefore, a natural question which is to be answered is the question which asks to find the modal logic which characterizes provability in theories extending PA.

A first step towards a solution of this problem was found by Löb [30] who offered sufficient conditions to prove (a formalized version of) the second incompleteness theorem. In terms of modal logic, we may formulate principles in reminiscence to Löb's conditions as follows:

(i) all propositional tautologies;

(ii) modus ponens and necessitation;

(iii) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$;

(iv) $\Box\varphi \to \Box\Box\varphi$.

Furthermore, the formalized version of Löb's theorem offers another propositional principle which can be shown to be independent of the above ones:

(v) $\Box(\Box\varphi \to \varphi) \to \Box\varphi$.

With the corresponding principles formulated in the language of arithmetic, one can prove a formalized version of the second incompleteness theorem by substituting $\bot$ for φ in (v). The logic consisting of the axioms and rules above is nowadays called GL (for Gödel and Löb). It is well-known that the axiom scheme $\Box\varphi \to \Box\Box\varphi$ is derivable from the others [11]. A landmark result of Solovay [39] reveals that GL axiomatizes the notion of provability in sufficiently strong and sound theories. In this narrower sense, the study of provability logics is concerned with the investigation of modal logics which axiomatize properties of provability predicates of an arithmetical theory T. We have a degree of freedom in this characterization in the sense that the question which theory is capable of proving these properties of T is left open. Most importantly, we might be concerned which modal properties T can prove about its *own* provability predicate(s).

Returning to the case of GL, let us state Solovay's celebrated results more precisely. Let T be an axiomatizable theory extending PA. An *arithmetical realization* is a function which assigns sentences to propositional variables. Let f be an arithmetical realization. A *T-interpretation f_T under f* is defined for all modal formulas inductively as follows:

(i) $f_T(\bot) = \bot$;

(ii) $f_T(p) = f(p)$, where p is a propositional variable;

(iii) $f_T(\varphi \to \psi) = f_T(\varphi) \to f_T(\psi)$ and $f_T(\neg\varphi) = \neg f_T(\varphi)$ i.e., $f_T(\cdot)$ commutes with the propositional connectives;

(iv) $f_T(\Box\varphi) = \Box_T f_T(\varphi)$.

We may state Solovay's first arithmetical completeness result as follows.

Theorem 2.6.1. *Let T be an axiomatizable extension of* PA *which is Σ_1-sound. Then,*

$$\mathsf{GL} \vdash \varphi \iff T \vdash f_T(\varphi), \text{ for all arithmetical realizations } f.$$

The logic GL enjoys many desirable properties like Craig interpolation, a fixed point property, the existence of a sequent calculus formulation where the cut rule is admissible, and a well-studied Kripke semantics admitting the finite model property [11, 1, 14].

Solovay also showed that the modal logic which axiomatizes the universally true principles concerning provability in PA is a decidable extension of GL. This extension, denoted by S, consists of all theorems of GL, the additional axiom scheme

(vi) $\Box\varphi \to \varphi$,

while dropping the necessitation rule, i.e., the sole rules of inference are modus ponens and substitution. This is necessary, for otherwise we could derive $\mathsf{S} \vdash \Box\bot \to \bot$ and so $\mathsf{S} \vdash \Box(\Box\bot \to \bot)$, whence $\mathsf{S} \vdash \Box\bot$ and $\mathsf{S} \vdash \bot$. Solovay's second theorem may then be stated as follows.

Theorem 2.6.2. *Let T be an axiomatizable extension of* PA *which is sound. Then,*

$$\mathsf{S} \vdash \varphi \iff \mathbb{N} \models f_T(\varphi), \text{ for all arithmetical realizations } f.$$

Japaridze's GLP

Ever since the landmark results of Solovay, researchers have sought for investigating modal principles of other forms of provability. Boolos [10] investigated the logic of *ω-provability* in PA and showed that this logic coincides with GL for sound theories extending PA.

Recall that a theory T is called ω-consistent if there is no sentence $\exists x\, \varphi(x)$ such that $T \vdash \neg\varphi(\overline{n})$ for all $n \in \omega$, but $T \vdash \exists x\, \varphi(x)$. A sentence φ is *ω-provable in T* if $T + \neg\varphi$ is ω-inconsistent.

It is well-known that the notion of ω-provability in PA coincides with the notion of being *provable in* PA *by one application of the ω-rule* [11, 10], i.e., provability in the theory

$$\mathsf{PA}' \coloneqq \mathsf{PA} + \{\forall x\, \varphi(x) \mid \forall n \in \omega \colon \mathsf{PA} \vdash \varphi(\overline{n})\}.$$

(See also Artemov and Beklemishev [1].) For suppose that φ is ω-provable in PA. Then, by the deduction theorem, there is a formula $\psi(x)$ such that $\mathsf{PA} \vdash \neg\varphi \to \neg\psi(\overline{n})$ for all $n \in \omega$, but $\mathsf{PA} \vdash \neg\varphi \to \exists x\, \psi(x)$. It follows that $\mathsf{PA} \vdash (\varphi \vee \forall x\, \neg\psi(x)) \to \varphi$ and thus $\mathsf{PA} \vdash \forall x\, (\neg\varphi \to \neg\psi(x)) \to \varphi$ and so φ is provable in PA by one application of the ω-rule. Conversely, suppose that φ is provable in PA. Then there are formulas $\psi_1(x), \dots, \psi_k(x)$ such that

$$\mathsf{PA} + \{\forall x\, \psi_1(x), \dots, \forall x\, \psi_k(x)\} \vdash \varphi,$$

and $\mathsf{PA} \vdash \psi_i(\overline{n})$ for $i = 1, \dots, k$ and all $n \in \omega$. Therefore,

$$\mathsf{PA} \vdash \forall x\, (\psi_1(x) \wedge \cdots \wedge \psi_k(x)) \to \varphi.$$

Now it is easy to see that $\mathsf{PA} + \neg\varphi$ is ω-inconsistent.

Interest arises in the modal logic which contains operators for both formalized ω-provability and standard provability in the Hilbertian sense. Let $[0]$ and $[1]$ be modal operators which are interpreted as provability and ω-provability in PA, respectively and let $[0]_\omega(\alpha)$ and $[1]_\omega(\alpha)$ be their according formalizations in the language of arithmetic (i.e., we set $[0]_\omega \coloneqq \Box_{\mathsf{PA}}$). These notions of provability can be formalized such that both modalities $[0]$ and $[1]$ satisfy the postulates of GL. Moreover, one can show that

$$\mathsf{PA} \vdash [0]_\omega\varphi \to [1]_\omega\varphi,$$

for all arithmetical sentences φ. Furthermore,

$$\mathsf{PA} \vdash \neg[0]_\omega\varphi \to [1]_\omega\neg[0]_\omega\varphi$$

can also be established for every arithmetical sentence φ. This *bimodal* logic of provability and ω-provability is thus axiomatized by the following axiom schemes and rules of inference:

(i) all propositional tautologies;

(ii) axioms of GL for $[0]$ and $[1]$;

(iii) $[0]\varphi \to [1]\varphi$;

(iv) $\langle 0\rangle\varphi \to [1]\langle 0\rangle\varphi$;

(v) modus ponens, [0]-, and [1]-necessitation.[5]

(Here, for $n = 0, 1$, $\langle n \rangle := \neg[n]\neg$ is the dual of $[n]$.) The question whether this logic is arithmetically sound and complete in PA for the interpretation in arithmetic as discussed above was answered positively by Japaridze [25]. Japaridze showed even more: he introduced modalities $[n]$ for every natural number n and assigned to $[n]$ the arithmetical interpretation

"provable under n nested applications of the ω-rule."

That is, the modalities [0], [1], [2], etc. receive the interpretation as (formalized) provability in PA, PA', PA'', and so on [1]. The resulting polymodal logic is called GLP and is axiomatized by the following axiom schemes and rules of inference:

(i) all propositional tautologies;

(ii) axioms of GL for $[n]$ $(n \geq 0)$;

(iii) $[m]\varphi \to [n]\varphi$, for $m < n$ (*monotonicity*);

(iv) $\langle m \rangle\varphi \to [n]\langle m \rangle\varphi$, for $m < n$;

(v) modus ponens and $[n]$-necessitation, for $n \geq 0$.

Formulas in the language of GLP are called *polymodal formulas*. Let $[n]_\omega$ be a formalization of provability in PA under n nested applications of the ω-rule (cf. also Section 3.3). For all polymodal formulas, define a PA-interpretation f_{PA} under an arithmetical realization f as usual, except that

$$f_{\mathsf{PA}}([n]\varphi) = [n]_\omega f_{\mathsf{PA}}(\varphi).$$

Japaridze's results then reads as follows:

Theorem 2.6.3. *Let φ be a polymodal formula. Then,*

$$\mathsf{GLP} \vdash \varphi \iff \mathsf{PA} \vdash f_{\mathsf{PA}}(\varphi), \textit{ for all arithmetical realizations } f.$$

Ignatiev [24] extended the results of Japaridze and showed that GLP is arithmetically complete with respect to a very general class of arithmetical interpretations. We have already encountered one such admissible interpretation for GLP in Section 2.6. Let T be a sound theory. Again, define a T-interpretation f_T for all polymodal formulas under an arithmetical realization f as usual, except that we stipulate

$$f_T([n]\varphi) = [n]_T f_T(\varphi).$$

That is, the modalities $[n]$ are interpreted as n-provability in T. (The broader class will be examined during our treatment of the arithmetical completeness of our many-sorted

[5] By $[n]$-*necessitation* we just mean the rule $\varphi/[n]\varphi$.

variant of GLP in Chapter 3.) It follows from Ignatiev's results that, for T being sound, GLP is also arithmetically sound and complete for T under this interpretation.

Ignatiev also obtained many other results. He showed that GLP enjoys nice properties like Craig interpolation, a fixed point property, and that the closed fragment of GLP (i.e., the class of formulas with no occurrences of propositional variables) has a universal model based on the ordinal ε_0. It is easy to show that GLP is not sound and complete for any class of Kripke frames. To cope with that, Ignatiev identified a weaker logic than GLP which is sound and complete with respect to a decent class of Kripke frames. It it then a corollary of the Ignatiev's arithmetical completeness theorem for GLP that GLP has a natural translation into that weaker logic. Beklemishev [5] also isolated a subsystem J of GLP which arises from GLP if we drop monotonicity and add the axiom schemes

(vii) $[m]\varphi \to [n][m]\varphi$, if $m \leq n$;

(viii) $[m]\varphi \to [m][n]\varphi$, if $m \leq n$.

In contrast to GLP, the logic J is sound and complete with respect to a nice class of Kripke frames and Beklemishev [6] provided a proof of the arithmetical completeness theorem for GLP which is based on the logic J and is closer to the arithmetical completeness proof for GL.

CHAPTER 3

The Logics GLP^* and J^*

In this section we introduce our logics GLP^* and J^* which are many-sorted variants of Japaridze's GLP and Beklemishev's J. Section 3.2 contains basic definitions of GLP^* and J^*. We continue in Section 3.3 with an exposition of the arithmetical interpretation of GLP^*. Section 3.4 treats Kripke semantics for J^*. In particular, we show that J^* is sound and complete with respect to a nice class of Kripke models. In Section 3.5 we show that GLP^* is arithmetically complete with respect to our arithmetical interpretation. The proof is an extension of the one provided by Beklemishev [6]. Afterwards, we discuss some extensions and corollaries of this theorem.

3.1 Motivation

As pointed out in the introduction, Beklemishev [2] proposed an approach to ordinal analysis based on the notion of *graded provability algebra*. Consider a theory T and let $\mathcal{L}_T$ be the set of sentences factorized by provable equivalence in T, i.e., by the relation defined by

$$\varphi \sim \psi \iff_{df} T \vdash \varphi \leftrightarrow \psi.$$

Let $\{\varphi\}$ denote the equivalence class of φ under this equivalence relation. We can equip the set of all equivalence classes with the usual operations $\wedge$, $\vee$, $\neg$, and the ordering

$$\{\varphi\} \leq \{\psi\} \iff_{df} T \vdash \varphi \to \psi.$$

From an algebraic point of view, this makes the structure $\mathcal{L}_T$ to a Boolean algebra (called the *Lindenbaum algebra of* T) whose minimal element $\bot$ denotes the class of refutable sentences of T, while $\top$ denotes the class of provable sentences of T [1, 3]. A Boolean algebra $\mathcal{B}$ is called *atomless* if

$$\forall x, y\, (x < y \Rightarrow \exists z \in \mathcal{B} \colon x < z < y).$$

It can be shown that if T is a consistent axiomatizable extension of a very weak fragment of PA, then $\mathcal{L}_T$ is a countable atomless Boolean algebra [3]. Furthermore, it is known that all countable atomless Boolean algebras are pairwise isomorphic. Therefore, we can draw the conclusion that the structure $\mathcal{L}_T$ is expressively too weak to gain any meaningful proof-theoretic information for T by investigating $\mathcal{L}_T$.

Let T be an axiomatizable extension of PA. For each $n \geq 0$, the formula $[n]_T$ defines an operator on the equivalence classes of $\mathcal{L}_T$:

$$[n]_T \colon \{\varphi\} \longmapsto \{[n]_T\varphi\}.$$

Note that $[n]_T$ is well-defined on the equivalence classes since $T \vdash \varphi \leftrightarrow \psi$ implies $T \vdash [n]_T\varphi \leftrightarrow [n]_T\psi$. Call the structure $\mathcal{L}_T$ enriched by $[n]_T$ the *n-provability algebra of T* and denote it by $\mathcal{M}_T^n$. The structure $\mathcal{M}_T^0 = \langle \mathcal{L}_T, \Box_T \rangle$ was first considered by Magari [31] and is therefore called *Magari algebra of T*. Terms in the language of $\mathcal{M}_T^0$ correspond to modal formulas and identities in T can be understood to be the *provability logic of T*. Note that an arbitrary algebraic identity $\psi(\vec{p}) = \chi(\vec{p})$ reduces to $\psi(\vec{p}) \leftrightarrow \chi(\vec{p}) = \top$. In algebraic terms, Solovay's first theorem reads as

$$\mathsf{GL} \vdash \varphi(\vec{p}) \iff \mathcal{M}_T^0 \models \forall \vec{p}\,(\varphi(\vec{p}) = \top),$$

for any $\varphi(\vec{p})$ and any Σ_1-sound axiomatizable extension T of PA.

In order to study proof-theoretic properties of PA, Beklemishev [2] introduces the algebra $\mathcal{M}_T^\infty = \langle \mathcal{L}_T, [0]_T, [1]_T, \ldots \rangle$. Identities of $\mathcal{M}_T^\infty$ correspond to polymodal formulas. Japaridze's result (together with the generalization by Ignatiev) establishes that

$$\mathsf{GLP} \vdash \varphi(\vec{p}) \iff \mathcal{M}_T^\infty \models \forall \vec{p}\,(\varphi(\vec{p}) = \top),$$

for all $\varphi(\vec{p})$ and any sound axiomatizable extension T of PA.

$\mathcal{M}_T^\infty$ provides a rather abstract view on T and its extension. For example, consider the theory EA[1] which contains a function symbol for exponentiation and, apart from that, differs from PA in the fact that the induction axiom is restricted to bounded formulas. By a theorem of Kreisel and Lévy [28], PA is embeddable into $\mathcal{M}_{\mathsf{EA}}^\infty$ as a filter generated by $\{\langle n \rangle_{\mathsf{EA}} \top \mid n < \omega\}$ [2].

$\mathcal{M}_T^\infty$ implicitly contains an additional structure, namely it can be divided into subsets $P_0 \subset P_1 \subset \cdots \subseteq \mathcal{M}_T^\infty$, which respectively correspond to $\Pi_1, \Pi_2, \ldots$ sentences, i.e., sentences are classified according to their natural quantifier complexity. We know that $\bigcup_{i \geq 0} P_i = \mathcal{M}_T^\infty$. Beklemishev calls this family of subsets *stratification of* $\mathcal{M}_T^\infty$. The algebra $\mathcal{M}_T^\infty$ together with its stratification gives rise to a many-sorted algebra (called *graded provability algebra of* T) which is formulated over a language with sorted variables p_i^n, where the index n indicates that the variable p_i^n ranges over elements from P_n, that is, Π_{n+1}-sentences [1]. It is the logic induced by this very many-sorted algebra we will study in the sequel by modal logical means of investigation. For more details on provability algebras, we refer the reader to Artemov and Beklemishev [1], Beklemishev [3, 2], and their many references.

[1]For *Elementary Arithmetic*, see Beklemishev [3].

3.2 Basics

From now on, we assume that every propositional variable p is assigned a *sort* α such that $0 \leq \alpha \leq \omega$. We use $p, q, \ldots$ as metavariables which range over propositional variables. A *signature* is a set $\Lambda \subseteq \omega + 1$.

Definition 3.2.1. Let Φ be a set of propositional variables and Λ a signature. We define $L^*(\Phi, \Lambda)$, the (*many-sorted*) *formulas* (*over* Φ *and* Λ) and their corresponding *sorts*, inductively as follows:

(i) $\bot$ and $\top$ are formulas of sort 0.

(ii) If $p \in \Phi$ is a propositional variable of sort α, then p is a formula of sort α.

(iii) If φ and ψ are formulas of sorts α and β, then $(\varphi \vee \psi)$ and $(\varphi \wedge \psi)$ are formulas of sort $\max\{\alpha, \beta\}$.

(iv) If φ is a formula of sort α, then $\neg\varphi$ is a formula of sort $\alpha + 1$.

(v) If φ is a formula (of any sort) and $\alpha \in \Lambda$, then $\langle\alpha\rangle\varphi$ is a formula of sort α.

We denote the sort of any formula φ by $|\varphi|$. ⊣

Furthermore, it will be notationally convenient for us to write p^α to designate that p is a variable of sort α. We call $\langle\alpha\rangle$, for $\alpha \in \Lambda$, *modal operator* or *modality.* Overloading this notion, we will sometimes also call α in $\langle\alpha\rangle$ modal operator or modality. We define $\mathbb{V} \coloneqq \bigcup_{\alpha \leq \omega}\{Var_\alpha\}$, where for $\alpha \leq \omega$ we set $Var_\alpha \coloneqq \{p_0^\alpha, p_1^\alpha, \ldots\}$, i.e., $\mathbb{V}$ contains a countably infinite supply of variables of each sort α such that $\alpha \neq \beta$ implies $Var_\alpha \cap Var_\beta = \varnothing$. Unless stated otherwise, we use $p, q, \ldots$ as metavariables which range over elements from $\{p_0, p_1, \ldots\}$. Hence, in this notation, p^α denotes a variable of sort α. We denote by L^*_Λ the set $L^*(\mathbb{V}, \Lambda)$. We abbreviate L^*_ω by L^*.

As usual, we introduce abbreviations $\varphi \rightarrow \psi \coloneqq \neg\varphi \vee \psi$, $\varphi \leftrightarrow \psi \coloneqq (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$, and $[\alpha]\varphi \coloneqq \neg\langle\alpha\rangle\neg\varphi$. We omit parentheses whenever possible and assign $\langle\alpha\rangle$, $[\alpha]$, and $\neg$ the highest, while $\rightarrow$ and $\leftrightarrow$ the least binding priority.

Definition 3.2.2. A *general substitution* is a map $\tau\colon \mathbb{V} \rightarrow L^*_\Lambda$. Given any substitution τ, we extend τ inductively to a function $\cdot^\tau\colon L^*_\Lambda \rightarrow L^*_\Lambda$ in the following way:

$$\begin{aligned} \top^\tau &= \top, \quad \bot^\tau = \bot \\ p^\tau &= \tau(p), \quad \text{for } p \in \mathbb{V}, \\ (\neg\varphi)^\tau &= \neg\varphi^\tau, \\ (\varphi \wedge \psi)^\tau &= \varphi^\tau \wedge \psi^\tau, \\ (\varphi \vee \psi)^\tau &= \varphi^\tau \vee \psi^\tau, \\ (\langle\alpha\rangle\varphi)^\tau &= \langle\alpha\rangle\varphi^\tau, \quad \text{for } \alpha \in \Lambda. \end{aligned}$$

We call τ simply *substitution* if for every variable p^α we have that $\tau(p^\alpha) = \varphi$ implies $|\varphi| \leq \alpha$. We say that ψ is a (*general*) *substitution instance of* φ if there is a (general) substitution τ such that $\psi = \varphi^\tau$. ⊣

We denote by $(p_1/\chi_1, \ldots, p_n/\chi_n)$ the general substitution τ where $\tau(p_i) = \chi_i$ for $i = 1, \ldots, n$ and $\tau(q) = q$ for $q \neq p_1, \ldots, p_n$. We write $\varphi(p_1/\chi_1, \ldots, p_n/\chi_n)$ for φ^τ and often omit the propositional variables if they are clear from context. In this notation, we also often denote a formula φ which contains propositional variables among $p_1, \ldots, p_n$ by $\varphi(p_1, \ldots, p_n)$.

An *axiom scheme* is a formula $\Phi(p_1, \ldots, p_n)$ which is a representative for all its substitution instances $\Phi(\varphi_1, \ldots, \varphi_n)$. We usually write axiom schemes as formulas $\Phi(\varphi_1, \ldots, \varphi_n)$, where each φ_i ($i = 1, \ldots, n$) is intended to range over all formulas (or over a specific class of formulas).

We exhibit each logic $\mathcal{L}$ as a list of *axiom schemes* and certain *rules of inference*. An *$\mathcal{L}$-proof* is then defined as usual, i.e., an $\mathcal{L}$-proof is a finite sequence $\varphi_1, \ldots, \varphi_n$ of formulas such that for every φ_i ($1 \leq i \leq n$) we have that φ_i is either an axiom or φ_i results from an application of some rule of inference from some $\varphi_{j_1}, \ldots, \varphi_{j_k}$, where $j_1, \ldots, j_k < i$. We say that φ is *provable in $\mathcal{L}$*, *$\mathcal{L}$-provable*, or a *theorem* of $\mathcal{L}$ (notation: $\mathcal{L} \vdash \varphi$) if there is an $\mathcal{L}$-proof $\varphi_1, \ldots, \varphi_n$ such that $\varphi_n = \varphi$. We say that a logic $\mathcal{L}'$ *extends* $\mathcal{L}$ if every theorem of $\mathcal{L}$ is also a theorem of $\mathcal{L}'$.

When we consider *many-sorted modal logics over* Λ in the sequel, we mean any set of formulas $\mathcal{L}_\Lambda$ over a signature Λ which (i) at least contains all propositional tautologies, (ii) is closed under substitutions as defined above, (iii) is closed under modus ponens, (iv) is closed under the rule $\varphi \to \psi / \langle\alpha\rangle\varphi \to \langle\alpha\rangle\psi$ for each $\alpha \in \Lambda$, and (v) contains the axioms $\neg\langle\alpha\rangle\neg\top$ and $\langle\alpha\rangle(\varphi \vee \psi) \to (\langle\alpha\rangle\varphi \vee \langle\alpha\rangle\psi)$ for each $\alpha \in \Lambda$. When we denote logics, the subscript "Λ" in $\mathcal{L}_\Lambda$ indicates that $\mathcal{L}_\Lambda$ is a logic over Λ.

GLP^* and J^*

Definition 3.2.3. Let Λ be a signature. The logic GLP^*_Λ is given by the following axiom schemes (the modalities range over Λ):

(i) all propositional tautologies;

(ii) $\langle\alpha\rangle(\varphi \vee \psi) \to (\langle\alpha\rangle\varphi \vee \langle\alpha\rangle\psi)$; $\quad \neg\langle\alpha\rangle\neg\top$;

(iii) $\langle\alpha\rangle\varphi \to \langle\alpha\rangle(\varphi \wedge \neg\langle\alpha\rangle\varphi)$ (*Löb's axiom*);

(iv) $\langle\alpha\rangle\varphi \to \langle\beta\rangle\varphi$, for $\beta < \alpha$ (*monotonicity*);

(v) $\langle\alpha\rangle\varphi \to \varphi$, if $|\varphi| \leq \alpha$ ($\Sigma_{\alpha+1}$-*completeness*).

GLP^*_Λ is closed under the rules of inference (i) modus ponens and (ii) for each $\alpha \leq \omega$, if $\varphi \to \psi$ then infer $\langle\alpha\rangle\varphi \to \langle\alpha\rangle\psi$. We denote the logic GLP^*_ω by GLP^*. ⊣

Remark. Note that GLP is usually axiomatized using the connective $[n]$ as a primitive. However, regarding the sorts of formulas using $\langle n\rangle$ instead seems to be more natural due to our intended arithmetical interpretation which focuses on Π_n-axiomatized concepts.

Furthermore, note that the results of this section concerning GLP^* also make sense if we disregard variables of sort ω, i.e., if we consider the many-sorted polymodal logic which strictly captures the notion of graded provability algebra as described in the introductory section of this chapter. However, introducing variables of sort ω is convenient with foresight of Chapter 4.

Note that GLP^*_Λ is indeed a logic as defined in the above sense. In particular, if $\mathsf{GLP}^*_\Lambda \vdash \varphi$ for $\varphi \in L^*_\Lambda$, then there is a proof $\chi_1, \ldots, \chi_n$ in GLP^*_Λ such that $\chi_n = \varphi$. Clearly, for any substitution τ, we have that $\chi_1^\tau, \ldots, \chi_n^\tau$ is a proof of φ^τ, since substitutions (according to our definition) respect the sorts of variables. The following basic properties will be used without any explicit mention.

Proposition 3.2.4. *Suppose $\mathcal{L}_\Lambda \vdash \varphi \leftrightarrow \psi$. Then, $\mathcal{L}_\Lambda \vdash \chi(p/\varphi) \leftrightarrow \chi(p/\psi)$ for any χ.*

Proof. By induction on χ. If $\chi = p$ then clearly $\mathcal{L}_\Lambda \vdash \varphi \leftrightarrow \psi$ by assumption. If $\chi = q$ for some $q \neq p$, then $\mathcal{L}_\Lambda \vdash q \leftrightarrow q$ by propositional logic. The same holds for the case where $\chi = \top$ or $\chi = \bot$.

Assume $\chi = \chi_1 \wedge \chi_2$ for some χ_1, χ_2. By inductive hypothesis we have $\mathcal{L}_\Lambda \vdash \chi_i(p/\varphi) \leftrightarrow \chi_i(p/\psi)$ for $i = 1, 2$, whence $\mathcal{L}_\Lambda \vdash \chi(p/\varphi) \leftrightarrow \chi(p/\psi)$ follows by purely propositional reasoning and the definition of substitution. The other propositional connectives are treated similarly.

Suppose $\chi = \langle\alpha\rangle\xi$ for some ξ. By inductive hypothesis, we have $\mathcal{L}_\Lambda \vdash \xi(p/\varphi) \leftrightarrow \xi(p/\psi)$, whence $\mathcal{L}_\Lambda \vdash \langle\alpha\rangle\xi(p/\varphi) \leftrightarrow \langle\alpha\rangle\xi(p/\psi)$ follows. □

Lemma 3.2.5. *Every logic $\mathcal{L}_\Lambda$ is closed under $[\alpha]$-necessitation, for each $\alpha \in \Lambda$.*

Proof. Suppose $\mathcal{L}_\Lambda \vdash \varphi$. Then $\mathcal{L}_\Lambda \vdash \neg\varphi \to \neg\top$, whence $\mathcal{L}_\Lambda \vdash \langle\alpha\rangle\neg\varphi \to \langle\alpha\rangle\neg\top$. Thus, $\mathcal{L}_\Lambda \vdash \neg\langle\alpha\rangle\neg\top \to \neg\langle\alpha\rangle\neg\varphi$ and so $\mathcal{L}_\Lambda \vdash [\alpha]\varphi$. □

Lemma 3.2.6. *Let $\varphi_1, \ldots, \varphi_k$ be formulas and $\mathcal{L}_\Lambda$ a logic. For all $\alpha \in \Lambda$ we have*

$$\mathcal{L}_\Lambda \vdash [\alpha](\varphi_1 \wedge \cdots \wedge \varphi_k) \to ([\alpha]\varphi_1 \wedge \cdots \wedge [\alpha]\varphi_k).$$

Proof. For $i = 1, \ldots, k$, we obtain by $\mathcal{L}_\Lambda \vdash \varphi_1 \wedge \cdots \wedge \varphi_k \to \varphi_i$

$$\mathcal{L}_\Lambda \vdash \langle\alpha\rangle\neg\varphi_i \to \langle\alpha\rangle\neg(\varphi_1 \wedge \cdots \wedge \varphi_k),$$

whence $\mathcal{L}_\Lambda \vdash [\alpha](\varphi_1 \wedge \cdots \wedge \varphi_k) \to ([\alpha]\varphi_1 \wedge \cdots \wedge [\alpha]\varphi_k)$ by propositional logic. □

Lemma 3.2.7. *Let $\varphi_1, \ldots, \varphi_k$ be formulas and $\mathcal{L}_\Lambda$ a logic. For all $\alpha \in \Lambda$ we have*

$$\mathcal{L}_\Lambda \vdash \langle\alpha\rangle(\varphi_1 \vee \cdots \vee \varphi_k) \to (\langle\alpha\rangle\varphi_1 \vee \cdots \vee \langle\alpha\rangle\varphi_k).$$

Proof. By repeated application of the axiom $\langle\alpha\rangle(\varphi \vee \psi) \to (\langle\alpha\rangle\varphi \vee \langle\alpha\rangle\psi)$ and propositional logic. □

Note that $\mathsf{GLP}^*_\Lambda \vdash \langle\alpha\rangle\langle\alpha\rangle\varphi \to \langle\alpha\rangle\varphi$. Furthermore, if $\beta < \alpha$ then $\mathsf{GLP}^* \vdash \langle\alpha\rangle\neg[\beta]\varphi \to \neg[\beta]\varphi$, whence

$$\mathsf{GLP}^*_\Lambda \vdash [\beta]\varphi \to [\alpha][\beta]\varphi,$$

by propositional logic. Similarly, $\mathsf{GLP}^*_\Lambda \vdash \langle\alpha\rangle\neg\langle\beta\rangle\varphi \to \neg\langle\beta\rangle\varphi$ and thus

$$\mathsf{GLP}^*_\Lambda \vdash \langle\beta\rangle\varphi \to [\alpha]\langle\beta\rangle\varphi,$$

again by propositional logic.

Definition 3.2.8. The logic J^*_Λ is obtained from GLP^*_Λ by dropping the monotonicity axioms and adding the following additional scheme:

(vi) $\langle\beta\rangle\langle\alpha\rangle\varphi \to \langle\beta\rangle\varphi$, for $\beta < \alpha$.

J^*_Λ has the same rules of inference as GLP^*_Λ. We denote by J^* the logic J^*_ω. ⊣

For $\beta < \alpha$, we note that

$$\begin{aligned}\mathsf{GLP}^*_\Lambda \vdash \langle\beta\rangle\langle\alpha\rangle\varphi &\to \langle\beta\rangle\langle\beta\rangle\varphi \quad \text{(by monotonicity)}\\ &\to \langle\beta\rangle\varphi.\end{aligned}$$

We thus have the following:

Lemma 3.2.9. *The logic* GLP^*_Λ *extends* J^*_Λ*, i.e.,* $\mathsf{J}^*_\Lambda \vdash \varphi$ *implies* $\mathsf{GLP}^*_\Lambda \vdash \varphi$*.*

3.3 Arithmetical Interpretation

The following notion was originally suggested by Ignatiev [24].

Definition 3.3.1. Let T be an extension of PA. A *provability predicate of level n over T* is a formula $\mathsf{Prv}(x)$ with one free variable which satisfies the following conditions, for all sentences φ, ψ,

(i) Prv is Σ_{n+1} in T;

(ii) if $T \vdash \varphi$ then $T \vdash \mathsf{Prv}(\ulcorner\varphi\urcorner)$;

(iii) $T \vdash \mathsf{Prv}(\ulcorner\varphi \to \psi\urcorner) \to (\mathsf{Prv}(\ulcorner\varphi\urcorner) \to \mathsf{Prv}(\ulcorner\psi\urcorner))$;

(iv) if φ is a Σ_{n+1}-sentence, then $T \vdash \varphi \to \mathsf{Prv}(\ulcorner\varphi\urcorner)$.

We say that a provability predicate Prv is *sound* if, for all φ, $\mathbb{N} \models \mathsf{Prv}(\ulcorner\varphi\urcorner)$ implies that $\mathbb{N} \models \varphi$.

A sequence π of formulas $\mathsf{Prv}_0, \mathsf{Prv}_1, \ldots$ is called a *strong sequence of provability predicates over T* if there is a sequence of natural numbers $r_0 < r_1 < r_2 < \cdots$ such that for all $n \geq 0$,

(i) Prv_n is a provability predicate of level r_n over T;

(ii) $T \vdash \mathsf{Prv}_n(\ulcorner\varphi\urcorner) \to \mathsf{Prv}_{n+1}(\ulcorner\varphi\urcorner)$, for every sentence φ.

We denote by $|\pi_n|$ the level of Prv_n. ⊣

Given a strong sequence of provability predicates over T, we write $[n]_\pi$ for the n-th provability predicate of π and use the abbreviation $[n]_\pi\varphi$ for $[n]_\pi(\ulcorner\varphi\urcorner)$ if no confusion arises. As usual, $\langle n\rangle_\pi$ is defined to be the dual of $[n]_\pi$.

As Ignatiev [24] and Beklemishev [6], we want to mention two important examples of strong sequences of provability predicates over T, for T extending PA.

In Section 2.5 we introduced the notion of n-provability, i.e., formalized provability in the theory $T+\mathsf{Th}_{\Pi_n}(\mathbb{N})$. In fact, every predicate of the sequence $[0]_T,[1]_T,\dots$ satisfies the conditions of Definition 3.3.1 by virtue of our treatment of $[n]_T$ in Chapter 2. Note that, for each $n\geq 0$, $[n]_T$ is of level n. Furthermore, this sequence is easily seen to be a strong sequence of provability predicates over T, since every Π_n-sentence is provably equivalent to a Π_{n+1}-sentence by introducing dummy quantifiers.

The second strong sequence of provability predicates we want to mention is that which arises from the closure under the n-fold application of the ω-rule in PA [24, 6]. Formally, define $[0]_\omega := \Box_{\mathsf{PA}}$ and

$$[n+1]_\omega(\alpha) := \exists\beta\,(\forall x\,[n]_\omega\beta(\dot{x})\wedge[n]_\omega(\forall x\,\beta(x)\to\alpha)),\quad \text{for } n\geq 0.$$

For $n\geq 0$, the predicate $[n]_\omega$ is a Σ_{2n+1}-formula. It can be shown that the sequence $[0]_\omega,[1]_\omega,\dots$ defines a strong sequence of provability predicates over PA, where $[n]_\omega$ has level $2n$ (cf. also Boolos [10] for more details on ω-provability).

Definition 3.3.2. Let π be a sequence of strong provability predicates over T. An (*arithmetical*) *realization* (*over* π) is a function f_π which maps formulas from L^* to sentences in the language of arithmetic such that the following conditions are satisfied:

(i) $f_\pi(\top)=\top$;[2] $f_\pi(\bot)=\bot$;

(ii) for every propositional variable p^α, $f_\pi(p^\alpha)$ is a $\Pi_{|\pi_n|+1}$-sentence in case $n=\alpha<\omega$;

(iii) f_π commutes with the propositional connectives;

(iv) $f_\pi(\langle n\rangle\varphi)=\langle n\rangle_\pi f_\pi(\varphi)$, for $n<\omega$.

We say that $f_\pi(\varphi)$ is the *translation of* φ *under* f_π. ⊣

Clearly, a realization over π only depends on the assignment of sentences to propositional variables. Note that, for any φ and any realization f_π, we have

$$T\vdash f_\pi([n]\varphi)\leftrightarrow[n]_\pi f_\pi(\varphi).$$

Lemma 3.3.3. *Let π be a strong sequence of provability predicates over T and let f_π be a realization. For all many-sorted formulas φ we have that $f_\pi(\varphi)$ is provably equivalent to an arithmetical $\Pi_{|\pi_k|+1}$-sentence, where $k=|\varphi|$.*

Proof. By an easy induction on φ. The base case holds by definition. Furthermore, for $k\geq 0$, $[k]_\pi$ is a $\Sigma_{|\pi_k|+1}$-sentence. The induction step then follows by simple closure properties of Π_n-sentences (see Proposition 2.2.5). □

[2]As usual, we use $\top$ to denote a valid statement in the language of arithmetic, e.g., $\top:=\neg\bot$.

Since provability predicate $[n]_\pi$ from π is a Σ_k-sentence for some $k > 0$, we can associate (in analogy to the standard Gödelian provability predicate) a predicate $\mathsf{Prf}_n(\alpha, y)$ which expresses the statement "y codes a proof of α" and

$$T \vdash \mathsf{Prv}_n(\alpha) \leftrightarrow \exists y\, \mathsf{Prf}_n(\alpha, y).$$

We say that Prf_n is the *proof relation of* Prv_n and stress that Prf_n is chosen in such a way, such that every number y codes a proof of at most one formula and that every provable formula has arbitrarily long proofs. Intuitively, since a proof is coded as a finite sequence in the Hilbertian sense, given any proof of φ, any proof containing redundant axioms will also witness the provability of φ. These properties can be achieved in such a way as to hold provably in T, i.e.,

$$T \vdash \mathsf{Prf}_n(\alpha, y) \land \mathsf{Prf}_n(\beta, y) \to \alpha = \beta,$$
$$T \vdash \mathsf{Prf}_n(\alpha, y) \to \exists z > y\, \mathsf{Prf}_n(\alpha, z).$$

We can already conclude that GLP^* is arithmetically sound.

Proposition 3.3.4. *If* $\mathsf{GLP}^* \vdash \varphi$*, then* $T \vdash f_\pi(\varphi)$ *for all realizations* f_π*.*

Proof. By induction the length of a derivation of φ. For the base case, the propositional tautologies and axioms (ii) of Definition 3.2.3 are clear. Note that $T \vdash f_\pi(\langle n\rangle\varphi) \to f_\pi(\varphi)$ follows by $\Sigma_{|\pi_n|+1}$-completeness and Lemma 3.3.3: since $f_\pi(\varphi)$ is provably equivalent to a $\Pi_{|\pi_n|+1}$-sentence, we know that $\neg f_\pi(\varphi)$ is provably equivalent to a $\Sigma_{|\pi_n|+1}$-sentence, whence

$$T \vdash \neg f_\pi(\varphi) \to [n]_\pi \neg f_\pi(\varphi), \quad \text{(by } \Sigma_{|\pi_n|+1}\text{-completeness)}$$
$$T \vdash \neg [n]_\pi \neg f_\pi(\varphi) \to f_\pi(\varphi).$$

The soundness of Löb's axioms can be proved similarly to the formalized version of Löb's theorem (Theorem 2.4.10) and propositional logic. (Note that in our case Löb's axiom is formulated with $\langle n\rangle$ rather than $[n]$.) For the induction step, the soundness of the rules of inference is clear by the definition of strong provability predicates. □

3.4 Kripke Semantics

In this section, we are going to develop Kripke semantics for J^*. We show that J^* is complete for a decent class of Kripke models which will be exploited in the proof of the arithmetical completeness theorem for GLP^* in Section 3.5. In our elaboration, we closely follow the work of Beklemishev [5] and the standard methods known from the area of modal logics. For background information concerning Kripke semantics of modal logics in general, we refer the reader to Blackburn et al. [9].

Definition 3.4.1. A (*Kripke*) *frame* $\mathfrak{F}$ (*over* Λ) is a tuple $\mathfrak{F} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}\rangle$, where W is a non-empty set of *worlds* and R_α is a binary relation on W for all $\alpha \in \Lambda$ (called the *accessibility relations*). We say that $\mathfrak{F}$ is *finite* if W is finite and all but finitely many relations of $\{R_\alpha\}_{\alpha\in\Lambda}$ are empty. ⊣

Definition 3.4.2. A (*Kripke*) *model* $\mathfrak{K}$ (*over* Λ) is a tuple of the form $\mathfrak{K} = \langle \mathfrak{F}, [\![\cdot]\!] \rangle$, where $\mathfrak{F} = \langle W, \{R_\alpha\}_{\alpha \in \Lambda} \rangle$ is a frame over Λ and $[\![\cdot]\!] \colon L^*_\Lambda \to \mathcal{P}(W)$ is a function called *valuation* which maps many-sorted formulas to subsets of W such that the following conditions are satisfied:

(i) $[\![\bot]\!] = \varnothing$; $\quad [\![\top]\!] = W$;

(ii) $[\![\neg\varphi]\!] = W \setminus [\![\varphi]\!]$;

(iii) $[\![\varphi \wedge \psi]\!] = [\![\varphi]\!] \cap [\![\psi]\!]$;

(iv) $[\![\varphi \vee \psi]\!] = [\![\varphi]\!] \cup [\![\psi]\!]$;

(v) $[\![\langle\alpha\rangle\varphi]\!] = \{x \mid \exists y \colon xR_\alpha y \;\&\; y \in [\![\varphi]\!]\}$, for $\alpha \in \Lambda$.

We say that $\mathfrak{K}$ is *based on* $\mathfrak{F}$. ⊣

Note that, by definition of $[\alpha]\varphi$, we have

$$[\![[\alpha]\varphi]\!] = \{x \mid \forall y \colon xR_\alpha y \Rightarrow y \in [\![\varphi]\!]\}.$$

Furthermore, note that $[\![\cdot]\!]$ only depends on the assignment of subsets of W to propositional variables. Given any model $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!] \rangle$, it will be convenient to define a relation $\Vdash_{\mathfrak{K}}$ between worlds of $\mathfrak{K}$ and formulas by

$$x \Vdash_{\mathfrak{K}} \varphi \iff_{df} x \in [\![\varphi]\!].$$

We often omit the subscript of $\Vdash_{\mathfrak{K}}$ when $\mathfrak{K}$ is clear from context. Furthermore, we often write $\mathfrak{K}, x \Vdash \varphi$ instead of $x \Vdash_{\mathfrak{K}} \varphi$ and say that x *forces* φ (*in* $\mathfrak{K}$). The relation $\Vdash_{\mathfrak{K}}$ is called *forcing relation of* $\mathfrak{K}$.

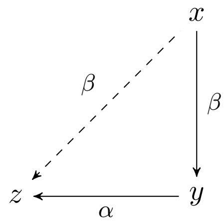
Definition 3.4.3. Let $\mathfrak{K} = \langle \mathfrak{F}, [\![\cdot]\!] \rangle$ be a model and φ a formula. We say that φ is *true at* a world x of $\mathfrak{K}$ if $x \in [\![\varphi]\!]$. The formula φ is (*globally*) *true in* $\mathfrak{K}$ (notation: $\mathfrak{K} \models \varphi$) if it is true at every world of $\mathfrak{K}$. Similarly, φ is *valid in* $\mathfrak{F}$ (notation: $\mathfrak{F} \models \varphi$) if it is true in every model based on $\mathfrak{F}$. ⊣

A relation $R \subseteq W \times W$ is said to be *conversely well-founded* if there is no infinite sequence $x_1, x_2, \ldots$ such that $x_1 R x_2 R \cdots$, i.e., if every R-increasing chain is finite.

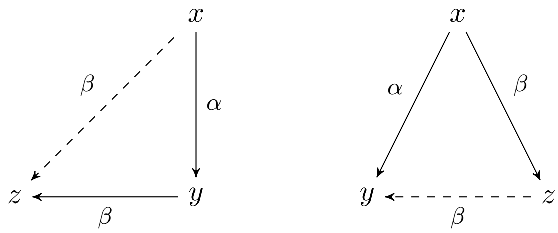
Lemma 3.4.4. *Let W be finite and $R \subseteq W \times W$ transitive. Then, R is conversely well-founded iff it is irreflexive.*

Proof. If xRx for some $x \in W$, then the set $\{x\}$ has no R-greatest element, whence it follows that R is not conversely well-founded. On the other hand, suppose that R is not conversely well-founded. Then there is an infinite R-increasing chain $x_1 R x_2 R \cdots$. Since W is finite, there are $i, j \geq 1$ such that $i \geq j$ and $x_i = x_j$. By transitivity and induction we obtain $x_j R x_i$, so R is not irreflexive. □

Definition 3.4.5. A frame $\mathfrak{F} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda} \rangle$ is called J^*_Λ-*frame* if the following conditions are satisfied:



(a) Condition (J2)



(b) Condition (J3)

Figure 3.1: Frame conditions of a J^*_Λ-frame, where $\alpha, \beta \in \Lambda$ such that $\beta < \alpha$. Dashed arrows represent relations which must exist provided the solid ones do.

(J1) R_α is transitive and conversely well-founded for $\alpha \in \Lambda$;

(J2) $\forall x, y, z\, (xR_\beta y \;\&\; yR_\alpha z \Rightarrow xR_\beta z)$ for $\beta < \alpha$;

(J3) $\forall x, y\, (xR_\alpha y \Rightarrow \forall z\, (xR_\beta z \Leftrightarrow yR_\beta z))$ for $\beta < \alpha$.

A *root* of a J^*_Λ-frame $\mathfrak{F} = \langle W, \{R_\alpha\}_{\alpha \in \Lambda}\rangle$ is a world $r \in W$ such that $\forall x \in W\, \exists \lambda \in \Lambda \colon rR_\lambda x$ or $r = x$. A frame which has a root is called *rooted.* ⊣

For a visualization of the conditions (J2) and (J3) see Figure 3.1. Note that item (J3) is equivalent to the conjunction of

$$\forall x, y, z\, (xR_\alpha y \;\&\; yR_\beta z \Rightarrow xR_\beta z) \text{ and } \forall x, y, z\, (xR_\alpha y \;\&\; xR_\beta z \Rightarrow yR_\beta z), \text{ for } \beta < \alpha.$$

$\mathfrak{F}$ is called *irreflexive* (*transitive*, *conversely well-founded*) if its accessibility relations have the corresponding property.

Definition 3.4.6. A J^*_Λ-*model* is a Kripke model based on a J^*_Λ-frame. Given a J^*_Λ-model $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha \in \Lambda}, [\![\cdot]\!]\rangle$, we call $\mathfrak{K}$

(i) *persistent* if for all $\alpha \in \Lambda$, all propositional variables p^β with $\beta \leq \alpha$, and all $x, y \in W$ we have

$$xR_\alpha y \text{ and } y \in [\![p^\beta]\!] \text{ imply } x \in [\![p^\beta]\!];$$

(ii) *strongly persistent* if $\mathfrak{K}$ is persistent and for all $\alpha \in \Lambda$, all propositional variables p^β with $\beta < \alpha$, and all $x, y \in W$ we have

$$xR_\alpha y \text{ and } y \notin [\![p^\beta]\!] \text{ imply } x \notin [\![p^\beta]\!].$$

We say that $\mathfrak{K}$ is *finite* if its underlying frame is finite and $[\![p]\!] \neq \varnothing$ only for finitely many variables p. Furthermore, $\mathfrak{K}$ is *rooted* if the frame it is based on is rooted. Likewise, $\mathfrak{K}$ is *irreflexive* (*transitive*, *conversely well-founded*) if its underlying frame has the corresponding property. ⊣

In case $\Lambda = \omega$, we drop the subscript "Λ" in the terms J^*_Λ-frame and J^*_Λ-model. Let $\mathcal{C}$ be a class of models. Recall that a logic $\mathcal{L}$ is *sound for* $\mathcal{C}$ if $\mathcal{L} \vdash \varphi$ implies $\mathfrak{K} \models \varphi$ for all $\mathfrak{K} \in \mathcal{C}$. $\mathcal{L}$ is *complete for* $\mathcal{C}$ if whenever $\mathfrak{K} \models \varphi$ for all $\mathfrak{K} \in \mathcal{C}$, then also $\mathcal{L} \vdash \varphi$. Soundness and completeness with respect to a class of frames is defined *mutatis mutandis*.

The remainder of this section will be devoted to the proof that J^*_Λ is sound and complete with respect to the class of finite and strongly persistent J^*_Λ-models.

Lemma 3.4.7. *Let $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!]\rangle$ be a J^*_Λ-model. Then, $\mathfrak{K}$ is strongly persistent iff for all formulas $\varphi \in L^*_\Lambda$ and all $\alpha \in \Lambda$ we have*

(i) *if $|\varphi| \leq \alpha$ then $xR_\alpha y$ and $y \in [\![\varphi]\!]$ imply $x \in [\![\varphi]\!]$;*

(ii) *if $|\varphi| < \alpha$ then $xR_\alpha y$ and $y \notin [\![\varphi]\!]$ imply $x \notin [\![\varphi]\!]$.*

Proof. The direction from right to left is clear. For the other direction, we proceed by induction on the number of propositional connectives which are not in the scope of any $\langle\alpha\rangle$. Let $\mathfrak{K}$ be strongly persistent. For the base case we distinguish two cases. Firstly, if $\varphi = p^\beta$ or $\varphi = \top$ or $\varphi = \bot$ is just the definition of $\mathfrak{K}$ being strongly persistent—note that $[\![\top]\!] = W$ and $[\![\bot]\!] = \varnothing$. Secondly, suppose $\varphi = \langle\alpha\rangle\psi$ for some ψ. Then $|\varphi| = \alpha$. So let $\lambda \geq \alpha$ for some $\lambda \in \Lambda$ and assume $xR_\lambda y$ and $y \in [\![\langle\alpha\rangle\psi]\!]$. Then $z \in [\![\psi]\!]$ for some $z \in W$ such that $yR_\alpha z$. Since $\lambda \geq \alpha$ and $\mathfrak{K}$ is J^*_Λ-model, we have $xR_\alpha z$, whence $x \in [\![\langle\alpha\rangle\psi]\!]$ follows. Suppose now $\lambda > \alpha$ and $xR_\lambda y$ such that $y \notin [\![\langle\alpha\rangle\psi]\!]$. Then for all $z \in W$ such that $yR_\alpha z$ we have $z \notin [\![\psi]\!]$. Now if $x \in [\![\langle\alpha\rangle\psi]\!]$ then $z \in [\![\psi]\!]$ for some z such that $xR_\alpha z$, whence we infer $yR_\alpha z$ (since $\mathfrak{K}$ is a J^*_Λ-model) and arrive at contradiction.

Assume $\varphi = \varphi_1 \wedge \varphi_2$ where $|\varphi| \leq \alpha$ and let $xR_\alpha y$ and $y \in [\![\varphi]\!]$, i.e., $y \in [\![\varphi_1]\!]$ and $y \in [\![\varphi_2]\!]$. By inductive hypothesis we know, as $|\varphi_1|, |\varphi_2| \leq \alpha$, that $x \in [\![\varphi_1]\!]$ and $x \in [\![\varphi_2]\!]$. Furthermore, if $|\varphi| < \alpha$ and $y \notin [\![\varphi]\!]$, then either $y \notin [\![\varphi_1]\!]$ or $y \notin [\![\varphi_2]\!]$. Since $|\varphi_1|, |\varphi_2| < \alpha$, we infer by inductive hypothesis that $x \notin [\![\varphi_1]\!]$ or $x \notin [\![\varphi_2]\!]$ as required. The case where $\varphi = \varphi_1 \vee \varphi_2$ is treated similarly.

Finally, suppose $\varphi = \neg\psi$ for some ψ and suppose $|\varphi| \leq \alpha$. Let $xR_\alpha y$ and $y \in [\![\varphi]\!]$. Then $|\psi| < \alpha$ and $y \notin [\![\psi]\!]$, whence by inductive hypothesis we obtain $x \notin [\![\psi]\!]$, i.e., $x \in [\![\varphi]\!]$ as desired. If $|\varphi| < \alpha$ and $y \notin [\![\varphi]\!]$, then $|\psi| < \alpha$ and $y \in [\![\psi]\!]$. By inductive hypothesis we then obtain $x \in [\![\psi]\!]$ and so $x \notin [\![\varphi]\!]$ which finishes this case. □

Lemma 3.4.8. *For all $\alpha \in \Lambda$, the axiom scheme $\langle\alpha\rangle\varphi \to \varphi$ (where $\varphi \in L^*_\Lambda$ such that $|\varphi| \leq \alpha$) is true in a J^*_Λ-model $\mathfrak{K}$, iff $\mathfrak{K}$ is strongly persistent.*

Proof. For the direction from left to right, suppose that $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!]\rangle$ is not strongly persistent. Suppose first that there are $x, y \in W$ such that $xR_\alpha y$ and $y \in [\![p^\lambda]\!]$, but $x \notin [\![p^\lambda]\!]$ for some $\lambda \leq \alpha$ and $\lambda, \alpha \in \Lambda$. Then clearly $\mathfrak{K}, x \not\Vdash \langle\alpha\rangle p^\lambda \to p^\lambda$. For the case where $xR_\alpha y$ and $y \notin [\![p^\lambda]\!]$ but $x \in [\![p^\lambda]\!]$ for some $\lambda < \alpha$ ($\lambda, \alpha \in \Lambda$), we similarly have $\mathfrak{K}, x \nVdash \langle\alpha\rangle\neg p^\lambda \to \neg p^\lambda$.

For the other direction, suppose that there is a ψ with $|\psi| \leq \alpha$ ($\alpha \in \Lambda$) such that $\mathfrak{K}, x \nVdash \langle\alpha\rangle\psi \to \psi$. Then $\mathfrak{K}, x \Vdash \langle\alpha\rangle\psi$ and $\mathfrak{K}, x \nVdash \psi$, and so there is a $y \in W$ such that $xR_\alpha y$ and $\mathfrak{K}, y \Vdash \psi$. Hence, $x \notin [\![\psi]\!]$ but $y \in [\![\psi]\!]$ and $|\psi| \leq \alpha$, whence Lemma 3.4.7 yields that $\mathfrak{K}$ is not strongly persistent. □

Notice that our notion of substitution is substantial here. Indeed, consider the formula $\varphi := \langle 0\rangle p^0 \to p^0$. It is clear that φ is true in every strongly persistent J^*_Λ-model (for an appropriate Λ), but the formula $\varphi' := \langle 0\rangle p^1 \to p^1$ is not. However, it is easy to see that φ' is not a substitution instance of φ in the sense of Definition 3.2.2. Hence, our notion of substitution is geared in order to retain truth in the class of strongly persistent J^*_Λ-models.

Proposition 3.4.9. J^*_Λ *is sound for the class of all strongly persistent* J^*_Λ*-models.*

Proof. The proof is a routine induction on proof length. Most of the axioms were handled by the previous statements. The axiom $\langle\alpha\rangle(\varphi \vee \psi) \to \langle\alpha\rangle\varphi \vee \langle\alpha\rangle\psi$ and the propositional ones are obvious. The fact that instances of Löb's axiom are true in all such models follows from the well-known fact that these schemes are valid in all frames which are transitive and conversely well-founded. For the induction step, modus ponens and $\varphi \to \psi / \langle\alpha\rangle\varphi \to \langle\alpha\rangle\psi$ are easy to check. We leave the details to the reader. □

Definition 3.4.10. Let Γ be a set of formulas from L^*_Λ. We say that Γ is $\mathcal{L}_\Lambda$*-consistent* if there are no $\varphi_1, \ldots, \varphi_n \in \Gamma$ such that $\mathcal{L}_\Lambda \vdash \varphi_1 \wedge \cdots \wedge \varphi_n \to \bot$. Otherwise, Γ is called $\mathcal{L}_\Lambda$*-inconsistent.* Let Σ be a set of formulas. Then $\Gamma \subseteq \Sigma$ is a *maximal* $\mathcal{L}_\Lambda$*-consistent subset of* Σ, if every Γ' such that $\Sigma \supseteq \Gamma' \supset \Gamma$ is $\mathcal{L}_\Lambda$-inconsistent. ⊣

We define an operator $\sim$, called *modified negation*, for all formulas φ as follows:

$$\sim\varphi = \begin{cases} \psi, & \text{if } \varphi = \neg\psi \text{ for some } \psi, \\ \neg\varphi, & \text{otherwise.} \end{cases}$$

For a set of formulas Δ from L^*_Λ, we set $\ell(\Delta) := \{\alpha \in \Lambda \mid \langle\alpha\rangle\varphi \in \Delta \text{ for some } \varphi\}$. We say that a set of formulas Δ is *adequate* if $\top \in \Delta$, it is closed under subformulas, modified negations, and the operations

$$\begin{aligned} \langle\alpha\rangle\varphi, \langle\beta\rangle\psi \in \Delta &\Rightarrow \langle\beta\rangle\varphi \in \Delta, \\ p^\lambda \in \Delta, \alpha \in \ell(\Delta) &\Rightarrow \langle\alpha\rangle p^\lambda \in \Delta, \quad \text{for all variables } p^\lambda \text{ and } \alpha \geq \lambda, \\ \neg p^\lambda \in \Delta, \alpha \in \ell(\Delta) &\Rightarrow \langle\alpha\rangle\neg p^\lambda \in \Delta, \quad \text{for all variables } p^\lambda \text{ and } \alpha > \lambda. \end{aligned}$$

We can easily convince ourselves that any finite set Γ can be extended to a finite adequate Γ' such that $\ell(\Gamma) = \ell(\Gamma')$. We denote the smallest such set by $Cl(\Gamma)$ and note that $Cl(\Gamma)$ is finite. Furthermore, it is easy to see that if Δ is adequate, then for every maximal consistent subset Γ of Δ we have (i) $\varphi \in \Gamma$ or $\sim\varphi \in \Gamma$ for every $\varphi \in \Delta$, (ii) $\varphi \to \psi, \varphi \in \Gamma$ implies $\psi \in \Gamma$, and (iii) $\varphi \vee \psi \in \Gamma$ iff $\varphi \in \Gamma$ or $\psi \in \Gamma$. The following fact is well-known.

Lemma 3.4.11. *Let $\mathcal{L}_\Lambda$ be a logic and Γ, Δ finite sets of formulas such that $\Gamma \subseteq Cl(\Delta)$, where Γ is $\mathcal{L}_\Lambda$-consistent. Then, there is a maximal $\mathcal{L}_\Lambda$-consistent $\Gamma' \subseteq Cl(\Delta)$ such that $\Gamma \subseteq \Gamma'$.*

Proof. Let $\varphi_1, \varphi_2, \dots, \varphi_n$ be an enumeration of $Cl(\Delta)$. Define $\Sigma_0 := \Gamma$ and, for $k < n$, construct sets Σ_{k+1} in the following way:

$$\Sigma_{k+1} := \begin{cases} \Sigma_k \cup \{\varphi_{k+1}\}, & \text{if } \Sigma_k \cup \{\varphi_{k+1}\} \text{ is } \mathcal{L}_\Lambda\text{-consistent,} \\ \Sigma_k \cup \{\sim\varphi_{k+1}\}, & \text{otherwise.} \end{cases}$$

Let $\Sigma^+ = \Sigma_n$. By induction on k, we easily see that Σ_k is $\mathcal{L}_\Lambda$-consistent for every $k \geq 0$. Hence, Σ^+ is $\mathcal{L}_\Lambda$-consistent as $\Sigma_k \subseteq \Sigma_{k+1}$, for $k = 1, \dots, n-1$. Furthermore, for every $\varphi \in Cl(\Delta)$ we either have $\varphi \in \Sigma^+$ or $\sim\varphi \in \Sigma^+$. It follows that $\Sigma^+ \subseteq Cl(\Delta)$ is a maximal $\mathcal{L}_\Lambda$-consistent set containing Γ. □

Let us now fix a finite adequate set Δ and assume that all modalities range within $\ell(\Delta)$. Let $\Lambda := \ell(\Delta)$ and define a Kripke frame $\mathfrak{F}_\Delta = \langle W, \{R_\alpha\}_{\alpha \in \Lambda}\rangle$, where

$$W := \{x \mid x \text{ is a maximal } \mathsf{J}^*_\Lambda\text{-consistent subset of } \Delta\},$$

for $\alpha \in \Lambda$ and $x, y \in W$, define $xR_\alpha y$ if the following conditions are satisfied:

(i) For any $\varphi \in y$, if $\langle\alpha\rangle\varphi \in \Delta$ then $\langle\alpha\rangle\varphi \in x$.

(ii) For any $\langle\alpha\rangle\varphi \in \Delta$, we have that $\langle\alpha\rangle\varphi \in y$ implies $\langle\alpha\rangle\varphi \in x$.

(iii) For any $\langle\beta\rangle\varphi \in \Delta$ such that $\beta < \alpha$, we have $\langle\beta\rangle\varphi \in x \iff \langle\beta\rangle\varphi \in y$.

(iv) There exists a formula $\langle\alpha\rangle\varphi \in \Delta$ such that $\langle\alpha\rangle\varphi \in x \setminus y$.

Lemma 3.4.12. *$\mathfrak{F}_\Delta$ is a finite J^*_Λ-frame.*

Proof. We first check the conditions of a J^*_Λ-frame. Obviously, $\mathfrak{F}_\Delta$ is finite. Hence, to establish well-foundedness of each R_α, it suffices to check irreflexivity. But this is guaranteed by item (iii). We first prove simultaneously that R_α is transitive and condition (J2) is satisfied, so suppose $xR_\beta y$ and $yR_\alpha z$ for $\beta \leq \alpha$. We show $xR_\beta z$ by checking the four conditions above. Suppose $\varphi \in z$ such that $\langle\beta\rangle\varphi \in \Delta$. By $yR_\alpha z$ and items (i) and (ii) we know that $\langle\beta\rangle\varphi \in y$, whence $\langle\beta\rangle\varphi \in x$ follows by $xR_\beta y$. Hence, item (i) is established. Let $\langle\beta\rangle\varphi \in \Delta$ such that $\langle\beta\rangle\varphi \in z$. Again, items (i) and (ii) and $yR_\alpha z$ yield $\langle\beta\rangle\varphi \in y$, whence $\langle\beta\rangle\varphi \in x$ follows by $xR_\beta y$. So item (ii) is verified. Let $\lambda < \beta$ and consider any $\langle\lambda\rangle\varphi \in \Delta$. We know

$$\langle\lambda\rangle\varphi \in x \iff \langle\lambda\rangle\varphi \in y \iff \langle\lambda\rangle\varphi \in z,$$

whence item (iii) follows. For item (iv), we know that there is an $\langle\alpha\rangle\psi \in y \setminus z$. But then $\langle\beta\rangle\psi \in x$ by item (i) and $xR_\beta y$. Hence, item (iv) is also verified, i.e., $xR_\beta z$ holds.

It remains to establish condition (J3). So first suppose $xR_\alpha y$ and $xR_\beta z$ for $\beta < \alpha$. We prove $yR_\beta z$. Indeed, if $\varphi \in z$ and $\langle\beta\rangle\varphi \in \Delta$, then $\langle\beta\rangle\varphi \in x$ by $xR_\beta z$, whence $\langle\beta\rangle\varphi \in y$ follows since $\beta < \alpha$ and $xR_\alpha y$. So item (i) holds. If $\langle\beta\rangle\varphi \in \Delta$ such that $\langle\beta\rangle\varphi \in z$, then $\langle\beta\rangle\varphi \in x$ since $xR_\beta z$, whence $\langle\beta\rangle\varphi \in y$ follows since $\beta < \alpha$ and $xR_\alpha y$. This proves item (ii). Let $\lambda < \beta$ and $\langle\lambda\rangle\varphi \in \Delta$. We have

$$\langle\lambda\rangle\varphi \in y \iff \langle\lambda\rangle\varphi \in x \iff \langle\lambda\rangle\varphi \in z.$$

Hence, item (iii) follows. For item (iv), we know that there is a $\langle\beta\rangle\psi \in \Delta$ such that $\langle\beta\rangle\psi \in x \setminus z$. Now, $\langle\beta\rangle\psi \in y$ by $xR_\alpha y$ and $\beta < \alpha$. Thus, $yR_\beta z$ follows.

Suppose now $xR_\alpha y$ and $yR_\beta z$. We show $xR_\beta z$. Let $\varphi \in z$ such that $\langle\beta\rangle\varphi \in \Delta$. By $yR_\beta z$, we know $\langle\beta\rangle\varphi \in y$, whence $\beta < \alpha$ and $xR_\alpha y$ give us $\langle\beta\rangle\varphi \in x$. Hence, item (i) is established. Consider any $\langle\beta\rangle\varphi \in \Delta$ such that $\langle\beta\rangle\varphi \in z$. Since $yR_\beta z$, we have $\langle\beta\rangle\varphi \in y$, whence $\alpha < \beta$ gives us again $\langle\beta\rangle\varphi \in x$, i.e., item (ii) holds. Now let $\lambda < \beta$ and consider a $\langle\lambda\rangle\varphi \in \Delta$ such that $\langle\lambda\rangle\varphi \in x$. We know

$$\langle\lambda\rangle\varphi \in x \iff \langle\lambda\rangle\varphi \in y \iff \langle\lambda\rangle\varphi \in z,$$

which entails item (iii). For item (iv), we know that there is a $\langle\beta\rangle\psi \in y \setminus z$. But then $\langle\beta\rangle\psi \in x$ as $\alpha < \beta$ and $xR_\alpha y$. This proves $xR_\beta z$. ☐

Lemma 3.4.13. *Let $\langle\alpha\rangle\varphi \in \Delta$ and x be a maximal J^*_Λ-consistent subset from Δ. Then $\langle\alpha\rangle\varphi \in x$ iff there exists a maximal J^*_Λ-consistent subset $y \subseteq \Delta$ such that $xR_\alpha y$ and $\varphi \in y$.*

Proof. For the direction from right to left, suppose $\langle n\rangle\varphi \notin x$. If $y \subseteq \Delta$ is a maximal J^*-consistent set such that $xR_\alpha y$, we clearly have $\varphi \notin y$ by definition of R_α.

For the other direction, assume $\langle\alpha\rangle\varphi \in x$. We will construct a maximal J^*_Λ-consistent $y \subseteq \Delta$ such that $\varphi \in y$ and $xR_\alpha y$. In the following, given any finite set Γ of formulas, we write $\Gamma^\wedge$ ($\Gamma^\vee$, respectively) for the conjunction (disjunction, respectively) of all formulas in Γ. Similarly, we write ${\sim}\Gamma$ for $\{{\sim}\gamma \mid \gamma \in \Gamma\}$ and $\langle\beta\rangle\Gamma$ for $\{\langle\beta\rangle\gamma \mid \gamma \in \Gamma\}$. Now let Σ be the union of the following sets of formulas (modalities range over Λ):

$$\Sigma_1 = \{\neg\langle\lambda\rangle\psi, {\sim}\psi \mid \neg\langle\alpha\rangle\psi \in x, \lambda \geq \alpha\}, \quad \Sigma_2 = \{\langle\beta\rangle\psi \mid \langle\beta\rangle\psi \in x, \beta < \alpha\},$$
$$\Sigma_3 = \{\neg\langle\beta\rangle\psi \mid \neg\langle\beta\rangle\psi \in x, \beta < \alpha\}, \qquad \Sigma_4 = \{\neg\langle\alpha\rangle\varphi, \varphi\}.$$

We claim that Σ is J^*_Λ-consistent. For if not then

$$\mathsf{J}^*_\Lambda \vdash (\Sigma_1^\wedge \wedge \Sigma_2^\wedge \wedge \Sigma_3^\wedge) \to (\varphi \to \langle\alpha\rangle\varphi)$$

and so by propositional logic

$$\mathsf{J}^*_\Lambda \vdash (\varphi \wedge \neg\langle\alpha\rangle\varphi) \to ({\sim}\Sigma_1^\vee \vee {\sim}\Sigma_2^\vee \vee {\sim}\Sigma_3^\vee) \tag{3.1}$$

and furthermore

$$\langle\alpha\rangle(\varphi \wedge \neg\langle\alpha\rangle\varphi) \to \langle\alpha\rangle({\sim}\Sigma_1^\vee \vee {\sim}\Sigma_2^\vee \vee {\sim}\Sigma_3^\vee).$$

By Löb's axiom, we know that

$$\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\varphi \to \langle\alpha\rangle(\varphi \wedge \neg\langle\alpha\rangle\varphi).$$

For $i = 1, 2, 3$, let $\Pi_i = \langle\alpha\rangle{\sim}\Sigma_i$. We have by propositional logic, (3.1), and Lemma 3.2.7

$$\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\varphi \to \Pi_1^\vee \vee \Pi_2^\vee \vee \Pi_3^\vee.$$

Now since x is maximal J^*_Λ-consistent and $\langle\alpha\rangle\varphi \in x$, we infer that $\chi \in x$ for some $\chi \in \Pi_1 \cup \Pi_2 \cup \Pi_3$. We distinguish three cases.

Case 1: $\chi \in \Pi_1$, i.e., $\chi = \langle\alpha\rangle\langle\lambda\rangle\psi$ for some $\lambda \geq \alpha$ or $\chi = \langle\alpha\rangle\psi$. In both cases $\neg\langle\alpha\rangle\psi \in x$ by construction. We have by axiom (vi) and $\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\langle\alpha\rangle\delta \to \delta$ for all δ, that

$$\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\langle\lambda\rangle\psi \to \langle\alpha\rangle\psi,$$

whence $\langle\alpha\rangle\psi \in x$ follows in both cases, contradicting $\neg\langle\alpha\rangle\psi \in x$.

Case 2: $\chi \in \Pi_2$, i.e., $\chi = \langle\alpha\rangle\neg\langle\beta\rangle\psi$ for some ψ and $\beta < \alpha$ such that $\langle\beta\rangle\psi \in x$. Since $\beta < \alpha$, we know by axiom (v) that

$$\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\neg\langle\beta\rangle\psi \to \neg\langle\beta\rangle\psi,$$

whence $\neg\langle\beta\rangle\psi \in x$, contradicting the consistency of x.

Case 3: Finally, suppose $\chi \in \Pi_3$, i.e., $\sigma = \langle\alpha\rangle\langle\beta\rangle\psi$ for some $\beta < \alpha$ such that $\neg\langle\beta\rangle\psi \in x$. By an easy application of axiom (v) we know that

$$\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\langle\beta\rangle\psi \to \langle\beta\rangle\psi,$$

whence we immediately obtain $\langle\beta\rangle\psi \in x$, contradiction.

We see that Σ is J^*_Λ-consistent. Hence, by Lemma 3.4.11, there exists a maximal J^*_Λ-consistent $y \supseteq \Sigma$. Furthermore, $xR_\alpha y$ and $\varphi \in y$ by construction of Σ. ☐

Now define a Kripke model $\mathfrak{K}_\Delta = \langle \mathfrak{F}_\Delta, [\![\cdot]\!]\rangle$, where

$$\mathfrak{K}_\Delta, x \Vdash p^\alpha \iff_{df} p^\alpha \in x,$$

for all variables $p^\alpha \in \Delta$ and $x \in W$.

Lemma 3.4.14. *$\mathfrak{K}_\Delta$ is a strongly persistent, finite J^*_Λ-model.*

Proof. Again, finiteness is immediate. Let $\alpha \in \ell(\Delta)$ and consider a propositional variable $p^\lambda \in \Delta$ such that $\lambda \leq \alpha$. Suppose $xR_\alpha y$ and $y \in [\![p^\lambda]\!]$. We show $x \in [\![p^\lambda]\!]$ by checking the conditions of Definition 3.4.6. Indeed, since $p^\lambda \in \Delta$ we have $\langle\alpha\rangle p^\lambda \in \Delta$. Furthermore, since x is maximal J^*_Λ-consistent and $xR_\alpha y$, it follows by Lemma 3.4.13 that $\langle\alpha\rangle p^\lambda \in x$, whence $p^\lambda \in x$ follows since $\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle p^\lambda \to p^\lambda$ and x is maximal J^*_Λ-consistent. Hence, $\mathfrak{K}_\Delta$ is persistent.

Now let $\lambda < \alpha$ and suppose $xR_\alpha y$ but $y \notin [\![p^\lambda]\!]$. If $p^\lambda \notin \Delta$ then $x \notin [\![p^\lambda]\!]$ by definition, so suppose $p^\lambda \in \Delta$. From $\neg p^\lambda \in \Delta$, we obtain $\langle\alpha\rangle\neg p^\lambda \in \Delta$, whence $xR_\alpha y$ gives us $\langle\alpha\rangle\neg p^\lambda \in x$ by Lemma 3.4.13. Since $\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\neg p^\lambda \to \neg p^\lambda$ and x is maximal J^*_Λ-consistent, we obtain $\neg p^\lambda \in x$, i.e., $x \notin [\![p^\lambda]\!]$ and thus strong persistence as desired. ☐

Lemma 3.4.15. *For all $\varphi \in \Delta$ we have $\mathfrak{K}_\Delta, x \Vdash \varphi$ iff $\varphi \in x$.*

Proof. We proceed by induction on φ. The base cases hold by definition—note that $\top \in x$ and $\bot \notin x$ for all maximal J^*_Λ-consistent $x \in W$. Suppose $\varphi = \psi_1 \wedge \psi_2$. Then $x \Vdash \varphi$ iff $x \Vdash \psi_1$ and $x \Vdash \psi_2$, which by inductive hypothesis is equivalent to $\psi_1, \psi_2 \in x$ and, since x is maximal J^*_Λ-consistent, holds iff $\psi_1 \wedge \psi_2 \in x$. The other propositional connectives are treated similarly. Finally, suppose $\varphi = \langle\alpha\rangle\psi$ for some $\alpha \in \ell(\Delta)$ and some ψ. Then if $x \Vdash \langle\alpha\rangle\psi$, there is a $y \in W$ such that $y \Vdash \psi$, whence by inductive hypothesis we obtain $\psi \in y$ and Lemma 3.4.13 gives us $\langle\alpha\rangle\psi \in x$. Conversely, if $\langle\alpha\rangle\psi \in x$ by Lemma 3.4.13 there exists a $y \in W$ such that $xR_\alpha y$ and $\psi \in y$, whence by inductive hypothesis $y \Vdash \psi$ and so $x \Vdash \langle\alpha\rangle\psi$ follows. □

We can now conclude completeness of J^*_Λ in a standard way.

Theorem 3.4.16. J^*_Λ *is complete for the class of finite strongly persistent* J^*_Λ*-models.*

Proof. Consider a formula $\varphi \in L^*_\Lambda$ and suppose $\mathsf{J}^*_\Lambda \nvdash \varphi$. Then $\{\sim\varphi\}$ is J^*_Λ-consistent. Consider the finite adequate $\Delta := Cl(\{\varphi\})$ and let $\Sigma := \ell(\Delta)$. Let $\mathfrak{K}_\Delta$ be the corresponding finite and strongly persistent J^*_Σ-model. Then, making use of Lemma 3.4.11, there is a maximal J^*_Σ-consistent $x \subseteq \Delta$ such that $\sim\varphi \in x$, whence Lemma 3.4.15 gives us $\mathfrak{K}_\Delta, x \nVdash \varphi$. Notice that $\Sigma \subseteq \Lambda$. Expand $\mathfrak{K}_\Delta$ to a J^*_Λ-model $\mathfrak{K}$ by setting $R_\alpha = \varnothing$ for all $\alpha \in \Lambda \setminus \Sigma$. Then it is immediate that $\mathfrak{K}$ is a finite and strongly persistent J^*_Λ-model such that $\mathfrak{K}, x \nVdash \varphi$. □

Corollary 3.4.17. J^*_Λ *is decidable for every* $\Lambda \subseteq \omega + 1$*.*

Furthermore, we can already conclude that J^* is conservative over its fragments.

Corollary 3.4.18. *Let* $\Lambda \subseteq \omega$*. For all* $\varphi \in L^*_\Lambda$*,*

$$\mathsf{J}^* \vdash \varphi \iff \mathsf{J}^*_\Lambda \vdash \varphi.$$

Proof. The direction from right to left is clear. For the other direction, suppose $\mathsf{J}^*_\Lambda \nvdash \varphi$. Then $\{\sim\varphi\}$ is J^*_Λ-consistent and a similar argument as in the proof of Theorem 3.4.16 yields a J^*-model $\mathfrak{K}$ and a world x such that $\mathfrak{K}, x \nVdash \varphi$. □

3.5 Arithmetical Completeness of GLP*

This section is devoted to a proof of the arithmetical completeness theorem for GLP^*. We closely follow the construction provided by Beklemishev [6] which is close to the original construction of Solovay for GL. We have no Kripke semantics for GLP^* at hand. Therefore, we aim at reducing GLP^* to J^*, which is reminiscent of Solovay's reduction of S to GL for the proof of the arithmetical completeness theorem for S [1, 11, 39]. To this end, for any many-sorted formula from L^*, we define formulas $M(\varphi)$ and $M^+(\varphi)$ as follows [6]. Let $\langle m_1\rangle\varphi_1, \langle m_2\rangle\varphi_2, \ldots, \langle m_s\rangle\varphi_s$ be an enumeration of all subformulas of φ of the form $\langle k\rangle\psi$ and let $n := \max_{i \leq s} m_i$. Define

$$M(\varphi) := \bigwedge_{\substack{1 \leq i \leq s \\ m_i < j \leq n}} (\langle j\rangle\varphi_i \to \langle m_i\rangle\varphi_i),$$

and, furthermore,

$$M^+(\varphi) \coloneqq M(\varphi) \wedge \bigwedge_{i \leq n} [i] M(\varphi).$$

By the monotonicity axioms, it is clear that $\mathsf{GLP}^* \vdash M^+(\varphi)$.

Before turning our attention to the arithmetical completeness proof, let us first restrict the class of models we have to consider for an unprovable formula in J^* (see also Beklemishev [6]).

Lemma 3.5.1. *For any $\varphi \in L^*$, if $\mathsf{J}^* \nvdash \varphi$ then there is a finite and strongly persistent J^*-model $\mathfrak{K}$ with root r such that $\mathfrak{K}, r \nVdash \varphi$.*

Proof. Assume that $\mathsf{J}^* \nvdash \varphi$ and let $\mathfrak{K}_0 = \langle W_0, \{R_n\}_{n<\omega}, [\![\cdot]\!]\rangle$ be a finite and strongly persistent Kripke model such that $\mathfrak{K}_0, x_0 \nVdash \varphi$. Define $\mathfrak{K} = \langle W, \{R_n\}_{n<\omega}, [\![\cdot]\!]\rangle$, where we set $y \in W$ iff $y = x_0$ or there is a sequence of elements $x_1, x_2, \ldots, x_{k+1}$ such that for some $n_0, n_1, \ldots, n_k$ we have

$$x_0 R_{n_0} x_1 R_{n_1} x_2 R_{n_2} \cdots R_{n_k} x_{k+1} = y.$$

Furthermore, let the valuation of $\mathfrak{K}$ agree with that of $\mathfrak{K}_0$ (on the corresponding nodes). We can easily convince ourselves that $\mathfrak{K}$ is a finite and strongly persistent J^*-model. Furthermore, we can easily prove by induction on ψ that

$$\forall x \in W \colon \mathfrak{K}, x \Vdash \psi \iff \mathfrak{K}_0, x \Vdash \psi.$$

Finally, we stipulate that x_0 is a root of $\mathfrak{K}$. Indeed, consider any $y \in W \setminus \{x_0\}$ and suppose

$$x_0 R_{n_0} x_1 R_{n_1} x_2 R_{n_2} \cdots R_{n_k} x_{k+1} = y.$$

By induction on k, by the property

$$u R_n v R_m w \implies u R_{\min\{n,m\}} w,$$

it follows that $x_0 R_s y$, where $s = \min\{n_0, n_1, \ldots, n_k\}$. □

Theorem 3.5.2. *Let T be a sound axiomatizable extension of* PA *and π a strong sequence of provability predicates over T of which every provability predicate is sound. Then, for all many-sorted formulas φ, the following statements are equivalent:*

(i) $\mathsf{GLP}^* \vdash \varphi$*;*

(ii) $\mathsf{J}^* \vdash M^+(\varphi) \to \varphi$*;*

(iii) $T \vdash f_\pi(\varphi)$*, for all realizations f_π.*

Proof. The direction from (ii) to (i) is immediate since GLP^* extends J^* and $\mathsf{GLP}^* \vdash M^+(\varphi)$. Furthermore, the direction from (i) to (iii) is the arithmetical soundness of GLP^* (Proposition 3.3.4). We show that (iii) implies (ii) by assuming the contrapositive, i.e., assume that $\mathsf{J}^* \nvdash M^+(\varphi) \to \varphi$. Then there is a finite and strongly persistent J^*-model $\mathfrak{K} = \langle W, \{R'_n\}_{n<\omega}, [\![\cdot]\!]\rangle$ with root r such that $\mathfrak{K}, r \Vdash M^+(\varphi)$ and $\mathfrak{K}, r \nVdash \varphi$. Without loss of generality, assume that $W = \{1, 2, \dots, N\}$ for some $N \geq 1$ and $r = 1$. We define a new model $\mathfrak{K}_0 = \langle W_0, \{R_n\}_{n<\omega}, [\![\cdot]\!]\rangle$, where

(i) $W_0 = \{0\} \cup W$;

(ii) $R_0 = \{(0, x) \mid x \in W\} \cup R'_0$;

(iii) $R_k = R'_k$, for $k > 0$;

(iv) $\mathfrak{K}_0, 0 \Vdash p \iff_{df} \mathfrak{K}, 1 \Vdash p$, for all variables p.

Notice that $\mathfrak{K}_0$ is still a finite and strongly persistent J^*-model such that $\mathfrak{K}_0, r \nVdash M^+(\varphi) \to \varphi$. Define the following auxiliary notions:

$$\begin{aligned}
R_k(x) &\coloneqq \{y \mid xR_ky\},\\
R^*_k(x) &\coloneqq \{y \mid y \in R_i(x),\ \text{for some } i \geq k\},\\
R^\circ_k(x) &\coloneqq R^*_k(x) \cup \bigcup\{R^*_k(z) \mid x \in R^*_{k+1}(z)\}.
\end{aligned}$$

We are now going to construct an arithmetical realization f_π such that $T \nvdash f_\pi(\varphi)$. Let m be the least number such that $R_m \neq \varnothing$ and $R_k = \varnothing$ for all $k > m$. We define Solovay functions $h_n \colon \omega \to W_0$ for all $n \leq m$ and use their properties to construct such an f_π which witnesses $T \nvdash f_\pi(\varphi)$. In the following, let $\mathsf{Prf}_0, \mathsf{Prf}_1, \dots, \mathsf{Prf}_n, \dots$ be the sequence of proof relations of the respective provability predicates $[0]_\pi, [1]_\pi, \dots, [n]_\pi, \dots$ over T.

Definition 3.5.3. For all $n \leq m$, define a function $h_n \colon \omega \to W_0$ as follows:

$$h_0(0) = 0 \text{ and } h_n(0) = \ell_{n-1}, \text{ for } n > 0;$$

$$h_n(x+1) = \begin{cases} y, & \text{if } h_n(x)R_nz \text{ and } \mathsf{Prf}_n(\ulcorner\neg S_z\urcorner, x),\\ h_n(x), & \text{otherwise.}\end{cases}$$

Let $\ell_k = x$ be a formalization of the statement that the function h_k (defined by H_k) has as its *limit at* x, i.e.,

$$\ell_k = x \iff_{df} \exists N_0\, \forall n \geq N_0\, H_k(n, x).$$

For $x \in W_0$, S_x denotes the sentence $\ell_m = \overline{x}$. ⊣

We now show that the concepts defined in the previous definition are well-defined in formal arithmetic. First of all, we need to construct formulas $H_k(x, y)$ for $k = 0, 1, \dots, m$ which define the corresponding functions h_k and provably satisfy the clauses of their definitions. Notice that $\ulcorner S_z\urcorner$ is a primitive recursive function of $\ulcorner H_m\urcorner$ and z. Let $\mathsf{notlim}(y, x)$ be a term for the function which, given the Gödel number of a formula $F(a, b)$ and an x, returns the Gödel number of the sentence which asserts that the function defined by $F(a, b)$ has no limit at $\overline{x}$. Hence, if $F(a, b)$ defines a function and has Gödel number n, then the value of $\mathsf{notlim}(\overline{n}, \overline{k})$ will be the Gödel number of $\neg\exists N_0\, \forall n \geq N_0\, F(n, \overline{k})$, asserting that the function defined by $F(a, b)$ has no limit at $\overline{k}$ (cf. Boolos [11]). Now let $A_0(w, x, y)$ be the arithmetical formula which naturally formalizes the following statement:

> There is a finite sequence s of length $a + 1$ such that $s_0 = 0$ and $s_x = y$ and the following conditions hold for all $a < x$:
>
> (i) Whenever $s_a = i$ for an $i \leq N$ and $\mathsf{Prf}_0(\mathsf{notlim}(w, j), a)$ for some j such that iR_0j, we have that $s_{a+1} = j$.
>
> (ii) Whenever $\neg\mathsf{Prf}_0(\mathsf{notlim}(w, \overline{j}), a)$ for all j such that iR_0j, we have that $s_{a+1} = s_a$.

For $k > 0$, let $A_k(w, l, x, y)$ be an arithmetical formula which expresses the following statement:

> There is a finite sequence s of length $a + 1$ such that $s_0 = l$ and $s_a = y$ and the following conditions hold for all $a < x$:
>
> (i) Whenever $s_a = i$ for an $i \leq N$ and $\mathsf{Prf}_k(\mathsf{notlim}(w, \overline{j}), a)$ for some j such that iR_kj, we have that $s_{a+1} = j$.
>
> (ii) Whenever $\neg\mathsf{Prf}_k(\mathsf{notlim}(w, \overline{j}), a)$ for all j such that iR_kj, we have that $s_{a+1} = s_a$.

By our assumptions on the predicates $\mathsf{Prf}_n(\alpha, y)$ (identical proof sequences cannot code proofs of different formulas), we can prove that $\exists! y\, A_0(w, x, y)$ by induction on x. The statement $\ell_0 = \overline{z}$ is expressible via $A_0(w, x, y)$ and ℓ_0 can then be shown to be unique. Suppose now that we have shown that ℓ_0 exists (which we will do below). Then we can use the formula $A_0(w, x, y)$ to express the statement $A_1(w, \ell_0, x, y)$. Similarly, we successively continue to obtain $A_{k+1}(w, \ell_k, x, y)$. In the end, we will obtain a formula $A'_m(w, x, y)$ from which we can infer by diagonalization that there is a formula $H_m(x, y)$ such that

$$T \vdash H_m(x, y) \leftrightarrow A'_m(\ulcorner H_m(x, y) \urcorner, x, y).$$

Performing converse substitutions then yields definitions of the formulas H_k for all $k < m$.

For a set $A \subseteq W_0$, we denote by $\ell_k \in A$ the sentence $\bigvee_{x \in A} \ell_k = \overline{x}$. Furthermore, given such an $A \subseteq W_0$, we use quantifiers to naturally abbreviate statements of the form

$$\bigwedge_{a \in A} \psi(\overline{a}), \quad \bigvee_{a \in A} \psi(\overline{a})$$

by $\forall a \in A\colon \psi(\overline{a})$ and $\exists a \in A\colon \psi(\overline{a})$ (or stylistic variations thereof), respectively.

Lemma 3.5.4. *For all $k \geq 0$,*

(i) $T \vdash \forall x\, \exists! w \in W_0\colon H_k(x, \overline{w})$;

(ii) $T \vdash \exists! w \in W_0\colon \ell_k = \overline{w}$;

(iii) $T \vdash \forall i, j\, \forall z \in W_0\, (i < j \wedge h_k(i) = \overline{z} \to h_k(j) \in R_k(z) \cup \{z\})$;

(iv) $T \vdash \forall z \in W_0\, (\exists x\, h_k(x) = \overline{z} \to \ell_m \in R_k^*(z) \cup \{z\})$.

Proof. Item (i) follows from our previous discussion. For (ii), note that uniqueness easily follows from (i). For existence, we prove that

$$T \vdash H_k(a, \overline{b}) \to \ell_k = \overline{b} \vee \ell_k \in R_k(b),$$

by induction on the converse of R_k. So suppose that for each $c \in R_k(b)$, we have

$$T \vdash H_k(a, \overline{c}) \to \ell_k = \overline{c} \vee \ell_k \in R_k(c).$$

By definition of H_k, we know that

$$T \vdash H_k(a, \overline{b}) \to \forall x \geq a\, (H_k(x, \overline{b}) \vee \exists w \in R_k(b) \colon H_k(x, \overline{w})).$$

By inductive hypothesis, we obtain

$$T \vdash H_k(a, \overline{b}) \to \forall x \geq a\, (H_k(x, \overline{b}) \vee \exists w \in R_k(b) \colon \ell_k = \overline{w} \vee \ell_k \in R_k(w)).$$

Hence,

$$T \vdash H_k(a, \overline{b}) \to \ell_k = \overline{b} \vee \exists w \in R_k(b) \colon \ell_k = \overline{w} \vee \ell_k \in R_k(w).$$

Since R_k is transitive, we obtain

$$T \vdash H_k(a, \overline{b}) \to \ell_k = \overline{b} \vee \ell_k \in R_k(b),$$

as required. We know that $T \vdash H_0(0, 0)$ and so (ii) follows for $k = 0$. Hence, ℓ_0 exists. By induction, we infer that $T \vdash H_k(0, \ell_{k-1})$ for all $k > 0$ and so (ii) is proved. Items (iii) and (iv) are immediate consequences of the definitions of the formulas H_k. □

Lemma 3.5.5. *The following conditions hold for the sentences S_x:*

(i) $T \vdash \bigvee_{x \in W_0} S_x$ *and* $T \vdash \neg(S_x \wedge S_y)$ *for all* $x \neq y$*;*

(ii) $T \vdash S_x \to \langle k \rangle_\pi S_y$*, for all* y *such that* $x R_k y$*;*

(iii) $T \vdash S_x \to [k]_\pi (\ell_m \in R_k^\circ(x))$*, for all* $x \neq 0$*;*

(iv) $\mathbb{N} \models S_0$.

Proof. Item (i) is just a special case of item (ii) of Lemma 3.5.4. Item (ii) is proved by formalizing the following argument in T. Assume S_x. Then we either have $\ell_k = x$ or $\ell_k \in R_{k+1}^*(x)$. By the properties of a J^*-model, in both cases it holds that $R_k(\ell_k) = R_k(x)$. Let n_0 be such that $\forall n \geq n_0 \colon h_k(n) = \ell_k$. Now consider a y such that $x R_k y$. Suppose $[k]_\pi \neg S_y$. Then there is an $n_1 \geq n_0$ such that $\mathsf{Prf}_k(\ulcorner \neg S_y \urcorner, n_1)$, whence $h_k(n_1) = y$ follows by definition of h_k, a contradiction.

For (iii), we formalize the following argument in T. Assume S_x, where $x \neq 0$ and let $z \in W_0$ be such that $\ell_k = \overline{z}$. By definition, we have $x \in R^*_{k+1}(z)$ or $x = z$. Hence, $R^*_k(z) \subseteq R^\circ_k(x)$ and, since this property is definable by a Δ_0-formula, $[k](R^*_k(z) \subseteq R^\circ_k(x))$. So,

$$[k]_\pi(\ell_m \in R^*_k(z)) \to [k]_\pi(\ell_m \in R^\circ_k(x)).$$

We know that $\exists n\, h_k(n) = z$ and, being a $\Sigma_{|\pi_k|+1}$-formula, we have

$$[k]_\pi(\exists n\, h_k(n) = \overline{z}).$$

But for any $w \in W_0$, we have

$$T \vdash \exists n\, h_k(n) = \overline{w} \to \ell_m \in R^*_k(w) \cup \{w\}.$$

Therefore,

$$T \vdash [k]_\pi(\exists n\, h_k(n) = \overline{w}) \to [k]_\pi(\ell_m \in R^*_k(w) \cup \{w\}).$$

Reasoning in T, we obtain that $[k]_\pi(\ell_m \in R^*_k(z) \cup \{z\})$. It remains to notice that $\overline{z} \neq 0$ since, by assumption, $x \neq 0$. But then $\exists n\, h_k(n) = \overline{z}$ implies that $[k]_\pi \neg S_z$ which means that $[k]_\pi(\ell_m \neq \overline{z})$. It follows that $[k]_\pi(\ell_m \in R^*_k(z))$ and therefore $[k]_\pi(\ell_m \in R^\circ_k(x))$ as required.

To establish (iv), we prove by induction on k that $\mathbb{N} \models \ell_k = 0$ for all $k \leq m$. For $k = 0$, if $\mathbb{N} \models \ell_0 = \overline{z}$ for some $z \neq 0$, then $[0]_\pi \neg S_z$ which by the soundness of $[k]_\pi$ yields $\ell_0 \neq \overline{z}$ in the standard model, a contradiction. The induction step is then based on a similar argument, taking into account that $h_{k+1}(0) = \ell_k = 0$ holds in the standard model. $\square$

Lemma 3.5.6. *For all $k < m$, provably in T,*

(i) *either $\ell_k = \ell_{k+1}$ or $\ell_k R_{k+1} \ell_{k+1}$;*

(ii) *if $k < n \leq m$ then either $\ell_k = \ell_n$ or $\ell_k R_j \ell_n$ for some $j \in (k, n]$.*

Proof. Item (i) is clear from our previous considerations. Item (ii) is proved by an external induction on n from (i). $\square$

Now we define a realization f_π as follows:

$$f_\pi \colon p \longmapsto \bigvee_{x \Vdash p} S_x.$$

In the following, we assume that we are given a natural arithmetization of the forcing relation for the model $\mathfrak{K}_0$ by bounded formulas.

Lemma 3.5.7. *For any variable p of sort $k \leq m$, provably in T,*

$$f_\pi(p) \iff \forall w \in W_0 \setminus [\![p]\!] \colon \forall x\, \neg H_k(x, \overline{w}).$$

Proof. For the direction from left to right, we reason in T as follows. Suppose $f_\pi(p)$ and, towards a contradiction, suppose that $\exists x\, h_k(x) = \overline{w}$ for some $w \in W_0$ such that $w \nVdash p$. By item (iv) of Lemma 3.5.4, we know that, provably in T,

$$\exists x\, h_k(x) = \overline{u} \implies S_u \vee \bigvee_{z \in R_k^*(u)} S_z,$$

for any $z \in W_0$. In particular, we infer that

$$S_w \vee \bigvee_{u \in R_k^*(w)} S_u.$$

Since $\mathfrak{K}_0$ is strongly persistent and $w \nVdash p$, we know that $u \nVdash p$ for all $u \in R_k^*(w)$. This contradicts $f_\pi(p)$ by item (i) of Lemma 3.5.5.

For the other direction, we reason in T as follows. Suppose the right-hand side of the equivalence. We certainly know that $\neg S_u$ for all $u \in W_0$ such that $u \nVdash p$. Now if $\ell_k = \ell_m$, then, by item (i) of Lemma 3.5.5, S_x for some $x \in W_0$ such that $x \Vdash p$ and we are thus finished. So suppose that $\ell_k \neq \ell_m$. We know that $\ell_k \in [\![p]\!]$, since $\forall x\, h_k(x) \neq \overline{w}$ for all $w \in [\![\neg p]\!]$. Assume now that $\ell_m \in [\![\neg p]\!]$. By Lemma 3.5.6, there must be a $j \in (k, m]$ such that $\ell_k R_j \ell_m$. By strong persistence, for any $x, y \in W_0$ such that $x R_j y$, it holds that

$$y \nVdash p \implies x \nVdash p.$$

Thus, $\ell_m \in [\![\neg p]\!]$ is impossible and therefore $\ell_m \in [\![p]\!]$ by item (i) of Lemma 3.5.5. $\square$

Lemma 3.5.8. *For every variable p^k, where $k < \omega$, $f_\pi(p^k)$ is $\Pi_{|\pi_k|+1}$ in T.*

Proof. Notice that $H_k(x, y)$ is $\Delta_{|\pi_k|+1}$ in T, since Prf_k is $\Pi_{|\pi_k|}$ in T and, moreover, $T \vdash \forall x \exists y!\, H_k(x, y)$. Now if $k > m$ then the sentence $f_\pi(p^k)$ is a disjunction of sentences which are $\Sigma_{|\pi_k|+2}$ in T. Since $T \vdash \exists! w \in W_0 \colon \ell_m = \overline{w}$ (item (i) of Lemma 3.5.5), we know that, provably in T,

$$f_\pi(p^k) \iff \bigvee_{x \Vdash p} S_x \iff \bigwedge_{x \nVdash p} \neg S_x,$$

i.e., $f_\pi(p^k)$ is $\Pi_{|\pi_k|+2}$ in T as required, since it is provably equivalent to a conjunction of sentences which are $\Pi_{|\pi_k|+2}$ in T.

If $k \leq m$, then by Lemma 3.5.7 we know that, provably in T,

$$f_\pi(p) \iff \forall w \in W_0 \setminus [\![p]\!] \colon \forall x\, \neg H_k(x, \overline{w}),$$

which is visibly $\Pi_{|\pi_k|+1}$ in T. $\square$

Therefore, f_π defines an arithmetical realization in the sense of Definition 3.3.2.

Lemma 3.5.9. *For each subformula χ of φ and each $x \neq 0$,*

(i) *if $\mathfrak{K}_0, x \Vdash \chi$ then $T \vdash S_x \to f_\pi(\chi)$;*

(ii) *if* $\mathfrak{K}_0, x \nVdash \chi$ *then* $T \vdash S_x \to \neg f_\pi(\chi)$.

Proof. We prove both statements simultaneously by induction on χ. If χ is a propositional variable, $\top$, or $\bot$, the claims follow by the definition of f_π and item (i) of Lemma 3.5.5.

Suppose that $\chi = \tau_1 \wedge \tau_2$, then $\mathfrak{K}_0, x \Vdash \tau_1 \wedge \tau_2$ implies that $\mathfrak{K}_0, x \Vdash \tau_i$ for $i = 1, 2$, whence by inductive hypothesis we infer $T \vdash S_x \to f_\pi(\tau_1)$ and $T \vdash S_x \to f_\pi(\tau_2)$ and thus $T \vdash S_x \to f_\pi(\tau_1 \wedge \tau_2)$ follows. If $\mathfrak{K}_0, x \nVdash \tau_1 \wedge \tau_2$, then either $\mathfrak{K}_0, x \nVdash \tau_1$ or $\mathfrak{K}_0, x \nVdash \tau_2$. By inductive hypothesis, we either have $T \vdash S_x \to \neg f_\pi(\tau_1)$ or $T \vdash S_x \to \neg f_\pi(\tau_2)$. Therefore, $T \vdash S_x \to \neg f_\pi(\tau_1) \vee \neg f_\pi(\tau_2)$ and so $T \vdash S_x \to \neg f_\pi(\tau_1 \wedge \tau_2)$. The other propositional connectives are treated similarly.

Suppose $\chi = \langle k \rangle \tau$ and assume $\mathfrak{K}_0, x \Vdash \langle k \rangle \tau$. Then there is a $y \in W_0 \setminus \{0\}$ such that $x R_k y$ and $\mathfrak{K}_0, y \Vdash \tau$. By inductive hypothesis, we have $T \vdash S_y \to f_\pi(\tau)$, whence

$$T \vdash \langle k \rangle_\pi S_y \to \langle k \rangle_\pi f_\pi(\tau).$$

By item (ii) of Lemma 3.5.5, we obtain that $T \vdash S_x \to \langle k \rangle_\pi S_y$ and so $T \vdash S_x \to \langle k \rangle_\pi f_\pi(\tau)$ as desired.

Suppose now that $\mathfrak{K}_0, x \nVdash \langle k \rangle \tau$. We prove that $\mathfrak{K}_0, y \nVdash \tau$ for all $y \in R_k^\circ(x)$. If $y \in R_k^\circ(x)$, then for some z we have $x \in R_{k+1}^*(z) \cup \{z\}$ and $y \in R_k^*(z)$. Clearly, for all $w \in R_k(x)$, we have $\mathfrak{K}_0, w \nVdash \tau$. Notice that $R_k(x) = R_k(z)$ and, therefore, $\mathfrak{K}_0, z \nVdash \langle k \rangle \tau$. Furthermore, $\mathfrak{K}_0, z \Vdash \langle j \rangle \tau \to \langle k \rangle \tau$ for all j such that $k < j \leq m$. Hence, $\mathfrak{K}_0, z \nVdash \langle j \rangle \tau$ for every such j. It follows that $\mathfrak{K}_0, y \nVdash \tau$ as desired. By inductive hypothesis, we know that $T \vdash S_y \to \neg f_\pi(\tau)$ for all $y \in R_k^\circ(x)$ and thus,

$$T \vdash \ell_m \in R_k^\circ(x) \to \neg f_\pi(\tau),$$

whence,

$$T \vdash [k]_\pi(\ell_m \in R_k^\circ(x)) \to [k]_\pi \neg f_\pi(\tau),$$

which, using item (iii) of Lemma 3.5.5, implies

$$T \vdash S_x \to [k]_\pi \neg f_\pi(\tau).$$

That is, $T \vdash S_x \to \neg \langle k \rangle_\pi f_\pi(\tau)$. □

In particular, $T \vdash S_r \to \neg f_\pi(\varphi)$. Furthermore, $T \vdash S_0 \to \neg[0]_\pi \neg S_r$. Now if $T \vdash f_\pi(\varphi)$ then $T \vdash \neg S_r$, whence $T \vdash [0]_\pi \neg S_r$ and so $T \vdash \neg S_0$, whence by the soundness of T, we get $\mathbb{N} \not\models S_0$ which contradicts item (iv) of Lemma 3.5.5. Therefore, $T \nvdash f_\pi(\varphi)$. □

As in the work of Beklemishev [6], one can obtain an arithmetical completeness theorem for a many-sorted truth provability logic. More precisely, let GLPS^* denote the logic which consists of the set of theorems of GLP^* extended by the schema $\varphi \to \langle n \rangle \varphi$ $(n \geq 0)$ and with modus ponens as its sole rule of inference. Let $\langle n_1 \rangle \varphi_1, \ldots, \langle n_s \rangle \varphi_s$ be an enumeration of all subformulas from φ of the form $\langle k \rangle \psi$. Let

$$H(\varphi) := \bigwedge_{i=1}^{s} (\varphi_i \to \langle n_i \rangle \varphi_i).$$

Theorem 3.5.10. *Let T be a sound axiomatizable extension of PA and π a strong sequence of provability predicates over T of which every provability predicate is sound. Then, for all many-sorted formulas φ, the following statements are equivalent:*

(i) $\mathsf{GLPS}^* \vdash \varphi$;

(ii) $\mathsf{GLP}^* \vdash H(\varphi) \to \varphi$;

(iii) $\mathbb{N} \models f_\pi(\varphi)$, *for all realizations* f_π.

Proof. The direction from (ii) to (i) is clear since $\mathsf{GLPS}^* \vdash H(\varphi)$. The direction from (i) to (iii) is easy to see since T is sound. We prove that (iii) implies (ii) again by assuming the contrapositive, i.e., suppose $\mathsf{GLP}^* \nvdash H(\varphi) \to \varphi$. Then $\mathsf{J}^* \nvdash M^+(H(\varphi) \to \varphi) \to (H(\varphi) \to \varphi)$ and so there is a J^*-model $\mathfrak{K} = \langle W, \{R_n\}_{n<\omega}, [\![\cdot]\!]\rangle$ which is strongly persistent and $\mathfrak{K}, r \Vdash M^+(\varphi) \wedge H(\varphi)$ but $\mathfrak{K}, r \nVdash \varphi$ for a root r. As before, identify $W = \{1, 2, \ldots, N\}$ for some $N \geq 1$ and let $r = 1$. Construct a model $\mathfrak{K}_0$ which is defined as in the proof of Theorem 3.5.2.

All the lemmas in the proof of Theorem 3.5.2 hold without change except that we need to supplement Lemma 3.5.9 by the following statement. The proof is as that of Beklemishev [6].

Lemma 3.5.11. *For each subformula χ of φ we have*

(i) *if* $\mathfrak{K}_0, 0 \Vdash \chi$ *then* $T \vdash S_x \to f_\pi(\chi)$;

(ii) *if* $\mathfrak{K}_0, 0 \nVdash \chi$ *then* $T \vdash S_x \to \neg f_\pi(\chi)$;

Proof. By induction on χ. We prove both statements simultaneously. The only difference to the proof of Lemma 3.5.9 is the proof of item (ii) in the case where $\chi = \langle k \rangle \tau$.

Suppose $\mathfrak{K}_0, 0 \nVdash \langle k \rangle \tau$. We know that $\mathfrak{K}_0, 0 \nVdash \varphi$ since $\mathfrak{K}_0, 0 \Vdash H(\varphi)$. Furthermore, since $\mathfrak{K}_0, 0 \Vdash M(\varphi)$, it holds that $\mathfrak{K}_0, 0 \nVdash \langle j \rangle \tau$ for all $j \geq k$. Therefore, $\mathfrak{K}_0, x \nVdash \tau$ for all $x \in R_k^*(0) \cup \{0\}$. By induction hypothesis and Lemma 3.5.9, we have

$$T \vdash \ell_m \in R_k^*(0) \cup \{0\} \to \neg f_\pi(\tau),$$

whence it follows that

$$T \vdash [k]_\pi(\ell_m \in R_k^*(0) \cup \{0\}) \to [k]_\pi \neg f_\pi(\tau).$$

Furthermore,

$$\begin{aligned} T \vdash S_0 &\to \exists n\, h_k(n) = 0 \\ &\to [k]_\pi(\exists n\, h_k(n) = 0). \end{aligned}$$

Thus, $T \vdash S_0 \to [k]_\pi(\ell_m \in R_k(0) \cup \{0\})$ by Lemma 3.5.4 and so $T \vdash S_0 \to [k]_\pi \neg f_\pi(\tau)$, i.e., $T \vdash S_0 \to \neg \langle k \rangle_\pi f_\pi(\tau)$ as required. □

Now $\mathfrak{K}_0, 0 \nVdash \varphi$ yields $T \vdash S_0 \to \neg f_\pi(\varphi)$ which by $\mathbb{N} \models S_0$ and soundness gives us $\mathbb{N} \not\models f_\pi(\varphi)$ as desired. □

Corollary 3.5.12. GLPS^* *is decidable.*

Notice that Theorem 3.5.2 yields a reduction from GLP^* to J^*. However, the formula $M^+(\varphi)$ is, in a sense, inconvenient since its size does not depend on the size of φ and, additionally, $M^+(\varphi)$ is not necessarily in the language of φ. We borrow a result from Beklemishev et al. [8] to improve upon that. Let $\langle m_1\rangle\varphi_1, \langle m_2\rangle\varphi_2, \ldots, \langle m_s\rangle\varphi_s$ be an enumeration of all subformulas of φ of the form $\langle k\rangle\psi$ such that $i < j$ implies $m_i \leq m_j$. Define

$$N(\varphi) \coloneqq \bigwedge_{\substack{1\leq i\leq s\\ i<j\leq s}} (\langle m_j\rangle\varphi_j \to \langle m_i\rangle\varphi_i).$$

Furthermore, let

$$N^+(\varphi) \coloneqq N(\varphi) \wedge \bigwedge_{1\leq i\leq s} [m_i]\varphi.$$

Lemma 3.5.13. *Let $\varphi \in L^*_\Lambda$, where $\Lambda \subseteq \omega$. Then,*

$$\mathsf{GLP}^*_\Lambda \vdash \varphi \iff \mathsf{J}^*_\Lambda \vdash N^+(\varphi) \to \varphi.$$

Proof. The direction from right to left is immediate, since $N^+(\varphi)$ is in L^*_Λ and $\mathsf{GLP}^*_\Lambda \vdash N^+(\varphi)$. For the other direction, suppose that $\mathsf{J}^*_\Lambda \nvdash N^+(\varphi) \to \varphi$. Then there is a finite and strongly persistent J^*_Λ-model $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!]\rangle$ with root r such that $\mathfrak{K}, r \Vdash N^+(\varphi)$ and $\mathfrak{K}, r \nVdash \varphi$. Expand $\mathfrak{K}$ to a J^*-model, call it $\mathfrak{K}'$, by setting $R_\alpha = \varnothing$ for all $\alpha \in \omega \setminus \Lambda$. Notice that $\mathfrak{K}'$ forces the same formulas from L^*_Λ at every of its point as $\mathfrak{K}$. In particular, $\mathfrak{K}', r \nVdash \varphi$. We show that $\mathfrak{K}', r \Vdash M^+(\varphi)$. Let $i \in \{1, \ldots, s\}$ and consider any j such that $m_i < j \leq n$. Now $\mathfrak{K}', r \Vdash \langle j\rangle\varphi_i$ only if $j = m_k$ for some $k = 1, \ldots, s$. In this case, $\mathfrak{K}', r \Vdash \langle m_i\rangle\varphi$ since $\mathfrak{K}', r \Vdash N^+(\varphi)$. Otherwise, if $j \neq m_k$ for all $k = 1, \ldots, s$, then trivially $\mathfrak{K}', r \Vdash \langle j\rangle\varphi_i \to \langle m_i\rangle\varphi_i$, since $\mathfrak{K}', r \nVdash \langle j\rangle\varphi_i$ due to the fact that $R_j = \varnothing$. Let $n \coloneqq \max_{i\leq s} m_i$ and consider any $i \leq n$. Similarly as before, $\mathfrak{K}', r \Vdash [i]M(\varphi)$ if $i = m_k$ for some $k = 1, \ldots, s$. If not then trivially $\mathfrak{K}', r \Vdash [i]M(\varphi)$. Hence, $\mathfrak{K}', r \Vdash M^+(\varphi)$ and so $\mathsf{J}^* \nvdash M^+(\varphi) \to \varphi$, whence $\mathsf{GLP}^* \nvdash \varphi$ and thus $\mathsf{GLP}^*_\Lambda \nvdash \varphi$ follows. ☐

Let φ be a formula from L^* and let $p_1, \ldots, p_k$ exhaust all variables from φ and let $\alpha_1, \ldots, \alpha_k$ be their respective sorts. Furthermore, let $\Theta \subseteq \omega$ be a set of modalities. Define

$$R_\Theta(\varphi) \coloneqq \bigwedge_{i=1}^{k} \bigwedge (\{\langle j\rangle p_i \to p_i \mid j \in \Theta, j \geq \alpha_i\} \cup \{\langle j\rangle\neg p_i \to \neg p_i \mid j \in \Theta, j > \alpha_i\})$$

and

$$R^+_\Theta(\varphi) \coloneqq R_\Theta(\varphi) \wedge \bigwedge_{j\in\Theta} [j]R_\Theta(\varphi).$$

Lemma 3.5.14. *Let $\varphi \in L^*_\omega$ and let Θ be the set of all modalities occurring in φ. Then,*

$$\mathsf{GLP}^* \vdash \varphi \iff \mathsf{GLP} \vdash R^+_\Theta(\varphi) \to \varphi.$$

Proof. The direction from right to left is immediate since $\mathsf{GLP}^* \vdash R^+_\Theta(\varphi)$ and GLP^* extends GLP. For the other direction, suppose $\mathsf{GLP} \nvdash R^+_\Theta(\varphi) \to \varphi$. It follows from results of Beklemishev [6] together with a result by Beklemishev et al. [8] that this implies

$$\mathsf{J} \nvdash N^+(R^+_\Theta(\varphi) \to \varphi) \to (R^+_\Theta(\varphi) \to \varphi). \tag{3.2}$$

Beklemishev [5] showed that J is complete with respect to the class of all J^*-models (there called J-models). So let $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha<\omega}, [\![\cdot]\!]\rangle$ be a J^*-model with root r such that

$$\mathfrak{K}, r \nVdash N^+(R^+_\Theta(\varphi) \to \varphi) \to (R^+_\Theta(\varphi) \to \varphi).$$

Therefore, $\mathfrak{K}, r \Vdash N^+(R^+_\Theta(\varphi) \to \varphi)$ and $\mathfrak{K}, r \Vdash R^+_\Theta(\varphi)$. Now it follows that $\mathfrak{K}, r \Vdash N^+(\varphi)$. Since $\mathfrak{K}, r \Vdash R^+_\Theta(\varphi)$ and $\mathfrak{K}$ is rooted, it is easy to see that $\mathfrak{K}$ is strongly persistent by the construction of $R^+_\Theta(\varphi)$. (Notice that $\mathfrak{K}$ can be chosen such that $R_\alpha = \varnothing$ for all $\alpha \notin \Theta$, since the formula depicted in (3.2) is in the language of φ.) Therefore, $\mathsf{J}^* \nvdash N^+(\varphi) \to \varphi$ and so $\mathsf{GLP}^* \nvdash \varphi$ follows. □

We say that a logic $\mathcal{L}$ has the *Craig interpolation property* if, whenever $\mathcal{L} \vdash \varphi \to \psi$, then there is an η containing only variables which are present in φ and ψ such that both $\mathcal{L} \vdash \varphi \to \eta$ and $\mathcal{L} \vdash \eta \to \psi$.

Corollary 3.5.15. GLP^* *has the Craig interpolation property.*

Proof. Suppose $\mathsf{GLP}^* \vdash \varphi \to \psi$. Let Θ be the set of all modalities from $\varphi \to \psi$. By Lemma 3.5.14, we have

$$\mathsf{GLP} \vdash R^+_\Theta(\varphi \to \psi) \to (\varphi \to \psi).$$

Note that $R^+_\Theta(\varphi \to \psi)$ is equivalent in GLP to $R^+_\Theta(\varphi) \wedge R^+_\Theta(\psi)$. Hence,

$$\mathsf{GLP} \vdash R^+_\Theta(\varphi) \wedge R^+_\Theta(\psi) \to (\varphi \to \psi),$$

whence by propositional logic

$$\mathsf{GLP} \vdash R^+_\Theta(\varphi) \wedge \varphi \to (R^+_\Theta(\psi) \to \psi).$$

Ignatiev [24] showed that GLP has the Craig interpolation property. Hence, there is an η containing only variables which occur in $R^+_\Theta(\varphi) \wedge \varphi$ and $R^+_\Theta(\psi) \to \psi$ such that

$$\mathsf{GLP} \vdash R^+_\Theta(\varphi) \wedge \varphi \to \eta \quad \text{and} \quad \mathsf{GLP} \vdash \eta \to (R^+_\Theta(\psi) \to \psi).$$

But $\mathsf{GLP}^* \vdash R^+_\Theta(\varphi)$ and $\mathsf{GLP}^* \vdash R^+_\Theta(\psi)$. Therefore, $\mathsf{GLP}^* \vdash \varphi \to \eta$ and $\mathsf{GLP}^* \vdash \eta \to \psi$. Note that η only contains variables which occur in φ and ψ, since $R^+_\Theta(\tau)$ contains exactly the variables from τ, for any formula τ. □

Corollary 3.5.16. *Deciding whether* $\mathsf{GLP}^* \vdash \varphi$ *is* PSPACE*-complete.*

Proof. Shapirovsky [34] showed that deciding whether $\mathsf{GLP} \vdash \varphi$ is complete for PSPACE. Thus, the claim follows by Lemma 3.5.14 and the fact that the size of $R^+_\Theta(\varphi)$ (where Θ is the set of modalities from φ) is polynomially bounded by the size of φ. □

CHAPTER 4

Many-Sorted Reflection Calculi

In this chapter we continue to study positive calculi which allow for a richer arithmetical interpretation than the full language of GLP^*. After defining our basic formalism in Section 4.1 we continue to define our arithmetical interpretation in Section 4.2. Section 4.3 treats Kripke semantics and Section 4.4 establishes the relationship between our many-sorted calculi and the positive fragment[1] of GLP^*. The arithmetical completeness of our positive calculus is proved in Section 4.5.

4.1 Motivation and Basics

Positive fragments of modal logics were first studied by Dunn [15].[2] Dashkov [13] brought the study of positive modal logics into the realm of provability logics which is motivated by the fact that an ordinal analysis proposed by Beklemishev [2] makes only use of the positive fragment of GLP. In particular, Dashkov showed that the positive fragment of GLP can be axiomatized by a purely positive calculus and that the question of theoremhood in this calculus can be decided in polynomial time. As pointed out by Dashkov, the restriction of GLP to the positive fragment allows one to interpret propositional variables as theories instead of single sentences. Furthermore, these theories need neither be finitely axiomatizable nor of bounded arithmetical complexity. This permits the introduction of new modal operators which axiomatize stronger properties than n-consistency for every $n < \omega$. More precisely, in the positive setting it is well-defined to define an operator $\langle \omega \rangle$ on which maps any theory T to the full uniform reflection scheme for T.

Recently, Beklemishev [7] investigated such positive calculi (which he calls *reflection calculi*) with an additional modality $\langle \omega \rangle$. He showed that the calculi he introduced are decidable in polynomial time and that they are arithmetically complete with respect

[1] If $\mathcal{L}$ is a modal logic, then its positive fragment is defined to be its theorems of the form $\varphi \to \psi$, where φ and ψ are positive formulas in the sense of Definition 4.1.1.

[2] Dunn actually identifies positive fragments of modal logics to also contain the connectives $\Box$ and $\vee$ besides those in Definition 4.1.1.

to the interpretation of $\langle\omega\rangle$ as the full uniform reflection schema in arithmetic. Propositional variables are there interpreted as primitive recursive enumerations of theories extending PA. We will continue this line of research and define a family of positive *many-sorted reflection calculi* in the sequel. In our elaboration, we mainly follow the papers of Beklemishev [7] and Dashkov [13].

Definition 4.1.1. Let $\Lambda \subseteq \omega + 1$ be a signature. (*Positive*) *formulas* (*over* Λ) and their associated *sorts* are defined inductively as follows:

(i) $\top$ is a positive formula of sort 0.

(ii) Every propositional variable p^α is a positive formula of sort α.

(iii) If A and B are positive formulas of sorts α and β then $(A \wedge B)$ is a positive formula of sort $\max\{\alpha, \beta\}$.

(iv) If A is a positive formula (of any sort) and $\alpha \in \Lambda$, then $\langle\alpha\rangle A$ is a positive formula of sort α.

For any positive formula A, we denote its sort by $|A|$. When considering positive formulas, we write αA instead of $\langle\alpha\rangle A$. ⊣

Definition 4.1.2. Let Λ be a signature. A *sequent* (*over* Λ) is an expression of the form $A \Rightarrow B$, where A and B are positive formulas over Λ. ⊣

Given any $\Lambda \subseteq \omega+1$, we denote by L^+_Λ the set of all positive formulas over Λ. Furthermore, we denote by L^+ the set of all positive formulas over $\omega+1$. The notion of (general) *substitution* is defined as in the case of unrestricted modal languages (cf. Definition 3.2.2). The notation we agreed upon directly carries over into the positive setting. Note in particular that substitution is defined as to respect the corresponding sorts. For a substitution τ and a sequent $\gamma = A \Rightarrow B$, we define $\gamma^\tau := A^\tau \Rightarrow B^\tau$.

The following axiom schemes and rules of inference are called *propositional axioms* and *propositional rules*, respectively.

(i) $A \Rightarrow A$; $A \Rightarrow \top$;

(ii) $A \wedge B \Rightarrow A$; $A \wedge B \Rightarrow B$;

(iii) if $A \Rightarrow B$ and $A \Rightarrow C$ then infer $A \Rightarrow B \wedge C$;

(iv) if $A \Rightarrow B$ and $B \Rightarrow C$ then infer $A \Rightarrow C$.

In this chapter, a (*positive*) *logic over* Λ will be a set of sequents over Λ which is closed under substitutions, is closed under all propositional rules, and under the rule

(v) if $A \Rightarrow B$ then infer $\alpha A \Rightarrow \alpha B$, for any $\alpha \in \Lambda$.

We again use the standard notation that the subscript "Λ" in $\mathcal{L}_\Lambda$ indicates that $\mathcal{L}_\Lambda$ is a positive logic over Λ.

Definition 4.1.3. The logic RC^*_Λ is given by the postulates (i) to (v) as well as the following axiom schemes and rules of inference (modalities range over Λ):

(vi) $\alpha A \Rightarrow A$, where $|A| \leq \alpha$ (α-*persistence*);

(vii) $\alpha A \Rightarrow \beta A$, for $\beta < \alpha$ (*monotonicity*);

(viii) $\alpha A \wedge B \Rightarrow \alpha(A \wedge B)$, where $|B| < \alpha$.

The logic RJ^*_Λ is obtained from RC^*_Λ by dropping monotonicity but adding the axiom scheme

(ix) $\beta\alpha A \Rightarrow \beta A$, for $\beta \leq \alpha$.

We set $\mathsf{RC}^* := \mathsf{RC}^*_{\omega+1}$ and $\mathsf{RJ}^* := \mathsf{RJ}^*_{\omega+1}$. ⊣

Let $\mathcal{L}_\Lambda$ be any logic. As usual, given a sequent γ, a *proof of* γ *in* $\mathcal{L}_\Lambda$ is a finite sequence of sequents $\gamma_1, \dots, \gamma_n$ such that $\gamma_n = \gamma$ and for $i = 1, \dots, n$, γ_i is either an axiom or follows from previous elements of the sequence by an application of a rule. In this case, γ is called *provable* (*in* $\mathcal{L}_\Lambda$), which we denote by $\mathcal{L}_\Lambda \vdash \gamma$. When exhibiting proofs in a logic $\mathcal{L}_\Lambda$, we often write $\mathcal{L}_\Lambda \vdash A_1 \Rightarrow A_2 \Rightarrow \cdots \Rightarrow A_n$ to express that the sequence $A_1 \Rightarrow A_2, A_2 \Rightarrow A_3, \dots, A_{n-1} \Rightarrow A_n$ is (part of) a proof in $\mathcal{L}_\Lambda$. In this notation, we usually refer to some previously derived results which will then be clear from context. Given any logic $\mathcal{L}_\Lambda$ and a set of formulas $\Gamma \subseteq L^+_\Lambda$, we write $\mathcal{L}_\Lambda \vdash \Gamma \Rightarrow B$ if there exist $A_1, \dots, A_n \in \Gamma$ such that $\mathcal{L}_\Lambda \vdash A_1 \wedge \cdots \wedge A_n \Rightarrow B$.[3] For logics as defined above, we have a statement related to Proposition 3.2.4.

Proposition 4.1.4. *Suppose* $\mathcal{L}_\Lambda \vdash A \Rightarrow B$ *and* $|A|, |B| \leq \alpha$. *Then,* $\mathcal{L}_\Lambda \vdash C(p^\alpha/A) \Rightarrow C(p^\alpha/B)$ *for any* $C \in L^+_\Lambda$.

Proof. As in the case of Proposition 3.2.4, by induction on C. □

Note that if $\beta < \alpha$, then $\mathsf{RJ}^*_\Lambda \vdash \alpha A \wedge \beta B \Rightarrow \alpha(A \wedge \beta B)$, for any formula B. Furthermore, note that $\mathsf{RC}^*_\Lambda \vdash \alpha\beta A \Rightarrow \beta A$, for $\alpha \leq \beta$. Furthermore, if $\alpha \geq \beta$ then $\mathsf{RC}^*_\Lambda \vdash \alpha A \Rightarrow \beta A$ by monotonicity (and $A \Rightarrow A$), whence $\mathsf{RC}^*_\Lambda \vdash \beta\alpha A \Rightarrow \beta\beta A \Rightarrow \beta A$. We thus immediately obtain:

Lemma 4.1.5. RC^*_Λ *extends* RJ^*_Λ.

Example 4.1.6. Let A and B be a formulas such that $|B| < \alpha$. We know that $\mathsf{RC}^* \vdash \alpha A \wedge B \Rightarrow \alpha(A \wedge B)$. But also $\mathsf{RC}^* \vdash \alpha B \Rightarrow B$. Hence, $\mathsf{RC}^* \vdash \alpha A \wedge \alpha B \Rightarrow \alpha(A \wedge B)$. Now if $|A| < \alpha$ and B is an arbitrary formula then

$$\mathsf{RC}^* \vdash \alpha B \wedge \alpha A \Rightarrow \alpha A \wedge \alpha B \Rightarrow \alpha(A \wedge B).$$

Hence, whenever $|A| < \alpha$ or $|B| < \alpha$ then $\mathsf{RC}^* \vdash \alpha A \wedge \alpha B \Rightarrow \alpha(A \wedge B)$. ⊣

[3] The empty conjunction is defined to be $\top$.

4.2 Arithmetical Interpretation

As in Beklemishev [7], propositional variables will be realized via *primitive recursive numerations* of theories extending PA. Recall that $\mathsf{Ax}_{\mathsf{PA}}(\alpha)$ denotes a bounded formula which arithmetically defines (the set of Gödel numbers of) the axioms of PA. Furthermore, recall that we assume that we let PA contain function symbols and predicate symbols for all primitive recursive functions and predicates, respectively.

Definition 4.2.1. A (*primitive recursive*) *numeration* is a bounded formula which arithmetically defines the Gödel numbers of the axioms of an extension of PA. We say that σ *numerates* S. A numeration σ numerates a Π_{n+1}-*axiomatized extension of* PA if

$$\mathsf{PA} \vdash \forall\alpha\,(\sigma(\alpha) \to \mathsf{Ax}_{\mathsf{PA}}(\alpha) \vee \alpha \in \Pi_{n+1}),$$

where the expression "$\alpha \in \Pi_{n+1}$" denotes a natural bounded formula (possibly with an additional parameter n) which expresses that α is the Gödel number of a Π_{n+1}-sentence (see Hájek and Pudlák [21]). ⊣

The notion of a numeration strongly coincides with the concepts from Chapter 2 where for a theory T, we considered a formula $\mathsf{Ax}_T(\alpha)$ which arithmetically defines the axioms of T. The formalized notion of theoremhood in T was then naturally formulated by a predicate $\Box_T(\alpha)$. Analogously, for a numeration σ we denote by $\Box_\sigma(\alpha)$ the standard formula (the *provability predicate of* σ) which arithmetically defines the theorems of the theory numerated by σ. For numerations σ and τ, we write $\sigma \Rightarrow_{\mathsf{PA}} \tau$ if

$$\mathsf{PA} \vdash \forall\alpha\,(\Box_\tau(\alpha) \to \Box_\sigma(\alpha)),$$

and $\sigma \Rightarrow \tau$ if

$$\mathbb{N} \models \forall\alpha\,(\Box_\tau(\alpha) \to \Box_\sigma(\alpha)),$$

i.e., if the theory numerated by τ proves every theorem of the theory numerated by σ. We assume that every numeration provably in PA numerates an extension of PA, i.e., $\tau \Rightarrow_{\mathsf{PA}} \mathsf{Ax}_{\mathsf{PA}}$ for any τ, i.e., that every theory numerated by some numeration provably extends PA. As usual, we write $\Box_\sigma\varphi$ instead of $\Box_\sigma(\ulcorner\varphi\urcorner)$ if no confusion arises.

We denote by $\mathsf{Con}(\sigma)$ the sentence $\neg\Box_\sigma\bot$. Furthermore, as we did in Section 2.5, we denote by $\mathsf{Con}_n(\sigma)$ the formula which expresses that the theory numerated by σ is n-consistent. (Recall that n-consistency is equivalent over PA to the uniform reflection principle for Π_{n+1}-formulas; see Section 2.5.) We stress that $\mathsf{Con}_0(\sigma)$ is provably equivalent in PA to $\mathsf{Con}(\sigma)$. Furthermore, $\mathsf{Con}_n(\sigma)$ can be formalized to be a formula depending on n. Hence, we often regard $\ulcorner\mathsf{Con}_n(\sigma)\urcorner$ as a definable term which depends on n. We use these facts without adhering to any special notation.

Given any arithmetical sentence φ, we denote by $\underline{\varphi}$ the numeration

$$\mathsf{Ax}_{\mathsf{PA}}(\alpha) \vee \alpha = \ulcorner\varphi\urcorner,$$

which numerates the theory $\mathsf{PA} + \varphi$. In this setting, for any numeration σ, $\underline{\mathsf{Con}}_n(\sigma)$ numerates the theory $\mathsf{PA} + \mathsf{Con}_n(\sigma)$. Let σ numerate S. The scheme

$$\mathsf{Con}_\omega(\sigma): \qquad \{\mathsf{Con}_n(\sigma) \mid n \in \omega\}$$

is equivalent over PA to the uniform reflection principle for S. Overloading notation, we denote by $\underline{\mathsf{Con}}_\omega(\sigma)$ a numeration which numerates the theory $\mathsf{PA} + \mathsf{Con}_\omega(\sigma)$.

Suppose that σ numerates a finite extension of PA of the form $\mathsf{PA} + \psi$ for a sentence ψ. Suppose $\underline{\psi} \Rightarrow_{\mathsf{PA}} \underline{\varphi}$. Then, since PA is sound, we obtain $\underline{\psi} \Rightarrow \underline{\varphi}$ and so $\mathsf{PA} + \psi \vdash \varphi$. Conversely, if $\mathsf{PA} + \psi \vdash \varphi$ then $\mathsf{PA} \vdash \Box_{\mathsf{PA}}(\psi \to \varphi)$, whence by the formalized deduction theorem $\mathsf{PA} \vdash \Box_{\underline{\psi}}\, \varphi$. Therefore, by an argument formalizable in PA, we also have $\underline{\psi} \Rightarrow_{\mathsf{PA}} \underline{\varphi}$ and so $\underline{\psi} \Rightarrow \underline{\varphi}$. Hence,

$$\mathsf{PA} + \psi \vdash \varphi \iff \underline{\psi} \Rightarrow_{\mathsf{PA}} \underline{\varphi} \iff \underline{\psi} \Rightarrow \underline{\varphi}.$$

In particular, if σ numerates a finite extension of PA of the form $\mathsf{PA} + \{\varphi_1, \dots, \varphi_n\}$ then in order to establish $\sigma \Rightarrow_{\mathsf{PA}} \underline{\psi}$, it is sufficient to establish $\mathsf{PA} + \{\varphi_1, \dots, \varphi_n\} \vdash \psi$.

Definition 4.2.2. An *arithmetical realization* is a function $\cdot^*$ from positive formulas to numerations such that the following conditions are satisfied:

(i) $\top^* = \mathsf{Ax}_{\mathsf{PA}}$;

(ii) for every propositional variable p of sort α, p^* is a numeration which numerates (1) a $\Pi_{\alpha+1}$-axiomatized extension of PA in case $\alpha < \omega$ and (2) an arbitrary extension of PA in case $\alpha = \omega$;

(iii) $(A \wedge B)^* = A^* \vee B^*$;

(iv) $(\alpha A)^* = \underline{\mathsf{Con}}_\alpha(A^*)$, for $\alpha \leq \omega$.

We say that A^* is the *translation of A under $\cdot^*$*. ⊣

Lemma 4.2.3. *Let $\cdot^*$ be an arithmetical realization and A a formula such that $|A| < \omega$. Then A^* numerates a $\Pi_{|A|+1}$-axiomatized extension of* PA.

Proof. By an easy induction on A. The cases for propositional variables and $\top$ are clear. For the induction step, notice that for $n < \omega$, $\mathsf{Con}_n(\sigma)$ provably belongs to Π_{n+1} for any numeration σ. Furthermore, provably in PA, if φ belongs to Π_m, then also to Π_n for $n > m$. Using these facts, the claim easily follows. □

Recall that $\mathsf{True}_{\Pi_{n+1}}(x)$ denotes a truth-definition for Π_{n+1}-formulas. In particular, $\mathsf{PA} \vdash \varphi \leftrightarrow \mathsf{True}_{\Pi_{n+1}}(\ulcorner\varphi\urcorner)$ for all Π_{n+1}-sentences φ. This fact can be formalized uniformly in n [7], i.e.,

$$\mathsf{PA} \vdash \forall n\, \forall \alpha \in \Pi_{n+1}\, \Box_{\mathsf{PA}}(\alpha \leftrightarrow \mathsf{True}_{\Pi_{n+1}}(\dot{\alpha})). \tag{4.1}$$

Lemma 4.2.4. *For $n \in \omega$, $\mathsf{Con}_n(\sigma)$ is provably equivalent in* PA *to*

$$\forall \alpha \in \Pi_{n+1}\, (\Box_\sigma(\alpha) \to \mathsf{True}_{\Pi_{n+1}}(\alpha)).$$

Proof. Let σ numerate S. We show that the formula presented above is equivalent to $\mathsf{RFN}_{\Pi_{n+1}}(S)$ over PA. Let $\varphi(x)$ be a Π_{n+1}-formula. Then,

$$\begin{aligned}\mathsf{PA} + \forall\alpha \in \Pi_{n+1}\,(\Box_\sigma(\alpha) \to \mathsf{True}_{\Pi_{n+1}}(\alpha)) \vdash \Box_\sigma\varphi(\dot{x}) &\to \mathsf{True}_{\Pi_{n+1}}(\ulcorner\varphi(\dot{x})\urcorner)\\ &\to \varphi(x).\end{aligned}$$

Conversely,

$$\begin{aligned}\mathsf{PA} + \mathsf{RFN}_{\Pi_{n+1}}(S) \vdash \alpha \in \Pi_{n+1} \wedge \Box_\sigma(\alpha) &\to \Box_\sigma\mathsf{True}_{\Pi_{n+1}}(\dot\alpha)\\ &\to \mathsf{True}_{\Pi_{n+1}}(\alpha),\end{aligned}$$

since $\mathsf{True}_{\Pi_{n+1}}(x)$ is a Π_{n+1}-formula. This proves the claim. ☐

The following lemma generalizes Corollary 2.5.2.

Lemma 4.2.5. *Let σ numerate S and φ be a Π_{n+1}-sentence. If $S \vdash \varphi$ then $\mathsf{PA} + \mathsf{Con}_n(\sigma) \vdash \varphi$. Moreover, this statement is formalizable in* PA *uniformly in n, i.e.,*

$$\mathsf{PA} \vdash \forall n\,\forall\alpha \in \Pi_{n+1}\,(\Box_\sigma(\alpha) \to \Box_{\underline{\mathsf{Con}}_n(\sigma)}(\alpha)).$$

Proof. The informal version is an easy consequence of Lemma 2.5.11. Indeed, $\mathsf{PA} + \mathsf{Con}_n(\sigma) \vdash \Box_\sigma\varphi \to \varphi$, whence by $\mathsf{PA} \vdash \Box_\sigma\varphi$ we obtain $\mathsf{PA} + \mathsf{Con}_n(\sigma) \vdash \varphi$ as desired.

For the formalized version, reason in PA as follows. Let $\alpha \in \Pi_{n+1}$ and suppose $\Box_\sigma(\alpha)$. We know $\Box_{\mathsf{PA}}(\dot\alpha \in \Pi_{n+1} \wedge \Box_\sigma(\dot\alpha))$. By the previous lemma, we know that

$$\Box_{\underline{\mathsf{Con}}_n(\sigma)}\forall\beta \in \Pi_{n+1}\,(\Box_\sigma(\beta) \to \mathsf{True}_{\Pi_{n+1}}(\beta)).$$

In particular,

$$\Box_{\underline{\mathsf{Con}}_n(\sigma)}(\dot\alpha \in \Pi_{n+1} \wedge \Box_\sigma(\dot\alpha) \to \mathsf{True}_{\Pi_{n+1}}(\dot\alpha)),$$

whence $\Box_{\underline{\mathsf{Con}}_n(\sigma)}\mathsf{True}_{\Pi_{n+1}}(\dot\alpha)$ and thus $\Box_{\underline{\mathsf{Con}}_n(\sigma)}(\alpha)$ follows by (4.1). ☐

Corollary 4.2.6. *Let σ be a numeration and $n < \omega$. Then $\underline{\mathsf{Con}}_n(\sigma) \Rightarrow_{\mathsf{PA}} \sigma$, whenever σ numerates a Π_{n+1}-axiomatized extension of* PA.

Proof. We reason in PA as follows. Suppose $\Box_\sigma(\varphi)$ and reason by induction on proof length of φ. The only interesting case is the case when $\varphi \in \Pi_{n+1}$. By the previous lemma, we obtain $\Box_{\underline{\mathsf{Con}}_n(\sigma)}(\varphi)$. Hence, $\underline{\mathsf{Con}}_n(\sigma) \Rightarrow_{\mathsf{PA}} \sigma$ as required. ☐

Corollary 4.2.7. *For any numeration σ, $\underline{\mathsf{Con}}_\omega(\sigma) \Rightarrow_{\mathsf{PA}} \sigma$.*

Proof. Note that

$$\begin{aligned}\mathsf{PA} \vdash \Box_\sigma(\alpha) &\to \exists n\,(\alpha \in \Pi_{n+1} \wedge \Box_\sigma(\alpha))\\ &\to \exists n\,\Box_{\underline{\mathsf{Con}}_n(\sigma)}(\alpha)\\ &\to \Box_{\underline{\mathsf{Con}}_\omega(\sigma)}(\alpha).\end{aligned}$$

Further notice that we used formalizations of the facts that every sentence is PA-provably equivalent to a Π_{n+1}-sentence for some n and that $\ulcorner\mathsf{Con}_n(\sigma)\urcorner$ can be constructed primitive recursively from the parameter n. ☐

Lemma 4.2.8. *Let φ be a Π_{m+1}-sentence and σ a numeration. For $m < n < \omega$ it holds that*

$$\mathsf{PA} \vdash \mathsf{Con}_n(\sigma) \wedge \varphi \to \mathsf{Con}_n(\sigma \vee \underline{\varphi}).$$

Proof. We reason in PA as follows. Suppose $\Box_{\sigma\vee\underline{\varphi}}(\psi)$ for $\psi \in \Pi_{n+1}$. Then $\Box_\sigma(\varphi \to \psi)$ by the formalized deduction theorem. We know that $\varphi \to \psi$ is a Π_{n+1}-sentence since $m < n$. Thus, if $\mathsf{Con}_n(\sigma)$ then also $\mathsf{True}_{\Pi_{n+1}}(\varphi \to \psi)$ and so $\mathsf{True}_{\Pi_{n+1}}(\varphi) \to \mathsf{True}_{\Pi_{n+1}}(\psi)$. But also $\varphi \in \Pi_{n+1}$ and so $\mathsf{True}_{\Pi_{n+1}}(\varphi)$, whence $\mathsf{True}_{\Pi_{n+1}}(\psi)$ follows as required. ☐

Corollary 4.2.9. *Let τ numerate a Π_{m+1}-axiomatized extension of* PA*. Then for any numeration σ,*

$$\underline{\mathsf{Con}_\omega}(\sigma) \vee \tau \Rightarrow_{\mathsf{PA}} \underline{\mathsf{Con}_\omega}(\sigma \vee \tau).$$

Proof. We show an informal version of this statement by an argument formalizable in PA. That is, we must show that for each n,

$$\mathsf{PA} + \underline{\mathsf{Con}_\omega}(\sigma) + \tau \vdash \underline{\mathsf{Con}_n}(\sigma \vee \tau).$$

We may assume $n > m$ and use the previous lemma. A formalization of the corresponding argument yields the proof. ☐

Proposition 4.2.10. RC^* *is arithmetically sound, i.e., if* $\mathsf{RC}^* \vdash A \Rightarrow B$ *then* $A^* \Rightarrow_{\mathsf{PA}} B^*$ *for every arithmetical realization* $\cdot^*$*.*

Proof. By induction on the length of a derivation of $A \Rightarrow B$. Most of the axioms have been handled by the previous lemmas and corollaries. The soundness of the propositional axioms and rules is also obvious. For monotonicity, it is clear that $\underline{\mathsf{Con}_\alpha}(\sigma) \Rightarrow_{\mathsf{PA}} \underline{\mathsf{Con}_\beta}(\sigma)$ for $\alpha > \beta$, since the strength of $\mathsf{Con}_\alpha(\sigma)$ increases with α. Suppose $A^* \Rightarrow_{\mathsf{PA}} B^*$ and let $n < \omega$. We can easily see that $\mathsf{PA} + \mathsf{Con}_n(A^*) \vdash \mathsf{Con}_n(B^*)$, since

$$\begin{aligned}\mathsf{PA} + \mathsf{Con}_n(A^*) \vdash \alpha \in \Pi_{n+1} \wedge \Box_{B^*}(\alpha) &\to \Box_{A^*}(\alpha)\\ &\to \mathsf{True}_{\Pi_{n+1}}(\alpha).\end{aligned}$$

Hence, $\underline{\mathsf{Con}_n}(A^*) \Rightarrow_{\mathsf{PA}} \underline{\mathsf{Con}_n}(B^*)$ follows. Formalizing this fact also establishes that if $A^* \Rightarrow_{\mathsf{PA}} B^*$ then $\underline{\mathsf{Con}_\omega}(A^*) \Rightarrow_{\mathsf{PA}} \underline{\mathsf{Con}_\omega}(B^*)$. ☐

4.3 Kripke Semantics

Since $L^+_\Lambda \subseteq L^*_\Lambda$ (i.e., the set of positive formulas is contained in the set of many-sorted formulas) for any signature Λ, the notion of Kripke model over Λ directly carries over into the positive setting.

Definition 4.3.1. Let Λ be a signature and $\mathfrak{K} = \langle \mathfrak{F}, [\![\cdot]\!] \rangle$ a Kripke model, where $\mathfrak{F} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda} \rangle$, and $x \in W$ a world. A sequent $A \Rightarrow B$ over Λ is *true at x* (notation: $\mathfrak{K}, x \Vdash A \Rightarrow B$) if $\mathfrak{K}, x \Vdash A$ implies $\mathfrak{K}, x \Vdash B$. $A \Rightarrow B$ is (*globally*) *true in* $\mathfrak{K}$ (notation: $\mathfrak{K} \models A \Rightarrow B$) if $\mathfrak{K}, x \Vdash A \Rightarrow B$ for all $x \in W$. Similarly, $A \Rightarrow B$ is *valid in* $\mathfrak{F}$ (notation: $\mathfrak{F} \models A \Rightarrow B$) if it is globally true in every model based on $\mathfrak{F}$. ⊣

Definition 4.3.2. We say that a Kripke frame $\mathfrak{F} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}\rangle$ is an RJ^*_Λ-*frame* if it satisfies the following conditions for all $\alpha, \beta \in \Lambda$:

(i) $\forall x, y, z\,(xR_\alpha y \,\&\, yR_\beta z \Rightarrow xR_\gamma z)$, for $\gamma = \min\{\alpha, \beta\}$;

(ii) $\forall x, y, z\,(xR_\alpha y \,\&\, xR_\beta z \Rightarrow yR_\beta z)$, for $\alpha > \beta$.

An RJ^*_Λ-*model* is a Kripke model based an an RJ^*_Λ-frame. ⊣

Definition 4.3.3. An RC^*_Λ-*frame* is an RJ^*_Λ-frame $\mathfrak{F} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}\rangle$ which is *monotone*, i.e, where $R_\alpha \subseteq R_\beta$ for all $\alpha, \beta \in \Lambda$ such that $\alpha > \beta$. An RC^*_Λ-*model* is a Kripke model based on an RC^*_Λ-frame. ⊣

In case $\Lambda = \omega + 1$, we drop the subscript "Λ" in the terms RJ^*_Λ-frame and RJ^*_Λ-model (similarly for RC^*). Recall that a Kripke model $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!]\rangle$ is

(i) *persistent* if for all $\alpha \in \Lambda$, all propositional variables p^β with $\beta \le \alpha$, and all $x, y \in W$ we have

$$xR_\alpha y \text{ and } y \in [\![p^\beta]\!] \text{ imply } x \in [\![p^\beta]\!];$$

(ii) *strongly persistent* if $\mathfrak{K}$ is persistent and for all $\alpha \in \Lambda$, all propositional variables p^β with $\beta < \alpha$, and all $x, y \in W$ we have

$$xR_\alpha y \text{ and } y \notin [\![p^\beta]\!] \text{ imply } x \notin [\![p^\beta]\!].$$

Lemma 4.3.4. *Let $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!]\rangle$ be an RJ^*_Λ-model. Then, $\mathfrak{K}$ is strongly persistent iff for all formulas $A \in L^+_\Lambda$ and all $\alpha \in \Lambda$ we have*

(i) *if $|A| \le \alpha$ then $xR_\alpha y$ and $y \in [\![A]\!]$ imply $x \in [\![A]\!]$;*

(ii) *if $|A| < \alpha$ then $xR_\alpha y$ and $y \notin [\![A]\!]$ imply $x \notin [\![A]\!]$.*

Proof. The direction from right to left is immediate. The other direction is proved *mutatis mutandis* as Lemma 3.4.7. □

Lemma 4.3.5. *Let $\Lambda \subseteq \omega + 1$ and $\alpha \in \Lambda$. The axiom schemes*

(i) *$\alpha A \Rightarrow A$, where $|A| \le \alpha$, and*

(ii) *$\alpha A \wedge B \Rightarrow \alpha(A \wedge B)$, where $|B| < \alpha$,*

are true in every strongly persistent RJ^*_Λ-model.

Proof. Item (i) is clear by virtue of Lemma 4.3.4. For (ii), let $|B| < \alpha$ and consider an RJ^*_Λ-model $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!]\rangle$. Let $x \in W$ and suppose $x \Vdash \alpha A \wedge B$. Then there is a $y \in W$ such that $xR_\alpha y$ and $y \Vdash A$. Now if $y \nVdash B$, then $x \nVdash B$ by Lemma 4.3.4. Hence, $y \Vdash \alpha(A \wedge B)$. □

Proposition 4.3.6. *Let $\Lambda \subseteq \omega + 1$.*

(i) *RJ^*_Λ is sound for the class of all strongly persistent RJ^*_Λ-models;*

(ii) RC^*_Λ *is sound for the class of all strongly persistent* RC^*_Λ*-models.*

Proof. In both cases by induction on the length of a derivation of a sequent. Most of the axioms are clear from our previous discussion. In particular, the previous lemmas handle the cases of the axioms where we express conditions on the sorts of formulas. We leave the details to the reader. □

Example 4.3.7. Let p and q be variables of sort α. We easily see that $\mathsf{RC}^* \nvdash \alpha p \wedge \alpha q \Rightarrow \alpha(p \wedge q)$. Indeed, consider the model $\mathfrak{K} = \langle \{a, b, c\}, \{R_\alpha\}_{\alpha \leq \omega}, [\![\cdot]\!] \rangle$, where

(i) $R_\alpha = \{(a, b), (a, c)\}$ and $R_\gamma = \varnothing$, for all $\gamma > \alpha$;

(ii) $R_\beta = R_\alpha \cup (\{b, c\} \times \{b, c\})$, for $\beta < \alpha$;

(iii) $[\![p]\!] = \{a, b\}$, $[\![q]\!] = \{a, c\}$;

We see that $\mathfrak{K}$ is a strongly persistent RC^*-model which falsifies $\alpha p \wedge \alpha q \Rightarrow \alpha(p \wedge q)$. By Proposition 4.3.6, we know that $\mathsf{RC}^* \nvdash \alpha p \wedge \alpha q \Rightarrow \alpha(p \wedge q)$. Combining this with Example 4.1.6, we obtain

$$\mathsf{RC}^* \vdash \alpha p \wedge \alpha q \Rightarrow \alpha(p \wedge q) \iff |p| < \alpha \text{ or } |q| < \alpha.$$

In particular, $\mathsf{RC}^* \vdash \omega p \wedge \omega q \Rightarrow \omega(p \wedge q)$ iff $|p| < \omega$ or $|q| < \omega$. ⊣

We continue now to prove that our calculi are complete with respect to certain classes of Kripke models. For a set of formulas Δ, we set $\ell(\Delta) \coloneqq \{\alpha \mid \alpha A \in \Delta \text{ for some } A\}$. We say that a set of formulas Δ is *adequate* if the following conditions are satisfied:

(i) $\top \in \Delta$;

(ii) Δ is closed under subformulas;

(iii) if $\beta A \in \Delta$ and $\beta < \alpha$ for some $\alpha \in \ell(\Delta)$, then $\alpha A \in \Delta$;

(iv) for any variable p^β, if $p^\beta \in \Delta$ and $\alpha \geq \beta$ for some $\alpha \in \ell(\Delta)$ then $\alpha p^\beta \in \Delta$.

It is clear that every finite set Γ can be extended to a finite adequate Γ' such that $\ell(\Gamma) = \ell(\Gamma')$.

Definition 4.3.8. Let Δ be an adequate set and $\Lambda \coloneqq \ell(\Delta)$. An $\mathcal{L}_\Lambda$*-theory in* Δ is a set $\Gamma \subseteq \Delta$ such that if $\mathcal{L}_\Lambda \vdash \Gamma \Rightarrow B$ and $B \in \Delta$, then $B \in \Gamma$. ⊣

Now fix a finite adequate Δ and let $\Lambda \coloneqq \ell(\Delta)$. Consider an arbitrary logic $\mathcal{L}_\Lambda$. We define a Kripke frame $\mathfrak{F}_\Delta = \langle W, \{R_\alpha\}_{\alpha \in \Lambda} \rangle$, where

$$W \coloneqq \{x \mid x \text{ is an } \mathcal{L}_\Lambda\text{-theory in } \Delta\},$$

and for $\alpha \in \ell(\Delta)$, we define $xR_\alpha y$ iff

(i) if $A \in y$ and $\alpha A \in \Delta$ then $\alpha A \in x$;

(ii) if $\beta A \in y$ and $\alpha A \in \Delta$ then $\min\{\alpha, \beta\}A \in x$;

(iii) if $\beta < \alpha$ and $\beta A \in x$ then $\beta A \in y$;

(iv) for all variables p^β, if $\beta < \alpha$ and $p^\beta \notin y$, then $p^\beta \notin x$;

Lemma 4.3.9. *If $\mathcal{L}_\Lambda$ extends RJ^*_Λ then $\mathfrak{F}_\Delta$ is an RJ^*_Λ-frame.*

Proof. Suppose $xR_\alpha y$ and $yR_\beta z$. We prove $xR_\gamma z$ for $\gamma = \min\{\alpha, \beta\}$. Indeed, if $A \in z$ and $\gamma A \in \Delta$, then $\gamma A \in y$, whence $\gamma A \in x$ since $xR_\alpha y$, i.e., item (i) is proved. Suppose $\delta A \in z$ and $\gamma A \in \Delta$. By adequacy, we know $\delta A \in \Delta$, whence $\min\{\beta, \delta\}A \in y$ follows since $yR_\beta z$. By $xR_\alpha y$ we know $\min\{\alpha, \min\{\beta, \delta\}\}A \in x$ and thus $\min\{\gamma, \delta\}A \in x$, since $\min\{\gamma, \delta\} = \min\{\alpha, \min\{\beta, \delta\}\}$. Hence, item (ii) is proved. Let $\delta < \gamma$ and $\delta A \in x$. Then $\delta A \in y$ as $\delta < \alpha$ and $xR_\alpha y$ and thus $\delta A \in z$ since $yR_\beta z$ and $\delta < \beta$. Thus, item (iii) is proved. For item (iv), if $\delta < \gamma$ and $p^\delta \notin z$ then $p^\delta \notin y$ as $yR_\beta z$ and $\delta < \beta$, whence $p^\delta \notin x$ by $xR_\alpha y$ and $\delta < \alpha$.

Suppose $xR_\alpha y$ and $xR_\beta z$ for $\beta < \alpha$. We show $yR_\beta z$. If $A \in z$ and $\beta A \in \Delta$, then $\beta A \in x$ since $xR_\beta z$, whence $\beta A \in y$ as $\beta < \alpha$ and $xR_\alpha y$, i.e., item (i) follows. Let $\gamma A \in z$ and $\beta A \in \Delta$. We show $\min\{\gamma, \beta\}A \in y$. Since $xR_\beta z$, we know $\min\{\gamma, \beta\}A \in x$. Now $\min\{\gamma, \beta\} < \alpha$, whence from $xR_\alpha y$ we obtain $\min\{\gamma, \beta\}A \in y$ which establishes item (ii). Suppose now $\gamma < \beta$ and $\gamma A \in y$. By $xR_\alpha y$ we get $\gamma A \in x$, since $\alpha A \in \Delta$ and $\gamma = \min\{\alpha, \gamma\}$. Since $xR_\beta z$ and $\gamma < \beta$ we infer $\gamma A \in z$ as required. Thus, item (iii) is proved. For item (iv), suppose $\gamma < \beta$ and let $p^\gamma \notin z$. Since $xR_\beta z$ by assumption, we know $p^\gamma \notin x$.

Now if $p^\gamma \in y$ then, by adequacy, we obtain $\alpha p^\gamma \in \Delta$, whence $xR_\alpha y$ gives us $\alpha p^\gamma \in x$. But since $\mathcal{L}_\Lambda \vdash x \Rightarrow \alpha p^\gamma \Rightarrow p^\gamma$, we infer $p^\gamma \in x$, a contradiction. Therefore $p^\gamma \notin y$ and item (iv) is established. ☐

Now define a model $\mathfrak{K}_\Delta = \langle \mathfrak{F}_\Delta, [\![\cdot]\!] \rangle$, where

$$\mathfrak{K}_\Delta, x \Vdash p^\alpha \iff_{df} p^\alpha \in x,$$

for all variables p^α and all $x \in W$.

Lemma 4.3.10. *Let $\mathcal{L}_\Lambda$ extend RJ^*_Λ. For all $A \in \Delta$ we have $\mathfrak{K}_\Delta, x \Vdash A$ iff $A \in x$.*

Proof. By induction on A. If A is a propositional variable or $\top$, the claim is immediate. Suppose $A = B \wedge C$. Then $x \Vdash B \wedge C$ iff $x \Vdash B$ and $x \Vdash C$ which holds by inductive hypothesis iff $B \in x$ and $C \in x$.

Suppose $A = \alpha B$ for some B and suppose $x \Vdash A$. Then $y \Vdash B$ for some $y \in W$ such that $xR_\alpha y$. By inductive hypothesis, we know $B \in y$, whence $\alpha B \in x$ follows by definition of R_α as $\alpha B \in \Delta$. Conversely, suppose $A \in x$. We show $\mathfrak{K}_\Delta, x \Vdash A$. Let

$$\Sigma_1 \coloneqq \{\beta C \mid \beta C \in x, \beta < \alpha\},$$
$$\Sigma_2 \coloneqq \{p^\beta \mid p^\beta \in x, \beta < \alpha\},$$

and let $y \coloneqq \{C \in \Delta \mid \mathcal{L}_\Lambda \vdash \Sigma_1, \Sigma_2, B \Rightarrow C\}$. By inductive hypothesis, we know $y \Vdash B$ as $B \in y$. We prove $xR_\alpha y$. Let $D \in y$ and $\alpha D \in \Delta$. Then $\mathcal{L}_\Lambda \vdash \Gamma_1, \Gamma_2, B \Rightarrow D$ for some finite $\Gamma_1 \subseteq \Sigma_1$, $\Gamma_2 \subseteq \Sigma_2$. We know

$$\begin{aligned}\mathcal{L}_\Lambda \vdash x &\Rightarrow \alpha B \wedge \bigwedge \Gamma_1 \wedge \bigwedge \Gamma_2 \\ &\Rightarrow \alpha(B \wedge \bigwedge \Gamma_1 \wedge \bigwedge \Gamma_2) \quad (\text{since } |\bigwedge \Gamma_2| < \alpha) \\ &\Rightarrow \alpha D.\end{aligned}$$

Thus, $\alpha D \in x$ as required. Let $\beta D \in y$ and $\alpha D \in \Delta$. Again, $\mathcal{L}_\Lambda \vdash \Gamma_1, \Gamma_2, B \Rightarrow \gamma D$ for some finite $\Gamma_1 \subseteq \Sigma_1$, $\Gamma_2 \subseteq \Sigma_2$. Now

$$\begin{aligned}\mathcal{L}_\Lambda \vdash x &\Rightarrow \alpha B \wedge \bigwedge \Gamma_1 \wedge \bigwedge \Gamma_2 \\ &\Rightarrow \alpha(B \wedge \bigwedge \Gamma_1 \wedge \bigwedge \Gamma_2) \quad (\text{since } |\bigwedge \Gamma_2| < \alpha) \\ &\Rightarrow \alpha\beta D \\ &\Rightarrow \min\{\alpha, \beta\} D.\end{aligned}$$

By adequacy, we have $\alpha D \in \Delta$ which together with $\beta \in \ell(\Delta)$ implies $\min\{\alpha, \beta\} D \in \Delta$, whence $\min\{\alpha, \beta\} D \in x$ follows. Let $\beta < \alpha$ and $\beta D \in x$. Then, by definition of Δ, we have $\beta D \in \Delta$ and hence $\beta D \in y$. Clearly, if $p^\beta \notin y$ for $\beta < \alpha$, then $p^\beta \notin x$. □

Lemma 4.3.11. *If $\mathcal{L}_\Lambda$ extends RJ^*_Λ then $\mathfrak{K}_\Delta$ is strongly persistent.*

Proof. Let $y \Vdash p^\beta$ and consider some $x \in W$ such that $xR_\alpha y$ for $\beta \leq \alpha$. We know $p^\beta \in y$, whence by adequacy $\alpha p^\beta \in \Delta$ and so $\alpha p^\beta \in x$. This yields $\mathcal{L}_\Lambda \vdash x \Rightarrow \alpha p^\beta \Rightarrow p^\beta$ by α-persistence, whence $p^\beta \in x$ and therefore $x \Vdash p^\beta$.

Now let $y \nVdash p^\beta$ and consider some $x \in W$ such that $xR_\alpha y$ for $\beta < \alpha$. We have $p^\beta \notin y$, whence by definition of R_α we get $p^\beta \notin x$ and thus $x \nVdash p^\beta$. □

Lemma 4.3.12. *If $\mathcal{L}_\Lambda$ extends RC^*_Λ then $\mathfrak{F}_\Delta$ is monotone, i.e., $\mathfrak{F}_\Delta$ is an RC^*_Λ-frame.*

Proof. Suppose $xR_\alpha y$ and let $\beta \in \Lambda$ be such that $\beta < \alpha$. We show that $xR_\beta y$. Let $A \in y$ and $\beta A \in \Delta$. By adequacy, we know that $\alpha A \in \Delta$ and $xR_\alpha y$ implies $\alpha A \in x$. Now $\beta = \min\{\alpha, \beta\}$ and so $xR_\alpha y$ implies in turn $\beta A \in x$ as required. This proves item (i). Let $\gamma A \in y$ and $\beta A \in \Delta$. We show $\min\{\beta, \gamma\} A \in x$. Indeed, by adequacy we know that $\alpha A \in \Delta$, whence $\min\{\gamma, \alpha\} A \in x$. In case $\beta < \gamma$, by the monotonicity axioms, we obtain $\mathcal{L}_\Lambda \vdash \min\{\gamma, \alpha\} A \Rightarrow \beta A$ and so $\beta A \in x$ as required. Thus, item (ii) is shown. For item (iii), Let $\gamma < \beta$ and $\gamma A \in x$. Then $\gamma < \alpha$, whence $\gamma A \in y$ follows by $xR_\alpha y$ as required. Finally, let $p^\gamma \notin y$ where $\gamma < \beta$. Obviously, $p^\gamma \notin x$ by $xR_\alpha y$. Hence, item (iv) holds. □

Theorem 4.3.13. RJ^*_Λ *is complete with respect to the class of all finite and strongly persistent RJ^*_Λ-models.*

Proof. Suppose $\mathsf{RJ}^*_\Lambda \nvdash A \Rightarrow B$. Consider an adequate Δ containing A and B and let $x := \{C \mid \mathsf{RJ}^*_\Lambda \vdash A \Rightarrow C, C \in \Delta\}$ and $\Sigma := \ell(\Delta)$. Consider the corresponding RJ^*_Σ-model $\mathfrak{K}_\Delta$ as before. By Lemma 4.3.10, we know $\mathfrak{K}_\Delta, x \Vdash A$, but $\mathfrak{K}_\Delta, x \nVdash B$. Now $\mathfrak{K}_\Delta$ is a finite RJ^*_Σ-model and $\Sigma \subseteq \Lambda$. We can expand $\mathfrak{K}_\Delta$ to a finite, strongly persistent RJ^*_Λ-model (call it $\mathfrak{K}$) by setting $R_\alpha := \varnothing$ for all $\alpha \in \Lambda \setminus \Sigma$. Then obviously $\mathfrak{K}, x \nVdash A \Rightarrow B$. □

Theorem 4.3.14. RC^* *is complete with respect to the class of all strongly persistent* RC^**-models.*

Proof. Suppose $\mathsf{RC}^* \nvdash A \Rightarrow B$. Consider an adequate Δ containing A and B and let $x := \{C \mid \mathsf{RC}^* \vdash A \Rightarrow C, C \in \Delta\}$. Let $\Sigma := \ell(\Delta)$ and $\mathfrak{K}_\Delta$ be the corresponding RC^*_Σ-model. By Lemma 4.3.10, we have $\mathfrak{K}_\Delta, x \nVdash A \Rightarrow B$. We can expand $\mathfrak{K}_\Delta$ to a (possibly infinite) RC^*-model $\mathfrak{K}$ for which we then have $\mathfrak{K}, x \nVdash A \Rightarrow B$. □

Note that we cannot prove completeness of RC^* with respect to the class of finite and strongly persistent RC^*-models as we did for RJ^*, since we cannot simply declare infinitely many relations to be empty due to monotonicity. However, for a finite Λ, the finite model property immediately follows by an argument analogous to that of the proof of Theorem 4.3.13.

Theorem 4.3.15. *Let $\Lambda \subseteq \omega + 1$ be finite. Then,* RC^*_Λ *is complete with respect to the class of all finite and strongly persistent* RC^*_Λ*-models.*

4.4 Positive Fragments of GLP^* and J^*

Let Λ be a signature. Recall that L^*_Λ denotes the set of all many-sorted formulas over $\mathbb{V}$ and Λ. By a *positive formula* (*over* Λ) we mean any formula from L^*_Λ which is built using only propositional variables, $\top$, $\wedge$, and $\langle\alpha\rangle$ ($\alpha \in \Lambda$). It is clear that the set of positive formulas from L^*_Λ equals L^+_Λ as defined before.

In Chapter 3 we showed that J^*_Λ is sound and complete for the class of all finite, irreflexive, and strongly persistent J^*_Λ-models. We have not treated irreflexive models of our positive calculi so far due to the absence of Löb's axioms in the positive setting. This will be done now in order to establish a correspondence between the positive fragments of J^*_Λ and GLP^*_Λ to the corresponding reflection calculi RJ^*_Λ and RC^*_Λ, respectively.

Let Δ be an adequate set and consider a positive logic $\mathcal{L}_\Lambda$, where $\Lambda = \ell(\Delta)$. As before, we canonically define a frame $\mathfrak{F}_\Delta = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}\rangle$, where W and the R_α ($\alpha \in \Lambda$) are defined as before, except that we additionally stipulate in the definition of $xR_\alpha y$ that

(v) there is an $\alpha A \in x$ such that $\alpha A \notin y$.

The proofs of Lemmas 4.3.9 and 4.3.10 only need to be extended to this new condition.

Lemma 4.4.1. *If $\mathcal{L}_\Lambda$ extends* RJ^*_Λ *then $\mathfrak{F}_\Delta$ is an* RJ^*_Λ*-frame.*

Proof. Suppose $xR_\alpha y$ and $yR_\beta z$. We prove that $xR_\gamma z$ for $\gamma = \min\{\alpha, \beta\}$ by showing item (v). Suppose first that $\gamma = \alpha$. There is a $\gamma A \in x$ such that $\gamma A \notin y$. Now if $\gamma A \in z$, then by $yR_\beta z$ we obtain $\gamma A \in y$ which is not the case. Therefore, $\gamma A \notin z$ as required. Suppose now $\gamma = \beta$. There is a $\gamma A \in y$ such that $\gamma A \notin z$. By $xR_\alpha y$ we obtain $\gamma A \in x$ as desired, i.e., item (v) is established.

Suppose $xR_\alpha y$ and $xR_\beta z$ for $\beta < \alpha$. We prove $yR_\beta z$. There is a $\beta A \in x$ such that $\beta A \notin z$. Since $\beta < \alpha$ and $xR_\alpha y$, we obtain $\beta A \in y$ as required. This gives us item (v). □

We define a model $\mathfrak{K}_\Delta = \langle \mathfrak{F}_\Delta, [\![\cdot]\!] \rangle$ as before by setting

$$\mathfrak{K}_\Delta, x \Vdash p^\alpha \iff_{df} p^\alpha \in x,$$

for all variables $p^\alpha \in \Delta$ and all $x \in W$.

Lemma 4.4.2. *There is no $\alpha \leq \omega$ and no positive formula A such that* $\mathsf{RJ}^* \vdash A \Rightarrow \alpha A$.

Proof. Suppose there were such an A and an $\alpha \leq \omega$. Let $\cdot^*$ be an arithmetical realization mapping all variables to the standard numeration $\mathsf{Ax}_{\mathsf{PA}}$. We then obtain $\mathsf{PA} \vdash A^* \Rightarrow \underline{\mathsf{Con}}_\alpha(A^*)$. Since A is a positive formula and PA is sound, A^* numerates a sound theory. But this means that A^* proves its own consistency, in contradiction to Gödel's second incompleteness theorem (cf. the formulation of Gödel's theorems by Feferman [16]). $\square$

Lemma 4.4.3. *Let $\mathcal{L}_\Lambda$ extend* RJ^*_Λ. *For all $A \in \Delta$ we have $\mathfrak{K}_\Delta, x \Vdash A$ iff $A \in x$.*

Proof. The proof is by induction on A. The cases can be proved as in the proof of Lemma 4.3.10. We only need to consider the case where $A = \alpha B$ for some $B \in \Delta$. Suppose $x \Vdash A$. then $y \Vdash B$ for some $y \in W$ such that $x R_\alpha y$. By inductive hypothesis, we know $B \in y$, whence $\alpha B \in x$ follows by definition of R_α as $\alpha B \in \Delta$. For the other direction, suppose $A \in x$. We prove $\mathfrak{K}_\Delta, x \Vdash A$. As before, let

$$\Sigma_1 := \{\beta C \mid \beta C \in x, \beta < \alpha\},$$
$$\Sigma_2 := \{p^\beta \mid p^\beta \in x, \beta < \alpha\},$$

and let $y := \{C \in \Delta \mid \mathcal{L}_\Lambda \vdash \Sigma_1, \Sigma_2, B \Rightarrow C\}$. By inductive hypothesis, we know $y \Vdash B$ as $B \in y$. We prove $x R_\alpha y$. This works exactly as in the proof of Lemma 4.3.10, except that we additionally need to take item (v) into account. We show that $\alpha B \notin y$. Indeed, if $\alpha B \in y$ then $\mathcal{L}_\Lambda \vdash \Gamma_1, \Gamma_2, B \Rightarrow \alpha B$ for some finite $\Gamma_1 \subseteq \Sigma_2$, $\Gamma_2 \subseteq \Sigma_2$, whence

$$\begin{aligned} \mathcal{L}_\Lambda \vdash y &\Rightarrow B \wedge \bigwedge \Gamma_1 \wedge \bigwedge \Gamma_2 \\ &\Rightarrow \alpha B \wedge \bigwedge \Gamma_1 \wedge \bigwedge \Gamma_2 \\ &\Rightarrow \alpha(B \wedge \bigwedge \Gamma_1 \wedge \bigwedge \Gamma_2) \quad (\text{since } |\bigwedge \Gamma_2| < \alpha), \end{aligned}$$

which contradicts Lemma 4.4.2. $\square$

Theorem 4.4.4. RJ^*_Λ *is complete with respect to the class of all irreflexive, finite, and strongly persistent* RJ^*_Λ*-models.*

Proof. Suppose $\mathsf{RJ}^*_\Lambda \nvdash A \Rightarrow B$. Consider an adequate Δ containing A and B and let $x := \{C \mid \mathsf{RJ}^*_\Lambda \vdash A \Rightarrow C, C \in \Delta\}$. Let $\Sigma := \ell(\Delta)$. Consider the corresponding RJ^*_Σ-model $\mathfrak{K}_\Delta$ with the appropriate properties as before. By Lemma 4.4.3, we know $\mathfrak{K}_\Delta, x \Vdash A$, but $\mathfrak{K}_\Delta, x \nVdash B$. Now $\mathfrak{K}_\Delta$ is a finite RJ^*_Σ-model and $\Sigma \subseteq \Lambda$. We can expand $\mathfrak{K}_\Delta$ to an irreflexive, finite, and strongly persistent RJ^*_Λ-model (call it $\mathfrak{K}$) by setting $R_\alpha := \varnothing$ for all $\alpha \in \Lambda \setminus \Sigma$. Then obviously $\mathfrak{K}, x \nVdash A \Rightarrow B$. $\square$

Notice that irreflexivity is incompatible with our notion of RC^*_Λ-model. Indeed, if $xR_\alpha y$ for some $\alpha > 0$, then xR_0y, whence yR_0y follows. However, irreflexive models are vital for Solovay constructions. We say that a model $\mathfrak{K}$ is Δ-*monotone* if for any $\beta A \in \Delta$ and any $\alpha \in \ell(\Delta)$ such that $\beta < \alpha$ it holds that, for all worlds x of $\mathfrak{K}$,

$$\mathfrak{K}, x \Vdash \alpha A \implies \mathfrak{K}, x \Vdash \beta A.$$

Lemma 4.4.5. *If $\mathcal{L}_\Lambda$ extends RC^*_Λ, then $\mathfrak{K}_\Delta$ is Δ-monotone.*

Proof. Let $\beta A \in \Delta$ and let $\alpha \in \ell(\Delta)$ be such that $\beta < \alpha$. Suppose that $\mathfrak{K}_\Delta, x \Vdash \alpha A$. Then $\alpha A \in x$ and $\beta A \in \Delta$ by the adequacy of Δ. Thus, $\mathcal{L}_\Lambda \vdash x \Rightarrow \alpha A \Rightarrow \beta A$ and since $\beta A \in \Delta$, we infer $\beta A \in x$ and so $\mathfrak{K}_\Delta, x \Vdash \beta A$ as desired. □

Theorem 4.4.6. *Let Δ be an adequate set, and $\Lambda = \ell(\Delta)$. Then there is a model $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\in\Lambda}, [\![\cdot]\!]\rangle$ such that*

(i) *$\mathfrak{K}$ is an RJ^*-model and $R_\alpha = \varnothing$, for every $\alpha \notin \Lambda$;*

(ii) *$\mathfrak{K}$ is finite, strongly persistent, irreflexive, and Δ-monotone;*

(iii) *for any RC^*_Λ-theory Γ in Δ, there exists a node $x \in W$ such that for any $A \in \Delta$,*

$$A \in \Gamma \iff \mathfrak{K}, x \Vdash A.$$

We now investigate the relationships between RC^*_Λ (RJ^*_Λ, respectively) and GLP^*_Λ (J^*_Λ, respectively) as done by Dashkov [13] in the single-sorted setting. GLP^*_Λ differs from J^*_Λ with respect to the monotonicity axioms. This is also the only difference between RJ^*_Λ and RC^*_Λ. Hence, we immediately obtain:

Lemma 4.4.7. *Let $\varphi, \psi \in L^+_\Lambda$, where $\Lambda \subseteq \omega + 1$. If $\mathsf{RC}^*_\Lambda \vdash \varphi \Rightarrow \psi$ then $\mathsf{GLP}^*_\Lambda \vdash \varphi \to \psi$.*

Proposition 4.4.8. *Let $\varphi, \psi \in L^+_\Lambda$. Then, $\mathsf{RJ}^*_\Lambda \vdash \varphi \Rightarrow \psi$ iff $\mathsf{J}^*_\Lambda \vdash \varphi \to \psi$.*

Proof. The direction from left to right is shown by induction on the length of the proof of $\varphi \Rightarrow \psi$. Most of the axioms are clear. For axiom (vi), notice that if $|\psi| < \alpha$, then $\mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\neg\psi \to \neg\psi$, whence $\mathsf{J}^*_\Lambda \vdash \psi \to [\alpha]\psi$ and so

$$\begin{aligned} \mathsf{J}^*_\Lambda \vdash \langle\alpha\rangle\varphi \wedge \psi &\to [\alpha]\psi \\ &\to \langle\alpha\rangle(\varphi \wedge \psi). \end{aligned}$$

The translations of the rules of inference are immediate.

For the other direction, suppose $\mathsf{RJ}^*_\Lambda \nvdash \varphi \Rightarrow \psi$. By Theorem 4.4.4, there is an irreflexive, finite, and strongly persistent RJ^*_Λ-model $\mathfrak{K}$ such that $\mathfrak{K}, x \nVdash \varphi \Rightarrow \psi$ for some world x of $\mathfrak{K}$. Obviously, since $\mathfrak{K}$ is also a J^*_Λ-model, we obtain $\mathsf{J}^*_\Lambda \nvdash \varphi \to \psi$. □

Lemma 4.4.9. *Let $\varphi \in L^*_\Lambda$, where $\Lambda \subseteq \omega$. Then, $\mathsf{GLP}^* \vdash \varphi$ iff $\mathsf{GLP}^*_\Lambda \vdash \varphi$.*

Proof. The direction from right to left is clear. We prove the other direction. Suppose $\mathsf{GLP}^* \vdash \varphi$. Then $\mathsf{J}^* \vdash N^+(\varphi) \to \varphi$ by Lemma 3.5.13 and, as $N^+(\varphi)$ is in L^*_Λ, we know by Corollary 3.4.18 that $\mathsf{J}^*_\Lambda \vdash N^+(\varphi) \to \varphi$, whence $\mathsf{GLP}^*_\Lambda \vdash N^+(\varphi) \to \varphi$ follows since GLP^*_Λ extends J^*_Λ. But $\mathsf{GLP}^*_\Lambda \vdash N^+(\varphi)$, so the claim follows by an application of modus ponens. □

Lemma 4.4.10. *Let $\Lambda \subseteq \omega$ and $\varphi, \psi \in L^+_\Lambda$. Then $\mathsf{GLP}^*_\Lambda \vdash \varphi \to \psi$ implies $\mathsf{RC}^*_\Lambda \vdash \varphi \Rightarrow \psi$.*

Proof. Suppose $\mathsf{RC}^*_\Lambda \nvdash \varphi \Rightarrow \psi$. Let a finite adequate Δ containing φ and ψ be given. Let $\Sigma := \ell(\Delta)$ and consider the corresponding model $\mathfrak{K}_\Delta = \langle W, \{R_\alpha\}_{\alpha\in\Sigma}, [\![\cdot]\!]\rangle$ with the properties of Theorem 4.4.6 such that $\mathfrak{K}_\Delta, x \nVdash \varphi \Rightarrow \psi$ for some $x \in W$. Now let $\eta := \varphi \to \psi$. Since $\mathfrak{K}_\Delta$ is Δ-monotone and $N^+(\eta)$ is in L^*_Λ, we know that $\mathfrak{K}_\Delta, x \Vdash N^+(\eta)$. Hence, $\mathfrak{K}_\Delta, x \nVdash N^+(\eta) \to \eta$ and so $\mathsf{J}^*_\Lambda \nvdash N^+(\eta) \to \eta$, whence $\mathsf{GLP}^*_\Lambda \nvdash \varphi \to \psi$ follows. □

In the following, we borrow some ideas from Beklemishev et al. [8]. Any signature Λ can be ordered according to the standard ordering of the ordinals $\alpha \in \omega + 1$. Hence, RC^*_Λ is in some sense a notational variant of the logic RC^*_λ, where λ denotes the order type of Λ. For any signature Λ, we denote by Λ_α the α-th element of Λ according to that ordering. Let $f : \Lambda \to \lambda$ be the (unique) order isomorphism from Λ into λ. Given any $U \subseteq \Lambda$, we denote by λ_U the order $f(U)$.

In the sequel, It will be convenient to assume that the set $\mathbb{V}$ properly contains a set $\overline{\mathbb{V}}$ such that for each sort $\alpha \leq \omega$, $\overline{\mathbb{V}}$ contains a countable infinite supply of variables of sort α. Hence, there is a bijection between $\mathbb{V} \setminus \overline{\mathbb{V}}$ and $\overline{\mathbb{V}}$ such that exactly the variables of sort α from $\mathbb{V} \setminus \overline{\mathbb{V}}$ are mapped to variables of sort α from $\overline{\mathbb{V}}$. Now if $p^\alpha \in \mathbb{V} \setminus \overline{\mathbb{V}}$, we denote the corresponding variable of $\overline{\mathbb{V}}$ by $\overline{p^\alpha}$. Conversely, if $q^\alpha \in \overline{\mathbb{V}}$ such that $\overline{p^\alpha} = q^\alpha$ for $p^\alpha \in \mathbb{V} \setminus \overline{\mathbb{V}}$, we set $\overline{q^\alpha} = p^\alpha$. Now, for any signature Λ of order type λ, we define a bijection $\pi_\Lambda : \mathbb{V} \to \overline{\mathbb{V}}$ as follows:

$$\pi_\Lambda : p^\alpha \longmapsto \begin{cases} \overline{p^\kappa}, \text{ where } \kappa = \Lambda_\alpha, & \text{if } \alpha < \lambda, \\ \overline{p^\alpha}, & \text{otherwise.} \end{cases}$$

Definition 4.4.11. Let Λ be a signature of order type λ. We define $\xi_\Lambda(\varphi)$ for all $\varphi \in L^*_\lambda$ recursively as follows:

(i) $\xi_\Lambda(\bot) = \bot$; $\xi_\Lambda(\top) = \top$;

(ii) $\xi_\Lambda(p^\alpha) = \pi_\Lambda(p^\alpha)$ for all propositional variables p^α;

(iii) $\xi_\Lambda(\cdot)$ commutes with the propositional connectives;

(iv) $\xi_\Lambda(\langle\alpha\rangle\varphi) = \langle\Lambda_\alpha\rangle\xi_\Lambda(\varphi)$.

Furthermore, we define $\xi^{-1}_\Lambda(\cdot)$ to be the inverse operation of $\xi_\Lambda(\cdot)$ such that (i) for all $\varphi \in L^*_\lambda$ we have $\varphi = \xi^{-1}_\Lambda(\xi_\Lambda(\varphi))$ and (ii) for all $\psi \in L^*_\Lambda$ we have $\psi = \xi_\Lambda(\xi^{-1}_\Lambda(\psi))$. ⊣

Note that if $U \subseteq \Lambda$ then for $\varphi \in L^*_\lambda$ and $\psi \in L^*_\Lambda$,

$$\varphi \in L^*_{\lambda_U} \implies \xi^{-1}_\Lambda(\varphi) \in L^*_U,$$
$$\xi_\Lambda(\psi) \in L^*_U \implies \psi \in L^*_{\lambda_U}.$$

Lemma 4.4.12. *Let λ be the order type of Λ and $U \subseteq \Lambda$. Let $\varphi \in L^*_{\lambda_U}$ be a formula and $\alpha \in \lambda_U$. Then, $|\varphi| \leq \alpha$ iff $|\xi_\Lambda(\varphi)| \leq \Lambda_\alpha$.*

Proof. We first prove the direction from left to right. We proceed by induction on the number of propositional connectives of φ which are not in the scope of any $\langle\gamma\rangle$. For the base case, if $\varphi = p^\beta$ for some $\beta \leq \alpha$, we have that $\beta < \lambda$, whence $\pi_\Lambda(p^\beta) = \overline{p^\kappa}$ follows for $\kappa = \Lambda_\beta$. Therefore, $|\xi_\Lambda(p^\kappa)| = \Lambda_\beta \leq \Lambda_\alpha$ since $\beta \leq \alpha$. Suppose $\varphi = \langle\beta\rangle\psi$ for some $\beta \leq \alpha$. Then $|\langle\beta\rangle\psi| = \beta$ and $|\xi_\Lambda(\langle\beta\rangle\psi)| = |\langle\Lambda_\beta\rangle\xi_\Lambda(\psi)| = \Lambda_\beta \leq \Lambda_\alpha$. The induction step for the propositional connectives is immediate as $\xi_\Lambda(\cdot)$ commutes with those connectives.

The direction from right to left is proved by induction on the number of propositional connectives of $\psi \coloneqq \xi_\Lambda(\varphi)$ which are not in the scope of any $\langle\gamma\rangle$. We know $\varphi = \xi_\Lambda^{-1}(\psi)$. Suppose $\psi = \overline{p^\beta}$ for some $\beta \leq \Lambda_\alpha$. Suppose first that $\beta \in \Lambda$. Then $\pi_\Lambda^{-1}(\overline{p^\beta}) = p^\gamma$ for some $\gamma < \lambda$ such that $\beta = \Lambda_\gamma$. Therefore, $\Lambda_\gamma \leq \Lambda_\alpha$, whence $\gamma \leq \alpha$ follows. Suppose now that $\beta \notin \Lambda$. Then, $\pi_\Lambda^{-1}(\overline{p^\beta}) = p^\beta$ and thus $|p^\beta| \leq \Lambda_\alpha$ by assumption. Suppose $\psi = \langle\beta\rangle\chi$ for some χ and $\beta \leq \Lambda_\alpha$. Certainly $\psi \in L^*_U$ and therefore $\beta \in U$. It follows that $\beta = \Lambda_\gamma$, for some $\gamma < \lambda$. We have $\varphi = \xi_\Lambda^{-1}(\psi) = \langle\gamma\rangle\xi_\Lambda^{-1}(\chi)$ and so $|\varphi| \leq \alpha$ as $\gamma \leq \alpha$. The induction step is again immediate. □

Lemma 4.4.13. *Let λ be the order type of Λ and $U \subseteq \Lambda$.*

(i) $\mathsf{GLP}^*_{\lambda_U} \vdash \varphi$ *iff* $\mathsf{GLP}^*_U \vdash \xi_\Lambda(\varphi)$*, for all* $\varphi \in L^*_{\lambda_U}$*;*

(ii) $\mathsf{GLP}^*_U \vdash \varphi$ *iff* $\mathsf{GLP}^*_{\lambda_U} \vdash \xi_\Lambda^{-1}(\varphi)$*, for all* $\varphi \in L^*_U$*.*

Proof. For item (i), we prove the direction from left to right by induction on proof length. The case when φ is a propositional axiom is clear, as $\xi_\Lambda(\cdot)$ commutes with the propositional connectives. Most of the other axioms are also clear. Consider the case where φ is of form $\langle\alpha\rangle\psi \to \psi$ and $|\psi| \leq \alpha$, where $\alpha \in \lambda_U$. By Lemma 4.4.12 we know $|\xi_\Lambda(\psi)| \leq \Lambda_\alpha$, i.e., $\langle\Lambda_\alpha\rangle\xi_\Lambda(\psi) \to \xi_\Lambda(\psi)$ is also an axiom. Since $\Lambda_\alpha \in U$, $\mathsf{GLP}^*_U \vdash \langle\Lambda_\alpha\rangle\xi_\Lambda(\psi) \to \xi_\Lambda(\psi)$ follows. For the induction step, consider the case where $\mathsf{GLP}^*_{\lambda_U} \vdash [\alpha]\psi$ for $\alpha \in \lambda_U$ and $[\alpha]\psi$ results from an application of $[\alpha]$-necessitation from ψ. By inductive hypothesis, we have $\mathsf{GLP}^*_U \vdash \xi_\Lambda(\psi)$ and, since $\Lambda_\alpha \in U$, we obtain $\mathsf{GLP}^*_U \vdash [\Lambda_\alpha]\xi_\Lambda(\psi)$. The case of modus ponens is immediate. The other direction and item (ii) are proved in an analogous way. □

We have a similar result for RC^*_Λ:

Lemma 4.4.14. *Let λ be the order type of Λ and $U \subseteq \Lambda$.*

(i) $\mathsf{RC}^*_{\lambda_U} \vdash A \Rightarrow B$ *iff* $\mathsf{RC}^*_U \vdash \xi_\Lambda(A) \Rightarrow \xi_\Lambda(B)$*;*

(ii) $\mathsf{RC}^*_U \vdash A \Rightarrow B$ *iff* $\mathsf{RC}^*_{\lambda_U} \vdash \xi_\Lambda^{-1}(A) \Rightarrow \xi_\Lambda^{-1}(B)$*.*

Theorem 4.4.15. *Let $\varphi \in L^*_\Lambda$, where $\Lambda \subseteq \omega + 1$. Then,* $\mathsf{GLP}^*_{\omega+1} \vdash \varphi$ *iff* $\mathsf{GLP}^*_\Lambda \vdash \varphi$*.*

Proof. The direction from right to left is again obvious. For the other direction, suppose

$\mathsf{GLP}^*_{\omega+1} \vdash \varphi$. Then there is a proof $\chi = \chi_1, \ldots, \chi_k$ in $\mathsf{GLP}^*_{\omega+1}$ such that $\chi_k = \varphi$. Let

$$S := \{\alpha \mid \langle\alpha\rangle \text{ occurs in } \chi_i, \text{ for some } i = 1, \ldots, k\},$$
$$U := \{\alpha \mid \langle\alpha\rangle \text{ occurs in } \varphi\}.$$

Obviously, $U \subseteq S$ and $U \subseteq \Lambda$. We know that U and S are finite and $\mathsf{GLP}^*_S \vdash \chi_i$ for $i = 1, \ldots, k$. Let $n \in \omega$ be the order type of S. By Lemma 4.4.13 (taking $U = S$), we know that

$$\mathsf{GLP}^*_S \vdash \chi_i \iff \mathsf{GLP}^*_n \vdash \xi_S^{-1}(\chi_i),$$

for $i = 1, \ldots, k$. Let $\psi := \xi_S^{-1}(\varphi)$. From $\mathsf{GLP}^*_n \vdash \psi$ we immediately infer $\mathsf{GLP}^* \vdash \psi$. Now consider the set of modalities $F := n_U$. We have that $\psi \in L^*_F$ and so by Lemma 4.4.9, $\mathsf{GLP}^*_F \vdash \psi$ as $F \subseteq \omega$. By Lemma 4.4.13, we obtain $\mathsf{GLP}^*_U \vdash \varphi$ and thus $\mathsf{GLP}^*_\Lambda \vdash \varphi$ follows since $U \subseteq \Lambda$. ☐

Corollary 4.4.16. *Let $A, B \in L^+_\Lambda$, where $\Lambda \subseteq \omega + 1$. Then,* $\mathsf{RC}^* \vdash A \Rightarrow B$ *iff* $\mathsf{RC}^*_\Lambda \vdash A \Rightarrow B$.

Proof. The direction from right to left is immediate. So suppose $\mathsf{RC}^* \vdash A \Rightarrow B$. Let $\Sigma := \ell(\{A, B\})$ and $n \in \omega$ be the order type of Σ. Obviously, $\Sigma \subseteq \Lambda$ and since $\mathsf{GLP}^* \vdash A \to B$, we infer $\mathsf{GLP}^*_\Sigma \vdash A \to B$ by Theorem 4.4.15. By Lemma 4.4.13, we obtain $\mathsf{GLP}^*_n \vdash \xi_\Sigma^{-1}(A) \to \xi_\Sigma^{-1}(B)$, whence $\mathsf{RC}^*_n \vdash \xi_\Sigma^{-1}(A) \Rightarrow \xi_\Sigma^{-1}(B)$ follows by Lemma 4.4.10. Therefore, $\mathsf{RC}^*_\Sigma \vdash A \Rightarrow B$ and thus $\mathsf{RC}^*_\Lambda \vdash A \Rightarrow B$ as desired. ☐

Thus, although we cannot prove that RC^* is complete with respect to the class of finite, strongly persistent, and irreflexive RC^*-models, we can always find an appropriate model of a finite fragment of RC^*. Furthermore, we can already conclude that RC^*_Λ axiomatizes the positive fragment of GLP^*_Λ.

Corollary 4.4.17. *Let $\varphi, \psi \in L^+_\Lambda$, where $\Lambda \subseteq \omega + 1$. Then,* $\mathsf{GLP}^*_\Lambda \vdash \varphi \to \psi$ *iff* $\mathsf{RC}^*_\Lambda \vdash \varphi \Rightarrow \psi$.

Proof. The direction from right to left is just Lemma 4.4.7. Suppose $\mathsf{GLP}^*_\Lambda \vdash \varphi \to \psi$ and let $\Sigma := \ell(\{\varphi, \psi\})$ and let $n \in \omega$ be the order type of Σ. It is clear that $\Sigma \subseteq \Lambda$. We have $\mathsf{GLP}^*_\Sigma \vdash \varphi \to \psi$ by Theorem 4.4.15 and so $\mathsf{GLP}^*_n \vdash \xi_\Sigma^{-1}(\varphi) \to \xi_\Sigma^{-1}(\psi)$. Therefore, $\mathsf{RC}^*_n \vdash \xi_\Sigma^{-1}(\varphi) \Rightarrow \xi_\Sigma^{-1}(\psi)$ by Lemma 4.4.10, whence $\mathsf{RC}^*_\Sigma \vdash \varphi \Rightarrow \psi$ follows. Thus, $\mathsf{RC}^*_\Lambda \vdash \varphi \Rightarrow \psi$ as required. ☐

In particular, $\mathsf{GLP}^*_{\omega+1} \vdash \varphi \to \psi$ iff $\mathsf{RC}^* \vdash \varphi \Rightarrow \psi$ for positive φ, ψ.

4.5 Arithmetical Completeness of RC*

In this section, we prove that RC^* is arithmetically complete with respect to the interpretation we defined in Section 4.2. Many of the results necessary for its proof were first obtained by Beklemishev [7] for a single-sorted variant of RC^*.

Theorem 4.5.1. *For every sequent $A \Rightarrow B$ over $\omega+1$, the following statements are equivalent:*

(i) $\mathsf{RC}^* \vdash A \Rightarrow B$;

(ii) $A^* \Rightarrow_{\mathsf{PA}} B^*$ *for every arithmetical realization* $\cdot^*$;

(iii) $A^* \Rightarrow B^*$ *for every arithmetical realization* $\cdot^*$.

Proof. For the direction from (iii) to (i) we proceed indirectly. Suppose $\mathsf{RC}^* \nvdash A \Rightarrow B$. Let Δ be a finite adequate set containing A and B and let $\Lambda := \ell(\Delta)$. Similarly as in the proof of Lemma 3.5.1, one can show that there is an RJ^*-model $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\leq\omega}, [\![\cdot]\!]\rangle$ with root r which satisfies the conditions of Theorem 4.4.6 such that $\mathfrak{K}, r \Vdash A$ and $\mathfrak{K}, r \nVdash B$. Without loss of generality, suppose $W = \{1, 2, \ldots, N\}$ for some $N \in \omega$ and $r = 1$. Extend $\mathfrak{K}$ to an RJ^*-model $\mathfrak{K}' = \langle W', \{R'_\alpha\}_{\alpha\leq\omega}, [\![\cdot]\!]'\rangle$, where

(i) $W' := W \cup \{0\}$;

(ii) $R'_0 := R_0 \cup \{(0, x) \mid x \in W\}$;

(iii) $R'_\alpha := R_\alpha$, for $\alpha \neq 0$;

(iv) $\mathfrak{K}', 0 \Vdash p \iff_{df} \mathfrak{K}, 1 \Vdash p$, for all variables $p \in \Delta$;

(v) $\mathfrak{K}', x \Vdash p \iff_{df} \mathfrak{K}, x \Vdash p$, for all $x \in W$ and all variables $p \in \Delta$.

Note that $\mathfrak{K}'$ is finite, strongly persistent, and irreflexive. Moreover, it is clear that $\mathfrak{K}', 1 \nVdash A \Rightarrow B$. For notational convenience, we denote $\mathfrak{K}'$ from now on by $\mathfrak{K} = \langle W, \{R_\alpha\}_{\alpha\leq\omega}, [\![\cdot]\!]\rangle$.

Recall that $[n]_{\mathsf{PA}}(\alpha)$ denotes a formula which formalizes that α is provable in PA from all true Π_{n+1}-sentences. For $n \geq 0$, let $\mathsf{Prf}_n(\alpha, y)$ be the proof relation of $[n]_{\mathsf{PA}}(\alpha)$. As in the proof of Theorem 3.5.2, we assume that information concerning the model $\mathfrak{K}$ is naturally encoded in arithmetic.

Definition 4.5.2. Let M be the maximal modality $m < \omega$ from Λ, provided there is such an m, and 0 otherwise. For all $n < \omega$, define a Solovay function $h_n \colon \omega \to W$ as follows:

$$h_n(0) = 0, \text{ and}$$

$$h_n(x+1) = \begin{cases} y, & \text{if } \exists i < n \colon h_i(x) \neq h_i(x+1) = y; \text{ otherwise} \\ z, & \text{if } \exists k \geq \max\{M, n\}\, \mathsf{Prf}_n(\ulcorner \ell_k \neq \overline{z} \urcorner, x) \\ & \text{and either } h_n(x) R_n z \text{ or } h_n(x) R_\omega z; \\ h_n(x), & \text{otherwise.} \end{cases}$$

Here we retain our convention that $\ell_k = x$ denotes that the limit of the function h_k equals x (see Definition 3.5.3). ⊣

Beklemishev [7] shows that there are formulas $H_0, H_1, \ldots, H_n$ such that, for $k = 1, \ldots, n$,

(i) H_k defines the graph of h_k in PA and $\mathsf{PA} \vdash \forall x \exists! y\, H_k(x,y)$. Thus, h_k is provably total in PA.

(ii) H_k is Δ_{k+1} in PA.

(iii) The function $\varphi\colon k \longmapsto \ulcorner H_k \urcorner$ is primitive recursive.

(iv) Each H_k provably satisfies the definition of h_k in Definition 4.5.2.

Note that, in particular, $\ulcorner \ell_n = \overline{x} \urcorner$ can be constructed primitive recursively from n and x.

Lemma 4.5.3. *For each $n, m \in \omega$, provably in* PA*,*

(i) $\exists! x \in W \colon \ell_n = \overline{x}$*;*

(ii) $\ell_n R_{n+1} \ell_{n+1}$ *or* $\ell_n R_\omega \ell_{n+1}$ *or* $\ell_n = \ell_{n+1}$*;*

(iii) *If $m < n$ then $\ell_n = \ell_m$ or $\ell_m R_\alpha \ell_n$, for some $\alpha \in (m,n] \cup \{\omega\}$.*

Proof. For (i), firstly notice that for all $n \in \omega$, h_n (provably) only moves along $R_n \cup R_\omega$. Secondly, $R_n \cup R_\omega$ is finite, transitive, and irreflexive for all $n \in \omega$. Uniqueness is clear, as PA proves $\forall x \exists! y\, H_n(x,y)$. For existence, we proceed by (external) induction on n. Let $S \coloneqq R_0 \cup R_\omega$ and $b \in W$. For the base case, we prove

$$\mathsf{PA} \vdash H_0(a, \overline{b}) \to \ell_0 = \overline{b} \vee \exists z \in S(b) \colon \ell_0 = \overline{z},$$

by induction on the converse of S. So suppose that for each $c \in S(b)$ we have

$$\mathsf{PA} \vdash H_0(a, \overline{c}) \to \ell_0 = \overline{c} \vee \exists z \in S(c) \colon \ell_0 = \overline{z}.$$

By definition of h_0, we know that

$$\mathsf{PA} \vdash H_0(a, \overline{b}) \to \forall x \geq a\, (H_0(x, \overline{b}) \vee \exists z \in S(b) \colon H_0(x, \overline{z})),$$

whence by inductive hypothesis we obtain

$$\mathsf{PA} \vdash H_0(a, \overline{b}) \to \forall x \geq a\, (H_0(x, \overline{b}) \vee \exists z \in S(b) \colon \ell_0 = \overline{z} \vee \exists w \in S(z) \colon \ell_0 = \overline{w}).$$

This is equivalent to

$$\mathsf{PA} \vdash H_0(a, \overline{b}) \to \ell_0 = \overline{b} \vee \exists z \in S(b) \colon \ell_0 = \overline{z} \vee \exists w \in S(z) \colon \ell_0 = \overline{w},$$

which by the transitivity of S is equivalent to

$$\mathsf{PA} \vdash H_0(a, \overline{b}) \to \ell_0 = \overline{b} \vee \exists z \in S(b) \colon \ell_0 = \overline{z}.$$

This proves the base case, since $\mathsf{PA} \vdash H_0(0,0)$. For the induction step, suppose that ℓ_{n-1} exists ($n > 0$). Then, from ℓ_{n-1} onward, h_n provably can only move along the relation $R_n \cup R_\omega$. A similar argument to that of the base case then shows that the limit of h_n exists. For (ii), notice that, provably, $\exists x\, h_{n+1}(x) = \ell_n$, i.e., h_{n+1} has to visit ℓ_n on its way to ℓ_{n+1}. Using this and the previously mentioned facts, we can prove (ii) by induction on n. Item (iii) is then obtained from (ii) by induction on n. $\square$

For all $n < \omega$, we define a formula $L_n(a,b)$ in our arithmetical language as follows:

$$L_n(a,b) \coloneqq \begin{cases} h_n(a) = b, & \text{if } n = 0; \\ h_n(a) = b \wedge \forall z \geq a\,(h_{n-1}(z) = h_{n-1}(a)), & \text{otherwise.} \end{cases}$$

Lemma 4.5.4. *For every $n \in \omega$, PA proves that*

$$L_n(a,b) \implies \forall i < n\, \forall x \geq a \colon h_i(x) = h_i(x+1).$$

Proof. For $n = 0$, the statement is clear. For $n > 0$, reason in PA as follows. Suppose $L_n(a,b)$ and assume to the contrary that there is an $i < n$ and an $x \geq a$ such that $h_i(x) \neq h_i(x+1)$. Since $L_n(a,b)$, we clearly have $i < n-1$. By definition of h_{n-1}, we infer $h_{n-1}(x+1) = h_i(x+1)$, whence $h_{n-1}(x) \neq h_i(x)$ follows. Let x_0 be the smallest $x \geq a$ such that $h_{n-1}(x) \neq h_i(x)$. We know $x_0 > 0$ and by $L_n(a,b)$ that $h_i(x_0 - 1) \neq h_i(x_0)$. But then $h_{n-1}(x_0) = h_i(x_0)$ by the definition of h_{n-1}, contradiction. ☐

Let $n < \omega$ and suppose (in PA) that $\exists x\, L_n(x, \bar{b})$ for some $b \in W$. It follows that for $k \geq n$, $\ell_k \in R_n^*(b) \cup \{b\}$, since no function h_m for $m < n$ makes a move beyond x. Hence, from b on, the function h_k must move along R_n^*.

For any positive formula A and any $n \geq 0$, we abbreviate by $\ell_n \Vdash A$ the statement $\bigvee\{\ell_n = \overline{x} \mid x \Vdash A\}$, while $\ell_n \nVdash A$ abbreviates $\bigwedge\{\ell_n \neq \overline{x} \mid x \Vdash A\}$. Note that Lemma 4.5.3 implies that for all positive formulas A, provably in PA either $\ell_n \Vdash A$ or $\ell_n \nVdash A$. Furthermore, these statements provably obey the usual forcing conditions which are already fulfilled in $\mathfrak{K}$:

Lemma 4.5.5. *Provably in PA,*

(i) *$\ell_n \Vdash \top$;*

(ii) *$\ell_n \Vdash A \wedge B$ iff $\ell_n \Vdash A$ and $\ell_n \Vdash B$.*

Furthermore,

(iii) *if $|A| \leq \alpha$ then PA proves that $\ell_m R_\alpha \ell_n$ and $\ell_n \Vdash A$ imply $\ell_m \Vdash A$;*

(iv) *if $|A| < \alpha$ then PA proves that $\ell_m R_\alpha \ell_n$ and $\ell_n \nVdash A$ imply $\ell_m \nVdash A$.*

Proof. Most of the items are clear. We only prove item (iv). Note that $\ell_n \nVdash A$ is, by virtue of Lemma 4.5.3, provably equivalent in PA to

$$\bigvee\{\ell_n = \overline{x} \mid x \nVdash A\}.$$

Now if $\ell_m R_\alpha \ell_n$ and $|A| < \alpha$, then by strong persistence of $\mathfrak{K}$ we see that this immediately implies $\ell_m \nVdash A$. ☐

Lemma 4.5.6. *Let $n \geq m$ and A be a formula such that $|A| \leq m$ or $m \geq M$. Then,*

$$\mathsf{PA} \vdash (\ell_n \Vdash A) \to (\ell_m \Vdash A).$$

Proof. Assume $n > m \geq M$. Since $R_n = \varnothing$, we must have $\ell_m R_\omega \ell_n$ by Lemma 4.5.3, whence the claim follows at once from Lemma 4.5.5. If $|A| \leq m$ and $n > m$, we have by Lemma 4.5.3 either $\ell_m = \ell_n$ or $\ell_m R_\alpha \ell_n$ for some $\alpha \in (m, n] \cup \{\omega\}$. The first case is clear, so suppose $\ell_m R_\alpha \ell_n$. The claim follows immediately from Lemma 4.5.5 since $|A| \leq m < \alpha$. ☐

We now again use the property of strong persistence in order to ensure that we are able to construct an arithmetical counter interpretation of the desired arithmetical complexity.

Lemma 4.5.7. *For all $n < \omega$ and all variables $p \in \Delta$ of sort $k \leq n$, provably in* PA*,*

$$\ell_n \Vdash p \iff \forall w \in W \setminus [\![p]\!] \colon \forall x \, \neg L_k(x, \overline{w}).$$

Proof. We reason in PA as follows. For the direction from left to right, suppose $\ell_n \Vdash p$ and suppose to the contrary that there is a $w \in W$ and an x such that $w \nVdash p$ and $L_k(x, \overline{w})$. By strong persistence, we know that $v \nVdash p$ for all $v \in R_k^*(w)$. Since $k \leq n$, we know that $\ell_n \in R_k^*(w) \cup \{w\}$, whence $\ell_n \in W \setminus [\![p]\!]$ follows. This contradicts the uniqueness of ℓ_n (cf. Lemma 4.5.3).

For the other direction, suppose $\forall w \in W \setminus [\![p]\!] \colon \forall x \, \neg L_k(x, \overline{w})$ and suppose further that $\ell_n \neq \overline{x}$ for all $x \in [\![p]\!]$. It follows that $\ell_n \in W \setminus [\![p]\!]$. Let $w \in W \setminus [\![p]\!]$; we prove $\ell_k \neq \overline{w}$. Consider an arbitrary x. If $k = 0$ then since $\neg L_k(x, \overline{w})$, we infer $h_k(x) \neq \overline{w}$ and we are done. Otherwise, we have $h_k(x) \neq \overline{w}$ or $\exists z \geq x \colon h_{k-1}(z) \neq h_{k-1}(x)$. In the former case we are finished, so suppose $h_k(x) = \overline{w}$. Since there is a z such that $h_{k-1}(z) \neq h_{k-1}(x) = \overline{w}$, we obtain by definition of h_k that $h_k(y) \neq h_k(y+1)$ for some $y \geq x$, whence $\ell_k \neq \overline{w}$ follows. Hence, $\ell_k \Vdash p$ and certainly $\ell_k \neq \ell_n$. It follows that $n > k$ and by Lemma 4.5.3 that $\ell_k R_\alpha \ell_n$ for some $\alpha \in (k, n] \cup \{\omega\}$. But this is in contradiction to the property of $\mathfrak{K}$ being strongly persistent, as $\ell_k \in [\![p]\!]$ and $\ell_n \in W \setminus [\![p]\!]$. ☐

If $\{\varphi_i \mid i \in I\}$ is a primitive recursive set of formulas, we denote by $[\varphi_i \mid i \in I]$ a numeration which numerates the theory $\mathsf{PA} + \{\varphi_i \mid i \in I\}$. We stipulate that, for a formula φ, the numeration $[\varphi]$ is the same as $\underline{\varphi}$. Furthermore, we write $\mathsf{Con}_n[\varphi_i \mid i \in I]$ instead of $\mathsf{Con}_n([\varphi_i \mid i \in I])$.

For any variable $p \in \Delta$ we set

$$p^* := [\ell_n \Vdash p \mid n \geq M].$$

Lemma 4.5.8. *For every variable $p \in \Delta$ of sort $k < \omega$, p^* numerates a Π_{k+1}-axiomatized extension of* PA*.*

Proof. Let p be a variable of sort $k < \omega$. Let $n \geq M$ and consider the sentence $\ell_n \Vdash p$. If $k \leq n$ then by Lemma 4.5.7, provably in PA,

$$\ell_n \Vdash p \iff \forall w \in W \setminus [\![p]\!] \colon \forall x \, \neg L_k(x, \overline{w}).$$

Notice that $L_k(a, b)$ is Σ_{k+1} in PA and so $\ell_n \Vdash p$ is provably equivalent to a Π_{k+1}-sentence.

If $k > n$ then notice that for any $x \in W$, by definition provably in PA,

$$\ell_n = \overline{x} \iff \exists N_0 \, \forall z > N_0 \, H_n(z, \overline{x}).$$

H_n is Δ_{n+1} in PA. It follows that $\ell_n = \overline{x}$ is Σ_{n+2} in PA and thus $\ell_n \neq \overline{x}$ is Π_{n+2} in PA. Since $\mathsf{PA} \vdash \exists! x \in W : \ell_n = \overline{x}$ (cf. Lemma 4.5.3), provably in PA,

$$\ell_n \Vdash p \iff \bigwedge\{\ell_n \neq \overline{x} \mid x \in W \setminus [\![p]\!]\}.$$

This proves the claim since $k = |p| \geq n + 1$. ☐

Therefore, $\cdot^*$ defines an arithmetical realization in the sense of Definition 4.2.2.

Lemma 4.5.9. *For all $n \in \omega$, $\ell_n = 0$ is true in the standard model of arithmetic.*

Proof. By Lemma 4.5.3, for every $n > M$ we either have $\ell_n R_\omega \ell_{n+1}$ or $\ell_{n+1} = \ell_n$. Since $\mathfrak{K}$ is finite and R_ω is transitive and irreflexive, there is a $z \in W$ and a k such that $\ell_m = \overline{z}$ is true for all $m \geq k$. Suppose $z \neq 0$ and consider the minimal m such that $\mathbb{N} \models \ell_m = \overline{z}$. The function h_m has to arrive at z via the second clause of the definition of h_m. So there is an $n \geq \max\{M, m\}$ such that $[m]_{\mathsf{PA}}(\ell_n \neq \overline{z})$ is true. Since PA is sound this means that $\ell_n \neq \overline{z}$ is true which contradicts our assumption. So there is a k such that $\ell_m = 0$ for all $m \geq k$. If $\ell_n \neq 0$ for some $n < k$ then by Lemma 4.5.3 we know that $\ell_n R_\alpha \ell_k$ for some $\alpha \in (n, k] \cup \{\omega\}$. But this is impossible since the node 0 has no incoming arcs. ☐

The following two main lemmas are from Beklemishev [7].

Lemma 4.5.10. *For any formula $A \in \Delta$,*

$$[\ell_n \Vdash A \mid n \geq M] \Rightarrow_{\mathsf{PA}} A^*.$$

Proof. By induction on A. The base cases where $A = \top$ or $A = p$ for some $p \in \Delta$ are trivial. Suppose $A = B \wedge C$. By inductive hypothesis we know

$$[\ell_n \Vdash B \mid n \geq M] \Rightarrow_{\mathsf{PA}} B^*,$$
$$[\ell_n \Vdash C \mid n \geq M] \Rightarrow_{\mathsf{PA}} C^*,$$

whence

$$[\ell_n \Vdash A \mid n \geq M] \Rightarrow_{\mathsf{PA}} B^*,$$
$$[\ell_n \Vdash A \mid n \geq M] \Rightarrow_{\mathsf{PA}} C^*,$$

and thus $\mathsf{PA} \vdash [\ell_n \Vdash A \mid n \geq M] \Rightarrow A^*$ follows.

Suppose $A = mB$ for some $m < \omega$. Then $A^* = \underline{\mathsf{Con}}_m(B^*)$ and so A^* numerates a finite extension of PA. Therefore it is sufficient to establish

$$\mathsf{PA} + \ell_M \Vdash A \vdash \underline{\mathsf{Con}}_m(B^*).$$

We know by the inductive hypothesis that

$$[\ell_n \Vdash B \mid n \geq M] \Rightarrow_{\mathsf{PA}} B^*,$$

and so

$$\mathsf{PA} + \mathsf{Con}_m[\ell_n \Vdash B \mid n \geq M] \vdash \mathsf{Con}_m(B^*).$$

By a formalized version of the compactness theorem, the formula $\mathsf{Con}_m[\ell_n \Vdash B \mid n \geq M]$ is equivalent to

$$\forall n \geq M\, \mathsf{Con}_m[\bigwedge_{k=M}^{n} \ell_k \Vdash B],$$

which is by Lemma 4.5.6 equivalent to

$$\forall n \geq M\, \mathsf{Con}_m[\ell_n \Vdash B].$$

To infer this sentence from $\ell_M \Vdash mB$, we reason in PA as follows. Suppose $\ell_M \Vdash mB$. Then there is a $w \in W$ such that $\ell_M R_m w$ and $w \Vdash B$. We know that $m \leq M$ and by Lemma 4.5.3, we either have $\ell_m R_\alpha \ell_M$ for some $\alpha \in (m, M] \cup \{\omega\}$ or $\ell_m = \ell_M$. Since $\mathfrak{K}$ is an RJ^*-model, we certainly have $\ell_m R_m w$. Suppose to the contrary that $\neg\mathsf{Con}_m[\ell_n \Vdash B]$ for some $n \geq M$. Then $[m]_{\mathsf{PA}}(\ell_n \nVdash B)$ and also $[m]_{\mathsf{PA}}(\ell_n \neq \overline{w})$. Consider an N_0 such that $\forall x \geq N_0\colon h_m(x) = \ell_m$. There exists a $y > x_0$ such that $\mathsf{Prf}_m(\ulcorner \ell_n \neq \overline{w} \urcorner, y)$, since there are arbitrarily long proofs. But then $h_m(y+1)$ is different from ℓ_m due to irreflexivity of $\mathfrak{K}$, a contradiction. This proves this case.

Suppose now that $A = \omega B$. We know that

$$[\mathsf{Con}_n(B^*) \mid n \geq M] \Rightarrow_{\mathsf{PA}} (\omega B)^*,$$

since the strength of $\mathsf{Con}_n(B^*)$ increases with n. Furthermore,

$$\begin{aligned} [\forall k \geq \overline{n}\, \mathsf{Con}_n[\ell_k \Vdash B] \mid n \geq M] &\Rightarrow_{\mathsf{PA}} [\forall k \geq M\, \mathsf{Con}_n[\ell_k \Vdash B] \mid n \geq M] \\ &\Rightarrow_{\mathsf{PA}} [\mathsf{Con}_n(B^*) \mid n \geq M], \end{aligned}$$

by the inductive hypothesis and Lemma 4.5.6. We prove

$$[\ell_n \Vdash \omega B \mid n \geq M] \Rightarrow_{\mathsf{PA}} [\forall k \geq \overline{n}\ \mathsf{Con}_n[\ell_k \Vdash B] \mid n \geq M],$$

by proving by an argument formalizable in PA that, for all $n \geq M$,

$$\mathsf{PA} + \ell_n \Vdash \omega B \vdash \forall k \geq \overline{n}\ \mathsf{Con}_n[\ell_k \Vdash B].$$

Let $n \geq M$ and suppose $\ell_n \Vdash \omega B$. Then $z \Vdash B$ for some $z \in W$ such that $\ell_n R_\omega z$ and z being different from ℓ_n. Now if $\exists k \geq \overline{n}\, [n]_{\mathsf{PA}}(\ell_k \nVdash B)$ then since $n \geq M$, we obtain that $\exists k \geq \max\{n, M\}\, [n]_{\mathsf{PA}}(\ell_k \neq z)$. But then h_n has to obtain a value different from ℓ_n by the second clause of the definition of h_n, a contradiction. ☐

Lemma 4.5.11. *For any formula $A \in \Delta$,*

$$\underline{\ell_0 \neq 0} \vee A^* \Rightarrow_{\mathsf{PA}} [\ell_n \Vdash A \mid n \geq M].$$

Proof. By induction on A. The cases $A = \top$ and $A = p$ for a variables p are immediate. The case of conjunction can be easily derived too. Suppose that $A = mB$ for $m < \omega$. Notice that $\ell_0 \neq 0$ is a Σ_1-sentence. Hence,

$$\begin{aligned}\mathsf{PA} + \ell_0 \neq 0 \wedge \mathsf{Con}_m(B^*) &\vdash \Box_{\mathsf{PA}}(\ell_0 \neq 0) \wedge \mathsf{Con}_m(B^*) \\ &\vdash \mathsf{Con}_m(\underline{\ell_0 \neq 0} \vee B^*) \\ &\vdash \mathsf{Con}_m[\ell_k \Vdash B \mid k \geq M] \\ &\vdash \forall k \geq M\, \mathsf{Con}_m[\ell_k \Vdash B].\end{aligned}$$

We prove that for all $n \geq M$,

$$\mathsf{PA} \vdash \ell_0 \neq 0 \wedge \ell_n \nVdash mB \to \exists k \geq M\, [m]_{\mathsf{PA}}(\ell_k \nVdash B).$$

We reason in PA as follows. Suppose $\ell_0 \neq 0$ and $\ell_n \nVdash mB$. We know by Lemma 4.5.3 that $\ell_m R_k \ell_n$ for some $k > m$, $\ell_m = \ell_n$, or $\ell_m R_\omega \ell_n$. In each case, since $\mathfrak{K}$ is an RJ^*_Λ-model, it holds that $\ell_m \nVdash mB$. Let $b := \ell_m$. We know that $\exists x\, L_m(x, b)$, whence $\ell_k \in R^*_m(b) \cup \{b\}$ follows for all $k \geq m$. Furthermore, $\exists x\, L_m(x, b)$ is easily seen to be a Σ_{m+1}-formula. Hence, $[m]_{\mathsf{PA}} \exists x\, L_m(x, \dot{b})$ and so

$$\forall k \geq m\, [m]_{\mathsf{PA}}(\ell_k \in R_m(\dot{b}) \cup \{\dot{b}\}). \tag{4.2}$$

We now claim that $\forall z \in R_m(b) \colon z \nVdash B$. For suppose otherwise, i.e., $z \Vdash B$ for some z such that $bR_\alpha z$ and $\alpha \geq m$. Then $b \Vdash \alpha B$, whence by Δ-monotonicity we have $b \Vdash mB$, contradicting $\ell_m \nVdash mB$. The formula $\forall z \in R_m(b) \colon z \nVdash B$ is bounded, hence

$$[m]_{\mathsf{PA}}(\forall z \in R_m(\dot{b}) \colon z \nVdash B). \tag{4.3}$$

We now prove that $\exists k \geq M\, [m]_{\mathsf{PA}}(\ell_k \neq \dot{b})$ which finishes this case by (4.2) and (4.3). Consider the minimal $i \leq m$ such that $\ell_i = \ell_m = b$. By $\ell_0 \neq 0$, we also have that $\ell_m = b \neq 0$. It follows that h_i can move to b only by virtue of the second clause of the definition of h_i. Therefore,

$$\exists k \geq \max\{M, i\}\, [i]_{\mathsf{PA}}(\ell_k \neq \dot{b}).$$

By $i \leq m \leq M$, it follows that

$$\exists k \geq M\, [m]_{\mathsf{PA}}(\ell_k \neq \dot{b}),$$

as required.

Suppose $A = \omega B$. It holds that

$$(\omega B)^* = [\mathsf{Con}_n(B^*) \mid n \in \omega].$$

By inductive hypothesis, we have

$$\underline{\ell_0 \neq 0} \vee B^* \Rightarrow_{\mathsf{PA}} [\ell_k \Vdash B \mid k \geq M],$$

whence

$$\underline{\mathsf{Con}}_n(\underline{\ell_0 \neq 0} \vee B^*) \Rightarrow_{\mathsf{PA}} \underline{\mathsf{Con}}_n[\ell_k \Vdash B \mid k \geq M],$$

follows for every $n \in \omega$. Therefore, for every $n \in \omega$,

$$\mathsf{PA} + \ell_0 \neq 0 + \mathsf{Con}_n(B^*) \vdash \forall k \geq M\, \mathsf{Con}_n[\ell_k \Vdash B].$$

Being formalizable uniformly in n, we obtain

$$\underline{\ell_0 \neq 0} \vee (\omega B)^* \Rightarrow_{\mathsf{PA}} [\forall k \geq M\, \mathsf{Con}_n[\ell_k \Vdash B] \mid n \geq M].$$

It remains to prove that

$$\underline{\ell_0 \neq 0} \vee [\forall k \geq M\, \mathsf{Con}_n[\ell_k \Vdash B] \mid n \geq M] \Rightarrow_{\mathsf{PA}} [\ell_n \Vdash \omega B \mid n > M],$$

since $[\ell_n \Vdash \omega B \mid n > M] \Rightarrow_{\mathsf{PA}} [\ell_n \Vdash \omega B \mid n \geq M]$. We show by an argument uniformly formalizable in n in PA that, for each $n > M$,

$$\mathsf{PA} + \ell_0 \neq 0 \wedge \ell_n \nVdash \omega B \vdash \exists k \geq M\, [n]_{\mathsf{PA}}(\ell_k \nVdash B).$$

We reason in PA as follows. Suppose that $n > M$ and $\ell_n \nVdash \omega D$. Let $b := \ell_n$ and consider the minimal $m \leq n$ such that $\ell_m = \ell_n = b$. We know by $\ell_0 \neq 0$ that $\ell_n = b \neq 0$. We first prove that

$$\forall k \geq \max\{m, M\}\, [n]_{\mathsf{PA}}(\ell_k \in R_\omega(\dot{b}) \cup \{\dot{b}\}). \tag{4.4}$$

We distinguish two cases. If $m > M$ then we infer $\exists x\, L_m(x, b)$ and $[m]_{\mathsf{PA}} \exists x\, L_m(x, b)$. As before, for $k \geq m > M$, it holds that $\ell_k \in R_\omega(b) \cup \{b\}$, whence $[m]_{\mathsf{PA}}(R_\omega(\dot{b}) \cup \{\dot{b}\})$ and so $\forall k \geq m\, [n]_{\mathsf{PA}}(\ell_k \in R_\omega(\dot{b}) \cup \{\dot{b}\})$ as desired.

Suppose now $m \leq M < n$. By the definition of h_n, it is easy to convince oneself that $\ell_M = b$. Thus $[n]_{\mathsf{PA}}(\ell_M = \dot{b})$ since $M < n$. Furthermore, for $k > M$, $\ell_M = b$ also entails that $\ell_k \in R_\omega(b) \cup \{b\}$, since $R^*_k(b) = R_\omega(b)$ and so h_k has to move on R_ω from b onward. Therefore, $\forall k \geq M\, [n]_{\mathsf{PA}}(\ell_k \in R_\omega(\dot{b}) \cup \{\dot{b}\})$ as before.

Notice that $\forall z \in R_\omega(b) \colon z \nVdash B$. As in the case before, from $\ell_k \in R_\omega(b)$, we thus easily infer that $\ell_k \nVdash B$. Hence by (4.4) we have

$$\forall k \geq \max\{m, M\}\, [n]_{\mathsf{PA}}(\ell_k \nVdash B \vee \ell_k = \dot{b}). \tag{4.5}$$

By $\ell_n = \ell_m = b \neq 0$ and the definition of h_m, we know that $\exists k \geq \max\{m, M\}\, [m]_{\mathsf{PA}}(\ell_k \neq \dot{b})$. Combining this with (4.5), we obtain $\exists k \geq \max\{m, M\}\, [n]_{\mathsf{PA}}(\ell_k \nVdash B)$ and so $\exists k \geq M\, [n]_{\mathsf{PA}}(\ell_k \nVdash B)$ as required. □

Now at the root $r = 1$ we have $\mathfrak{K}, r \Vdash A$ and $\mathfrak{K}, r \nVdash B$. Let σ be the numeration $[\ell_n = \overline{1} \mid n \geq M]$ and let S be the theory numerated by σ. By Lemma 4.5.10, we know that

$$\begin{aligned} \sigma &\Rightarrow_{\mathsf{PA}} [\ell_n \Vdash A \mid n \geq M] \\ &\Rightarrow_{\mathsf{PA}} A^*. \end{aligned}$$

By Lemma 4.5.11 we also have

$$\begin{aligned}\underline{\ell_0 \neq 0} \vee B^* &\Rightarrow_{\mathsf{PA}} [\ell_n \Vdash B \mid n \geq M] \\ &\Rightarrow_{\mathsf{PA}} [\ell_n \neq \overline{1} \mid n \geq M].\end{aligned}$$

Now if $A^* \Rightarrow B^*$ then $S \vdash \ell_M \neq \overline{1}$ and so S is inconsistent. Since $\mathsf{PA} \vdash \ell_n = \overline{1} \to \ell_m = \overline{1}$ for all $m \leq n$, we know that there is a PA-proof of $\ell_n \neq \overline{1}$ for some $n \geq M$. (For otherwise, $\mathsf{PA} \vdash S$ and so PA would be inconsistent too.) Therefore, the function h_0 has to take on a value different from 0 which is impossible since $\ell_0 = 0$ is true in the standard model. □

Example 4.5.12. Recall from Example 4.3.7 that $\mathsf{RC}^* \vdash \omega p \wedge \omega q \Rightarrow \omega(p \wedge q)$ iff $|p| < \omega$ or $|q| < \omega$. By Theorem 4.5.1 we infer that there are theories S, T extending PA such that

$$\mathsf{PA} + \mathsf{RFN}(S) + \mathsf{RFN}(T) \nvdash \mathsf{RFN}(S + T).$$

As also remarked by Beklemishev [7], note that both S and T must be both of unbounded arithmetical complexity, for suppose without loss of generality that S is a Π_{n+1}-axiomatized extension of PA. By Example 4.3.7 we know that $\mathsf{RC}^* \vdash \omega p \wedge \omega q \Rightarrow \omega(p \wedge q)$, whenever $|p| < \omega$. In particular, if $|p| = n$, we infer by Proposition 4.2.10 that

$$\mathsf{PA} + \mathsf{RFN}(S) + \mathsf{RFN}(T) \vdash \mathsf{RFN}(S + T).$$

Therefore, as remarked by Beklemishev [7], the use of infinite theories is necessary in the proof of Theorem 4.5.1. ⊣

For a strengthening of the insights due to the previous examples, we augment the proof of Theorem 4.5.1 by the following lemma which strengthens Lemma 4.5.6.

Lemma 4.5.13. *Let $n \geq m$ and A be a formula such that $|A| \leq m$. Then,*

$$\mathsf{PA} \vdash (\ell_m \Vdash A) \to (\ell_n \Vdash A).$$

Proof. Assume $n > m \geq |A|$ and, reasoning in PA, suppose $\ell_n \nVdash A$. By Lemma 4.5.3 we know that $\ell_m = \ell_n$ or $\ell_m R_\alpha \ell_n$ for some $\alpha \in (m, n] \cup \{\omega\}$. The case of $\ell_m = \ell_n$ is clear, so suppose $\ell_m R_\alpha \ell_n$. By Lemma 4.5.5 we then also have $\ell_m \nVdash A$ as $|A| \leq m < \alpha$. □

Although we know that the use of infinite theories is necessary by the previous example, we readily see that the following corollary can be read of the proof of Theorem 4.5.1.

Corollary 4.5.14. *Let A and B be formulas such that $\mathsf{RC}^* \nvdash A \Rightarrow B$. Then there exists an arithmetical realization $\cdot^*$ such that*

(i) *for any variable p occurring in A or B, p^* numerates a finite extension of PA in case $|p| < \omega$;*

(ii) *$A^* \Rightarrow B^*$ does not hold.*

Proof. Consider the arithmetical interpretation $\cdot^*$ constructed in the proof of Theorem 4.5.1. Let p be a variable of sort n occurring in A or B. By definition,

$$p^* = [\ell_n \Vdash p \mid n \geq M].$$

We claim that p^* is provably equivalent in PA to $\ell_k \Vdash p$, where $k = \max\{M, n\}$. The fact that $\ell_k \Vdash p$ is implied by p^* in PA is immediate by the definition of p^*. We thus show that p^* follows from $\ell_k \Vdash p$ in PA. Suppose first that $n < M$. Let $m > M$ and, reasoning in PA, suppose $\ell_M \Vdash p$. We know that either $\ell_m \Vdash p$ or $\ell_m \nVdash p$. Now if $\ell_m \nVdash p$ then Lemma 4.5.13 yields $\ell_M \nVdash p$ which is impossible. Suppose now that $n \geq M$. For all j such that $M \leq j \leq n$ we know that PA proves that $\ell_n \Vdash p$ implies $\ell_j \Vdash p$ by Lemma 4.5.6. So suppose $m > n$ and assume $\ell_n \Vdash p$. Then either $\ell_m \Vdash p$ or $\ell_m \nVdash p$, where we see that the latter is impossible by Lemma 4.5.13 and considering the fact that $|p| = n < m$.

Hence, we can let

$$p^* := \underline{\ell_k \Vdash p}.$$

Notice that, as in the proof of Lemma 4.5.8, we can directly find a Π_{n+1}-sentence which is equivalent in PA to the sentence $\ell_k \Vdash p$. □

Bibliography

[1] Sergei N. Artemov and Lev D. Beklemishev. Provability Logic. In *Handbook of Philosophical Logic, 2nd ed.*, pages 229–403. Kluwer, 2004.

[2] Lev D. Beklemishev. Provability algebras and proof-theoretic ordinals, I. *Annals of Pure and Applied Logic*, 128(1-3):103–123, 2004.

[3] Lev D. Beklemishev. *Reflection Principles and Provability Algebras in Formal Arithmetic.* Logic Group preprint series. Department of Philosophy, Utrecht Univ., 2005.

[4] Lev D. Beklemishev. The Worm Principle. In Z. Chatzidakis, P. Koepke, and W. Pohlers, editors, *Lecture Notes in Logic 27. Logic Colloquium '02*, pages 75–95. AK Peters, 2006. Preprint: Logic Group Preprint Series 219, Utrecht Univ., March 2003.

[5] Lev D. Beklemishev. Kripke semantics for provability logic GLP. *Annals of Pure and Applied Logic*, 161(6):756–774, 2010.

[6] Lev D. Beklemishev. A simplified proof of arithmetical completeness theorem for provability logic GLP. *Proceedings of the Steklov Institute of Mathematics*, 274(1): 25–33, 2011.

[7] Lev D. Beklemishev. Positive provability logic for uniform reflection principles. *Annals of Pure and Applied Logic*, 165(1):82–105, 2014.

[8] Lev D. Beklemishev, David Fernández-Duque, and Joost J. Joosten. On provability logics with linearly ordered modalities. *Studia Logica*, 102(3):541–566, 2014.

[9] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic.* Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2002.

[10] George S. Boolos. Omega-consistency and the diamond. *Studia Logica*, 39(2–3): 237–243, 1980.

[11] George S. Boolos. *The Logic of Provability.* Cambridge University Press, 1993.

[12] George S. Boolos, John P. Burgess, and Richard C. Jeffrey. *Computability and Logic.* Cambridge University Press, fifth edition, 2007.

[13] Evgeny V. Dashkov. On the positive fragment of the polymodal provability logic GLP. *Mathematical Notes*, 91(3):318–333, 2012. Original Russsian text in: *Matematicheskie Zametki*, 91:(3):331–336, 2012.

[14] Dick de Jongh and Giorgi K. Japaridze. The Logic of Provability. In Samuel R. Buss, editor, *Handbook of Proof Theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 475–546. Elsevier, Amsterdam, 1998.

[15] J. Michael Dunn. Positive Modal Logic. *Studia Logica*, 55(2):301–317, 1995.

[16] Solomon Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, 49(1):35–92, 1960.

[17] Gerhard Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, 112(1):493–565, 1936.

[18] Gerhard Gentzen. Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie. 1938.

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

[20] Kurt Gödel. Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse Math. Kolloq.*, 4:39–40, 1933.

[21] Petr Hájek and Pavel Pudlák. *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag Berlin Heidelberg, 1998.

[22] Leon Henkin. A problem concerning provability. *The Journal of Symbolic Logic*, 17 (2):160, 1952.

[23] David Hilbert and Paul Bernays. *Grundlagen der Mathematik II*. Springer-Verlag, second edition, 1970.

[24] Konstantin N. Ignatiev. On Strong Provability Predicates and the Associated Modal Logics. *The Journal of Symbolic Logic*, 58(1):249–290, 03 1993.

[25] Giorgi K. Japaridze. The polymodal logic of provability. In *Intensional Logics and Logical Structure of Theories: Material from the fourth Soviet-Finnish Symposium on Logic, Telavi, May 20-24, 1985, Metsniereba, Tbilisi*, pages 16–48, 1988. In Russian.

[26] Stephen C. Kleene. *Introduction to Metamathematics*. North-Holland Publishing Company, 1952. Co-publisher: ISHI Press Int., 2009.

[27] Georg Kreisel. Wie die Beweistheorie zu ihren Ordinalzahlen kam und kommt. *Jahresbericht der DMV*, 78(4):177–223, 1977.

[28] Georg Kreisel and Azriel Lévy. Reflection principles and their use for establishing the complexity of axiomatic systems. *Zeitschrift f. math. Logik und Grundlagen d. Math.*, 14:97–142, 1968.

[29] Kenneth Kunen. *Set Theory*, volume 34 of *Studies in logic*. College Publications, 2011.

[30] Martin H. Löb. Solution of a problem of Leon Henkin. *The Journal of Symbolic Logic*, 20:115–118, 1955.

[31] Roberto Magari. The diagonalizable algebras (the algebraization of the theories which express Theor.:II). *Bollettino della Unione Matematica Italiana*, 12(4):117–125, 1975.

[32] Michael Rathjen. The realm of ordinal analysis. In S. Cooper and J. Truss, editors, *Sets and Proofs*, pages 219–279. Cambridge University Press, 1999.

[33] John B. Rosser. Extensions of Some Theorems of Gödel and Church. *The Journal of Symbolic Logic*, 1(3):87–91, 1936.

[34] Ilya Shapirovsky. PSPACE-decidability of Japaridze's polymodal logic. In *Advances in Modal Logic*, volume 7, pages 289–304, 2008.

[35] Joseph R. Shoenfield. *Mathematical Logic*. Addison-Wesley Series in Logic. Addison-Wesley, 1967.

[36] Craig Smoryński. The Incompleteness Theorems. In Jon Barwise, editor, *Handbook of Mathematical Logic*, volume 90 of *Studies in logic and the foundations of mathematics*, pages 821–865. North-Holland Pub. Co., 1977.

[37] Craig Smoryński. Modal Logic and Self-Reference. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic: Volume II: Extensions of Classical Logic*, pages 441–495. Reidel, Dordrecht, 1984.

[38] Craig Smoryński. *Self-Reference and Modal Logic*. Springer-Verlag, 1985.

[39] Robert M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 25(3-4):287–304, 1976.

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

[41] Richard Zach. Hilbert's program. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2014 edition, 2014.