Institute for Logic, Language and Computation

# UNDECIDABILITY IN DIAGONALIZABLE ALGEBRAS

V.Yu. Shavrukov

ILLC Prepublication Series
for Mathematical Logic and Foundations ML-93-13

University of Amsterdam

# The ILLC Prepublication Series

**1990** *Logic, Semantics and Philosophy of Language*

LP-90-01 Jaap van der Does A Generalized Quantifier Logic for Naked Infinitives
LP-90-02 Jeroen Groenendijk, Martin Stokhof Dynamic Montague Grammar
LP-90-03 Renate Bartsch Concept Formation and Concept Composition
LP-90-04 Aarne Ranta Intuitionistic Categorial Grammar
LP-90-05 Patrick Blackburn Nominal Tense Logic
LP-90-06 Gennaro Chierchia The Variablity of Impersonal Subjects
LP-90-07 Gennaro Chierchia Anaphora and Dynamic Logic
LP-90-08 Herman Hendriks Flexible Montague Grammar
LP-90-09 Paul Dekker The Scope of Negation in Discourse, towards a Flexible Dynamic Montague grammar
LP-90-10 Theo M.V. Janssen Models for Discourse Markers
LP-90-11 Johan van Benthem General Dynamics
LP-90-12 Serge Lapierre A Functional Partial Semantics for Intensional Logic
LP-90-13 Zhisheng Huang Logics for Belief Dependence
LP-90-14 Jeroen Groenendijk, Martin Stokhof Two Theories of Dynamic Semantics
LP-90-15 Maarten de Rijke The Modal Logic of Inequality
LP-90-16 Zhisheng Huang, Karen Kwast Awareness, Negation and Logical Omniscience
LP-90-17 Paul Dekker Existential Disclosure, Implicit Arguments in Dynamic Semantics

*Mathematical Logic and Foundations*

ML-90-01 Harold Schellinx Isomorphisms and Non-Isomorphisms of Graph Models
ML-90-02 Jaap van Oosten A Semantical Proof of De Jongh's Theorem
ML-90-03 Yde Venema Relational Games
ML-90-04 Maarten de Rijke Unary Interpretability Logic
ML-90-05 Domenico Zambella Sequences with Simple Initial Segments
ML-90-06 Jaap van Oosten Extension of Lifschitz' Realizability to Higher Order Arithmetic, and a Solution to a Problem of F. Richman
ML-90-07 Maarten de Rijke A Note on the Interpretability Logic of Finitely Axiomatized Theories
ML-90-08 Harold Schellinx Some Syntactical Observations on Linear Logic
ML-90-09 Dick de Jongh, Duccio Pianigiani Solution of a Problem of David Guaspari
ML-90-10 Michiel van Lambalgen Randomness in Set Theory
ML-90-11 Paul C. Gilmore The Consistency of an Extended NaDSet

*Computation and Complexity Theory*

CT-90-01 John Tromp, Peter van Emde Boas Associative Storage Modification Machines
CT-90-02 Sieger van Denneheuvel, Gerard R. Renardel de Lavalette A Normal Form for PCSJ Expressions
CT-90-03 Ricard Gavaldà, Leen Torenvliet, Osamu Watanabe, José L. Balcázar Generalized Kolmogorov Complexity in Relativized Separations
CT-90-04 Harry Buhrman, Edith Spaan, Leen Torenvliet Bounded Reductions
CT-90-05 Sieger van Denneheuvel, Karen Kwast Efficient Normalization of Database and Constraint Expressions
CT-90-06 Michiel Smid, Peter van Emde Boas Dynamic Data Structures on Multiple Storage Media, a Tutorial
CT-90-07 Kees Doets Greatest Fixed Points of Logic Programs
CT-90-08 Fred de Geus, Ernest Rotterdam, Sieger van Denneheuvel, Peter van Emde Boas Physiological Modelling using RL
CT-90-09 Roel de Vrijer Unique Normal Forms for Combinatory Logic with Parallel Conditional, a case study in conditional rewriting

*Other Prepublications*

X-90-01 A.S. Troelstra Remarks on Intuitionism and the Philosophy of Mathematics, Revised Version
X-90-02 Maarten de Rijke Some Chapters on Interpretability Logic
X-90-03 L.D. Beklemishev On the Complexity of Arithmetical Interpretations of Modal Formulae
X-90-04 Annual Report 1989
X-90-05 Valentin Shehtman Derived Sets in Euclidean Spaces and Modal Logic
X-90-06 Valentin Goranko, Solomon Passy Using the Universal Modality: Gains and Questions
X-90-07 V.Yu. Shavrukov The Lindenbaum Fixed Point Algebra is Undecidable
X-90-08 L.D. Beklemishev Provability Logics for Natural Turing Progressions of Arithmetical Theories
X-90-09 V.Yu. Shavrukov On Rosser's Provability Predicate
X-90-10 Sieger van Denneheuvel, Peter van Emde Boas An Overview of the Rule Language RL/1
X-90-11 Alessandra Carbone Provable Fixed points in $I\Delta_0+\Omega_1$, revised version
X-90-12 Maarten de Rijke Bi-Unary Interpretability Logic
X-90-13 K.N. Ignatiev Dzhaparidze's Polymodal Logic: Arithmetical Completeness, Fixed Point Property, Craig's Property
X-90-14 L.A. Chagrova Undecidable Problems in Correspondence Theory
X-90-15 A.S. Troelstra Lectures on Linear Logic

**1991** *Logic, Semantics and Philosophy of Language*

LP-91-01 Wiebe van der Hoek, Maarten de Rijke Generalized Quantifiers and Modal Logic
LP-91-02 Frank Veltman Defaults in Update Semantics
LP-91-03 Willem Groeneveld Dynamic Semantics and Circular Propositions
LP-91-04 Makoto Kanazawa The Lambek Calculus enriched with Additional Connectives
LP-91-05 Zhisheng Huang, Peter van Emde Boas The Schoenmakers Paradox: Its Solution in a Belief Dependence Framework
LP-91-06 Zhisheng Huang, Peter van Emde Boas Belief Dependence, Revision and Persistence
LP-91-07 Henk Verkuyl, Jaap van der Does The Semantics of Plural Noun Phrases
LP-91-08 Víctor Sánchez Valencia Categorial Grammar and Natural Reasoning
LP-91-09 Arthur Nieuwendijk Semantics and Comparative Logic
LP-91-10 Johan van Benthem Logic and the Flow of Information

*Mathematical Logic and Foundations*

ML-91-01 Yde Venema Cylindric Modal Logic
ML-91-02 Alessandro Berarducci, Rineke Verbrugge On the Metamathematics of Weak Theories
ML-91-03 Domenico Zambella On the Proofs of Arithmetical Completeness for Interpretability Logic
ML-91-04 Raymond Hoofman, Harold Schellinx Collapsing Graph Models by Preorders
ML-91-05 A.S. Troelstra History of Constructivism in the Twentieth Century
ML-91-06 Inge Bethke Finite Type Structures within Combinatory Algebras
ML-91-07 Yde Venema Modal Derivation Rules
ML-91-08 Inge Bethke Going Stable in Graph Models
ML-91-09 V.Yu. Shavrukov A Note on the Diagonalizable Algebras of PA and ZF
ML-91-10 Maarten de Rijke, Yde Venema Sahlqvist's Theorem for Boolean Algebras with Operators
ML-91-11 Rineke Verbrugge Feasible Interpretability
ML-91-12 Johan van Benthem Modal Frame Classes, revisited

*Computation and Complexity Theory*

CT-91-01 Ming Li, Paul M.B. Vitányi Kolmogorov Complexity Arguments in Combinatorics
CT-91-02 Ming Li, John Tromp, Paul M.B. Vitányi How to Share Concurrent Wait-Free Variables
CT-91-03 Ming Li, Paul M.B. Vitányi Average Case Complexity under the Universal Distribution Equals Worst Case Complexity
CT-91-04 Sieger van Denneheuvel, Karen Kwast Weak Equivalence
CT-91-05 Sieger van Denneheuvel, Karen Kwast Weak Equivalence for Constraint Sets
CT-91-06 Edith Spaan Census Techniques on Relativized Space Classes
CT-91-07 Karen L. Kwast The Incomplete Database
CT-91-08 Kees Doets Levationis Laus
CT-91-09 Ming Li, Paul M.B. Vitányi Combinatorial Properties of Finite Sequences with high Kolmogorov Complexity
CT-91-10 John Tromp, Paul Vitányi A Randomized Algorithm for Two-Process Wait-Free Test-and-Set
CT-91-11 Lane A. Hemachandra, Edith Spaan Quasi-Injective Reductions
CT-91-12 Krzysztof R. Apt, Dino Pedreschi Reasoning about Termination of Prolog Programs

*Computational Linguistics*

CL-91-01 J.C. Scholtes Kohonen Feature Maps in Natural Language Processing
CL-91-02 J.C. Scholtes Neural Nets and their Relevance for Information Retrieval
CL-91-03 Hub Prüst, Remko Scha, Martin van den Berg A Formal Discourse Grammar tackling Verb Phrase Anaphora

**Institute for Logic, Language and Computation**
Plantage Muidergracht 24
1018TV Amsterdam
Telephone 020-525.6051, Fax: 020-525.5101

# UNDECIDABILITY IN DIAGONALIZABLE ALGEBRAS

V.Yu. Shavrukov
Department of Mathematics and Computer Science
University of Amsterdam

*ILLC Prepublications*
*for Mathematical Logic and Foundations*
ISSN 0928-3315
*Coordinating editor: Dick de Jongh*

received September 1993

# Undecidability in Diagonalizable Algebras

V. Yu. Shavrukov
Department of Mathematics and Computer Science
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
the Netherlands
volodya@fwi.uva.nl

September 13, 1993

**Abstract.** If a formal theory T is able to reason about its own syntax, then the diagonalizable algebra of a formal theory T is defined as its Lindenbaum sentence algebra endowed with a unary operator $\Box$ which sends a sentence $\varphi$ to the sentence $\Box\varphi$ asserting the provability of $\varphi$ in T. We prove that the elementary theories of diagonalizable algebras of a wide class of theories are undecidable and establish some related results.

## 0. Introduction

A *diagonalizable algebra* $\mathcal{D}$ is a Boolean algebra $\mathcal{A}$ together with an operator $\Box$ satisfying the following identities:

$$\Box(\alpha \to \beta) \leq \Box\alpha \to \Box\beta$$

$$\Box\alpha \leq \Box\Box\alpha$$

$$\Box(\Box\alpha \to \alpha) = \Box\alpha$$

$$\Box\top = \top,$$

where $\top$ is the unit of $\mathcal{A}$. These were introduced by Magari [13] with the following example in mind:

Take a formal theory T over classical propositional logic containing some arithmetic and consider $\mathcal{A}_{\mathrm{T}}$, the *Lindenbaum sentence algebra* of T. $\mathcal{A}_{\mathrm{T}}$ consists of classes of T-provably equivalent sentences (i.e. formulas without free variables). The Boolean operations on $\mathcal{A}_{\mathrm{T}}$ are induced by propositional connectives from the language of T in the obvious way. The operator $\Box_{\mathrm{T}}$, which turns $\mathcal{A}_{\mathrm{T}}$ into a diagonalizable algebra $\mathcal{D}_{\mathrm{T}}$ called the *diagonalizable algebra of* T, comes from the *provability predicate* of T. This provability predicate is a formula $\mathrm{Pr}_{\mathrm{T}}(x)$ in the language of T expressing that $x$ is a gödelnumber of a formula provable in T. One takes $\Box_{\mathrm{T}} : \varphi \mapsto \mathrm{Pr}_{\mathrm{T}}(\ulcorner\varphi\urcorner)$, where $\ulcorner\varphi\urcorner$ is the gödelnumber of the sentence $\varphi$. $\mathrm{Pr_T}(x)$ is constructed in such a way as to most faithfully represent the inductive definition of (Hilbert-style) provability from the axioms of an effectively presented formal system (see Feferman [9, § 4] for a detailed description of the provability predicate), so that the theory T is itself able to follow many of our arguments about provability in T, translating them as those about $\mathrm{Pr_T}(x)$. This circumstance takes care of $\mathcal{D}_{\mathrm{T}}$'s being, as has been claimed, a diagonalizable algebra, which can be restated as saying that for all T-sentences $\alpha$ and $\beta$

$$\vdash \mathrm{Pr_T}(\ulcorner \alpha \rightarrow \beta \urcorner) \rightarrow (\mathrm{Pr_T}(\ulcorner \alpha \urcorner) \rightarrow \mathrm{Pr_T}(\ulcorner \beta \urcorner))$$

$$\vdash \mathrm{Pr_T}(\ulcorner \alpha \urcorner) \rightarrow \mathrm{Pr_T}(\ulcorner \mathrm{Pr_T}(\ulcorner \alpha \urcorner) \urcorner) \qquad (\Sigma\text{-completeness})$$

$$\vdash \mathrm{Pr_T}(\ulcorner \mathrm{Pr_T}(\ulcorner \alpha \urcorner) \rightarrow \alpha \urcorner) \rightarrow \mathrm{Pr_T}(\ulcorner \alpha \urcorner) \qquad (\text{Löb's Theorem})$$

$$\vdash \alpha \Rightarrow \vdash \mathrm{Pr_T}(\ulcorner \alpha \urcorner).$$

($\vdash$ stands for provability in T and, further on, will stand for that in the formal theory currently under consideration. In parentheses the names by which the corresponding principles will in future be referred to are given.) See Solovay [26] or Smoryński [24, Theorem 0.6.18(i)] for verifications of the four stated facts, which together go by the name of *Löb's Derivability Conditions.*

The present paper is entirely devoted to $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$, the first order theory of the diagonalizable algebras of formal theories T. We present two (not too drastically) different proofs of the undecidability of $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$ for T coming from a reasonably large variety of formal theories, such prominent natural examples as Peano arithmetic PA and Zermelo–Fraenkel set theory ZF included. The first proof establishes the nonarithmeticity of $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$ for $\Sigma_1$-sound T. (T is $\Sigma_1$-*sound* if every sentence of the form $\mathrm{Pr_T}(\ulcorner \varphi \urcorner)$ proved by T is true, or, equivalently, if T proves no false $\Sigma_1$-sentences; T is $\Sigma_1$-*ill* otherwise.) The second proof gives mere undecidability, but works for a still larger class of formal theories and gives a sharper upper bound on undecidable quantifier alternations. Our theorems are answers to question(s) found in e.g. Montagna [18] and, more recently, in Artemov & Beklemishev [3].

Among earlier results concerning (un)decidability questions related to diagonalizable algebras, one should mention Solovay's Theorem [26], which shows the equational theory of $\mathcal{D}_{\mathrm{T}}$ for T $\Sigma_1$-sound to coincide with a decidable modal logic **L**. Later investigations by Artemov [2] and Visser [27] adjust Solovay's discovery to any formal theory fulfilling minimal strength conditions. Smoryński [22] strengthens Solovay's results upto the decidability of $\mathrm{Th}_{\bar{\forall}}\mathcal{D}_{\mathrm{T}}$, the universal theory of $\mathcal{D}_{\mathrm{T}}$. For these T, $\mathrm{Th}_{\bar{\forall}}\mathcal{D}_{\mathrm{T}}$ does not depend on the particular choice of one. In Section 4 we indicate how to get the same for other kinds of formal theories. Artemov & Beklemishev [3] obtain decidability and undecidability results for first order theories of a number of individual diagonalizable algebras. It has also been known that the first order theory of the whole variety of diagonalizable algebras is hereditarily undecidable (Montagna [19] and Smoryński [23]).

The paper is organized as follows: Section 1 constructs a parameter-free first order definition of the set $\{\Box^n\Box\bot\}_{n\in\omega}$ in diagonalizable algebras of $\Sigma_1$-sound theories. This definability is put to use in Section 2, where it serves to extract the nonarithmeticity of the first order theories of these algebras. In Section 3 we are going to show that the mechanics of a special class of Post canonical systems are in a manner of way reflected in the structure of the diagonalizable algebra of any theory of *infinite credibility extent*, i.e. any theory satisfying $\vdash \Box^n\Box\bot$ for no $n \in \omega$. Finally, Section 4 harvests the undecidability of $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$ for theories T of infinite credibility extent as well as the undecidability of these theories' opinions on their own diagonalizable algebras. New questions generated by our answers are scattered around the paper in hope for potential researchers.

Throughout the sequel we shall use $\Box\varphi$ in place of $\mathrm{Pr}_{\mathrm{T}}(\ulcorner\varphi\urcorner)$ for T-sentences $\varphi$. Löb's Derivability Conditions, when written out using this convention, take on a more compact form

$$\vdash \Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$$

$$\vdash \Box\alpha \to \Box\Box\alpha$$

$$\vdash \Box(\Box\alpha \to \alpha) \to \Box\alpha$$

$$\vdash \alpha \;\Rightarrow\; \vdash \Box\alpha,$$

in which shape they could be forgivably mistaken for the extrapropositional axioms and rules of the modal logic **L** (cf. Smoryński [24, Chapter 1, **L**=**PRL**]). A consequence of this typographical coincidence is the wide-ranging utility of modal-logical methods in the study of diagonalizable algebras of formal theories. In Section 2 we shall sample the flavour of such applications.

To add to confusion, we shall even slightly readjust our way of presenting diagonalizable algebraic expressions, writing the more suggestive $\vdash \alpha$ for $\alpha = \top$; $\vdash \alpha \to \beta$ for $\alpha \leq \beta$; etc. Please note that our conventions deal away with many distinctions between diagonalizable algebraic, provabilistic, and modal-logical notation. Apart from the obvious drawbacks, this notational manoeuvre may merit some appreciation for promoting unity, for every once in a while the reader of this paper will be encountering arguments about the diagonalizable algebra $\mathcal{D}_{\mathrm{T}}$, involving (and occasionally dipping within) a formal theory T, that appeal to his/her knowledge of **L** for substantiation of certain claims.

In the context of formal theories T this unified notation may, for the purposes of our exposition, be treated as part of the generally more expressive vocabulary of T. A key feature of T is that it is able to talk first order arithmetic, which is needed to carry out the gödelnumbering of the syntax of T in the first place, for otherwise $\mathcal{D}_{\mathrm{T}}$ would not be a well-defined object. Thus we shall assume either that the language of T physically contains symbols for arithmetical operations, or that a particular interpretation of the arithmetical language in T is given. (We shall later specify exactly how much arithmetic T should know.) In order to grease the interaction between the arithmetic part of the language of T and the chosen modal logic-like format for its pronouncements on $\mathcal{D}_{\mathrm{T}}$, we shall somewhat relax the orthodoxy of that format.

The first step in this direction is to allow quantification to percolate inside (scopes of) $\Box$'s as in the expression $\forall x\, \Box \forall y\, \varphi(x, y)$, whose meaning should be transparent once one recalls that $\Box$ stands for the provability predicate of T. Next, in formal theories it is possible to quantify over iterates $\Box^x$ of the provability predicate, which legitimizes expressions like $\Diamond \forall x\, \Box^x \Box \bot$ ($\Diamond$ is short for $\neg\Box\neg$), with the understanding that $\Box^0\varphi$ is the same as $\varphi$. Furthermore, we shall write $\Box_y\varphi$ ($\Box_{\leq y}\varphi$) for the T-formula expressing that $\varphi$ has a T-proof of gödelnumber (smaller than or equal to) $y$. Their 'informal' analogues $\vdash_n \varphi$ ($\vdash_{\leq n} \varphi$) are used to convey to the reader messages of similar content. A very useful schema now easily formulated is known under the name of the *Small*

*Reflection Principle*:

$$\vdash \forall \varphi, x \, \Box(\Box_{\leq x}\varphi \rightarrow \varphi).$$

(Note, incidentally, that we do not hesitate to quantify over sentences occurring within the scope of a $\Box$ which is itself within the scope of the quantifier in question.) The Small Reflection Principle is a formalization of the obvious fact for any natural number $n$ and any sentence $\varphi$, either $\varphi$ itself or the fact that no $m \leq n$ is the gödelnumber of a T-proof of $\varphi$ is provable in T.

The requirements on the strength of the formal theories T in this paper are as follows: It is certainly safe to presuppose that T contains PA. In fact, the theory $I\Sigma_1$ suffices throughout, although not all of our arguments intended to formalize in T do so in $I\Sigma_1$ straightforwardly. Furthermore, it is even possible to obtain all our Theorems and Corollaries for theories extending just $I\Delta_0$+exp. This would, however, necessitate extensive modifications in our constructions as well as arguments along the lines of Zambella [29].

The theories mentioned can be looked up in Hájek & Pudlák [10]. The definition of the hierarchy $\{\Sigma_n, \Pi_n\}_{n\in\omega}$ of arithmetical formulas, to whose lower levels we are going to refer, is also found there.

Underlying the heuristics of almost every construction in this paper is the mental picture of $\mathcal{D}^*_{\mathrm{T}}$, the *dual space* of $\mathcal{D}_{\mathrm{T}}$ (which is not to say that familiarity with this object is formally necessary for understanding our arguments). The dual space $\mathcal{D}^*$ of a diagonalizable algebra $\mathcal{D}$ is the Stone dual space of the Boolean structure of $\mathcal{D}$ (i.e. the topological space of ultrafilters on $\mathcal{D}$ (= maximal consistent extensions of T in case $\mathcal{D} = \mathcal{D}_{\mathrm{T}}$) with sets of the form $\{ x \mid \alpha \in x \}$, where $\alpha \in \mathcal{D}$, constituting an open basis for the accompanying topology; clopens $\alpha^*$ in $\mathcal{D}^*$ correspond to elements $\alpha$ of $\mathcal{D}$) serving as domain for a binary relation $R$ defined by

$$x \, R \, y \quad \text{iff} \quad \forall \alpha{\in}\mathcal{D} \, (\Box\alpha \in x \;\Rightarrow\; \alpha \in y).$$

Boolean operations in $\mathcal{D}$ correspond to set-theoretical ones in $\mathcal{D}^*$. The operator $\Box$ of $\mathcal{D}$ is mimicked in $\mathcal{D}^*$ by the operator $\Box^*$ on the power set of $\mathcal{D}^*$. For $X \subseteq \mathcal{D}^*$ we have

$$\Box^* X = \{ x \in \mathcal{D}^* \mid \forall y{\in}\mathcal{D}^* \, (x \, R \, y \;\Rightarrow\; y \in X) \}.$$

$\Box^* X$ is clopen if so is $X$. The upshot is that $\Box^*\alpha^* = (\Box\alpha)^*$ for all $\alpha \in \mathcal{D}$ (cf. Abashidze [1] or Magari [14]). See Montagna [17] for a breathtaking glimpse inside $\mathcal{D}^*_{\mathrm{PA}}$.

We shall be referring to an element $y$ of $\mathcal{D}^*$ as lying *(R-)above* an element $x$ if $x \, R \, y$. This spatial orientation suggests the definition of $d(x)$, the *(R-)depth* of an ultrafilter $x \in \mathcal{D}^*$. $d(x)$ is an ordinal defined as the supremum of $d(y)+1$ over all $y \in \mathcal{D}^*$ satisfying $x \, R \, y$. Clearly, only the ultrafilters in the well-founded part of $R^{-1}$ enjoy a well-defined depth. The well-founded part of $R^{-1}$ can be shown to coincide with the *finite (R-)depth part* of $\mathcal{D}^*$, i.e. the open set $\{ x \in \mathcal{D}^* \mid d(x) < \omega \}$.

Another circumstance contributing to the relevance of the earlier mentioned modal logic **L** to the study of diagonalizable algebras of formal theories T is that its Kripke models, that are indispensable in modal logic proper, come, for many practical purposes, close to being factors of $\mathcal{D}^*_{\mathrm{T}}$. Many arguments shedding light on the structure of $\mathcal{D}_{\mathrm{T}}$ can be seen as demonstrating that $\mathcal{D}^*_{\mathrm{T}}$ 'factors', in a weaker or stronger sense, onto a particular (class of) Kripke model(s).

The first example of such an argument is found in Solovay [26], whence the method of *Solovay functions* originated. These functions provide a systematic way of constructing, given a Kripke model $\boldsymbol{K}$, particular T-sentences $\xi_a$ corresponding to nodes $a$ of $\boldsymbol{K}$ such that the dual space of the subalgebra generated within $\mathcal{D}_{\mathrm{T}}$ by the sentences $\xi_a$ shares with $\boldsymbol{K}$ some of its properties.

Solovay functions have also been successfully applied to the study of algebraic behaviour of formal predicates other than the provability predicate, accumulating a rich variety of tricks for attaining diverse goals. The constructions in the present paper borrow heavily from this arsenal. We employ two Solovay functions very close in form and spirit to the ones devised by Berarducci [5] and Dzhaparidze [7] and [8] to deal with problems (originating) in interpretability logic. This, in the author's view, shows that investigations into other predicates of metatheoretical extraction might be not irrelevant to our better understanding of the provability predicate.

# 1. Defining the natural numbers

This Section is devoted to first order defining the set $\{\Box^n\Box\bot\}_{n\in\omega}$ in the diagonalizable algebras of $\Sigma_1$-sound theories T. We therefore fix such a theory T for the whole of the Section. On default, T is the formal theory we are dealing with in various definitions and lemmas.

1.1. DEFINITION. We define two predicates in the language of the first order theory of diagonalizable algebras:

$$\sigma \in B \equiv \exists\varphi \vdash \sigma \leftrightarrow \Box\varphi$$

$$\sigma \in T \equiv \sigma \in B \quad \& \ \forall\xi \Big( (\vdash \Box\bot \to \xi \ \& \ \forall\tau{\in}B\, (\vdash \tau \to \xi \Rightarrow \vdash \Box\tau \to \xi)) \Rightarrow \vdash \sigma \to \xi \Big).$$

Clearly, $\sigma \in B$ expresses that $\sigma$ is of the form $\Box\varphi$. In the dual space of a diagonalizable algebra, $\sigma$ corresponds then to an $R$-upwards closed clopen $\sigma^*$. Note that in any diagonalizable algebra the set $B$ is closed under conjunction.

We shall be referring to elements of $B$ as *box* elements (or sentences) and to those in $T$ as *top-box* ones. The prefix "top-" hints at the fact that in the diagonalizable algebra of any theory S, a top-box sentence corresponds to a clopen lying entirely within the finite $R$-depth part of the dual space $\mathcal{D}^*_{\mathrm{S}}$, as we shall shortly see.

1.2. LEMMA. *In any diagonalizable algebra, if* $\vdash \tau \to \Box^n\Box\bot$ *for some* $n \in \omega$ *and* $\tau$ *is a box sentence, then* $\tau \in T$. *In particular,* $\Box^n\Box\bot \in T$ *for all* $n \in \omega$.

PROOF. First, it is clear that if $\sigma \in B$, $\tau \in T$ and $\vdash \sigma \to \tau$, then $\sigma \in T$. Second, one easily verifies by induction on $n \in \omega$ that $\Box^n\Box\bot \in T$. ∎

1.3. LEMMA. *In the diagonalizable algebra of any theory* S, $\tau \in T$ *if and only if* $\tau$ *is a box sentence and there is an* $n \in \omega$ *s.t.* $\vdash \tau \to \Box^n\Box\bot$.

PROOF. The (if) direction follows from Lemma 1.2.

(only if). Following Lindström [12], we consider the sentence $\xi$ defined, with the help of self-reference, as follows:

$$\xi \equiv \exists x \left(\Box^x \Box\bot \wedge \forall\varphi \left(\Box_{\leq x}(\Box\varphi \to \xi) \to \neg\Box\varphi\right)\right).$$

We are going to show that for an arbitrary box sentence $\sigma \equiv \Box\psi$ one has $\vdash \sigma \to \xi$ iff $\vdash \sigma \to \Box^n\Box\bot$ for some $n \in \omega$.

Suppose $\vdash \Box\psi \to \Box^n\Box\bot$. Reason in S:

Assume $\Box\psi$ and $\neg\xi$. By the Small Reflection Principle we have that for all sentences $\varphi$, $\Box_{\leq n}(\Box\varphi \to \xi)$ implies $\Box\varphi \to \xi$ and hence $\neg\Box\varphi$. Thus we have $\forall\varphi\,(\Box_{\leq n}(\Box\varphi \to \xi) \to \neg\Box\varphi)$ and $\Box^n\Box\bot$ (this follows from $\Box\psi$), which, taken together, imply $\xi$.

Therefore, $\vdash \Box\psi \to \xi$ as was to be shown.

Conversely, suppose $\vdash \Box\psi \to \xi$. For some $n \in \omega$ we then have $\vdash \Box_n(\Box\psi \to \xi)$. Reason in S:

Assume $\Box\psi$ and hence $\xi$, which says that there is an $y$ s.t. $\Box^y\Box\bot$ and $\forall\varphi\,(\Box_{\leq y}(\Box\varphi \to \xi) \to \neg\Box\varphi))$. We can not have $y > n$ for then we would have $\Box_{\leq y}(\Box\psi \to \xi)$ implying $\neg\Box\psi$ contrary to the assumption. Thus we have $y \leq n$ and, in particular, $\Box^n\Box\bot$ by $\Sigma$-completeness.

We have just inferred $\vdash \sigma \to \Box^n\Box\bot$ as we said we would.

Now note that one has $\vdash \Box\bot \to \xi$ and, for any box sentence $\tau$, $\vdash \tau \to \xi$ implies $\vdash \tau \to \Box^m\Box\bot$ for an appropriate $m \in \omega$, hence $\vdash \Box\tau \to \Box^{m+1}\Box\bot$, hence $\vdash \Box\tau \to \xi$. Therefore, by the definition of $T$, $\sigma \in T$ implies $\vdash \sigma \to \xi$ which, as we have seen, entails $\vdash \sigma \to \Box^m\Box\bot$.

This completes the proof. ∎

We proceed to introduce more first order diagonalizable algebraic abbreviations.

1.4. DEFINITION.

$$S(\alpha;\mu,\tau) \equiv \tau \in B \;\&\; \vdash \mu \to \Box(\tau \to \alpha)$$
$$\&\; \forall\sigma{\in}B \vdash \mu \to (\Box(\sigma \to \alpha) \to \Box(\sigma \to \tau))$$

$$Q(\alpha,\varepsilon;\tau) \equiv \exists\mu\,(\nvdash \neg\mu \;\&\; \vdash \mu \to \varepsilon \;\&\; S(\alpha;\mu,\tau))$$

$$\nu \in N \equiv \nu \in T$$
$$\&\; \forall\alpha,\varepsilon \left(\left(Q(\alpha,\varepsilon;\Box\bot) \;\&\; \forall\tau{\in}T\,(Q(\alpha,\varepsilon;\tau) \Rightarrow Q(\alpha,\varepsilon;\Box\tau))\right) \Rightarrow Q(\alpha,\varepsilon;\nu)\right)$$

$S(\alpha;\mu,\tau)$ translates, roughly, as saying that $\mu$ is of the opinion that $\tau$ is the weakest box sentence provably implying $\alpha$, while $Q(\alpha,\varepsilon;\tau)$ says that such an opinion is consistent with $\varepsilon$. The content of $\nu \in N$ is that the top-box sentence $\nu$ is contained in any set of the form $\{\,\rho \in T \mid Q(\alpha,\varepsilon;\rho)\,\}$ once this set contains $\Box\bot$ and is closed under $\Box$. In diagonalizable algebras of $\Sigma_1$-sound theories, the formula $\nu \in N$ is intended to single out sentences of the form $\Box^n\Box\bot$ with $n \in \omega$. While one direction is rather trivial, the other will take us the rest of this Section: We shall have to show that an appropriate choice of $\alpha$ and $\varepsilon$ can prevent unwanted sentences from satisfying $Q(\alpha,\varepsilon;\cdot)$.

Here is what one should know about $S(\cdots)$:

1.5. LEMMA. *In any diagonalizable algebra,*

(a) *If* $\vdash \lambda \to \mu$ *and* $S(\alpha;\mu,\tau)$, *then* $S(\alpha;\lambda,\tau)$.

(b) *If* $\sigma \in B$, $S(\alpha;\mu,\tau)$ *and* $\vdash \mu \to \Box(\sigma \leftrightarrow \tau)$, *then* $S(\alpha;\mu,\sigma)$.

(c) *If* $S(\alpha;\mu,\tau)$ *and* $S(\alpha;\mu,\sigma)$, *then* $\vdash \mu \to \Box(\sigma \leftrightarrow \tau)$.

PROOF. Both (a) and (b) are quite obvious.

(c). $S(\alpha;\mu,\tau)$ implies $\vdash \mu \to \Box(\tau \to \alpha)$, and, since $\tau \in B$, $S(\alpha;\mu,\sigma)$ implies $\vdash \mu \to (\Box(\tau\to\alpha)\to\Box(\tau\to\sigma))$. Thus $\vdash \mu\to\Box(\tau\to\sigma)$. The converse $\vdash \mu\to\Box(\sigma\to\tau)$ is symmetric. ∎

1.6. LEMMA. *In any diagonalizable algebra, for all* $n \in \omega$, $\Box^n\Box\bot \in N$.

PROOF. Immediate from Lemma 1.2 and the definition of $N$ by induction on $n$. ∎

To prove that sentences of the form $\Box^n\Box\bot$ exhaust the set $N$ of $\mathcal{D}_{\mathrm{T}}$, we shall construct a Solovay function. Since our purpose is to take good care of certain top-box sentences' being or not being (seen by other sentences as) the weakest box sentences implying yet another given sentence (see Definition 1.4, esp. $S(\cdots)$), it is hardly surprising that the definition of our Solovay function occasionally almost quotes from Definition 5.6.3 of Dzhaparidze [7] and Definition 8.1.3 of Dzhaparidze [8] which create functions intended to gain thorough control over the behaviour w.r.t. $\Sigma_1$ sentences of the sentences arising from the function constructed. Dzhaparidze's Solovay functions are descendants of the one in Berarducci [5], who deals with relative interpretability between extensions of the ground theory, a relation reducing, in certain cases, to $\Pi_1$ conservativity.

1.7. DEFINITION. Until Proposition 1.15, fix a true $\Pi_1$ sentence $\pi$. Working within T, we define a recursive function $H_\pi$, ranging over the set $\{a_i\}_{i\in\omega} \cup \{d_i\}_{i\in\omega-\{0\}} \cup \{e_i\}_{i\in\omega} \cup \{0\}$, where the elements indicated are assumed to be pairwise distinct. We use the usual abbreviation $L_\pi{=}f$ for $\exists x\,\forall y{\geq}x\, H_\pi(y){=}f$.

We set

$$H_\pi(0) = 0.$$

The value of $H_\pi(x+1)$ is defined by Cases:

Case A. $H_\pi(x) = 0$ and $\Box_x(\pi \to L_\pi{\neq}e_i)$ for some $i \in \omega$.

$$H_\pi(x+1) = e_i.$$

Case B. $H_\pi(x) = e_i$ and for some sentence $\varphi$ and $y < x$ we have $\Box_x\varphi$, $\Box_y(\Box\varphi\to(\pi \to L_\pi{\neq}e_{i+1}))$ and $H_\pi(y) = 0$.

$$H_\pi(x+1) = e_{i+1}.$$

C a s e C. $H_\pi(x) = e_i$, Case B is not the case and $\Box_x L_\pi \neq a_j$ for some $j \in \omega$.

$$H_\pi(x+1) = a_j.$$

C a s e D. $H_\pi(x) = a_{i+1}$ and for some sentence $\varphi$ and $y < x$ we have $\Box_x \varphi$, $\Box_y(\Box\varphi \to L_\pi \neq d_{i+1})$ and $H_\pi(y) = e_i$.

$$H_\pi(x+1) = d_{i+1}.$$

C a s e E. $H_\pi(x) = a_i$ or $H_\pi(x) = d_i$, Case D is not the case and $\Box_x L_\pi \neq a_j$ with $j < i$.

$$H_\pi(x+1) = a_j.$$

C a s e F. None of the preceding Cases takes place.

$$H_\pi(x+1) = H_\pi(x).$$

This completes the definition of $H_\pi$.

Finally, we define two sentences:

$$\alpha_\pi \equiv \exists i\, L_\pi = a_i, \quad \text{and} \quad \varepsilon_\pi \equiv \pi \wedge \exists i\, L_\pi = e_i.$$

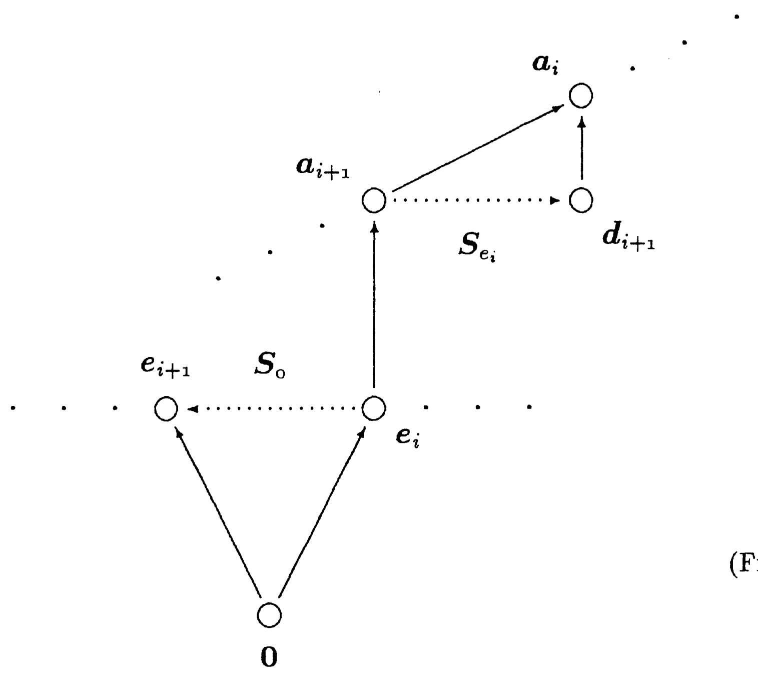
The sentence $\pi$ indexing $H_\pi$ is the parameter that we later are going to vary to demonstrate the versatility of the predicate $Q(\alpha, \varepsilon; \cdot)$ by substituting $\alpha_\pi$ and $\varepsilon_\pi$ for $\alpha$ and $\varepsilon$ respectively. A similar, although much more accurate treatment of arbitrary true $\Pi_1$ sentences within a Solovay function forerunning ours is found in Beklemishev [4, II].

Figure 1.1 is intended to help the reader visualize the behaviour of $H_\pi$. Note that this picture is in fact that of an infinite Kripke frame with the straight arrows playing the role of the familiar accessibility relation $r$ associated with the $\Box$ operator. Transitions of the function $H_\pi$ along these arrows correspond to Cases A, C and E of Definition 1.7, while those along the dotted arrows are due to Cases B and D.

The relations $S_0$ and $S_{e_i}$ come from interpretability logic (see De Jongh & Veltman [11]). The way $H_\pi$ goes about the relations $S_0$ and $S_{e_i}$ (in Cases B and D resp.) is reminiscent of the Solovay functions appearing in Dzhaparidze [7] and [8] and especially in Zambella [28, Section 5]. Our $H_\pi$ pursues much the same ends as in Zambella [28], but, since our situation is simpler in that we deal with an individual, albeit infinite, frame of a not too intricate structure, we are able to leave out some of the complications in Zambella [28]. In particular, the convergence of $H_\pi$ is, in our case, due to rather trivial reasons.

One superficially confusing distinction between our and many preceding similar constructions is that we do not describe the behaviour of $H_\pi$ in the general terms of relations $R$ and $S_x$, advocating rather an individual approach to particular nodes of the frame. (Note, for example, that $S_0$ and $S_{e_i}$ are not treated by $H_\pi$ in exactly the same way.) This is because we only handle a single frame rather than a class of those and because we have got an extra complication with the sentence $\pi$.

Next we proceed along the well-throdden path of demonstrating the adequacy of a Solovay function for our purposes. These are, however, somewhat different from many earlier cases (i.e. proving completeness theorems for various logics). Thus, we only need




(Figure 1.1)

'commutation lemmas' in a very modest degree of generality. In particular, we do not, as many authors do, take care of nested occurrences of modal operators other than $\Box$.

1.8. LEMMA. (a) $\vdash \exists f \in \{a_i\}_{i\in\omega} \cup \{d_i\}_{i\in\omega-\{0\}} \cup \{e_i\}_{i\in\omega} \cup \{0\}\, L_\pi = f$.

(b) *For all* $k \in \omega$, $H_\pi(k) = 0$.

(c) $\nvdash \pi \rightarrow L_\pi = e_0$.

PROOF. (a). We prove that $H_\pi$ reaches a limit value. If it ever leaves 0 for $\{e_i\}_{i\in\omega}$ then there is an $x$ s.t. $H_\pi(x) \neq 0$ and, while $H_\pi$ remains among the $e$'s, it can change its value at most $x$ times, for, by inspection of Case B of Definition 1.7, each change of value requires a different proof $<x$.

Leaving the $e$'s, $H_\pi$ finds itself among $a$'s and $d$'s. Each two moves of the function diminish the subscript of its value by at least 1 (see Cases D and E). Clearly, this can not go on forever.

(b). Suppose $H_\pi$ leaves 0. Then, by (a), it arrives at its limit value $L_\pi = f \neq 0$, for which there has to be a proof of $L_\pi \neq f$ or of $\pi \rightarrow L_\pi \neq f$, possibly from a true sentence of the form $\Box\varphi$ as a hypothesis. By $\Sigma$-completeness, we then have $\vdash \pi \rightarrow L_\pi \neq f$ in either case. Now, since our theory T has been assumed to be $\Sigma_1$-( and hence $\Pi_2$-)sound, $\pi \rightarrow L_\pi \neq f$ is true. We have chosen $\pi$ true, so $L_\pi \neq f$, which is a contradiction that leaves $H_\pi$ safely at 0.

(c). Immediate from (b) on inspection of Case A. ∎

In Lemmas 1.9–10 we show that $L_\pi$ respects the $R$-depth of the nodes $\{a_i\}_{i\in\omega}$ and $\{d_i\}_{i\in\omega-\{0\}}$.

1.9. LEMMA. (a) $\vdash \forall i\,(L_\pi{=}a_i \rightarrow \Box\exists j{<}i\, L_\pi{=}a_j)$.

(b) $\vdash \forall i\,(L_\pi{=}a_i \vee L_\pi{=}d_i \rightarrow. \forall j{<}i\, \Diamond L_\pi{=}a_j)$.

(c) $\vdash \forall i,j\,(L_\pi{=}e_i \rightarrow \Diamond L_\pi{=}a_j)$.

(d) $\vdash \forall j\,(L_\pi{=}0 \rightarrow \Diamond L_\pi{=}a_j)$.

PROOF. (a). We reason in T. Assume $L_\pi = a_i$ and fix an $x$ s.t. $H_\pi(x) = a_i$, so that one also has $\Box H_\pi(x){=}a_i$, and hence $\Box\exists j{\le}i\,(L_\pi{=}a_j \vee L_\pi{=}d_j)$ for $a_j$'s and $d_j$'s with $j \le i$ are the only places that $H_\pi$ can go from $a_i$. On the other hand, we must have $\Box L_\pi{\ne}a_i$ for $H_\pi$ to get to $a_i$ in the first place. Suppose $0 < j \le i$ and reason inside $\Box$:

> If $L_\pi = d_j$ then $\Box_y(\Box\varphi \rightarrow L_\pi{\ne}d_j)$ holds for some provable $\varphi$ and $y$ satisfying $H_\pi(y) = e_{j-1}$. Clearly, $y < x$ so that one has $\Box_{<x}(\Box\varphi \rightarrow L_\pi{\ne}d_j)$ implying $\Box\varphi{\rightarrow}L_\pi{\ne}d_j$ by Small Reflection. Since $\varphi$ is actually provable, we have $L_\pi \ne d_j$, a contradiction.

Thus $\Box L_\pi{\ne}d_j$ whenever $0 < j \le i$. We are left with $\Box\exists j{<}i\, L_\pi{=}a_j$ as required.

(b) is immediate on inspection of Case E of Definition 1.7.

(c). See Case C.

(d). Assuming, in T, $L_\pi = 0$, one has $\Diamond L_\pi{=}e_0$ by inspection of Case A, whence $\Diamond\Diamond L_\pi{=}a_j$ follows by (c) for any $j \in \omega$. $\Diamond L_\pi{=}a_j$ follows then by $\Sigma$-completeness. ∎

1.10. LEMMA. (a) $\vdash \forall i\,(\exists j{\le}i\, L_\pi{=}a_j \rightarrow \Box^i\Box\bot)$.

(b) $\vdash \forall i\,(\Box^i\Box\bot \rightarrow \exists j{\le}i\,(L_\pi{=}a_j \vee L_\pi{=}d_j))$.

(c) $\vdash \varepsilon_\pi \rightarrow \forall i\, \Diamond^i\Diamond\top$.

PROOF. (a) is proved by formal induction on $i$. For $i = 0$ the claim is immediate by Lemma 1.9(a). Assume it to hold for $i$ and consider $i + 1$:

$$\begin{aligned}
\vdash \exists j{\le}i{+}1\, L_\pi = a_j &\rightarrow \exists j{\le}i{+}1\, \Box\exists k{<}j\, L_\pi = a_k && \text{(by Lemma 1.9(a))}\\
&\rightarrow \Box\exists k{\le}i\, L_\pi = a_k\\
&\rightarrow \Box\Box^i\Box\bot && \text{(by IH)}\\
&\rightarrow \Box^{i+1}\Box\bot && \text{q.e.d.}
\end{aligned}$$

(b). Again, induction on $i$. For $i = 0$ one uses Lemma 1.9(b)–(d) to infer $\Box\bot \rightarrow L_\pi{=}a_0$. Here is the induction step:

$$\begin{aligned}
\vdash \Box^{i+1}\Box\bot &\rightarrow \Box\exists j{\le}i\,(L_\pi = a_j \vee L_\pi = d_j) && \text{(by IH)}\\
&\rightarrow \Box L_\pi \ne a_{i+1}\\
&\rightarrow \exists j{\le}i{+}1\,(L_\pi = a_j \vee L_\pi = d_j)\\
& && \text{(by Lemmas 1.8(a) and 1.9(b)–(d)) q.e.d.}
\end{aligned}$$

(c). Observe:

$$\begin{aligned}
\vdash \varepsilon_\pi &\to \exists i\, L_\pi = e_i && \text{(by the definition of } \varepsilon_\pi) \\
&\to \forall j\, \Diamond L_\pi = a_j && \text{(by Lemma 1.9(c))} \\
&\to \forall j{>}0\, \Diamond\Diamond^{j-1}\Diamond\top && \text{(by (b))} \\
&\to \forall j\, \Diamond^j \Diamond\top && \text{q.e.d.} \ \blacksquare
\end{aligned}$$

Lemmas 1.11–14 establish that the set $\{\, \rho \in T \mid Q(\alpha_\pi, \varepsilon_\pi; \rho)\,\}$ contains $\Box\bot$ and is closed under $\Box$.

1.11. Lemma. *If a box sentence $\rho$ is consistent with $L_\pi{=}e_i$ over* T *for some $i \in \omega$, then $\rho$ is consistent over* T *with $\pi \wedge L_\pi{=}e_{i+1}$.*

Proof. Suppose $\rho \equiv \Box\varphi$ is <u>not</u> consistent with $\pi \wedge L_\pi{=}e_{i+1}$: $\vdash \Box\varphi \to (\pi \to L_\pi{\neq}e_{i+1})$. It follows that $\vdash \Box_k(\Box\varphi{\to}(\pi{\to}L_\pi{\neq}e_{i+1}))$ for some $k \in \omega$. Note that, by Lemma 1.8(b), we have $H_\pi(k) = 0$. Reason in T:

Assume $\Box\varphi$. Suppose that $L_\pi = e_i$. Pick an $x$ s.t. $\Box_x\varphi$ and $H_\pi(x) = e_i$. By instructions of Case B of Definition 1.7, we get $H_\pi(x{+}1) = e_{i+1}$. Hence $L_\pi \neq e_i$.

So we have established $\vdash \Box\varphi \to L_\pi{\neq}e_i$ which contradicts our assumption. Therefore, $\Box\varphi$ must be consistent with $\pi \wedge L_\pi{=}e_{i+1}$. ∎

1.12. Lemma. (a) $\vdash \forall i\, (\exists j{\geq}i\, L_\pi{=}e_j \to \Box(\Box^i\Box\bot \to \alpha_\pi))$.

(b) *$S(\alpha_\pi; L_\pi{=}e_i, \Box^i\Box\bot)$ holds for all $i \in \omega$.*

Proof. (a). Our argument takes place in T. Fix an $i \in \omega$ and let $L_\pi = e_j$ for some $j \geq i$. Let $x$ be s.t. $H_\pi(x) = e_j$. Reason inside $\Box$:

Assume $L_\pi = d_k$ with $0 < k \leq i$. This can only happen if $\Box_y(\Box\varphi{\to}L_\pi{\neq}d_k)$ with $H_\pi(y) = e_{k-1}$ and $\varphi$ provable. Clearly, we have $y < x$ so that $\Box_{<x}(\Box\varphi{\to}L_\pi{\neq}d_k)$, whence, by Small Reflection, $\Box\varphi \to L_\pi{\neq}d_k$ and, therefore, $L_\pi \neq d_k$ since $\Box\varphi$ is true.

Thus $\Box\forall k\, (L_\pi{=}d_k \to k > i)$. By Lemma 1.10(b) we get $\Box(\Box^i\Box\bot \to \exists k{\leq}i\, (L_\pi{=}a_k \vee L_\pi{=}d_k))$, so by the above it follows that $\Box(\Box^i\Box\bot{\to}\exists k{\leq}i\, L_\pi{=}a_k)$. Ergo $\Box(\Box^i\Box\bot{\to}\alpha_\pi)$ q.e.d.

(b). We fix an arbitrary sentence $\varphi$ and argue in T:

Suppose $L_\pi = e_i$ and $\Box(\Box\varphi \to \alpha_\pi)$ so that $\Box_y(\Box\varphi \to L_\pi{\neq}d_{i+1})$ for some $y$. We may assume $H_\pi(y) = e_i$. Reason inside $\Box$:

Assume $\Box\varphi$ and suppose $L_\pi = a_{i+1}$. Since $\varphi$ must be provable by arbitrarily large proofs, there is an $x$ s.t. $\Box_x\varphi$ and $H_\pi(x) = a_{i+1}$. Obviously, $x > y$. But then, since $\Box_y(\Box\varphi \to L_\pi{\neq}d_{i+1})$ and $H_\pi(y)$ satisfies the conditions of Case D of Definition 1.7 for $x$, we would have $H_\pi(x{+}1) = d_{i+1}$, contradicting $L_\pi = a_{i+1}$. Thus $L_\pi \neq a_{i+1}$.

The above argument formalizes in T to the effect that $\Box(\Box\varphi \to L_\pi{\neq}a_{i+1})$. Since, having assumed $\Box\varphi$, we have $\Box\Box\varphi$ by $\Sigma$-completeness, this implies $\Box L_\pi{\neq}a_{i+1}$, which by Lemma 1.9(b) entails $L_\pi \neq a_k$ for all $k > i+1$ and, since we have established $\Box\varphi \to L_\pi{\neq}a_{i+1}$, also for all $k > i$.

Since $\Box\varphi$ implies $\alpha_\pi$, we have that $\exists k\, L_\pi{=}a_k$ and hence, by the preceding argument, $\exists k \leq i\, L_\pi = a_k$. By Lemma 1.10(a), this implies $\Box^i \Box\bot$.

Thus $\Box(\Box\varphi \to \Box^i\Box\bot)$.

Thus for all $\varphi$ there holds $\vdash L_\pi = e_i \to (\Box(\Box\varphi \to \alpha_\pi) \to (\Box(\Box\varphi \to \Box^i\Box\bot))$. To get the required $S(\alpha_\pi; L_\pi = e_i, \Box^i\Box\bot)$, we put this together with the obvious $\Box^i\Box\bot \in B$ and $\vdash L_\pi = e_i \to \Box(\Box^i\Box\bot \to \alpha_\pi)$, the latter fact being implied by clause (a). ■

1.13. Lemma. *If $\tau \in T$ and $Q(\alpha_\pi, \varepsilon_\pi; \tau)$, then there exists an $i \in \omega$ s.t. $\Box(\tau \leftrightarrow \Box^i\Box\bot)$ is consistent over* T *with $\pi \wedge L_\pi = e_i$.*

Proof. Since $\tau \in T$, we have by Lemma 1.3 that $\vdash \tau \to \Box^n\Box\bot$ for some $n \in \omega$. $Q(\alpha_\pi, \varepsilon_\pi; \tau)$ means that there is an irrefutable $\mu$ formally implying $\varepsilon_\pi$ and s.t. $S(\alpha_\pi; \mu, \tau)$ holds. One therefore has $\vdash \mu \to (\Box(\Box^{n+1}\Box\bot \to \alpha_\pi) \to \Box(\Box^{n+1}\Box\bot \to \tau))$ by the definition of $S(\cdots)$. Reason in T:

Assume $\mu$ and $\exists i > n\, L_\pi = e_i$. From Lemma 1.12(a) we have $\Box(\Box^{n+1}\Box\bot \to \alpha_\pi)$. Now, $\mu$ implies $\Box(\Box^{n+1}\Box\bot \to \alpha_\pi) \to \Box(\Box^{n+1}\Box\bot \to \tau)$, so that $\Box(\Box^{n+1}\Box\bot \to \tau)$, whence, due to the way we have chosen $n$ to be related to $\tau$, one has $\Box(\Box^{n+1}\Box\bot \to \Box^n\Box\bot)$. By Löb's Theorem it follows that $\Box^{n+1}\Box\bot$ contradicting, by Lemma 1.10(c), $\varepsilon_\pi$ and hence also $\mu$.

Thus $\vdash \mu \to \forall i > n\, L_\pi \neq e_i$. Recall that $\mu$, through $\varepsilon_\pi$, provably implies $\exists i\, L_\pi = e_i$, so that $\vdash \mu \to \exists i \leq n\, L_\pi = e_i$. Since $\mu$ is irrefutable, we can fix an $i \leq n$ s.t. $\mu$ is consistent with $L_\pi = e_i$. Let $\lambda \equiv \mu \wedge L_\pi = e_i$. Applying Lemma 1.5(a), we get $S(\alpha_\pi; \lambda, \tau)$. On the strength of Lemma 1.12(b), there holds $S(\alpha_\pi; L_\pi = e_i, \Box^i\Box\bot)$. By Lemma 1.5(a), this entails $S(\alpha_\pi; \lambda, \Box^i\Box\bot)$, whence, by Lemma 1.5(c), one has $\vdash \lambda \to \Box(\tau \leftrightarrow \Box^i\Box\bot)$. Since $\lambda$ is irrefutable and, clearly, $\vdash \lambda \to \varepsilon_\pi$, we are done. ■

1.14. Lemma. (a) $Q(\alpha_\pi, \varepsilon_\pi; \Box\bot)$.

(b) *For all $\tau \in T$, $Q(\alpha_\pi, \varepsilon_\pi; \tau)$ implies $Q(\alpha_\pi, \varepsilon_\pi; \Box\tau)$.*

Proof. (a). Consider $\mu \equiv \pi \wedge L_\pi = e_0$. We have $\vdash \mu \to \varepsilon_\pi$ and $\nvdash \neg\mu$ by Lemma 1.8(c). By Lemmas 1.12(b) and 1.5(a), one has $S(\alpha_\pi; \pi \wedge L_\pi = e_0, \Box\bot)$. Thus $Q(\alpha_\pi, \varepsilon_\pi; \Box\bot)$ is established.

(b). Take $\tau \in T$ and suppose $Q(\alpha_\pi, \varepsilon_\pi; \tau)$. We have by Lemmas 1.13 and 1.11 that for some $i \in \omega$ the sentence $\Box(\tau \leftrightarrow \Box^i\Box\bot)$ is consistent over T with $\pi \wedge L_\pi = e_{i+1}$. This means that there is a sentence $\mu$ irrefutable in T s.t. $\vdash \mu \to . \pi \wedge L_\pi = e_{i+1}$ and $\vdash \mu \to \Box(\tau \leftrightarrow \Box^i\Box\bot)$, which implies $\vdash \mu \to \Box(\Box\tau \leftrightarrow \Box^{i+1}\Box\bot)$ by $\Sigma$-completeness. Now, since $S(\alpha_\pi; L_\pi = e_{i+1}, \Box^{i+1}\Box\bot)$ holds by Lemma 1.12(b), we also have $S(\alpha_\pi; \mu, \Box^{i+1}\Box\bot)$ by Lemma 1.5(a). Therefore, by virtue of Lemma 1.5(b), one gets $S(\alpha_\pi; \mu, \Box\tau)$. Since we clearly have $\vdash \mu \to \varepsilon_\pi$, clause (b) is through. ■

We are now in a position to show that the formula $\nu \in N$ is not satisfied in $\mathcal{D}_{\mathrm{T}}$ by sentences $\nu$ not of the form $\Box^i\Box\bot$.

1.15. Proposition. *For all sentences $\nu$, we have $\nu \in N$ iff $\vdash \nu \leftrightarrow \Box^i\Box\bot$ for some $i \in \omega$.*

Proof. (if) was established in Lemma 1.6. We concentrate on (only if).

Suppose, for a contradiction, that $\vdash \nu \leftrightarrow \Box^i\Box\bot$ is not the case for any $i \in \omega$, and yet $\nu \in N$, that is $\nu \in T$ and

$$\forall\alpha,\varepsilon\left(\left(Q(\alpha,\varepsilon;\Box\bot)\ \&\ \forall\tau{\in}T\left(Q(\alpha,\varepsilon;\tau) \Rightarrow Q(\alpha,\varepsilon;\Box\tau)\right)\right) \Rightarrow Q(\alpha,\varepsilon;\nu)\right).$$

By our assumptions on $\nu$, $\pi \equiv \forall j\,\neg\Box(\nu \leftrightarrow \Box^j\Box\bot)$ is a true $\Pi_1$ sentence. Consider the sentences $\alpha_\pi$ and $\varepsilon_\pi$ corresponding to $\pi$ by Definition 1.7. By Lemma 1.14 one has

$$Q(\alpha_\pi,\varepsilon_\pi;\Box\bot)\ \&\ \forall\tau{\in}T\left(Q(\alpha_\pi,\varepsilon_\pi;\tau) \Rightarrow Q(\alpha_\pi,\varepsilon_\pi;\Box\tau)\right)$$

and, therefore, $Q(\alpha_\pi,\varepsilon_\pi;\nu)$ should hold. By Lemma 1.13, this implies that $\pi$ is consistent with $\Box(\nu \leftrightarrow \Box^i\Box\bot)$ for some $i \in \omega$ which is clearly absurd.

The contradiction settles our Proposition. ∎

Our argument can be visualized in the dual space $\mathcal{D}^*_{\mathrm{T}}$ as follows:

The clopen $\varepsilon_\pi{}^*$ lies well below the finite $R$-depth part of $\mathcal{D}^*_{\mathrm{T}}$. An observer inside $\varepsilon_\pi{}^*$, looking $R$-upwards, will see the clopen $\alpha_\pi{}^*$ as a vertical slate almost completely lying within the finite depth part of $\mathcal{D}^*_{\mathrm{T}}$. On this slate notches can be discerned in the following way: Suppose a nonempty clopen $\mu^* \subseteq \varepsilon_\pi{}^*$, from its $R$-upwards viewpoint, sees $\tau^*$ as the largest box clopen contained in $\alpha_\pi{}^*$ (that is, $S(\alpha_\pi;\mu,\tau)$ holds). Suppose further that $\tau^*$ is a top-box clopen, i.e. it is contained in the finite depth part of $\mathcal{D}^*_{\mathrm{T}}$. Let us then say that $\tau^*$ is, from the viewpoint of $\varepsilon_\pi{}^*$, a notch on $\alpha_\pi{}^*$ (this is equivalent to $Q(\alpha_\pi,\varepsilon_\pi;\tau)$).

Our construction provides for $(\Box\bot)^*$'s being a notch, and for the closure of the collection of top-box notches under $\Box^*$ (Lemma 1.14).

Note that the sentences $\Box^n\Box\bot$ correspond to clopens $\{\, x \in \mathcal{D}^*_{\mathrm{T}} \mid d(x) \leq n \,\}$. These are stripes at the very $R$-top of $\mathcal{D}^*_{\mathrm{T}}$ that are $n{+}1$ elements $R$-thick. (One should not take this too literally: There are maximal $R$-chains within $(\Box^n\Box\bot)^*$ containing less than $n{+}1$ elements.) $\varepsilon_\pi{}^*$ knows that each of them notches $\alpha_\pi{}^*$. Moreover, if any nonempty subclopen $\mu^*$ of $\varepsilon_\pi{}^*$ observes a top-box notch, then it is guaranteed that a nonempty subclopen of $\mu^*$ does not see any difference between this notch and one of $(\Box^n\Box\bot)^*$'s (this is the content of Lemma 1.13).

If a top-box sentence $\nu$ fails to equate to any of the $\Box^n\Box\bot$'s, then the clopen $\nu^*$ does not match any of these stripes. Intersecting $(\exists i\, L_\pi{=}e_i)^*$ with $\pi^* = (\forall j\,\neg\Box(\nu{\leftrightarrow}\Box^j\Box\bot))^*$ is a way to focus the resulting clopen $\varepsilon_\pi{}^*$'s attention on the part of $\nu^*$ that does not level up to any of the $(\Box^n\Box\bot)^*$'s. No nonempty subclopen of $\varepsilon_\pi{}^*$ will then think that $\nu^*$ is the same as any of these. In this way $\nu$ fails to satisfy $Q(\alpha_\pi,\varepsilon_\pi;\cdot)$ and hence finds itself outside $N$.

## 2. Representing arithmetical operations

Having established in Proposition 1.15 that the set $\{\Box^n\Box\bot\}_{n\in\omega}$ is elementarily definable in diagonalizable algebras of $\Sigma_1$-sound theories, there are several ways we can use this circumstance to show that the first order theories of these diagonalizable algebras are undecidable. We are going to indicate two approaches that have been known to specialists and elaborate on a third one of our own that affords a proof of the nonarithmeticity of the theories in question.

First we introduce some notation and state a lemma that will be needed here as well as in Section 3.

2.1. DEFINITION. For each $n \leq 0$ we define a term $\Downarrow^n$ in the language of diagonalizable algebra theory:

$$\Downarrow^0\varphi = \varphi \quad \text{and} \quad \Downarrow^{n+1}\varphi = \neg\Box^n\varphi \wedge \Box^{n+1}\varphi.$$

In the dual space of a diagonalizable algebra, $\Downarrow^{n+1^*}X$ is an $R$-antichain which, in case of an $R$-upwards closed set (in particular, a box clopen) $X$, finds itself $n$ $R$-steps below $X$.

2.2. LEMMA. *In any theory one has that for all sentences $\varphi$,*

(a) $\vdash \forall x, y > 0\,(\Downarrow^x\varphi \wedge \Downarrow^y\varphi \rightarrow .\ x = y)$.

(b) $\vdash \forall x > 0\,(\Box\neg\Downarrow^x\varphi \rightarrow \Box^x\varphi)$.

(c) $\vdash \forall x > 0\,\forall y\,(\Downarrow^x\Box^y\varphi \leftrightarrow \Downarrow^{x+y}\varphi)$.

PROOF. (a). Working within a theory S, observe that, if both $\Downarrow^x\varphi$ and $\Downarrow^y\varphi$ hold, then $x$ can not be smaller than $y$, for $\Box^x\varphi$ and $\neg\Box^{y-1}\varphi$ are otherwise incompatible by $\Sigma$-completeness.

(b). This is a disguised instance of Löb's Theorem:

$$\begin{aligned}\vdash \Box\neg\Downarrow^x\varphi &\leftrightarrow \Box(\Box^x\varphi \rightarrow \Box^{x-1}\varphi)\\ &\rightarrow \Box^x\varphi.\end{aligned}$$

(c). One only has to carefully count the $\Box$'s. ∎

The elementary definability of $\{\Box^n\Box\bot\}_{n\in\omega}$ in the diagonalizable algebra of a $\Sigma_1$-sound theory affords a first order definition of the domain of its $\bot$-generated subalgebra, which answers a question in Artemov & Beklemishev [3].

2.3. COROLLARY. *For T a $\Sigma_1$-sound theory, the (domain of the) $\bot$-generated subalgebra of $\mathcal{D}_{\mathrm{T}}$ is first order definable in the language of diagonalizable algebra theory.*

COMMENT. This subalgebra is defined by, e.g., the following formula:

$$\begin{aligned}\xi \in C \equiv \exists\nu{\in}N\,(\vdash\neg\nu\rightarrow\xi \text{ or } \vdash\neg\nu\rightarrow\neg\xi)\ \&\ (\vdash\Box\bot\rightarrow\xi \text{ or } \vdash\Box\bot\rightarrow\neg\xi)\\ \&\ \forall\nu{\in}N\,(\vdash\Downarrow\nu\rightarrow\xi \text{ or } \vdash\Downarrow\nu\rightarrow\neg\xi),\end{aligned}$$

as can be inferred from Proposition 1.15 and the fact that this subalgebra, considered as a Boolean algebra, is atomic and that its atoms are precisely sentences of the form $\Downarrow^n\Box\bot$, $n \in \omega$, which is the content of Corollary 2.5 in Artemov & Beklemishev [3]. In fact, this subalgebra is (isomorphic to) the free diagonalizable algebra on no generators. ∎

In anticipation of Corollary 2.3, Artemov & Beklemishev [3] enrich the language of diagonalizable algebra theory with a unary predicate $\xi \in C$, which they interpret as distinguishing the $\bot$-generated subalgebra, and prove the hereditary undecidability (i.e. the undecidability of the theory as well as that of any of its subtheories) of $\mathrm{Th}(\mathcal{D}_\mathrm{T}, C)$ for an arbitrary $\Sigma_1$-sound theory T (Theorem 3 of that paper). This is done by elementarily defining, relative to a parameter, each finite partial order in $(\mathcal{D}_\mathrm{T}, C)$. (In fact, their proof only appeals to the definabity from $C$ of the set $A = \{\Downarrow^n \Box\bot\}_{n\in\omega}$.) By our Corollary 2.3 the undecidability of $\mathrm{Th}\,\mathcal{D}_\mathrm{T}$ follows.

In dual spaces, elements of $A$ correspond to maximal $R$-antichains of constant finite depth.

Another way to exploit the set $A$ for undecidability results has been suggested by Domenico Zambella: Define the following two binary relations in first order diagonalizable algebraian:

$$\zeta \preceq \xi \;\equiv\; \forall\alpha{\in}A\,(\vdash \alpha \to \zeta \;\Rightarrow\; \vdash \alpha \to \xi)$$

$$\zeta \simeq \xi \;\equiv\; \zeta \preceq \xi \;\&\; \xi \preceq \zeta.$$

Clearly, $\preceq$ is a preorder and $\simeq$ is the corresponding equivalence relation. One easily establishes that, in theories of infinite credibility extent, for each r.e. set $V$ there exists an arithmetical sentence $\upsilon$ s.t. $\vdash \Downarrow^n \Box\bot \to \upsilon$ iff $n \in V$, all $n \in \omega$ (see e.g. Shavrukov [20, Lemma 11.7(a)]). Moreover, the set of $n$ s.t. $\vdash \Downarrow^n \Box\bot \to \upsilon$ is r.e. for any sentence $\upsilon$, for so are the theories we are involved with. Therefore, the structure induced by $\preceq$ on the $\simeq$-equivalence classes is isomorphic to the lattice $\mathcal{E}$ of recursively enumerable sets under inclusion. Soare [25, Theorem XVI.2.2] claims that Herrmann and Harrington have established the undecidability of the first order theory of $\mathcal{E}$. By the above isomorphism, this undecidability also hits the diagonalizable algebras of $\Sigma_1$-sound theories.

Here is the promised third approach. The idea is to treat the set $N = \{\Box^n \Box\bot\}_{n\in\omega}$ as the domain of the standard model of arithmetic with $\boldsymbol{i} : n \mapsto \Box^n \Box\bot$ the intended isomorphism. The missing bit of work we still have to do is to represent arithmetical predicates and operations. For equality, zero and successor this is trivial:

$$n = m \quad \text{iff} \quad \vdash \boldsymbol{i}(n) \leftrightarrow \boldsymbol{i}(m),$$
$$\boldsymbol{i}(0) = \Box\bot, \quad \text{and}$$
$$\boldsymbol{i}(Sn) = \Box\boldsymbol{i}(n).$$

Representing $+$ and $\times$ will require the use of parameters. Namely, we shall construct a tuple of arithmetical sentences that in a certain sense codes the $\boldsymbol{i}$-images of $+$ and $\times$.

2.4. Definition. Let us define a particular diagonalizable algebraic polynomial:

$$\mathrm{C}(r_1, r_2, r_3; q_1, q_2, q_3) \;\equiv\; \Box(r_1 \leftrightarrow q_1) \wedge \Box(r_2 \leftrightarrow q_2) \to\!.\; \Box(r_3 \leftrightarrow q_3).$$

2.5. Proposition. *There exist sentences $\varsigma_1, \varsigma_2, \varsigma_3$ and $\varpi_1, \varpi_2, \varpi_3$ s.t. for all $n, m, k \in \omega$*

(a) $\vdash \mathrm{C}(\vec{\varsigma}; \Box^n\Box\bot, \Box^m\Box\bot, \Box^k\Box\bot)$ *iff* $n + m = k$, *and*

(b) $\vdash \mathrm{C}(\vec{\varpi}; \Box^n\Box\bot, \Box^m\Box\bot, \Box^k\Box\bot)$ *iff* $n \cdot m = k$.

The proof of this Proposition will rely heavily on certain results from Shavrukov [20] and Zambella [29] characterizing diagonalizable algebras embeddable into those of $\Sigma_1$-sound theories. We shall now quote the relevant definitions and facts.

Modal formulas are built up from propositional letters using Boolean connectives and the unary operator $\Box$. They are, essentially, diagonalizable algebraic terms with propositional letters for variables.

Take a finite tuple $\vec{p}$ of propositional letters. A *(Kripke $\vec{p}$-) model* $\boldsymbol{K}$ is a tuple $(K, r, a, V)$ with $r$ a treelike irreflexive partial order on $K$ and $a$ the root of $K$ w.r.t. this order. $V$ is a subset of $K \times \vec{p}$ which gives rise to a *forcing relation* between elements of $K$ (= *nodes*) and modal formulas in $\vec{p}$:

$k$ forces $p$ iff $(k,p) \in V$,

$k$ forces $\neg$D iff $k$ does not force D,

$k$ forces D $\wedge$ E iff $k$ forces both D and E,

$k$ forces $\Box$D iff $h$ forces D for any $h \in K$ with $k \; r \; h$.

D *holds* at $k$ is another way to say $k$ forces D. If D holds at each node of a model $\boldsymbol{K}$ then $\boldsymbol{K}$ is said to be a *model* of or just to *model* D.

We shall be drawing pictures of models with arrows between nodes standing for the 'immediate predecessor' relation corresponding to $r$. The presence of a propositional letter near a node in the picture will indicate that this letter is forced at that node.

Recall the modal logic **L** mentioned in the Introduction.

2.6. Fact (Completeness Theorem for **L**; Segerberg, cf. Smoryński [24, Theorem 2.2.6]). *A modal formula* $\mathrm{A}(\vec{p})$ *in propositional letters* $\vec{p}$ *is derivable in* **L** ($\mathbf{L} \vdash \mathrm{A}(\vec{p})$) *iff* $\mathrm{A}(\vec{p})$ *is modelled in every* $\vec{p}$*-model iff* $\mathrm{A}(\vec{p})$ *holds at the root of every* $\vec{p}$*-model.* ∎

A model $(K, r, a, V)$ is a *proper cone* of the model $(H, s, b, W)$ if $a \in H$, $K = \{a\} \cup \{\, c \in H \mid a \; s \; c \,\}$, $r = s\lceil K$ and $V = W\lceil K$.

2.7. Fact (Shavrukov [20, Theorem 7.1, Lemmas 5.13 and 5.15]; Zambella [29, (proof of) Theorem 1]). *Let* T *be a* $\Sigma_1$*-sound theory. Suppose* $\mathcal{A}(\vec{r})$ *is an r.e. collection of modal formulas in finitely many variables* $\vec{r}$ *s.t.*

(i) *the conjunction of any finite subset of* $\mathcal{A}(\vec{r})$ *is irrefutable in* **L**, *and*

(ii) *for any two* $\vec{r}$*-models* $\boldsymbol{K}_1$ *and* $\boldsymbol{K}_2$ *there is an* $\vec{r}$*-model* $\boldsymbol{H}$ *s.t.* $\boldsymbol{K}_1$ *and* $\boldsymbol{K}_2$ *are isomorphic to proper cones of* $\boldsymbol{H}$ *and if both* $\boldsymbol{K}_1$ *and* $\boldsymbol{K}_2$ *model some element of* $\mathcal{A}(\vec{r})$ *then so does* $\boldsymbol{H}$. *(The readers of Shavrukov [20] should recognize this condition as saying (something stronger than) that the conjunction of any finite subcollection of* $\mathcal{A}(\vec{r})$ *is steady.)*

*Then there is a tuple* $\vec{\varrho}$ *of arithmetical sentences s.t. in (the diagonalizable algebra of)* T *we have for any modal formula* $\mathrm{B}(\vec{r})$ *that* $\vdash \mathrm{B}(\vec{\varrho})$ *if and only if* $\mathbf{L} \vdash \Box^+ \bigwedge\!\!\!\bigwedge \mathcal{G}(\vec{r}) \rightarrow \Box^+ \mathrm{B}(\vec{r})$ *for some finite subset* $\mathcal{G}(\vec{r})$ *of* $\mathcal{A}(\vec{r})$ *(*$\Box^+\varphi$ *is short for* $\varphi \wedge \Box\varphi$*). (This translates into the language of [20] as embeddability into* $\mathcal{D}_{\mathrm{T}}$ *of the factor of the free diagonalizable* *algebra on the generators $\vec{r}$ modulo the $\tau$-filter corresponding to $\mathcal{A}(\vec{r})$.)* ∎

2.8. PROOF of Proposition 2.5. We only handle clause (a), for (b) can be verified in a very similar manner.

Let the recursive set $\mathcal{A}(\vec{s})$ consist of the following modal formulas:

$$\mathrm{A}_{n,m}(\vec{s}) \;\equiv\; \mathrm{C}(\vec{s}; \Box^n\Box\bot, \Box^m\Box\bot, \Box^{n+m}\Box\bot)$$

with $n, m$ ranging over $\omega$. In order to obtain the required tuple $\vec{\varsigma}$ we would like to apply Fact 2.7 to which end we check that $\mathcal{A}(\vec{s})$ meets conditions (i) and (ii) of that Fact.

For (i) just note that the $\vec{s}$-model in Figure 2.1 (all variables from $\vec{s}$ are taken to be false at both nodes) models $\mathrm{A}_{n,m}(\vec{s})$ for any $n, m \in \omega$.

(Figure 2.1)

Turning to (ii), imagine any two $\vec{s}$-models $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ grafted just above the lower node of the $\vec{s}$-model in Figure 2.1 to form a new model $\boldsymbol{H}$. We claim that if both grafts happen to model $\mathrm{A}_{n,m}(\vec{s})$ for some $n, m \in \omega$ then the same is true of $\boldsymbol{H}$. Indeed, $\mathrm{A}_{n,m}(\vec{s})$ holds at the new nodes of the model by assumption. That same formula holds at the higher old node because the latter forces $\Box(\varsigma_3 \leftrightarrow \Box^{n+m}\Box\bot)$, the succedent of $\mathrm{A}_{n,m}(\vec{s})$. Finally, the bottom node forces $\mathrm{A}_{n,m}(\vec{s})$ since it does not force $\Box(\varsigma_1 \leftrightarrow \Box^n\Box\bot)$ (nor $\Box(\varsigma_2 \leftrightarrow \Box^m\Box\bot)$ for that matter).

By Fact 2.7 this shows that there are arithmetical sentences $\varsigma_1, \varsigma_2, \varsigma_3$ s.t.

$$\vdash \mathrm{C}(\vec{\varsigma}; \Box^n\Box\bot, \Box^m\Box\bot, \Box^k\Box\bot) \qquad \text{iff}$$

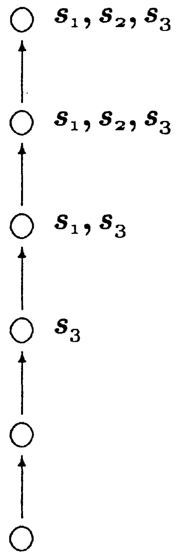
$$\mathbf{L} \vdash \Box^+ \bigwedge_{n,m\leq l} \mathrm{A}_{n,m}(\vec{s}) \rightarrow \Box^+\mathrm{C}(\vec{s}; \Box^n\Box\bot, \Box^m\Box\bot, \Box^k\Box\bot) \quad \text{for some } l \in \omega,$$

which immediately implies that if $n+m = k$ then $\vdash \mathrm{C}(\vec{\varsigma}; \Box^n\Box\bot, \Box^m\Box\bot, \Box^k\Box\bot)$ is indeed the case.

To establish the converse, we have to show under the assumption $n + m \neq k$ that

$$\mathbf{L} \vdash \Box^+ \bigwedge_{n,m\leq l} \mathrm{A}_{n,m}(\vec{s}) \rightarrow \Box^+\mathrm{C}(\vec{s}; \Box^n\Box\bot, \Box^m\Box\bot, \Box^k\Box\bot)$$

holds for no $l \in \omega$. We take a particular case $n = 2$ and $m = 1$. Consider the following $\vec{s}$-model in Figure 2.2. Observe that this is a model of every formula $\mathrm{A}_{i,j}(\vec{s})$, that is $\mathrm{A}_{i,j}(\vec{s})$ holds at each node of this model for any $i, j \in \omega$. On the other hand, it is easily seen that $k = 3 = 2 + 1 = n + m$ is the only value of $k$ s.t. the bottom node of this model forces $\mathrm{C}(\vec{s}; \Box^2\Box\bot, \Box^1\Box\bot, \Box^k\Box\bot)$. The proof is now easily completed by applying the Completeness Theorem for $\mathbf{L}$. ∎




(Figure 2.2)

2.9. THEOREM. *For* T *a* $\Sigma_1$*-sound theory,* $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$ *is mutually interpretable with true first order arithmetic.*

PROOF. For any $\Sigma_1$-sound theory, $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$ is straightforwardly interpretable in true arithmetic by gödelnumbering all the objects in question (cf. Montagna [18]).

Let us consider an inverse interpretation $[\cdot]$ of the variant of arithmetical language with ternary predicates in place of binary function symbols for addition and multiplication. $[\cdot]$ is defined by relativizing quantifiers to the set $N$ of Definition 1.4 and the following translation of terms and atomic formulas:

$$[0] \equiv \Box\bot$$

$$[Sx] \equiv \Box[x]$$

$$[t_1 = t_2] \equiv \vdash [t_1] \leftrightarrow [t_2]$$

$$[t_1 + t_2 = t_3] \equiv \forall \vec{\varsigma}, \vec{\varpi}\, \big(E(\vec{\varsigma}; \vec{\varpi}) \Rightarrow \vdash \mathrm{C}(\vec{\varsigma}; [t_1], [t_2], [t_3])\big)$$

$$[t_1 \cdot t_2 = t_3] \equiv \forall \vec{\varsigma}, \vec{\varpi}\, \big(E(\vec{\varsigma}; \vec{\varpi}) \Rightarrow \vdash \mathrm{C}(\vec{\varpi}; [t_1], [t_2], [t_3])\big),$$

where $E(\vec{\varsigma}; \vec{\varpi})$ is the following predicate:

$$E(\vec{\varsigma}; \vec{\varpi}) \equiv \forall \nu, \mu \in N\, \exists! \kappa \in N \vdash \mathrm{C}(\vec{\varsigma}; \nu, \mu, \kappa)\ \&\ \forall \nu, \mu \in N\, \exists! \kappa \in N \vdash \mathrm{C}(\vec{\varpi}; \nu, \mu, \kappa)$$
$$\&\ \forall \nu \in N \vdash \mathrm{C}(\vec{\varsigma}; \nu, \Box\bot, \nu)\ \&\ \forall \nu, \mu, \kappa \in N\, \big(\vdash \mathrm{C}(\vec{\varsigma}; \nu, \mu, \kappa) \Rightarrow \vdash \mathrm{C}(\vec{\varsigma}; \nu, \Box\mu, \Box\kappa)\big)$$
$$\&\ \forall \nu \in N \vdash \mathrm{C}(\vec{\varpi}; \nu, \Box\bot, \Box\bot)$$
$$\&\ \forall \nu, \mu, \kappa, \lambda \in N\, \Big(\big(\vdash \mathrm{C}(\vec{\varpi}; \nu, \mu, \kappa)\ \&\ \vdash \mathrm{C}(\vec{\varsigma}; \kappa, \nu, \lambda)\big) \Rightarrow \vdash \mathrm{C}(\vec{\varpi}; \nu, \Box\mu, \lambda)\Big).$$

$E(\vec{\varsigma}; \vec{\varpi})$ obviously says that $\vdash \mathrm{C}(\vec{\varsigma}; \cdot, \cdot, \cdot)$ and $\vdash \mathrm{C}(\vec{\varpi}; \cdot, \cdot, \cdot)$ are functional and that the clauses of the recursive definitions of addition and multiplication in terms of zero ($\Box\bot$) and successor function ($\Box$) hold for these formulas respectively.

Mimicking computations using these recursive definitions, one easily proves that once $E(\vec{\varsigma};\vec{\varpi})$ holds in some diagonalizable algebra of infinite credibility extent, the formulas $\vdash \mathrm{C}(\vec{\varsigma};\cdot,\cdot,\cdot)$ and $\vdash \mathrm{C}(\vec{\varpi};\cdot,\cdot,\cdot)$ do indeed correctly represent $+$ and $\times$ on the superscripts of elements of the set $\{\Box^n\Box\bot\}_{n\in\omega}$ which, by Proposition 1.15, is itself singled out by the formula $\nu \in N$ in diagonalizable algebras of $\Sigma_1$-sound theories. Proposition 2.5 is now all that is needed to see that $[\,\cdot\,]$ is the sought after interpretation. ∎

2.10. Corollary. *The first order theory of the diagonalizable algebra of any $\Sigma_1$-sound theory* T *(or even that of any collection of diagonalizable algebras of such) is not arithmetic.*

Proof. This is so because Theorem 2.9 interprets true arithmetic into $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$ by an interpretation which does not depend on T. ∎

In connection with Corollary 2.10 it is perhaps worth pointing out that it is not known whether $\mathrm{Th}\,\mathcal{D}_{\mathrm{T}}$ is at all influenced by particular choices of a $\Sigma_1$-sound theory T, although nonisomorphic diagonalizable algebras are indeed found among the ones of theories from this class (cf. Shavrukov [21]).

## 3. Simulating monogenic normal canonical systems

Our proof of Proposition 1.15 depended in an essential way on the unprovability of the sentence $\pi \to L_\pi{=}e_0$ in the formal theory under consideration and, ultimately, on the $\Sigma_1$-soundness of that theory (see Lemma 1.8). In this Section we are going to deal with an arbitrary theory T of infinite credibility extent which we now fix for this whole Section. Under these circumstances we can no longer count on the good behaviour of T w.r.t. $\Pi_2$ sentences and the proof of Proposition 1.15 is generally no longer valid for theories from this wider class. In $\mathcal{D}^*_{\mathrm{T}}$, the reason why we can not just repeat the construction of Section 1 is that it is no longer clear why the mere truth of the sentence $\forall j\, \neg\Box(\nu \leftrightarrow \Box^j\Box\bot)$ should guarantee that the clopen $\varepsilon^* \cap (\forall j\, \neg\Box(\nu \leftrightarrow \Box^j\Box\bot))^*$ is nonempty.

I do not know whether the diagonalizable algebraic formula $\nu \in N$ defines in $\mathcal{D}_{\mathrm{S}}$ the set $\{\Box^n\Box\bot\}_{n\in\omega}$ for every theory S of infinite credibility extent.

What we are able to do using a technique similar to the one featured in Section 1 is to translate into the structure of $\mathcal{D}_{\mathrm{T}}$ computations performed by monogenic normal canonical systems.

A *monogenic normal canonical system* (*mcsystem* for short) $\mathbf{t}$ consists of a finite alphabet $\Lambda = \{\ell_0, \ldots, \ell_N\}$, a non-null word $\boldsymbol{t}_0$, called the *axiom* of $\mathbf{t}$, in this alphabet, and a finite collection of *productions*. The latter are expressions of the form $\boldsymbol{g}\$ \to \$\boldsymbol{h}$, where $\boldsymbol{g}$ and $\boldsymbol{h}$ are words in $\Lambda$, $\boldsymbol{g}$ non-null, and \$ is a special symbol outside $\Lambda$. The *monogeneity condition* states that the multiset $\{\, \boldsymbol{g} \mid \boldsymbol{g}\$ \to \$\boldsymbol{h}$ is a production of $\mathbf{t}\,\}$ is prefix-free. Minsky's book [16, Part III] contains a detailed exposition of monogenic normal as well as other kinds of canonical systems.

The purpose of a mcsystem $\mathbf{t}$ is to produce a *monologue*, which is a sequence $(\boldsymbol{t}_i)_{i\in\bar{\omega}}$

of words in $\Lambda$, where $\tilde{\omega}$ stands for some ordinal $\leq\omega$, $\boldsymbol{t}_0$ is the axiom of $\mathbf{t}$, and for each word $\boldsymbol{t}_{i+1}$ in the monologue we have that there are words $\boldsymbol{g}$, $\boldsymbol{f}$ and $\boldsymbol{h}$ s.t. $\boldsymbol{t}_i = \boldsymbol{gf}$, $\boldsymbol{t}_{i+1} = \boldsymbol{fh}$ and $\boldsymbol{g}\$ \to \$\boldsymbol{h}$ is a production of $\mathbf{t}$. Finally, if $\tilde{\omega}$ is finite and $\boldsymbol{t}_K$ is the last word of the monologue then no production of $\mathbf{t}$ should be *applicable* to $\boldsymbol{t}_K$, i.e. $\boldsymbol{t}_K$ can not be represented as $\boldsymbol{gf}$ for any production $\boldsymbol{g}\$ \to \$\boldsymbol{h}$ of $\mathbf{t}$. Note that the monogeneity condition ensures that at most one production is applicable to any word, and hence the monologue of a mcsystem is determined uniquely by the identity of that system. Words in the monologue are said to be *produced* by $\mathbf{t}$. The mcsystem is *mortal* or *immortal* according to whether $\tilde{\omega}$ is finite or infinite.

Owing to the fact that the words $\boldsymbol{t}_i$ and $\boldsymbol{t}_{i+1}$ in the monologue may overlap considerably, there is a more economic way to write down the monologue of a mcsystem.

We consider three sequences $(t_i)_{i\in\bar{\omega}-\{0\}}$, $(r_i)_{i\in\tilde{\omega}}$ and $(s_i)_{i\in\tilde{\omega}}$, the first of symbols in $\Lambda$, the second and the third of natural numbers $\in\bar{\omega}+1$. $(t_i)_{i\in\bar{\omega}-\{0\}}$ is called the *tape* of $\mathbf{t}$ and $(r_i)_{i\in\tilde{\omega}}$ and $(s_i)_{i\in\tilde{\omega}}$ the *delimiting sequences.* The words of the monologue are written consequtively on the tape with a suffix of each word overlapping with the identical prefix of the following one, and the delimiting sequences tell us where each word begins and ends. Formally, for $i \in \tilde{\omega}$, we have $\boldsymbol{t}_i = t_{r_i} \cdots t_{s_i}$ (if $\boldsymbol{t}_i$ is the null word, then $r_i = s_i + 1$), and if $\boldsymbol{g}\$ \to \$\boldsymbol{h}$ is the (unique) production of $\mathbf{t}$ that bridges $\boldsymbol{t}_i$ and $\boldsymbol{t}_{i+1}$, then $\boldsymbol{g} = t_{r_i} \cdots t_{r_{i+1}-1}$ and $\boldsymbol{h} = t_{s_i+1} \cdots t_{s_{i+1}}$ so that $\boldsymbol{t}_{i+1} = t_{r_{i+1}} \cdots t_{s_{i+1}}$. (The overlapping part $t_{r_{i+1}} \cdots t_{s_i}$ is then understood to correspond to \$.) $\bar{\omega} \leq \omega$ is the ordinal just large enough to record all the words in the monologue, so that $\bar{\omega} = \omega$ if $\tilde{\omega} = \omega$, and $\max\bar{\omega} = s_K$ if $\max\tilde{\omega} = K$.

Two words are called *compatible* if one of them is a prefix of the other. $|\boldsymbol{e}|$ denotes the length of (= the number of letters in) the word $\boldsymbol{e}$. $(\boldsymbol{e})_i$ is the $i$th letter in that word. Each non-null word begins with its 1st letter.

Let us fix a mcsystem $\mathbf{t}$ along with the accompanying monologue, tape and delimiting sequences for this Section.

On pages to follow we shall quite often be reasoning about $\mathbf{t}$ within the formal theory T. We assume that the formalization of the objects related to $\mathbf{t}$ is honest and coherent so that T is aware of simple facts of $\mathbf{t}$'s life like e.g. $s_i = r_i - 1 + |\boldsymbol{t}_i|$.

Let us now try to explain our intentions from the viewpoint of the dual space $\mathcal{D}^*_{\mathrm{T}}$.

As in Section 1 (see the digression immediately after Proposition 1.15), a prominent role will be played by a certain clopen $\varepsilon^*$ which will this time observe $R$-above itself two slates corresponding to clopens $\alpha^*$ and $(\alpha \vee \beta)^*$. The single slate $\alpha_\pi{}^*$ of Section 1 was designed to represent the natural numbers. Here $\alpha^*$ and $\alpha^* \cup \beta^*$ are going to represent the tape of the mcsystem $\mathbf{t}$. These two clopens are notched by box clopens in technically the same way as in Section 1. Notches on $\alpha^*$ relate to the delimiting sequence $(r_i)_{i\in\tilde{\omega}}$ and those on $(\alpha \vee \beta)^*$ to $(s_i)_{i\in\tilde{\omega}}$.

Imagine a nonempty clopen $\mu^* \subseteq \varepsilon^*$ distinguishing two notches $\tau^*$ and $\sigma^*$ on $\alpha^*$ and $(\alpha \vee \beta)^*$ respectively. In accordance with our intended interpretation, we shall then think of the clopen $(\neg\tau \wedge \sigma)^*$ as delimiting a certain word $\boldsymbol{e}$ on $\mathbf{t}$'s tape. Since words are made of letters, we shall take account of these by associating to each letter $\ell \in \Lambda$ a particular clopen $\lambda_\ell{}^*$. The situation $(\Downarrow^i\tau)^* \subseteq \lambda_\ell{}^* \cap \tau^*$ will signal that $(\boldsymbol{e})_i = \ell$.

With these intuitions we shall be able to first order express that, for example, the word $\ell_1\ell_2$ occurs on the tape of $\mathbf{t}$ by saying that there is a nonempty clopen $\mu^* \subseteq \varepsilon^*$ observing a pair $(\tau^*, (\Box\Box\tau)^*)$ of notches with the antichains $(\Downarrow\tau)^*$ and $(\Downarrow^2\tau)^*$ contained in the clopens $\lambda_{\ell_1}{}^*$ and $\lambda_{\ell_2}{}^*$ respectively.

Note that our present aspirations are only limited to talking about the behaviour of clopens $(\lambda_\ell{}^*)_{\ell\in\Lambda}$ w.r.t. one another involving a fixed finite difference in $R$-depth between antichains, whereas in Sections 1 we succeeded with $\Sigma_1$-sound theories in first order defining membership in $\{\Box^n\Box\bot\}_{n\in\omega}$, which is a property concerning arbitrary finite $R$-depth of clopens. This restraint on ambitions owes to the now less manageable nature of the relation $R$ of the dual space: $R^{-1}$ in $\mathcal{D}_\mathrm{S}$ is directed if and only if S is $\Sigma_1$-sound.

The ability to express that a given word is produced by a given mcsystem will eventually lead us to undecidability results. What we have to do next, however, is to develop a language in which one can comprehensibly talk about a mcsystem to $\mathrm{Th}\,\mathcal{D}_\mathrm{T}$.

3.1. DEFINITION. We define two elementary diagonalizable algebraic predicates. Below, $S(\cdots)$ is taken as specified by Definition 1.4, $\vec{\lambda}$ stands for a collection of elements of a diagonalizable algebra indexed by letters in $\Lambda$, and $e$ is a (meta)variable for words in $\Lambda$. For the $\Downarrow^i$ operator, confer Definition 2.1.

$$R(\alpha,\beta;\mu,\tau,\sigma) \equiv S(\alpha;\mu,\tau) \;\&\; S(\alpha\vee\beta;\mu,\sigma)$$

$$W^e(\vec{\lambda};\mu,\tau) \equiv \mathop{\&}_{1\le i\le|e|} \vdash\mu\to\Box(\Downarrow^i\tau\to\lambda_{(e)_i})$$

$R(\alpha,\beta;\mu,\tau,\sigma)$ expresses that $\mu^*$ detects the box clopens $\tau^*$ and $\sigma^*$ as notches on $\alpha^*$ and $\alpha^*\cup\beta^*$ respectively, these two clopens delimiting a space on the tape in which we shall be trying to read a $\Lambda$-word. $\mu^*$ is of the opinion that a word $e$ is read below $\tau^*$ if $W^e(\vec{\lambda};\mu,\tau)$ holds, that is, if $\mu^*$ thinks that the antichain finding itself $i$ $R$-steps below $\tau^*$ is entirely contained in $\lambda_{(e)_i}{}^*$ whenever $0<i\le|e|$. $W^e(\cdots,\tau)$ will, in practice, only be applied to box elements $\tau$.

3.2. LEMMA. *In any diagonalizable algebra,*

(a) *If* $\vdash\kappa\to\mu$ *and* $R(\alpha,\beta;\mu,\tau,\sigma)$, *then* $R(\alpha,\beta;\kappa,\tau,\sigma)$.

(b) *If* $\rho\in B$, $R(\alpha,\beta;\mu,\tau,\sigma)$ *and* $\vdash\mu\to\Box(\rho\leftrightarrow\tau)$, *then* $R(\alpha,\beta;\mu,\rho,\sigma)$.

(c) *If* $\rho\in B$, $R(\alpha,\beta;\mu,\tau,\sigma)$ *and* $\vdash\mu\to\Box(\rho\leftrightarrow\sigma)$, *then* $R(\alpha,\beta;\mu,\tau,\rho)$.

(d) *If* $R(\alpha,\beta;\mu,\tau,\sigma)$ *and* $R(\alpha,\beta;\mu,\rho,\pi)$, *then* $\vdash\mu\to\Box(\rho\leftrightarrow\tau)$ *and* $\vdash\mu\to\Box(\pi\leftrightarrow\sigma)$.

PROOF. All clauses follow straightforwardly from Lemma 1.5. ∎

3.3. LEMMA. *In any diagonalizable algebra,*

(a) *If* $\vdash\kappa\to\mu$ *and* $W^e(\vec{\lambda};\mu,\tau)$, *then* $W^e(\vec{\lambda};\kappa,\tau)$.

(b) *If* $W^e(\vec{\lambda};\mu,\tau)$ *and* $\vdash\mu\to\Box(\sigma\leftrightarrow\tau)$, *then* $W^e(\vec{\lambda};\mu,\sigma)$.

(c) $W^{ef}(\vec{\lambda};\mu,\tau)$ *if and only if* $W^e(\vec{\lambda};\mu,\tau)$ *and* $W^f(\vec{\lambda};\mu,\Box^{|e|}\tau)$.

PROOF. (a) and (b) are quite obvious; to establish (c), one applies Lemma 2.2(c). ∎

We go on to display the diagonalizable algebraic formulas intended to embed the given mcsystem $\mathbf{t}$ into $\mathrm{Th}\,\mathcal{D}_\mathrm{T}$.

3.4. DEFINITION. Below, $\boldsymbol{e}$, $\boldsymbol{g}$ and $\boldsymbol{h}$ are arbitrary words in $\Lambda$, and $\vec{\lambda}$ is taken as in Definition 3.1. $B$ and $T$ are described in Definition 1.1.

$$X^{\boldsymbol{e}}(\alpha,\beta,\varepsilon,\vec{\lambda};\tau) \equiv \exists\mu\,(\nvdash\neg\mu \ \&\ \vdash\mu\rightarrow\varepsilon \ \&\ R(\alpha,\beta;\mu,\tau,\Box^{|\boldsymbol{e}|}\tau) \ \&\ W^{\boldsymbol{e}}(\vec{\lambda};\mu,\tau))$$

$$O^{\boldsymbol{e}}(\alpha,\beta,\varepsilon,\vec{\lambda}) \equiv \exists\tau{\in}T\, X^{\boldsymbol{e}}(\alpha,\beta,\varepsilon,\vec{\lambda};\tau)$$

$$P^{\boldsymbol{g}\$\rightarrow\$\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda}) \equiv \forall\tau,\sigma{\in}T\,\forall\rho{\in}B$$

$$\Big(\exists\mu\,(\nvdash\neg\mu \ \&\ \vdash\mu\rightarrow\varepsilon \ \&\ R(\alpha,\beta;\mu,\tau,\sigma) \ \&\ W^{\boldsymbol{g}}(\vec{\lambda};\mu,\tau) \ \&\ \vdash\mu\rightarrow\rho)$$

$$\Rightarrow \exists\nu\,(\nvdash\neg\nu \ \&\ \vdash\nu\rightarrow\varepsilon \ \&\ R(\alpha,\beta;\nu,\Box^{|\boldsymbol{g}|}\tau,\Box^{|\boldsymbol{h}|}\sigma) \ \&\ W^{\boldsymbol{h}}(\vec{\lambda};\nu,\sigma) \ \&\ \vdash\nu\rightarrow\rho)\Big)$$

$$D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda}) \equiv X^{\boldsymbol{t}_0}(\alpha,\beta,\varepsilon,\vec{\lambda};\Box\bot) \ \&\ \mathop{\&}_{\substack{\text{all productions}\\ \boldsymbol{g}\$\rightarrow\$\boldsymbol{h}\text{ of t}}} P^{\boldsymbol{g}\$\rightarrow\$\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda})$$

$$M^{\boldsymbol{e}}_{\mathbf{t}} \equiv \forall\alpha,\beta,\varepsilon,\vec{\lambda}\,(D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda}) \Rightarrow O^{\boldsymbol{e}}(\alpha,\beta,\varepsilon,\vec{\lambda}))$$

$X^{\boldsymbol{e}}(\alpha,\beta,\varepsilon,\vec{\lambda};\tau)$ asserts that there is a nonempty clopen $\mu^* \subseteq \varepsilon^*$ relative to which $(\tau^*,(\Box^{|\boldsymbol{e}|}\tau)^*)$ is a pair of notches on the tape which delimits the word $\boldsymbol{e}$. $O^{\boldsymbol{e}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ just says that there is such a pair of top-box notches.

The predicate $P^{\boldsymbol{g}\$\rightarrow\$\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ is very similar to

$$\forall\tau,\sigma{\in}T\,\forall\boldsymbol{f}\,(X^{\boldsymbol{g}\boldsymbol{f}}(\alpha,\beta,\varepsilon,\vec{\lambda};\tau) \Rightarrow X^{\boldsymbol{f}\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda};\Box^{|\boldsymbol{g}|}\tau)),$$

which is very much in tune with the effect of a production $\boldsymbol{g}\$ \rightarrow \$\boldsymbol{h}$. The quantifier $\forall\boldsymbol{f}$ over words in $\Lambda$ is, of course, not allowed in the first order language of diagonalizable algebras. Therefore, $W^{\boldsymbol{f}}(\vec{\lambda};\mu,\Box^{|\boldsymbol{g}|}\tau)$, which has the form $\vdash\mu\rightarrow\rho$ for a certain box element $\rho$, is replaced by this expression and quantification over $\boldsymbol{f}$ by that over box elements $\rho$. Thus $P^{\boldsymbol{g}\$\rightarrow\$\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ says something ever so slightly stronger than what we actually need.

$D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda})$: the axiom $\boldsymbol{t}_0$ of $\mathbf{t}$ is written at the very beginning of the tape and all productions of $\mathbf{t}$ are operative. We like to think of $D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ as describing $\mathbf{t}$ w.r.t. the parameters $\alpha,\beta,\varepsilon,\vec{\lambda}$.

The diagonalizable algebraic sentence $M^{\boldsymbol{e}}_{\mathbf{t}}$ claims that the word $\boldsymbol{e}$ occurs in the monologue once the parameters $\alpha,\beta,\varepsilon,\vec{\lambda}$ match the description of $\mathbf{t}$. Observe that $M^{\boldsymbol{e}}_{\mathbf{t}}$ is constructed effectively from $\mathbf{t}$ and $\boldsymbol{e}$.

We would like to verify that $M^{\boldsymbol{e}}_{\mathbf{t}}$ holds in $\mathcal{D}_{\mathrm{T}}$ if and only if $\mathbf{t}$ actually produces $\boldsymbol{e}$. Our strategy is to show, in one direction, that $D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ describes $\mathbf{t}$ in sufficient detail to infer $O^{\boldsymbol{e}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ for every word $\boldsymbol{e}$ in the monologue of $\mathbf{t}$ (Lemma 3.5). The other direction establishes that we can avoid every other word by choosing appropriate parameters $\alpha_{\mathbf{t}},\beta_{\mathbf{t}},\varepsilon_{\mathbf{t}},\vec{\lambda}_{\mathbf{t}}$ in $\mathcal{D}_{\mathrm{T}}$ (Definition 3.6–Proposition 3.18).

3.5. LEMMA. *If* $\mathbf{t}$ *produces* $\boldsymbol{e}$ *then* $M^{\boldsymbol{e}}_{\mathbf{t}}$ *holds in any diagonalizable algebra.*

PROOF. If $\mathbf{t}$ produces $\boldsymbol{e}$ then $\boldsymbol{e} = \boldsymbol{t}_k$ for some $k \in \tilde{\omega}$. We fix parameters $\alpha,\beta,\varepsilon,\vec{\lambda}$ and prove by induction on $k \in \tilde{\omega}$ that

$$D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda}) \Rightarrow X^{\boldsymbol{t}_k}(\alpha,\beta,\varepsilon,\vec{\lambda};\Box^{r_k-1}\Box\bot),$$

whence $D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda}) \Rightarrow O^{\boldsymbol{t}_k}(\alpha,\beta,\varepsilon,\vec{\lambda})$ readily follows.

We proceed under the hypothesis $D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda})$.

$X^{\boldsymbol{t}_0}(\alpha,\beta,\varepsilon,\vec{\lambda};\Box^{r_0-1}\Box\bot)$ follows from this hypothesis by the definition of $D_{\mathbf{t}}(\cdots)$ since $r_0 - 1 = 0$.

Turning to $k+1 \in \tilde{\omega}$, let us suppose that $\boldsymbol{t}_{k+1}$ is obtained from $\boldsymbol{t}_k$ by a production $\boldsymbol{g}\$ \rightarrow \$\boldsymbol{h}$ of $\mathbf{t}$ so that $\boldsymbol{t}_k = \boldsymbol{gf}$ and $\boldsymbol{t}_{k+1} = \boldsymbol{fh}$ for some word $\boldsymbol{f}$. By IH, we have $X^{\boldsymbol{t}_k}(\alpha,\beta,\varepsilon,\vec{\lambda};\Box^{r_k-1}\Box\bot)$ which proclaims the existence of an element $\mu \not\equiv \bot$ (i.e. $\nvdash \neg\mu$) s.t.

$$\vdash \mu \rightarrow \varepsilon \;\;\&\;\; R(\alpha,\beta;\mu,\Box^{r_k-1}\Box\bot,\Box^{|\boldsymbol{t}_k|}\Box^{r_k-1}\Box\bot) \;\;\&\;\; W^{\boldsymbol{t}_k}(\vec{\lambda};\mu,\Box^{r_k-1}\Box\bot).$$

Since $\boldsymbol{t}_k = \boldsymbol{gf}$, $W^{\boldsymbol{t}_k}(\vec{\lambda};\mu,\Box^{r_k-1}\Box\bot)$ implies, by Lemma 3.3(c), both $W^{\boldsymbol{g}}(\vec{\lambda};\mu,\Box^{r_k-1}\Box\bot)$ and $W^{\boldsymbol{f}}(\vec{\lambda};\mu,\Box^{|\boldsymbol{g}|}\Box^{r_k-1}\Box\bot)$. This latter fact can be rewritten as $\vdash \mu \rightarrow \Box \bigwedge\!\!\!\bigwedge_{1\le j\le|\boldsymbol{f}|}(\Downarrow^j \Box^{|\boldsymbol{g}|+r_k-1}\Box\bot \rightarrow \lambda_{(\boldsymbol{f})_j})$. Note that $|\boldsymbol{t}_k| + r_k - 1 = s_k$ and $|\boldsymbol{g}| + r_k - 1 = r_{k+1} - 1$.

Now, since $\boldsymbol{g}\$ \rightarrow \$\boldsymbol{h}$ is a production of $\mathbf{t}$, $D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ includes the formula $P^{\boldsymbol{g}\$\rightarrow\$\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda})$ among its conjuncts. In the preceding paragraph we have in effect established that

$$\nvdash \neg\mu \;\;\&\;\; \vdash \mu \rightarrow \varepsilon \;\;\&\;\; R(\alpha,\beta;\mu,\tau,\sigma) \;\;\&\;\; W^{\boldsymbol{g}}(\vec{\lambda};\mu,\tau) \;\;\&\;\; \vdash \mu \rightarrow \rho,$$

the antecedent of the matrix of $P^{\boldsymbol{g}\$\rightarrow\$\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda})$, holds for $\tau \equiv \Box^{r_k-1}\Box\bot$, $\sigma \equiv \Box^{s_k}\Box\bot$, $\rho \equiv \Box \bigwedge\!\!\!\bigwedge_{1\le j\le|\boldsymbol{f}|}(\Downarrow^j \Box^{r_{k+1}-1}\Box\bot \rightarrow \lambda_{(\boldsymbol{f})_j})$ and the element $\mu$ whose existence is asserted by $X^{\boldsymbol{t}_k}(\alpha,\beta,\varepsilon,\vec{\lambda};\Box^{r_k-1}\Box\bot)$. Therefore, as follows from $P^{\boldsymbol{g}\$\rightarrow\$\boldsymbol{h}}(\alpha,\beta,\varepsilon,\vec{\lambda})$, there is an element $\nu \not\equiv \bot$ s.t.

$$\vdash \nu \rightarrow \varepsilon \;\;\&\;\; R(\alpha,\beta;\nu,\Box^{|\boldsymbol{g}|}\Box^{r_k-1}\Box\bot,\Box^{|\boldsymbol{h}|}\Box^{s_k}\Box\bot) \;\;\&\;\; W^{\boldsymbol{h}}(\vec{\lambda};\nu,\Box^{s_k}\Box\bot)$$
$$\&\;\; \vdash \nu \rightarrow \Box \bigwedge\!\!\!\bigwedge_{1\le j\le|\boldsymbol{f}|} (\Downarrow^j \Box^{r_{k+1}-1}\Box\bot \rightarrow \lambda_{(\boldsymbol{f})_j}).$$

The last conjunct rewrites as $W^{\boldsymbol{f}}(\vec{\lambda};\nu,\Box^{r_{k+1}-1}\Box\bot)$ and, since $s_k = |\boldsymbol{f}|+r_{k+1}-1$, the last two give $W^{\boldsymbol{fh}}(\vec{\lambda};\nu,\Box^{r_{k+1}-1}\Box\bot)$ by Lemma 3.3(c). Recalling that $|\boldsymbol{g}|+r_k-1 = r_{k+1}-1$, $|\boldsymbol{h}| + s_k = s_{k+1} = |\boldsymbol{t}_{k+1}| + r_{k+1} - 1$ and $\boldsymbol{fh} = \boldsymbol{t}_{k+1}$, we obtain

$$\exists\nu\,(\nvdash \neg\nu \;\;\&\;\; \vdash \nu \rightarrow \varepsilon \;\;\&\;\; R(\alpha,\beta;\nu,\Box^{r_{k+1}-1}\Box\bot,\Box^{|\boldsymbol{t}_{k+1}|}\Box^{r_{k+1}-1}\Box\bot)$$
$$\&\;\; W^{\boldsymbol{t}_{k+1}}(\vec{\lambda};\nu,\Box^{r_{k+1}-1}\Box\bot)),$$

or $X^{\boldsymbol{t}_{k+1}}(\alpha,\beta,\varepsilon,\vec{\lambda};\Box^{r_{k+1}-1}\Box\bot)$ as we have pledged to show. ∎

To reverse the implication of Lemma 3.5 for $\mathcal{D}_{\mathrm{T}}$ we define another Solovay function similar to the one of Definition 1.7.

3.6. Definition. Here $H_{\mathbf{t}}$ ranges over the set $\{a_i\}_{i\in\omega} \cup \{b_i\}_{i\in\omega-\{0\}} \cup \{c_i\}_{i\in\omega-\{0\}} \cup \{e_i\}_{i\in\omega} \cup \{0\}$. We abbreviate $\exists x\,\forall y{\ge}x\, H_{\mathbf{t}}(y){=}f$ by $L_{\mathbf{t}}{=}f$.

As usual, we have

$$H_{\mathbf{t}}(0) = 0.$$

Next we define the value of $H_{\mathbf{t}}(x+1)$:

C a s e A. $H_t(x) = 0$ and $\Box_x L_t{\neq}e_i$ for some $i \in \tilde{\omega}$.

$H_t(x+1) = e_i$.

C a s e B. $H_t(x) = e_i$, $i+1 \in \tilde{\omega}$ and for some sentence $\varphi$ and $y < x$ we have $\Box_x\varphi$, $\Box_y(\Box\varphi \rightarrow L_t{\neq}e_{i+1})$ and $H_t(y) = 0$.

$H_t(x+1) = e_{i+1}$.

C a s e C. $H_t(x) = e_i$, Case B is not the case and $\Box_x L_t{\neq}a_j$ with $j \leq s_k + 1$ for some $k \in \tilde{\omega}$ s.t. $H_t(k \dot{-} i) = 0$.

$H_t(x+1) = a_j$.

C a s e D. $H_t(x) = a_{r_i}$ and for some sentence $\varphi$ and $y < x$ we have $\Box_x\varphi$, $\Box_y(\Box\varphi \rightarrow L_t{\neq}b_{r_i})$ and $H_t(y) = e_i$.

$H_t(x+1) = b_{r_i}$.

C a s e E. $H_t(x) = a_{s_i+1}$ or $H_t(x) = b_{s_i+1}$, Case D is not the case and for some sentence $\varphi$ and $y < x$ we have $\Box_x\varphi$, $\Box_y(\Box\varphi \rightarrow L_t{\neq}c_{s_i+1})$ and $H_t(y) = e_i$.

$H_t(x+1) = c_{s_i+1}$.

C a s e F. $H_t(x) = a_i$ or $H_t(x) = b_i$ or $H_t(x) = c_i$, Cases D–E do not hold and $\Box_x L_t{\neq}a_j$ with $j < i$.

$H_t(x+1) = a_j$.

C a s e G. None of the preceding Cases takes place.

$H_t(x+1) = H_t(x)$.

The function $H_t$ is thus successfully defined.

We also give special names to three sentences:

$$\alpha_t \equiv \exists i\, L_t = a_i, \quad \beta_t \equiv \exists i\, L_t = b_i, \quad \text{and} \quad \varepsilon_t \equiv \exists i\, L_t = e_i.$$

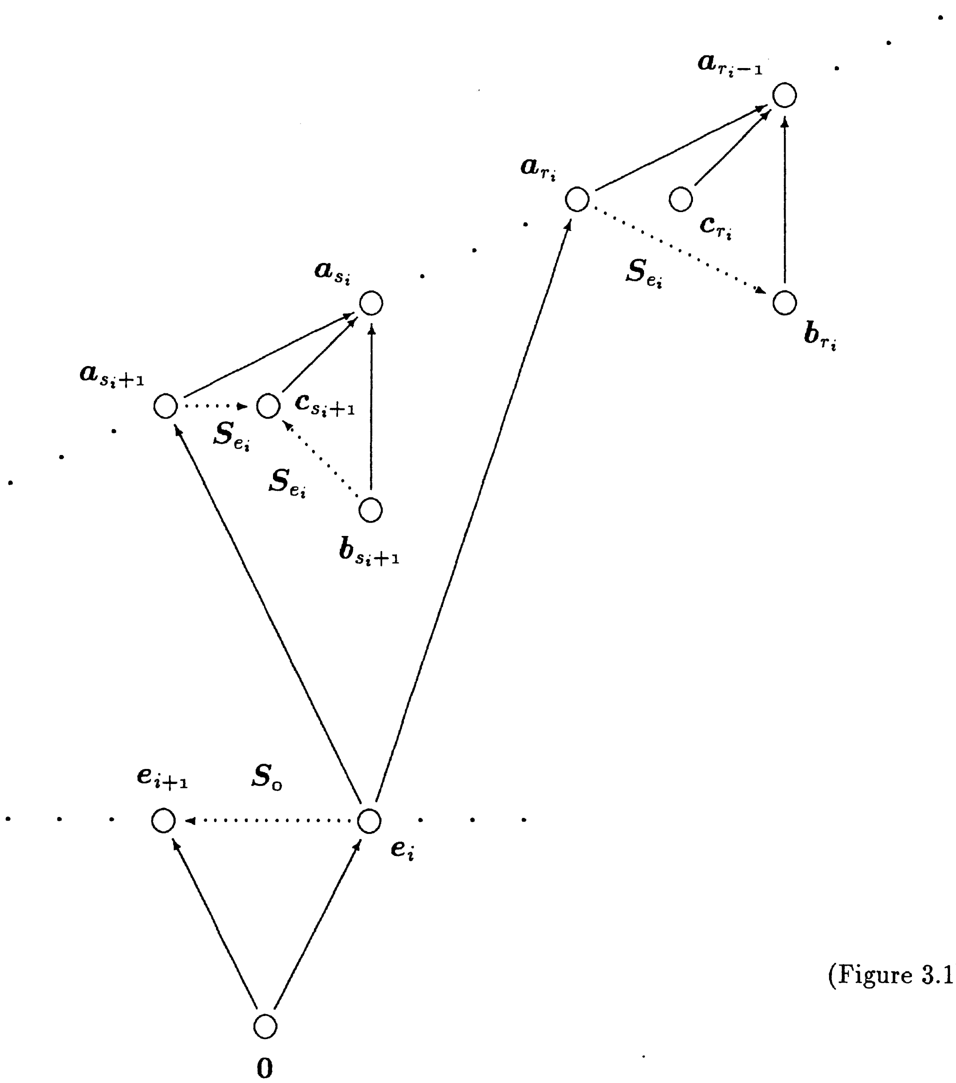
Note that the region accessible to $H_t$ from 0 and $\{e_i\}_{i\in\omega}$ depends on the size of **t**'s monologue (Cases A–C). The extra complication in Case C, i.e. the requirement $H_t(k \dot{-} i) = 0$, is designed to keep $H_t$ at 0 for the standard period of its life. What goes on in Cases D–F is fairly analogous to Cases D–E of Definition 1.7, although now we have got two series $\{b_i\}_{i\in\omega-\{0\}}$, $\{c_i\}_{i\in\omega-\{0\}}$ of auxiliary nodes in place of one.

Figure 3.1 shows what things look like if one tries to partially grasp the definition of $H_t$ in a Kripke frame. Cases D and E correspond to transitions along the $S_{e_i}$ arrows.

Now come the lemmas. While the convergence of $H_t$ does not present a problem, we shall have to exercise a little patience before we can claim that $H_t$ is, at standard arguments, constantly =0 (Lemma 3.10).

3.7. LEMMA. $\vdash \exists f{\in}\{a_i\}_{i\in\omega}{\cup}\{b_i\}_{i\in\omega-\{0\}}{\cup}\{c_i\}_{i\in\omega-\{0\}}{\cup}\{e_i\}_{i\in\omega}{\cup}\{0\}\ L_t{=}f$.

PROOF. The proof proceeds parallel to that of Lemma 1.8(a). The relevant observation here is that from $a_i$ the function $H_t$ can go to $b_i$ and then to $c_i$ whereafter, to keep on moving, it will have to go to $a_j$ with $j < i$. ∎




(Figure 3.1)

Lemmas 3.8–9 are largely similar to Lemmas 1.9–10 of Section 1.

3.8. LEMMA. (a) $\vdash \forall i\,(L_{\mathrm{t}}{=}a_i \vee L_{\mathrm{t}}{=}b_i \vee L_{\mathrm{t}}{=}c_i \rightarrow .\ \Box \exists j{<}i\, L_{\mathrm{t}}{=}a_j)$.

(b) *If* $H_{\mathrm{t}}(k) = e_i$ *for some* $i$ *and* $k$, *then there exists an* $m \in \tilde{\omega}$ *s.t.* $\vdash \exists j{\leq}s_m{+}1\,(L_{\mathrm{t}}{=}a_j \vee L_{\mathrm{t}}{=}b_j \vee L_{\mathrm{t}}{=}c_j)$. *Moreover, this fact formalizes in* T.

(c) $\vdash \forall i\,(L_{\mathrm{t}}{=}a_i \vee L_{\mathrm{t}}{=}b_i \vee L_{\mathrm{t}}{=}c_i \rightarrow .\ \forall j{<}i\, \Diamond L_{\mathrm{t}}{=}a_j)$.

(d) *If* $k \in \tilde{\omega}$ *and* $H_{\mathrm{t}}(k) = 0$, *then* $\vdash \forall i\,(L_{\mathrm{t}}{=}e_i \rightarrow \forall j{\leq}s_k{+}1\, \Diamond L_{\mathrm{t}}{=}a_j)$, *and this formalizes in* T.

(e) $\vdash L_t{=}0 \rightarrow \Diamond L_t{=}e_0$.

PROOF. (a) and (c) are similar to (a) and (b) of Lemma 1.9 respectively.

(b). Assume $H_t(k) = e_i$. Reason in T:

Suppose $L_t = e_j$ for some $j$. Then there has to exist a proof $<k$ of $L_t{\neq}e_j$ from a true box sentence. By the Small Reflection Principle, $L_t \neq e_j$. Therefore, $H_t$ ends up among $a$'s, $b$'s and $c$'s. The only way for $H_t$ to get there away from the $e$'s is via Case C of Definition 3.6. Hence there is an $x$ with $H_t(x) = e_n$, $\Box_x L_t{\neq}a_j$ for some $j \leq s_m{+}1$, where $m \in \tilde{\omega}$ and $H_t(m \dot{-} n) = 0$. Thus $m - n < k$. Note also that $n - i \leq k$, for to get to $e_n$ from $e_i$, $H_t$ has to make $n - i$ moves, each corresponding to a different proof $<k$. Therefore $m < k + n \leq 2k + i$. Thus $H_t(x{+}1) = a_j$ with $j \leq s_m + 1$ where $2k + i > m \in \tilde{\omega}$. No matter how $H_t$ moves from then on, the subscript of its value can not decrease. So, $\exists j{\leq}s_m{+}1\,(L_t{=}a_j \vee L_t{=}b_j \vee L_t{=}c_j)$.

This shows that one can put $m = 2k + i$.

(d). Assume $k \in \tilde{\omega}$ and $H_t(k) = 0$. Reason in T:

Suppose $L_t = e_i$ and $j \leq s_k + 1$. If we had $\Box_x L_t{\neq}a_j$ for some $x$ s.t. $H_t(x) = e_i$, then, since $k \dot{-} i \leq k$ and so $H_t(k \dot{-} i) = 0$, instructions of Case C would bring $H_t(x{+}1)$ to $a_j$ contradicting $L_t{=}e_j$.

Thus $\vdash \forall i\,(L_t{=}e_i \rightarrow \forall j{\leq}s_k{+}1\,\Diamond L_t{=}a_j)$ q.e.d.

(e) is clear on inspection of Case A because $0 \in \tilde{\omega}$. ∎

3.9. LEMMA. (a) $\vdash \forall i\,(\exists j{\leq}i\,(L_t{=}a_j \vee L_t{=}b_j \vee L_t{=}c_j) \rightarrow \Box^i\Box\bot)$.

(b) $\vdash \forall i{\in}\tilde{\omega}\,(L_t{=}e_i \rightarrow \forall j{\leq}s_i{+}1\,\Box(\Box^j\Box\bot \rightarrow \exists k{\leq}j\,(L_t{=}a_k \vee L_t{=}b_k \vee L_t{=}c_k)))$.

(c) *If* $H_t(k) \neq 0$ *then* $\vdash \Box^{s_m+1}\Box\bot$ *for some* $m \in \tilde{\omega}$.

PROOF. (a) is analogous to Lemma 1.10(a).

(b). Working in T, assume $L_t = e_i$ for some $i \in \tilde{\omega}$. We show $\Box(\Box^j\Box\bot \rightarrow \exists k{\leq}j\,(\cdots))$ by induction on $j \leq s_i + 1$. For $j = 0$ this follows by Lemma 3.8(c) and the formalized version of clause (b) of the same Lemma. To carry out the induction step, reason inside $\Box$:

We have that $\Box^{j+1}\Box\bot$ implies $\Box\exists k{\leq}j\,(L_t{=}a_k \vee L_t{=}b_k \vee L_t{=}c_k)$ by the IH. In particular, $\Box L_t{\neq}a_{k+1}$. Therefore, by Lemma 3.8(c), $L_t{=}a_m \vee L_t{=}b_m \vee L_t{=}c_m$ can only hold for $m \leq j + 1$.

Thus we have seen that $\Box^{j+1}\Box\bot \rightarrow \forall k(L_t{=}a_k \vee L_t{=}b_k \vee L_t{=}c_k \rightarrow .\, k \leq j + 1)$. Applying Lemma 3.8(b) formalized, we get $\Box(\Box^{j+1}\Box\bot \rightarrow \exists k{\leq}j{+}1\,(L_t{=}a_k \vee L_t{=}b_k \vee L_t{=}c_k))$ as required.

(c). Suppose $H_t(k) \neq 0$ holds for some $k \in \omega$. Take the minimal such $k$. We then have that $H_t(k) = e_i$ for an appropriate $i$, whence it follows by Lemma 3.8(b) that $\vdash \exists j{\leq}s_m{+}1\,(L_t{=}a_j \vee L_t{=}b_j \vee L_t{=}c_j)$ for some $m \in \tilde{\omega}$. By (a) of the present Lemma one gets $\vdash \Box^{s_m+1}\Box\bot$ q.e.d. ∎

3.10. LEMMA. (a) *For all* $k \in \omega$, $H_t(k) = 0$.

(b) $\nvdash L_t \neq e_0$.

(c) $\vdash \varepsilon_{\rm t} \to \Diamond^{s_i+1}\Diamond\top$ *for all* $i \in \tilde{\omega}$.

(d) *If* t *is mortal and* $K = \max\tilde{\omega}$, *then* $\vdash \Diamond^{s_K+2}\Diamond\top \to L_{\rm t}{=}0$.

PROOF. (a). If $H_{\rm t}(k) \neq 0$ then, by Lemma 3.9(c), $\vdash \Box^{s_m+1}\Box\bot$ for some $m \in \tilde{\omega}$, contradicting the infinite credibility extent of T.

(b) follows from (a) by Lemma 3.8(c).

(c). By Lemma 3.8(d) and clause (a) of the present Lemma we have $\vdash \varepsilon_{\rm t} \to \Diamond L_{\rm t}{=}a_{s_i+1}$ for all $i \in \tilde{\omega}$. By Lemma 3.9(b), there holds $\vdash L_{\rm t}{=}a_{s_i+1} \to \Diamond^{s_i}\Diamond\top$, so $\vdash \varepsilon_{\rm t} \to \Diamond^{s_i+1}\Diamond\top$.

(d). To prove this, we have to formalize the proof of (a) in T. The only ingredient of the proof that fails to formalize is that $\nvdash \Box^{s_m+1}\Box\bot$ for all $m \in \tilde{\omega}$. However, since $K = \max\tilde{\omega}$, this aspect is captured by the antecedent $\Diamond^{s_K+2}\Diamond\top$. ∎

Lemmas 3.11–13 are analogous to Lemmas 1.11–13.

3.11. LEMMA. *For all box sentences* $\rho$ *and all* $i$ *s.t.* $i+1 \in \tilde{\omega}$, *if* $\rho$ *is consistent over* T *with* $L_{\rm t}{=}e_i$, *then* $\rho$ *is consistent over* T *with* $L_{\rm t}{=}e_{i+1}$.

PROOF. Very similar to Lemma 1.11. ∎

3.12. LEMMA. (a) $\vdash \forall i{\in}\tilde{\omega}\,(\exists j{\geq}i\, L_{\rm t}{=}e_j \to \Box(\Box^{r_i-1}\Box\bot \to \alpha_{\rm t}))$.

(b) *For all* $i \in \tilde{\omega}$, $R(\alpha_{\rm t}, \beta_{\rm t}; L_{\rm t}{=}e_i, \Box^{r_i-1}\Box\bot, \Box^{s_i}\Box\bot)$ *holds.*

PROOF. (a). We work in T. Pick an $i \in \tilde{\omega}$ and imagine $L_{\rm t} = e_j$ for some $j \geq i$. Since $H_{\rm t}$ can only get to $e_k$ if $k \in \tilde{\omega}$ (see Cases A and B of Definition 3.6), we have that $j \in \tilde{\omega}$. Let $x$ be s.t. $H_{\rm t}(x) = e_j$ and argue inside $\Box$:

> Assume $L_{\rm t} = b_k$ or $L_{\rm t} = c_k$ with $0 < k < r_i$. This can only happen if $\Box_y(\Box\varphi \to L_{\rm t}{\neq}b_k)$ (or $\Box_y(\Box\varphi \to L_{\rm t}{\neq}c_k)$ respectively) with $H_{\rm t}(y) = e_m$, where $k = r_m$ (or $k = s_m + 1$), and $\varphi$ provable. Since $r_i > r_m$ ($s_i + 1 \geq r_i > s_m + 1$), we have that $j \geq i > m$. Therefore, $y < x$ so that $\Box_{<x}(\Box\varphi \to L_{\rm t}{\neq}b_k)$ ($\Box_{<x}(\Box\varphi \to L_{\rm t}{\neq}c_k)$) whence, by Small Reflection, $L_{\rm t} \neq b_k$ ($L_{\rm t} \neq c_k$ respectively), which is a contradiction.

So, $\Box\forall k\,(L_{\rm t}{=}b_k \vee L_{\rm t}{=}c_k \to\!. \; k{\geq}r_i)$. Since $r_i - 1 < s_i + 1 \leq s_j + 1$, one has $\Box(\Box^{r_i-1}\Box\bot \to \exists k{<}r_i\,(L_{\rm t}{=}a_k \vee L_{\rm t}{=}b_k \vee L_{\rm t}{=}c_k))$ by Lemma 3.9(b). It therefore follows that $\Box(\Box^{r_i-1}\Box\bot \to \exists k{<}r_i\, L_{\rm t}{=}a_k)$ implying $\Box(\Box^{r_i-1}\Box\bot \to \alpha_{\rm t})$ q.e.d.

(b). For an arbitrary $i \in \tilde{\omega}$, we have to prove

$$S(\alpha_{\rm t}; L_{\rm t}{=}e_i, \Box^{r_i-1}\Box\bot) \;\&\; S(\alpha_{\rm t} \vee \beta_{\rm t}; L_{\rm t}{=}e_i, \Box^{s_i}\Box\bot).$$

We clearly have $\Box^{r_i-1}\Box\bot, \Box^{s_i}\Box\bot \in B$, and $\vdash L_{\rm t}{=}e_i \to \Box(\Box^{r_i-1}\Box\bot \to \alpha_{\rm t})$ has already been established in clause (a). $\vdash L_{\rm t}{=}e_i \to \Box(\Box^{s_i}\Box\bot \to\!. \; \alpha_{\rm t} \vee \beta_{\rm t})$ is verified in perfect analogy with (a) (for $0 < k \leq s_i$, one follows the parenthetical line in the formalized part of the argument in (a) to show $\vdash L_{\rm t}{=}e_i \to \Box\forall k\,(L_{\rm t}{=}c_k \to k{>}s_i)$). Thus we only have to verify

$$\vdash L_{\mathrm{t}} = e_i \rightarrow \big(\Box(\sigma \rightarrow \alpha_{\mathrm{t}}) \rightarrow \Box(\sigma \rightarrow \Box^{r_i - 1}\Box\bot)\big) \quad \text{and}$$

$$\vdash L_{\mathrm{t}} = e_i \rightarrow \big(\Box(\sigma \rightarrow . \, \alpha_{\mathrm{t}} \vee \beta_{\mathrm{t}}) \rightarrow \Box(\sigma \rightarrow \Box^{s_i}\Box\bot)\big)$$

for any box sentence $\sigma$. Again, since these two facts are (proven) very similar(ly), we only do the second one.

Fix $\sigma \equiv \Box\varphi$ and reason in T:

Assume $L_{\mathrm{t}} = e_i$ and $\Box(\Box\varphi \rightarrow . \, \alpha_{\mathrm{t}} \vee \beta_{\mathrm{t}})$, so that $\Box_y(\Box\varphi \rightarrow L_{\mathrm{t}} \neq c_{s_i+1})$ for some $y$ s.t. $H_{\mathrm{t}}(y) = e_i$ and step inside $\Box$:

Assume $\Box\varphi$ and suppose $L_{\mathrm{t}} = a_{s_i+1}$ or $L_{\mathrm{t}} = b_{s_i+1}$. Then Case E of Definition 3.6 will, on encountering a proof of $\varphi$, bring $H_{\mathrm{t}}$ to $c_{s_i+1}$, contradicting both $L_{\mathrm{t}}{=}a_{s_i+1}$ and $L_{\mathrm{t}}{=}b_{s_i+1}$. Therefore $L_{\mathrm{t}} \neq a_{s_i+1}$ and $L_{\mathrm{t}} \neq b_{s_i+1}$.

Formalizing this, we get $\Box(\Box\varphi \rightarrow . \, L_{\mathrm{t}}{\neq}a_{s_i+1} \wedge L_{\mathrm{t}}{\neq}b_{s_i+1})$ implying $\Box(L_{\mathrm{t}}{\neq}a_{s_i+1} \wedge L_{\mathrm{t}}{\neq}b_{s_i+1})$ since $\varphi$ is, by assumption, provable. By Lemma 3.8(c), this results in $L_{\mathrm{t}} \neq a_k$ and $L_{\mathrm{t}} \neq b_k$ for all $k > s_i + 1$ and, taking into account the earlier argument, also for all $k > s_i$.

Since $\Box\varphi$ implies $\alpha_{\mathrm{t}} \vee \beta_{\mathrm{t}}$, one has $\exists k \, (L_{\mathrm{t}}{=}a_k \vee L_{\mathrm{t}}{=}b_k)$, whence, by the above, $\exists k{\leq}s_i \, (L_{\mathrm{t}}{=}a_k \vee L_{\mathrm{t}}{=}b_k)$. By Lemma 3.9(a) this entails $\Box^{s_i}\Box\bot$.

So, $\Box(\sigma \rightarrow \Box^{s_i}\Box\bot)$.

Thus, $\vdash L_{\mathrm{t}} = e_i \rightarrow (\Box(\sigma \rightarrow . \, \alpha_{\mathrm{t}} \vee \beta_{\mathrm{t}}) \rightarrow \Box(\sigma \rightarrow \Box^{s_i}\Box\bot))$ is established.

The proof of the Lemma is now complete. ∎

3.13. LEMMA. *If $\mu$ is an irrefutable sentence s.t. $\vdash \mu \rightarrow \varepsilon_{\mathrm{t}}$ and $R(\alpha_{\mathrm{t}}, \beta_{\mathrm{t}}; \mu, \tau, \sigma)$ holds for some top-box sentences $\tau$ and $\sigma$, then there exists an $i \in \tilde{\omega}$ s.t. $\mu$ is consistent over* T *with $L_{\mathrm{t}}{=}e_i$.*

PROOF. If **t** is mortal then for $K = \max\tilde{\omega}$ we clearly have $\vdash \forall i \, (L_{\mathrm{t}}{=}e_i \rightarrow i{\leq}K)$ since $H_{\mathrm{t}}$ can only get to $e_i$ if $i \in \tilde{\omega}$ and, therefore, since $\vdash \mu \rightarrow \varepsilon_{\mathrm{t}}$, one has $\vdash \mu \rightarrow \exists i{\leq}K \; L_{\mathrm{t}}{=}e_i$.

If **t** is immortal then $\sup_{k \in \tilde{\omega}} r_k - 1 = \omega$ and hence, by Lemma 1.3, there is a $k \in \tilde{\omega}$ s.t. $\vdash \tau \rightarrow \Box^{r_k - 1}\Box\bot$. Note that from $R(\alpha_{\mathrm{t}}, \beta_{\mathrm{t}}; \mu, \tau, \sigma)$ we have $S(\alpha_{\mathrm{t}}; \mu, \tau)$ which implies $\vdash \mu \rightarrow (\Box(\Box^{r_k}\Box\bot \rightarrow \alpha_{\mathrm{t}}) \rightarrow \Box(\Box^{r_k}\Box\bot \rightarrow \tau))$. Reason in T:

Assume $\mu$ and $\exists i{>}k \; L_{\mathrm{t}}{=}e_i$. From Lemma 3.12(a) we have $\Box(\Box^{r_i - 1}\Box\bot \rightarrow \alpha_{\mathrm{t}})$ and hence, by $\Sigma$-completeness, $\Box(\Box^{r_k}\Box\bot \rightarrow \alpha_{\mathrm{t}})$, for $k < i$. By $\mu$ we have $\Box(\Box^{r_k}\Box\bot \rightarrow \tau)$. By our assumption on $\tau$, there holds $\Box(\tau \rightarrow \Box^{r_k - 1}\Box\bot)$ and so $\Box(\Box^{r_k}\Box\bot \rightarrow \Box^{r_k - 1}\Box\bot)$ which, by Löb's Theorem, gives $\Box^{r_k}\Box\bot$, contradicting Lemma 3.10(c).

Therefore, $\vdash \mu {\rightarrow} \exists i{\leq}k \; L_{\mathrm{t}}{=}e_i$ holds for an appropriate $k \in \tilde{\omega}$ regardless of the lifespan of **t**. Since $\mu$ is irrefutable, it should be consistent with $L_{\mathrm{t}}{=}e_i$ for some $i \leq k$, q.e.d. ∎

Next we fix the last remaining parameters in our construction.

3.14. DEFINITION. For each letter $\ell$ of $\Lambda$ we define

$$\lambda_\ell \;\equiv\; \exists x {\in} \tilde{\omega} {-} \{0\} \, (\Downarrow^x \Box\bot \wedge t_x = \ell).$$

Thus an element $x$ of $\mathcal{D}^*_{\mathrm{T}}$ of finite depth $i$ is in $\lambda_\ell{}^*$ if and only if $t_i = \ell$.

$\vec{\lambda}_{\mathrm{t}}$ will henceforth stand for $\lambda_{\ell_1}, \ldots, \lambda_{\ell_N}$, where $\ell_1, \ldots, \ell_N$ is the tuple listing $\Lambda$.

Lemmas 3.15–17 establish that the parameters $\alpha_t, \beta_t, \varepsilon_t, \vec{\lambda}_t$ satisfy $D_t(\cdots)$, that is, they satisfactorily code the computation executed by $\mathbf{t}$.

3.15. LEMMA. *For all* $\ell \in \Lambda$, $\vdash \forall x {\in} \bar{\omega}{-}\{0\}\,(\Downarrow^x \Box\bot \to (\lambda_\ell \leftrightarrow t_x{=}\ell))$.

PROOF. Immediate from the definition of $\lambda_\ell$ and Lemma 2.2(a). ∎

3.16. LEMMA. (a) *For all* $i \in \tilde{\omega}$ *and all sentences* $\mu$ *there holds* $W^{\boldsymbol{t}_i}(\vec{\lambda}_t; \mu, \Box^{r_i-1}\Box\bot)$.

(b) *If* $\nvdash \neg\mu$, $\vdash \mu \to \Diamond^{s_i+1}\Diamond\top$ *for all* $i \in \tilde{\omega}$, $m \in \bar{\omega}$ *and* $W^{\boldsymbol{e}}(\vec{\lambda}_t; \mu, \Box^m\Box\bot)$, *then* $m + |\boldsymbol{e}| \in \bar{\omega}$.

(c) *If* $\nvdash \neg\mu$, $\vdash \mu \to \Diamond^{s_i+1}\Diamond\top$ *for all* $i \in \tilde{\omega}$, $m \in \bar{\omega}$, $W^{\boldsymbol{e}}(\vec{\lambda}_t; \mu, \Box^m\Box\bot)$ *and* $W^{\boldsymbol{f}}(\vec{\lambda}_t; \mu, \Box^m\Box\bot)$, *then the words* $\boldsymbol{e}$ *and* $\boldsymbol{f}$ *are compatible.*

PROOF. (a). By Lemma 3.15 we have $\vdash \forall x {\in} \bar{\omega}{-}\{0\}\,(\Downarrow^x\Box\bot \to \lambda_{t_x})$. In particular, $\vdash \Downarrow^k\Box\bot \to \lambda_{t_k}$ whenever $r_i \le k \le s_i$. Since $\boldsymbol{t}_i = t_{r_i}\cdots t_{s_i}$, Lemma 2.2(c) gives $\vdash \Downarrow^j\Box^{r_i-1}\Box\bot \to \lambda_{(\boldsymbol{t}_i)_j}$ for all $j$ s.t. $1 \le j \le |\boldsymbol{t}_i|$, which implies $\vdash \mu \to \Box(\Downarrow^j\Box^{r_i-1}\Box\bot \to \lambda_{(\boldsymbol{t}_i)_j})$, or, in other words, $W^{\boldsymbol{t}_i}(\vec{\lambda}_t; \mu, \Box^{r_i-1}\Box\bot)$, q.e.d.

(b). We can clearly assume that $\boldsymbol{e}$ is non-null. Suppose $m + |\boldsymbol{e}| \notin \bar{\omega}$. The membership of a natural number in $\bar{\omega}$ can only fail if $\mathbf{t}$ is mortal so that $\bar{\omega} < \omega$. Recall that then $\max\bar{\omega} = s_K$, where $K = \max\tilde{\omega}$. Thus one has $\vdash s_K{+}1 \notin \bar{\omega}$ which entails $\vdash \Downarrow^{s_K+1}\Box\bot \to \forall x{\in}\bar{\omega}\,\neg\Downarrow^x\Box\bot$ by Lemma 2.2(a), hence $\vdash \Downarrow^{s_K+1}\Box\bot \to \neg\bigvee\!\!\!\bigvee_{\ell\in\Lambda}\lambda_\ell$ and $\vdash \mu \to \Box(\Downarrow^{s_K+1}\Box\bot \to \neg\bigvee\!\!\!\bigvee_{\ell\in\Lambda}\lambda_\ell)$. Now, since $m + |\boldsymbol{e}| \ge s_K + 1$ and $m \le s_K$ by assumption, we can find an $n$ s.t. $1 \le n \le |\boldsymbol{e}|$ and $m + n = s_K + 1$. From $W^{\boldsymbol{e}}(\vec{\lambda}_t; \mu, \Box^m\Box\bot)$ we have $\vdash \mu \to \Box(\Downarrow^n\Box^m\Box\bot \to \lambda_{(\boldsymbol{e})_n})$. By Lemma 2.2(c), this rewrites as $\vdash \mu \to \Box(\Downarrow^{s_K+1}\Box\bot \to \lambda_{(\boldsymbol{e})_n})$. Together with the preceding argument, this gives $\vdash \mu \to \Box\neg\Downarrow^{s_K+1}\Box\bot$ implying, by Lemma 2.2(b), $\vdash \mu \to \Box^{s_K+1}\Box\bot$, which contradicts our assumptions on $\mu$. Therefore, $m + |\boldsymbol{e}| \in \bar{\omega}$.

(c). By (b), we have that $m + |\boldsymbol{e}|, m + |\boldsymbol{f}| \in \bar{\omega}$. If $\boldsymbol{e}$ and $\boldsymbol{f}$ were not compatible then there would exist an $i$, $1 \le i \le \min\{|\boldsymbol{e}|, |\boldsymbol{f}|\}$ s.t. $(\boldsymbol{e})_i \ne (\boldsymbol{f})_i$. Note that $m + i \in \bar{\omega}$. Now, $W^{\boldsymbol{e}}(\vec{\lambda}_t; \mu, \Box^m\Box\bot)$ and $W^{\boldsymbol{f}}(\vec{\lambda}_t; \mu, \Box^m\Box\bot)$ imply that $\vdash \mu \to \Box(\Downarrow^i\Box^m\Box\bot \to . \lambda_{(\boldsymbol{e})_i} \wedge \lambda_{(\boldsymbol{f})_i})$ which, by Lemma 3.15 and since $(\boldsymbol{e})_i \ne (\boldsymbol{f})_i$, implies $\vdash \mu \to \Box\neg\Downarrow^i\Box^m\Box\bot$ whence, by Lemma 2.2(b), $\vdash \mu \to \Box^{m+i}\Box\bot$. Since $m + i \in \bar{\omega}$ and so $m + i < s_k + 1$ for an appropriate $k \in \tilde{\omega}$, we get $\vdash \mu \to \Box^{s_k+1}\Box\bot$ contrary to the assumption on $\mu$. Thus $\boldsymbol{e}$ and $\boldsymbol{f}$ have to be compatible. ∎

3.17. LEMMA. $D_t(\alpha_t, \beta_t, \varepsilon_t, \vec{\lambda}_t)$ *holds. Moreover, if* $\mathbf{t}$ *is mortal and* $K = \max\tilde{\omega}$, *then this formalizes in* T *under the hypothesis* $\Diamond^{s_K+2}\Diamond\top$.

PROOF. First we have to establish $X^{\boldsymbol{t}_0}(\alpha_t, \beta_t, \varepsilon_t, \vec{\lambda}_t; \Box\bot)$. Consider the sentence $\mu \equiv L_t{=}e_0$. By Lemma 3.10(b), $\mu$ is irrefutable in T and, since $r_0 - 1 = 0$ and $s_0 = |\boldsymbol{t}_0|$, Lemma 3.12(b) gives $R(\alpha_t, \beta_t; \mu, \Box\bot, \Box^{|\boldsymbol{t}_0|}\Box\bot)$. Further, we clearly have $\vdash \mu \to \varepsilon_t$ and $W^{\boldsymbol{t}_0}(\vec{\lambda}_t; \mu, \Box\bot)$ by Lemma 3.16(a). Thus $X^{\boldsymbol{t}_0}(\alpha_t, \beta_t, \varepsilon_t, \vec{\lambda}_t; \Box\bot)$ holds.

Second, we check $P^{\boldsymbol{g}\$\to\$\boldsymbol{h}}(\alpha_t, \beta_t, \varepsilon_t, \vec{\lambda}_t)$ for an arbitrary production $\boldsymbol{g}\$ \to \$\boldsymbol{h}$ of $\mathbf{t}$. Suppose $\mu$ is an irrefutable sentence s.t. $\vdash \mu \to \varepsilon_t$, $R(\alpha_t, \beta_t; \mu, \tau, \sigma)$ for some top-box $\tau$

and $\sigma$, $W^{g}(\vec{\lambda}_t;\mu,\tau)$ and $\vdash \mu \rightarrow \rho$ for some $\rho \in B$. By Lemma 3.13, $\mu$ is consistent with $L_t{=}e_i$ for some $i \in \tilde{\omega}$. Let $\kappa \equiv \mu \wedge L_t{=}e_i$. By virtue of Lemmas 3.2(a) and 3.3(a), $\kappa$ satisfies all properties that have been assumed of $\mu$. Note that by Lemma 3.10(c), $\vdash \kappa \rightarrow \Diamond^{s_i+1}\Diamond\top$. Further, by Lemmas 3.12(b), 3.16(a), 3.2(a) and 3.3(a) we have

(1) $R(\alpha_t,\beta_t;\kappa,\Box^{r_i-1}\Box\bot,\Box^{s_i}\Box\bot)$ and

(2) $W^{t_i}(\vec{\lambda}_t;\kappa,\Box^{r_i-1}\Box\bot)$.

From (1) it follows by Lemma 3.2(d) that

$$\vdash \kappa \rightarrow . \Box(\tau \leftrightarrow \Box^{r_i-1}\Box\bot) \wedge \Box(\sigma \leftrightarrow \Box^{s_i}\Box\bot).$$

Hence, by Lemma 3.3(b), one has

(3) $W^{g}(\vec{\lambda}_t;\kappa,\Box^{r_i-1}\Box\bot)$.

Lemma 3.16(c) infers from (2) and (3) that $\boldsymbol{t}_i$ must be compatible with $\boldsymbol{g}$. Obviously, we can not have $|\boldsymbol{t}_i| < |\boldsymbol{g}|$ for then, by the monogeneity condition, no production of $\mathbf{t}$ applies to $\boldsymbol{t}_i$ and hence $s_i = \max\tilde{\omega}$. By Lemma 3.16(b) we could not then have (3). Thus $|\boldsymbol{g}| \leq |\boldsymbol{t}_i|$ and, therefore, $\boldsymbol{g}\$ \rightarrow \$\boldsymbol{h}$ is applicable to $\boldsymbol{t}_i$ and hence is the production responsible for the transition from $\boldsymbol{t}_i$ to $\boldsymbol{t}_{i+1}$ so that $i+1 \in \tilde{\omega}$. Since $\kappa$ is irrefutable, $\vdash \kappa {\rightarrow} .\rho \wedge \Box(\tau \leftrightarrow \Box^{r_i-1}\Box\bot) \wedge \Box(\sigma \leftrightarrow \Box^{s_i}\Box\bot)$ and $\vdash \kappa \rightarrow L_t{=}e_i$, $\kappa$ testifies to the consistency of $L_t{=}e_i$ with $\rho \wedge \Box(\cdots) \wedge \Box(\cdots)$. Now, the latter, being a conjunction of three such, is a box sentence. Therefore, by Lemma 3.11, it is consistent with $L_t{=}e_{i+1}$. In other words, there exists an irrefutable sentence $\nu$ s.t. $\vdash \nu \rightarrow L_t{=}e_{i+1}$ (note that this implies $\vdash \nu \rightarrow \varepsilon_t$), $\vdash \nu \rightarrow \rho$ and

(4) $\vdash \nu \rightarrow . \Box(\tau \leftrightarrow \Box^{r_i-1}\Box\bot) \wedge \Box(\sigma \leftrightarrow \Box^{s_i}\Box\bot)$,

whence by $\Sigma$-completeness one has

(5) $\vdash \nu \rightarrow . \Box(\Box^{|\boldsymbol{g}|}\tau \leftrightarrow \Box^{|\boldsymbol{g}|}\Box^{r_i-1}\Box\bot) \wedge \Box(\Box^{|\boldsymbol{h}|}\sigma \leftrightarrow \Box^{|\boldsymbol{h}|}\Box^{s_i}\Box\bot)$.

Next, by Lemmas 3.12(b), 3.16(a), 3.2(a) and 3.3(a),

(6) $R(\alpha_t,\beta_t;\nu,\Box^{r_{i+1}-1}\Box\bot,\Box^{s_{i+1}}\Box\bot)$ and

(7) $W^{t_{i+1}}(\vec{\lambda}_t;\nu,\Box^{r_{i+1}-1}\Box\bot)$.

Since $|\boldsymbol{g}| + r_i - 1 = r_{i+1} - 1$ and $|\boldsymbol{h}| + s_i = s_{i+1}$, Lemma 3.2(b) and (c) apply to (5) and (6) to produce

$$R(\alpha_t,\beta_t;\nu,\Box^{r_{i+1}-1}\Box\bot,\Box^{|\boldsymbol{g}|}\tau,\Box^{|\boldsymbol{h}|}\sigma).$$

Furthermore, observe that since $\boldsymbol{t}_{i+1} = \boldsymbol{fh}$, (7) implies $W^{h}(\vec{\lambda}_t;\nu,\Box^{|\boldsymbol{f}|}\Box^{r_{i+1}-1}\Box\bot)$ through Lemma 3.3(c). Now, recall that $|\boldsymbol{f}| + r_{i+1} - 1 = |\boldsymbol{g}| + |\boldsymbol{f}| + r_i - 1 = |\boldsymbol{t}_i| + r_i - 1 = s_i$ and, therefore, $W^{h}(\vec{\lambda}_t;\nu,\Box^{s_i}\Box\bot)$. Finally, $W^{h}(\vec{\lambda}_t;\nu,\sigma)$ follows from (4) by Lemma 3.3(b).

We have thus inferred that the sentence $\nu$ satisfies all the requirements needed to verify $P^{g\$\rightarrow\$h}(\alpha_t,\beta_t,\varepsilon_t,\vec{\lambda}_t)$.

The only point in our proof that does not formalize in T is our appeal to Lemma 3.10(b) in the very beginning of the proof. Under the hypothesis $\Diamond^{s_K+2}\Diamond\top$ for $\mathbf{t}$ mortal and $K = \max\tilde{\omega}$, this can however be remedied within T by Lemmas 3.10(d) and 3.8(e). ∎

We approach the denouement of this Section's story, Proposition 3.18. It also points out that our considerations on $\mathcal{D}_{\mathrm{T}}$ have not been completely lost on the formal theory T itself.

3.18. Proposition. (a) *If* $\mathbf{t}$ *produces* $e$ *then* $M_{\mathbf{t}}^{e}$ *holds in* $\mathcal{D}_{\mathrm{T}}$. *Moreover,* $\vdash M_{\mathbf{t}}^{e}$.

(b) *If* $\mathbf{t}$ *does not produce* $e$ *then* $M_{\mathbf{t}}^{e}$ *does not hold in* $\mathcal{D}_{\mathrm{T}}$. *If, in addition,* $\mathbf{t}$ *is mortal, then* $\nvdash M_{\mathbf{t}}^{e}$.

Proof. (a) follows from Lemma 3.5: Since $M_{\mathbf{t}}^{e}$ holds in any diagonalizable algebra and T verifies that $\Box$ satisfies all the axioms of the diagonalizable algebra theory (cf. Montagna [18]), $M_{\mathbf{t}}^{e}$ is provable in T.

(b). Suppose $M_{\mathbf{t}}^{e}$ holds, that is, one has $D_{\mathbf{t}}(\alpha,\beta,\varepsilon,\vec{\lambda}) \Rightarrow O^{e}(\alpha,\beta,\varepsilon,\vec{\lambda})$ for any choice of parameters. By Lemma 3.17, we have $D_{\mathbf{t}}(\alpha_{\mathbf{t}},\beta_{\mathbf{t}},\varepsilon_{\mathbf{t}},\vec{\lambda}_{\mathbf{t}})$ and, therefore, $O^{e}(\alpha_{\mathbf{t}},\beta_{\mathbf{t}},\varepsilon_{\mathbf{t}},\vec{\lambda}_{\mathbf{t}})$ must hold.

This means that there exists a top-box sentence $\tau$ and an irrefutable $\mu$ s.t.

$$\vdash \mu \rightarrow \varepsilon_{\mathbf{t}} \;\&\; R(\alpha_{\mathbf{t}},\beta_{\mathbf{t}};\mu,\tau,\Box^{|e|}\tau) \;\&\; W^{e}(\vec{\lambda}_{\mathbf{t}};\mu,\tau).$$

By Lemmas 3.13, 3.2(a) and 3.3(a), we may assume $\vdash \mu \rightarrow L_{\mathbf{t}}{=}e_i$ for a certain $i \in \tilde{\omega}$. By Lemmas 3.12(b) and 3.2(d) we then have $\vdash \mu \rightarrow. \Box(\tau \leftrightarrow \Box^{r_i-1}\Box\bot) \wedge \Box(\Box^{|e|}\tau \leftrightarrow \Box^{s_i}\Box\bot)$ whence $W^{e}(\vec{\lambda}_{\mathbf{t}};\mu,\Box^{r_i-1}\Box\bot)$ follows by Lemma 3.3(c). Combining this with $W^{\boldsymbol{t}_i}(\vec{\lambda}_{\mathbf{t}};\mu,\Box^{r_i-1}\Box\bot)$, which we may count on by Lemma 3.16(a), we get that $e$ is compatible with $\boldsymbol{t}_i$ by virtue of Lemma 3.16(c). Now note that we have

$$\begin{aligned}\vdash \mu \rightarrow \Box(\Box^{|e|}\Box^{r_i-1}\Box\bot &\leftrightarrow \Box^{|e|}\tau)\\ &\leftrightarrow \Box^{s_i}\Box\bot)\\ &\leftrightarrow \Box^{|\boldsymbol{t}_i|}\Box^{r_i-1}\Box\bot)\end{aligned}$$

for $s_i = |\boldsymbol{t}_i| + r_i - 1$. If $|e|$ failed to be equal to $|\boldsymbol{t}_i|$ then we would have

$$\begin{aligned}&\vdash \mu \rightarrow \Box(\Box^{\max\{|e|,|\boldsymbol{t}_i|\}}\Box^{r_i-1}\Box\bot \rightarrow \Box^{\min\{|e|,|\boldsymbol{t}_i|\}}\Box^{r_i-1}\Box\bot)\\ &\qquad \rightarrow \Box^{\min\{|e|,|\boldsymbol{t}_i|\}+1}\Box^{r_i-1}\Box\bot \qquad \text{(by ($\Sigma$-completeness and) Löb's Theorem).}\end{aligned}$$

Now, $\min\{|e|,|\boldsymbol{t}_i|\} + 1 + r_i - 1 \leq \max\{|e|,|\boldsymbol{t}_i|\} + r_i - 1 \in \tilde{\omega}$ by Lemma 3.16(b), whereas $\vdash \mu \rightarrow \varepsilon_{\mathbf{t}}$ and $\vdash \varepsilon_{\mathbf{t}} \rightarrow \Diamond^{j}\Diamond\top$ for any $j \in \tilde{\omega}$ by Lemma 3.10(d), so that $\vdash \mu \rightarrow \Diamond^{\min\{|e|,|\boldsymbol{t}_i|\}+1}\Diamond^{r_i-1}\Diamond\bot$, which is a contradiction. This shows that $|e| = |\boldsymbol{t}_i|$ and, since these words are compatible, $e = \boldsymbol{t}_i$ so that $e$ is actually produced by $\mathbf{t}$ as was to be shown.

If $\mathbf{t}$ is mortal then the above argument can be formalized in T, for then we only need finitely many instances of $\vdash \varepsilon_{\mathbf{t}} \rightarrow \Diamond^{j}\Diamond\top$. Thus we have $\vdash O^{e}(\alpha_{\mathbf{t}},\beta_{\mathbf{t}},\varepsilon_{\mathbf{t}},\vec{\lambda}_{\mathbf{t}}) \rightarrow \bigvee_{i\in\tilde{\omega}} e{=}\boldsymbol{t}_i$. Therefore, in the case that $e$ is not produced by $\mathbf{t}$, we get $\vdash \neg O^{e}(\alpha_{\mathbf{t}},\beta_{\mathbf{t}},\varepsilon_{\mathbf{t}},\vec{\lambda}_{\mathbf{t}})$, for the absence of a word in a finite monologue is verifiable in T. By Lemma 3.17, however, there holds $\vdash \Diamond^n \Diamond \top \rightarrow D_t(\alpha_t, \beta_t, \varepsilon_t, \vec{\lambda}_t)$ for an appropriate $n \in \omega$. Thus $\vdash M^e_t \rightarrow \Box^n \Box \bot$ and we can not have $\vdash M^e_t$ by the assumption on the credibility extent of T, q.e.d. ∎

## 4. Wordworks

In Proposition 3.18 we have seen that the question whether or not an arbitrary mcsystem produces a given word can be reformulated in terms of validity of first order statements in diagonalizable algebras of theories T of infinite credibility extent. Starting from this we shall diagnose the undecidability of $\mathrm{Th}\,\mathcal{D}_\mathrm{T}$, for mcsystems are known to be a universal breeding ground for undecidability phenomena in that they can, in a certain sense, simulate any computational process, as was first established by Minsky [15]. A particularly sharp version of this result is given in Cocke & Minsky [6] complete with a remarkably transparent proof reproduced in Minsky's book [16]. Unfortunately, we are not able to benefit from every aspect of this accomplishment, and we therefore only state their potent theorem in a rather general and simplified form.

4.1. FACT (Minsky [15, §2]; Cocke & Minsky [6]; see also Minsky [16, Theorem and Corollary 14.6–1]). *To every deterministic (say, turingmachine) computation $C$ we can effectively associate a mcsystem $\mathbf{t}_C$ in an alphabet $\Lambda_C$, a one-one total recursive function $f_C$ from the configuration space of $C$ to words in $\Lambda_C$, and its effective 'inverse' $g_C$ which, given a word $e$ in $\Lambda_C$, tells whether there exists a configuration $s$ with $f_C(s) = e$ and, if so, finds this $s$. These objects are related to $C$ in the following way:*

(i) *$C$ terminates iff $\mathbf{t}_C$ is mortal, and*

(ii) *$C$ reaches a configuration $s$ iff $\mathbf{t}_C$ produces $f_C(s)$.* ∎

In order to utilize Proposition 3.18 and Fact 4.1 for obtaining the promised undecidability results, we have to make precise agreements on the nature of computational devices we shall be dealing with. We opt for the usual Turing machines with a number of ridiculous restrictions on the kind of configurations accepted as legitimate output.

The class of *good old* Turing machines consists of the usual single two-way infinite tape Turing machines with two distinct distinguished *starting* and *halting* states. Good old Turing machines present the results of their computations by arriving at their halting state with the read/write head positioned at the leftmost non-blank square which marks off the beginning of the answer. That the imposition of this particular format is harmless can be gleaned from Minsky [16, chapter 6]. The point is that in each good old Turing machine to every natural number there corresponds a unique configuration that is considered to output that natural number. We give the name $(\varphi_i)_{i\in\omega}$ to the numbering of the class of unary recursive functions by (gödelnumbers of) good old Turing machines.

4.2. THEOREM. *The first order theory of any class of diagonalizable algebras containing that of a theory of infinite credibility extent is undecidable.*

PROOF. Fix a nonrecursive r.e. set $U$ and let $(u_n)_{n\in\omega}$ be its effective repetition-free enumeration. Define a recursive function $h$ by putting

$$h(x) \simeq \begin{cases} u_{i+1} & \text{if } x = u_i, \\ \uparrow & \text{otherwise.} \end{cases}$$

We now design the following Turing machine $k$: Take the good old Turing machine computing $h$ and identify its starting and halting states under the name of *enumerating* state, then feed it $u_0$ as input. (Note that $k$ is not good old.)

Observe that $k$ just iterates $h$ and therefore goes on and on enumerating $U$ with transition through the enumerating state indicating the finding of the next element of $U$. Therefore, the question "Does $k$ reach the enumerating state with the number $z$ written on its tape?" is equivalent to "$z \in ?\ U$".

By Fact 4.1, find a mcsystem $\mathbf{t}$ simulating the behaviour of $k$ and call "$z \in U$" the word corresponding to $k$'s finding itself in the enumerating state and $z$ written on $k$'s tape. Then, clearly, $\mathbf{t}$ produces "$z \in U$" iff $z \in U$.

Recalling Propositions 3.5 and 3.18(b), we see that if $z \in U$ then the sentence $M_{\mathbf{t}}^{\text{“}z \in U\text{”}}$ holds in any diagonalizable algebra and $M_{\mathbf{t}}^{\text{“}z \in U\text{”}}$ does not hold in the diagonalizable algebra of any theory of infinite credibility extent unless $z \in U$. Since $U$ is not recursive, the proof is complete. ∎

4.3. COROLLARY. *For* T *a theory of infinite credibility extent,* Th $\mathcal{D}_{\mathrm{T}}$ *is hereditarily undecidable.*

PROOF. Recall that the theory of all diagonalizable algebras is finitely axiomatized and use Theorem 4.2. ∎

Now that we know Th $\mathcal{D}_{\mathrm{T}}$ to be undecidable for certain theories T and recalling that first order statements about $\mathcal{D}_{\mathrm{T}}$ are straightforwardly formalizable in T itself (cf. Montagna [18] or note that in Sections 1 and 3 we have been using such formalizations all along), one might ask whether this is the case with these theories as seen by T, that is, whether or not the theory

$$\mathrm{Th}^{\mathrm{T}}\mathcal{D}_{\mathrm{T}} = \{\,\text{diagonalizable algebraic sentences } Z \mid \mathrm{T} \vdash \text{‘}\mathcal{D}_{\mathrm{T}} \models Z\text{’}\,\}$$

is decidable. For T = PA, Montagna [18] does so. The following Theorem gives an answer for all formal theories T of infinite credibility extent, which, for T $\neq$ PA, is generally not a straightforward consequence of Corollary 4.3 because T may very well happen to prove statements $Z$ about $\mathcal{D}_{\mathrm{T}}$ that are not, in the real world, valid in this structure. One example of such $Z$ is the diagonalizable algebraic sentence $B = T$, with $B$ and $T$ as in Definition 1.1.

4.4. THEOREM. *For any theory* T *of infinite credibility extent,* $\mathrm{Th}^{\mathrm{T}}\mathcal{D}_{\mathrm{T}}$ *is hereditarily undecidable.*

PROOF. We shall do something similar to the proof of Theorem 4.2.

As in that proof, we fix a nonrecursive r.e. set $U$ and an effective repetition-free enumeration $(u_n)_{n\in\omega}$ of it and agree additionally that $0 \notin U$.

Let us now devise a computable numbering of a certain class of computations $(\psi_i)_{i\in\omega}$ by Turing machines with prescribed input like the one we employed in the proof of Theorem 4.2. The whole of the good old Turing machine $k$ computing $\varphi_k$ along with its states and instructions is subsumed within the Turing machine executing the computation $\psi_k$. Within the new machine, the starting state of $k$ bears the name of *enumerating* state and the halting state of $k$ that of *contemplating* state. On top of that, there is at least one new state called the *stop* state. The new machines start off in the contemplating and finish, if so luck has it, in the stop state.

This is how $\psi_k$ works: Contemplate the number written on the tape. If it is 0, then pass on to the stop state and finish your activities. Otherwise, assume the enumerating state and run as $k$ does. On reaching the halting state of $k$, which now is the contemplating state, start all over again.

Essentially, $\psi_k$ performs the iteration of $\varphi_k$ starting with $u_0$ as long as 0 does not crop up among the iterated values, in which case it halts.

By Fact 4.1, to every $k$ we can effectively associate a mcsystem $\mathbf{t}_k$ mimicking the computation $\psi_k$ in the sense of that Fact. "$z \in U$" will stand for the word, effective in $z$, corresponding to the enumerating state of our machine for computing $\psi_k$ and $z$ on the tape.

Invoking the Recursion Theorem, we call into existence a particular index $k$ s.t.

$$\varphi_k(x) \simeq \begin{cases} u_{i+1} & \text{if } x = u_i, \text{ and whenever } \vdash_{\leq i} M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}} \text{ one has } z \in \{u_0, \ldots, u_i\}, \\ 0 & \text{if } x = u_i, \text{ and } \vdash_{\leq i} M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}} \text{ for some } z \notin \{u_0, \ldots, u_i\}, \\ \uparrow & \text{otherwise.} \end{cases}$$

C l a i m . *For all $z$ and $i$, $\vdash_{\leq i} M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}}$ implies $z \in \{u_0, \ldots, u_i\}$.*

Suppose this were not the case: $\vdash_{\leq i} M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}}$ and $z \notin \{u_0, \ldots, u_i\}$ and $i$ is the minimal s.t. this happens. Consider the computation $\psi_k$: Since $u_j \neq 0$, we have $\varphi_k(u_j) = u_{j+1}$ for all $j < i$ and, by our assumption, $\varphi_k(u_i) = 0$. By the minimality condition on $i$ this means that $\psi_k$ passes through the enumerating state exactly $i+1$ times with $u_0, \ldots, u_i$ showing up then on the tape, whereafter it proceeds to the stop state and grinds to a halt. By Fact 4.1, $\mathbf{t}_k$ is then mortal. By Proposition 3.18(b) we therefore have $\nvdash M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}}$ for, by our assumptions, $z \notin \{u_0, \ldots, u_i\}$ and hence the word "$z \in U$" is not produced by $\mathbf{t}_k$. But we have assumed $\vdash M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}}$.

The contradiction settles the Claim.

We are now adequately equipped to see that $\vdash M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}}$ if and only if $z \in U$. One direction is an immediate consequence of the Claim. For the opposite direction note that since, as follows from the Claim, $\varphi_k(u_i) = u_{i+1}$ for all $i \in \omega$, we have that $\psi_k$ eventually enumerates all elements of $U$, whence by Fact 4.1 we have that $z \in U$ implies that $\mathbf{t}_k$ produces "$z \in U$" which, by Proposition 3.18(a), implies in its turn $\vdash M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}}$. This shows that $\mathrm{Th}^{\mathrm{T}}\mathcal{D}_{\mathrm{T}}$ is undecidable.

The heredity of this undecidability follows, as in Theorem 4.2 and Corollary 4.3 by recalling that, by Proposition 3.5, $\mathbf{t}_k$'s producing "$z \in U$" implies that $M_{\mathbf{t}_k}^{\text{“}z \in U\text{”}}$ holds in any diagonalizable algebra, and that the diagonalizable algebra theory is finitely axiomatized. ∎

A bookkeeper's analysis of our constructions, which the reader is invited to follow, shows that the undecidability of Corollary 4.3 (as well as that of Theorem 4.4) already strikes at the $\vec{\forall}\vec{\exists}\vec{\forall}\vec{\exists}$ level of quantifier alternation in diagonalizable algebraic sentences. Indeed, the sets $B$ and $T$ of Definition 1.1 are $\exists$ and $\forall\exists$ definable respectively; the formula $W^{e}(\cdots)$ of Definition 3.1 is quantifier-free; formulas $S(\cdots)$ and $R(\cdots)$ of Definitions 1.4 and 3.1 are $\vec{\exists}\&\vec{\forall}$. In Definition 3.4 we have that $O^{e}(\cdots)$ is $\vec{\exists}\vec{\forall}\vec{\exists}$, and $X^{e}(\cdots)$ is $\vec{\exists}\vec{\forall}$, $P^{g\$\to\$h}(\cdots)$ and $D_{t}(\cdots)$ are $\vec{\forall}\vec{\exists}\vec{\forall}$. (The most complex aspect in $O^{e}(\cdots)$ and $P^{g\$\to\$h}(\cdots)$ is the restriction of certain quantifiers to $T$.) Finally, the complexity of sentences $M^{e}_{t}$, whose validity was shown to be undecidable, is $\vec{\forall}\vec{\exists}\vec{\forall}\vec{\exists}$.

In the opposite direction, we have Corollary 3.12 of Smoryński [22] to Lemma 5.4 of Solovay [26] stating that, for $\Sigma_1$-sound theories T, $\mathrm{Th}_{\vec{\forall}}\mathcal{D}_{\mathrm{T}}$ is decidable. The situation for other kinds of theories is the same:

4.5. PROPOSITION. $\mathrm{Th}_{\vec{\forall}}\mathcal{D}_{\mathrm{T}}$ *is decidable for each theory* T.

PROOF-SKETCH (for readers of Shavrukov [20]). We only treat $\Sigma_1$-ill theories T and we rather look at $\mathrm{Th}_{\vec{\exists}}\mathcal{D}_{\mathrm{T}}$. A natural number $n+1$ is the *credibility extent* of T if $n$ is the minimal s.t. $\vdash \Box^{n}\Box\bot$ holds in $\mathcal{D}_{\mathrm{T}}$. The *height* of an arbitrary diagonalizable algebra is defined in exactly the same way.

After a few straightforward manipulations the question whether $\mathcal{D}_{\mathrm{T}} \models Z$ for $\vec{\exists}$ diagonalizable algebraic sentences boils down to those $Z$ of the form

$$\exists\vec{\xi}\,(\vdash \mathrm{P}(\vec{\xi}) \;\&\; \mathop{\&}_{i} \nvdash \mathrm{U}_i(\vec{\xi})),$$

where $\mathrm{P}(\vec{x})$ and $\mathrm{U}_i(\vec{x})$ are modal formulas. Consider the factor $\mathcal{P}$ of the free diagonalizable algebra on the generators $\vec{x}$ modulo the $\tau$-filter corresponding to the formula

$$\mathrm{P}^{*}(\vec{x}) = \begin{cases} \mathrm{P}(\vec{x}) & \text{if T is of infinite credibility extent,} \\ \mathrm{P}(\vec{x}) \wedge \Box^{n}\Box\bot & \text{if } n+1 \text{ is the credibility extent of T.} \end{cases}$$

Using Corollary 2.14 of Shavrukov [20] it is effectively verifiable whether the height of $\mathcal{P}$ matches the credibility extent of T, which signals embeddability of $\mathcal{P}$ into $\mathcal{D}_{\mathrm{T}}$ and is a necessary condition for the sentence $Z$ to hold in $\mathcal{D}_{\mathrm{T}}$. If this is indeed the case then, in order to make sure that $\mathcal{D}_{\mathrm{T}} \models Z$, one only has to check that

$$\mathbf{L} \vdash \Box^{+}\mathrm{P}^{*}(\vec{x}) \to \Box^{+}\mathrm{U}_i(\vec{x})$$

takes place for no $i$. ∎

All in all, this leaves a comfortably large gap for further investigations into exactly how many quantifier changes one needs to get undecidability.

4.6. CONJECTURE. $\vec{\forall}\vec{\exists}$ is decidable and $\vec{\forall}\vec{\exists}\vec{\forall}$ is not.

The reader will certainly have also noticed that the question of complexity of the first order theories of diagonalizable algebras of $\Sigma_1$-ill theories of infinite credibility extent is left wide open. In particular, it is not known to the author whether any or all of these theories are arithmetic.

Neither am I aware of any information whatsoever on the question of decidability of first order theories of diagonalizable algebras of formal theories of finite credibility extent apart from the trivial case of credibility extent 1.

## References


[1] M. A. Abashidze. O nekotorykh svoĭstvakh algebr Magari. *Logiko-semanticheskie issledovaniya.* Metsniereba, Tbilisi 1981, 111–127.

[2] S. N. Artemov. O modal'nykh logikakh, aksiomatiziruyushchikh dokazuemost'. *Izvestiya Akademii Nauk SSSR. Seriya matematicheskaya* 49 (1985) 1123–1154 (English translation: On modal logics axiomatizing provability. *Mathematics of the USSR — Izvestiya* 27 (1986) 401–429).

[3] S. N. Artemov & L. D. Beklemishev. On propositional quantifiers in provability logic. *Notre Dame Journal of Formal Logic*, to appear.

[4] L. D. Beklemishev. On bimodal provability logics for $\Pi_1$-axiomatized extensions of arithmetical theories. *Annals of Pure and Applied Logic*, to appear.

[5] A. Berarducci. The interpretability logic of Peano arithmetic. *The Journal of Symbolic Logic* 55 (1990) 1059–1089.

[6] J. Cocke & M. Minsky. Universality of Tag systems with $P = 2$. *Journal of the Association for Computing Machinery* 11 (1964) 15–20.

[7] G. Dzhaparidze. The logic of linear tolerance. *Studia Logica* 51 (1992) 249–277.

[8] G. Dzhaparidze. A generalized notion of weak interpretability and the corresponding modal logic. *Annals of Pure and Applied Logic* 61 (1993) 113–160.

[9] S. Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae* 49 (1960) 35–92.

[10] P. Hájek & P. Pudlák. *Metamathematics of First-Order Arithmetic.* Springer-Verlag, Berlin 1993.

[11] D. de Jongh & F. Veltman. Provability logics for relative interpretability. *Mathematical Logic* (P. P. Petkov, ed.) Plenum Press, New York 1990, 31–42.

[12] P. Lindström. On partially conservative sentences and interpretability. *Proceedings of the American Mathematical Society* 91 (1984) 436–443.

[13] R. Magari. The diagonalizable algebras (the algebraization of the theories which express Theor.: II). *Bollettino della Unione Matematica Italiana. Serie IV* 12 (1975) Supplemento al fasc. 3, 117–125.

[14] R. Magari. Representation and duality theory for diagonalizable algebras (the algebraization of theories which express Theor; IV). *Studia Logica* 34 (1975) 305–313.

[15] M. L. Minsky. Recursive unsolvability of Post's problem of "Tag" and other topics in theory of Turing machines. *Annals of Mathematics. Second Series* 74 (1961) 437–455.

[16] M. L. Minsky. *Computation: Finite and Infinite Machines.* Prentice-Hall, Englewood Cliffs 1967.

[17] F. Montagna. On the diagonalizable algebra of Peano Arithmetic. *Bollettino della Unione Matematica Italiana. Serie V* 16–B (1979) 795–812.

[18] F. Montagna. Interpretations of the first-order theory of diagonalizable algebras in Peano arithmetic. *Studia Logica* 39 (1980) 347–354.

[19] F. Montagna. The undecidability of the first-order theory of diagonalizable algebras. *Studia Logica* 39 (1980) 355–359.

[20] V. Yu. Shavrukov. Subalgebras of diagonalizable algebras of theories containing arithmetic. *Dissertationes Mathematicae* 323 (1993).

[21] V. Yu. Shavrukov. A note on the diagonalizable algebras of PA and ZF. *Annals of Pure and Applied Logic* 61 (1993) 161–173.

[22] C. Smoryński. Fixed point algebras. *Bulletin (New Series) of the American Mathematical Society* 6 (1982) 317–356.

[23] C. Smoryński. The finite inseparability of the first-order theory of diagonalisable algebras. *Studia Logica* 41 (1982) 347–349.

[24] C. Smoryński. *Self-Reference and Modal Logic.* Springer-Verlag, New York 1985.

[25] R. I. Soare. *Recursively Enumerable Sets and Degrees. A Study of Computable Functions and Computably Generated Sets.* Springer-Verlag, Berlin 1987.

[26] R. M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics* 25 (1976) 287–304.

[27] A. Visser. The provability logics of recursively enumerable theories extending Peano arithmetic at arbitrary theories extending Peano arithmetic. *Journal of Philosophical Logic* 13 (1984) 97–113.

[28] D. Zambella. On the proofs of arithmetical completeness for interpretability logic. *Notre Dame Journal of Formal Logic* 33 (1992) 542–551.

[29] D. Zambella. Shavrukov's theorem on the subalgebras of diagonalizable algebras for theories containing $I\Delta_0$+exp. *Notre Dame Journal of Formal Logic*, to appear.

# The ILLC Prepublication Series

*Other Prepublications*

X-91-01 Alexander Chagrov, Michael Zakharyaschev The Disjunction Property of Intermediate Propositional Logics
X-91-02 Alexander Chagrov, Michael Zakharyaschev On the Undecidability of the Disjunction Property of Intermediate Propositional Logics
X-91-03 V. Yu. Shavrukov Subalgebras of Diagonalizable Algebras of Theories containing Arithmetic
X-91-04 K.N. Ignatiev Partial Conservativity and Modal Logics
X-91-05 Johan van Benthem Temporal Logic
X-91-06 Annual Report 1990
X-91-07 A.S. Troelstra Lectures on Linear Logic, Errata and Supplement
X-91-08 Giorgie Dzhaparidze Logic of Tolerance
X-91-09 L.D. Beklemishev On Bimodal Provability Logics for $\Pi_1$-axiomatized Extensions of Arithmetical Theories
X-91-10 Michiel van Lambalgen Independence, Randomness and the Axiom of Choice
X-91-11 Michael Zakharyaschev Canonical Formulas for K4. Part I: Basic Results
X-91-12 Herman Hendriks Flexibele Categoriale Syntaxis en Semantiek: de proefschriften van Frans Zwarts en Michael Moortgat
X-91-13 Max I. Kanovich The Multiplicative Fragment of Linear Logic is NP-Complete
X-91-14 Max I. Kanovich The Horn Fragment of Linear Logic is NP-Complete
X-91-15 V. Yu. Shavrukov Subalgebras of Diagonalizable Algebras of Theories containing Arithmetic, revised version
X-91-16 V.G. Kanovei Undecidable Hypotheses in Edward Nelson's Internal Set Theory
X-91-17 Michiel van Lambalgen Independence, Randomness and the Axiom of Choice, Revised Version
X-91-18 Giovanna Cepparello New Semantics for Predicate Modal Logic: an Analysis from a standard point of view
X-91-19 Papers presented at the Provability Interpretability Arithmetic Conference, 24-31 Aug. 1991, Dept. of Phil., Utrecht University
Annual Report 1991

1992 *Logic, Semantics and Philosophy of Langauge*

LP-92-01 Víctor Sánchez Valencia Lambek Grammar: an Information-based Categorial Grammar
LP-92-02 Patrick Blackburn Modal Logic and Attribute Value Structures
LP-92-03 Szabolcs Mikulás The Completeness of the Lambek Calculus with respect to Relational Semantics
LP-92-04 Paul Dekker An Update Semantics for Dynamic Predicate Logic
LP-92-05 David I. Beaver The Kinematics of Presupposition
LP-92-06 Patrick Blackburn, Edith Spaan A Modal Perspective on the Computational Complexity of Attribute Value Grammar
LP-92-07 Jeroen Groenendijk, Martin Stokhof A Note on Interrogatives and Adverbs of Quantification
LP-92-08 Maarten de Rijke A System of Dynamic Modal Logic
LP-92-09 Johan van Benthem Quantifiers in the world of Types
LP-92-10 Maarten de Rijke Meeting Some Neighbours (a dynamic modal logic meets theories of change and knowledge representation)
LP-92-11 Johan van Benthem A note on Dynamic Arrow Logic
LP-92-12 Heinrich Wansing Sequent Caluli for Normal Modal Propositional Logics
LP-92-13 Dag Westerståhl Iterated Quantifiers
LP-92-14 Jeroen Groenendijk, Martin Stokhof Interrogatives and Adverbs of Quantification

*Mathematical Logic and Foundations*

ML-92-01 A.S. Troelstra Comparing the theory of Representations and Constructive Mathematics
ML-92-02 Dmitrij P. Skvortsov, Valentin B. Shehtman Maximal Kripke-type Semantics for Modal and Superintuitionistic Predicate Logics
ML-92-03 Zoran Marković On the Structure of Kripke Models of Heyting Arithmetic
ML-92-04 Dimiter Vakarelov A Modal Theory of Arrows, Arrow Logics I
ML-92-05 Domenico Zambella Shavrukov's Theorem on the Subalgebras of Diagonalizable Algebras for Theories containing $I\Delta_0$ + EXP
ML-92-06 D.M. Gabbay, Valentin B. Shehtman Undecidability of Modal and Intermediate First-Order Logics with Two Individual Variables
ML-92-07 Harold Schellinx How to Broaden your Horizon
ML-92-08 Raymond Hoofman Information Systems as Coalgebras
ML-92-09 A.S. Troelstra Realizability
ML-92-10 V.Yu. Shavrukov A Smart Child of Peano's

*Compution and Complexity Theory*

CT-92-01 Erik de Haas, Peter van Emde Boas Object Oriented Application Flow Graphs and their Semantics
CT-92-02 Karen L. Kwast, Sieger van Denneheuvel Weak Equivalence: Theory and Applications
CT-92-03 Krzysztof R. Apt, Kees Doets A new Definition of SLDNF-resolution

*Other Prepublications*

X-92-01 Heinrich Wansing The Logic of Information Structures
X-92-02 Konstantin N. Ignatiev The Closed Fragment of Dzhaparidze's Polymodal Logic and the Logic of $\Sigma_1$ conservativity
X-92-03 Willem Groeneveld Dynamic Semantics and Circular Propositions, revised version
X-92-04 Johan van Benthem Modeling the Kinematics of Meaning
X-92-05 Erik de Haas, Peter van Emde Boas Object Oriented Application Flow Graphs and their Semantics, revised version

1993 *Logic, Semantics and Philosophy of Langauge*

LP-93-01 Martijn Spaan Parallel Quantification
LP-93-02 Makoto Kanazawa Dynamic Generalized Quantifiers and Monotonicity
LP-93-03 Nikolai Pankrat'ev Completeness of the Lambek Calculus with respect to Relativized Relational Semantics
LP-93-04 Jacques van Leeuwen Identity, Quarrelling with an Unproblematic Notion
LP-93-05 Jaap van der Does Sums and Quantifiers
LP-93-06 Paul Dekker Updates in Dynamic Semantics
LP-93-07 Wojciech Buszkowski On the Equivalence of Lambek Categorial Grammars and Basic Categorial Grammars
LP-93-08 Zisheng Huang, Peter van Emde Boas Information Acquisition from Multi-Agent resources; abstract

*Mathematical Logic and Foundations*

ML-93-01 Maciej Kandulski Commutative Lambek Categorial Grammars
ML-93-02 Johan van Benthem, Natasha Alechina Modal Quantification over Structured Domains
ML-93-03 Mati Pentus The Conjoinablity Relation in Lambek Calculus and Linear Logic
ML-93-04 Andreja Prijatelj Bounded Contraction and Many-Valued Semantics
ML-93-05 Raymond Hoofman, Harold Schellinx Models of the Untyped $\lambda$-calculus in Semi Cartesian Closed Categories
ML-93-06 J. Zashev Categorial Generalization of Algebraic Recursion Theory
ML-93-07 A.V. Chagrov, L.A. Chagrova Algorithmic Problems Concerning First-Order Definability of Modal Formulas on the Class of All Finite Frames
ML-93-08 Raymond Hoofman, Ieke Moerdijk Remarks on the Theory of Semi-Functors
ML-93-09 A.S. Troelstra Natural Deduction for Intuitionistic Linear Logic
ML-93-10 Vincent Danos, Jean-Baptiste Joinet, Harold Schellinx The Structure of Exponentials: Uncovering the Dynamics of Linear Logic Proofs
ML-93-11 Lex Hendriks Inventory of Fragments and Exact Models in Intuitionistic Propositional Logic
ML-93-12 V.Yu. Shavrukov Remarks on Uniformly Finitely Precomplete Positive Equivalences
ML-93-13 V.Yu. Shavrukov Undecidability in Diagonizable Algebras
ML-93-14 Dick de Jongh, Albert Visser Embeddings of Heyting Algebras

*Compution and Complexity Theory*

CT-93-01 Marianne Kalsbeek The Vanilla Meta-Interpreter for Definite Logic Programs and Ambivalent Syntax
CT-93-02 Sophie Fischer A Note on the Complexity of Local Search Problems
CT-93-03 Johan van Benthem, Jan Bergstra Logic of Transition Systems
CT-93-04 Karen L. Kwast, Sieger van Denneheuvel The Meaning of Duplicates in the Relational Database Model
CT-93-05 Erik Aarts Proving Theorems of the Lambek Calculus of Order 2 in Polynomial Time
CT-93-06 Krzysztof R. Apt Declarative programming in Prolog

*Computaional Linguistics*

CL-93-01 Noor van Leusen, László Kálmán The Interpretation of Free Focus
CL-93-02 Theo M.V. Janssen An Algebraic View On Rosetta

*Other Prepublications*

X-93-01 Paul Dekker Existential Disclosure, revised version
X-93-02 Maarten de Rijke What is Modal Logic?
X-93-03 Michiel Leezenberg Gorani Influence on Central Kurdish: Substratum or Prestige Borrowing
X-93-04 A.S. Troelstra (editor) Metamathematical Investigation of Intuitionistic Arithmetic and Analysis, Corrections to the First Edition
X-93-05 A.S. Troelstra (editor) Metamathematical Investigation of Intuitionistic Arithmetic and Analysis, Second, corrected Edition
X-93-06 Michael Zakharyashev Canonical Formulas for K4. Part II: Cofinal Subframe Logics