THE JOURNAL OF SYMBOLIC LOGIC
Volume 88, Number 4, December 2023

AXIOMATIZATIONS OF PEANO ARITHMETIC: A TRUTH-THEORETIC VIEW

ALI ENAYAT AND MATEUSZ ŁEŁYK

Abstract. We employ the lens provided by formal truth theory to study axiomatizations of Peano Arithmetic (PA). More specifically, let Elementary Arithmetic (EA) be the fragment $\mathsf{I}\Delta_0 + \mathsf{Exp}$ of PA, and let $\mathsf{CT}^-[\mathsf{EA}]$ be the extension of EA by the commonly studied axioms of compositional truth CT^-. We investigate both local and global properties of the family of first order theories of the form $\mathsf{CT}^-[\mathsf{EA}] + \alpha$, where α is a particular way of expressing "PA is true" (using the truth predicate). Our focus is dominantly on two types of axiomatizations, namely: (1) schematic axiomatizations that are deductively equivalent to PA and (2) axiomatizations that are proof-theoretically equivalent to the canonical axiomatization of PA.

§1. Introduction. Logicians have long known that different sets of axioms can have the same deductive closure and yet their arithmetizations might exhibit marked differences, e.g., by Craig's trick every recursively enumerable set of axioms is deductively equivalent to a primitive recursive set of axioms. Feferman's pivotal paper [8] on the arithmetization of metamathematics revealed many other dramatic instances of this phenomenon relating to Peano Arithmetic. Let PA be the usual axiomatization of Peano Arithmetic obtained by augmenting Q (Robinson Arithmetic) with the induction scheme, and consider the theory that has come to be known as *Feferman Arithmetic*, which we will denote by FA. The axioms of FA are obtained by an infinite recursive process of "weeding out" applied to PA as follows: enumerate the proofs of PA until a proof of $0 = 1$ is arrived, and then discard the largest axiom used in deriving $0 = 1$; we then proceed to enumerate proofs using only axioms of PA smaller than the one discarded. If we arrive at another proof of $0 = 1$ from the reduced axiom system, we proceed in the same manner. By definition, FA consists of the axioms of PA that remain upon the completion of this recursive infinite process. Thus FA = PA in a sufficiently strong metatheory that can prove the consistency of PA.[1] However, the consistency of FA is built into its definition and PA can readily verify this fact; thus the equality of FA and PA is not provable in PA even though this equality is provable in a sufficiently strong metatheory.

In this paper we employ the lens provided by formal truth theory to study axiomatizations of PA. Our focus is on two types of axiomatizations, namely: (1) schematic

Received August 17, 2021.
2020 *Mathematics Subject Classification.* Primary 03F30, 03F25, Secondary 03F25, 03C62.
Key words and Phrases. Peano arithmetic, axiomatic theories of truth, axiomatization, schemes, conservativity.

[1]Recall that the consistency of PA is provable within Zermelo–Fraenkel set theory ZF; indeed the consistency proof can be carried out in the small fragment of second order arithmetic obtained by augmenting ACA_0 with the induction scheme for Σ^1_1-formulae.

© The Author(s), 2022. Published by Cambridge University Press on behalf of The Association for Symbolic Logic. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

0022-4812/23/8804-0010
DOI:10.1017/jsl.2022.83

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

axiomatizations that are deductively equivalent to PA, and (2) axiomatizations that are proof-theoretically equivalent to the canonical axiomatization of PA. More specifically, let Elementary Arithmetic (EA) be the fragment $\mathsf{I}\Delta_0 + \mathsf{Exp}$ of PA, and $\mathsf{CT}^-[\mathsf{EA}]$ be the extension of EA by the commonly studied axioms of compositional truth CT^- (as in Definition 3). *We investigate the family of first order theories of the form* $\mathsf{CT}^-[\mathsf{EA}] + \alpha$, *where* α *either uses a schematic description of* PA *to express "*PA *is true," or* α *uses a proof-theoretically equivalent formulation of* PA *to express "*PA *is true"* (in the sense of Definition 16).

Several problems can be posed about the aforementioned finitely axiomatized theories of the form $\mathsf{CT}^-[\mathsf{EA}] + \alpha$, the most prominent of which is the determination of their position with respect to the *Tarski Boundary*, i.e., the boundary that demarcates the territory of truth theories that are conservative over PA.[2] For example, the pioneering work of [14] shows that $\mathsf{CT}^-[\mathsf{EA}] + \alpha_1$ is on the conservative side of the Tarski Boundary, where α_1 is the sentence that expresses "each instance of the induction scheme is true" (see Definition 7). On the other hand, let

$$\mathsf{PA}^+ := \mathsf{PA} + \{\mathsf{Con}(\underline{n}) \mid n \in \omega\},$$

where $\mathsf{Con}(\underline{n})$ is the arithmetical sentence that expresses "there is no proof of inconsistency of PA whose code is below n" and ω is the set of natural numbers. It is easy to see that PA^+ is deductively equivalent to PA (provably in EA). However, if we consider a natural arithmetical definition of PA^+, call it $\delta(x)$, and then we choose α_2 to be the sentence

$$T[\delta] := \forall x(\delta(x) \to T(x)), \text{(where } T \text{ is the truth predicate)},$$

then $\mathsf{CT}^-[\mathsf{EA}] + \alpha_2$ is on the nonconservative side of the Tarski Boundary since $\mathsf{CT}^-[\mathsf{EA}] + \alpha_2$ can prove the consistency of PA.

We now briefly discuss the highlights of the paper. In Theorem 26 we show that the set Cons consisting of the (codes of) sentences α such that $\mathsf{CT}^-[\mathsf{EA}] + \alpha$ is conservative over PA is Π_2-complete; which shows, *a fortiori*, that the collection of sentences α such that $\mathsf{CT}^-[\mathsf{EA}] + \alpha$ is conservative over PA is not recursively enumerable. Another main result of the paper pertains to the strengthening CT_0 of $\mathsf{CT}^-[\mathsf{EA}]$ obtained by augmenting $\mathsf{CT}^-[\mathsf{EA}]$ with the scheme of Δ_0-induction (in the extended language containing the truth predicate). It is known that the arithmetical strength of CT_0 far surpasses that of PA, e.g., CT_0 can prove $\mathsf{Con}_{\mathsf{PA}}$, $\mathsf{Con}_{\mathsf{PA}+\mathsf{Con}_{\mathsf{PA}}}$, etc. (see Theorem 6). In Theorem 42 we show that given any r.e. extension U of PA such that $\mathsf{CT}_0 \vdash U$, there is an axiomatization δ of PA which is proof-theoretically equivalent to the usual axiomatization of PA and which has the property that the arithmetical consequences of the (finitely axiomatized) theory[3] $\mathsf{CT}^-[\mathsf{EA}] + T[\delta]$ coincides with the deductive closure of U. (Note that Theorem 6 provides us with an ample supply of theories U that Theorem 42 is applicable to.)

[2]We will refer to the conservative (respectively nonconservative) side of the Tarski Boundary as the region that is *above* (respectively *below*) the Tarski Boundary; this is in step with the traditional Lindenbaum algebra view, where $p \to q$ is translated to $p \leq q$.

[3]Throughout the whole text we systematically employ the word "theory" to refer to an arbitrary set of sentences. In particular theories in our sense need not be closed under logical consequence.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

Our other main results are *structural*. In Section 3.2, we focus on the collection $\mathsf{Sch}_{\mathsf{PA}}$ consisting of the scheme templates τ such that PA is deductively equivalent to the scheme generated by τ (see Definitions 7 and 22). For example, in Theorem 30 we show that from the point of view of relative interpretability, theories of the form $\mathsf{CT}^-[\mathsf{EA}] + T[\tau]$, where $\tau \in \mathsf{Sch}_{\mathsf{PA}}$ and $T[\tau]$ is the sentence asserting that every instance of τ is true, have no maximal element.[4] In the same section we also prove that the partially ordered set $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ is universal for countable partial orderings (in particular, it contains infinite antichains, and also contains a copy of the linearly ordered set $\mathbb{Q}$ of the rationals), where the partial ordering $\leq_{\mathsf{CT}^-}$ is defined by

$$\tau_1 \leq_{\mathsf{CT}^-} \tau_2 \text{ iff } \mathsf{CT}^-[\mathsf{EA}] \vdash T[\tau_1] \rightarrow T[\tau_2].$$

In Section 4.2, we prove similar results about the partial ordering $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$, where Δ is the collection of elementary presentations of PA that are proof-theoretically equivalent to (the canonical axiomatization of) PA. In particular we show that there is an embedding $\mathsf{CT}_0/\mathsf{PA} \hookrightarrow \langle \Delta, \leq_{\mathsf{CT}^-}\rangle$, where $\mathsf{CT}_0/\mathsf{PA}$ is the end segment of the Lindenbaum algebra of PA generated by the collection of arithmetical consequences of CT_0.

Finally, in Theorem 58 of the last section of the paper we give a precise description of the set sup PA consisting of arithmetical sentences that are provable in some theory of the form $\mathsf{CT}^-[\mathsf{EA}] + T[\delta]$, where $\delta(x)$ is an elementary formula (in the sense of Definition 2) that defines an axiomatization of PA in the standard model $\mathbb{N}$ of arithmetic.

Our results are motivated by (1) seeking a better understanding of the contours of the Tarski Boundary; (2) exploring the extent to which the statement "PA is true" is determinate in the context of the basic compositional truth theory $\mathsf{CT}^-[\mathsf{EA}]$, and (3) further investigating structural aspects of finite axiomatizations of infinite theories, a topic initiated in the work of Pakhomov and Visser [22].

§2. Preliminaries.

2.1. CT^-, $\mathbf{CT}_0$, and the Tarski Boundary.

Definition 1. Peano Arithmetic (PA) is the theory formulated in the language $\{0, S, +, \times\}$ whose axioms consist of the axioms of Robinson's Arithmetic Q together with the induction scheme. We will denote the standard model of arithmetic by $\mathbb{N}$ and its universe of discourse by ω.

Definition 2. Elementary Arithmetic (EA) is the fragment $\mathsf{I}\Delta_0 + \mathsf{Exp}$ of PA, where $\mathsf{I}\Delta_0$ is the induction scheme for Δ_0-formulae (i.e., formulae with only bounded quantifiers), and Exp asserts the totality of the function $\exp(x) = 2^x$ (it is well-known that the graph of exp can be described by a Δ_0-formula). An *elementary formula* is an arithmetical formula whose quantifiers are bounded by terms built from the function symbols S, $+$, $\times$, and exp. The family of *(Kalmár) elementary functions* is a distinguished subfamily of the primitive recursive functions.[5] It is well-known that the provably recursive functions of EA are precisely the elementary functions;

[4] Once again, we treat the interpreted theory as greater in this ordering.

[5] Elementary functions occupy the third layer (E_3) of the Grzegorczyk hierarchy of primitive recursive functions $\{E_n \mid n \in \omega\}$. It is often claimed that almost all number theoretical functions that arise in mathematical practice are elementary.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

and that a function f is elementary iff f is computable by a Turing machine with a multiexponential time bound.

DEFINITION 3. We say that B is a *base theory* if B is formulated in $\mathcal{L}_{\mathsf{PA}}$ with $\mathsf{B} \supseteq \mathsf{EA}$. We use $\mathcal{L}_T$ to refer to the language obtained by adding a unary predicate T to $\mathcal{L}_{\mathsf{PA}}$. $\mathsf{CT}^-[\mathsf{B}]$ is the theory extending B with the $\mathcal{L}_T$-sentences CT1 through CT5 below.

We follow the notational conventions from [5]. In particular $x \in \mathsf{ClTerm}_{\mathcal{L}_{\mathsf{PA}}}$ is the arithmetical formula that expresses "x is (the code of) a closed term of $\mathcal{L}_{\mathsf{PA}}$"; $x \in \mathsf{ClTermSeq}_{\mathcal{L}_{\mathsf{PA}}}$ is the arithmetical formula that expresses "x is (the code of) a sequence of closed terms of $\mathcal{L}_{\mathsf{PA}}$"; $x \in \mathsf{Form}_{\mathcal{L}_{\mathsf{PA}}}$ is the arithmetical formula that expresses "x is (the code of) a formula of $\mathcal{L}_{\mathsf{PA}}$"; $x \in \mathsf{Sent}_{\mathcal{L}_{\mathsf{PA}}}$ is the arithmetical formula that expresses "x is (the code of) a sentence of $\mathcal{L}_{\mathsf{PA}}$"; $x \in \mathsf{Var}$ expresses "x is (the code of) a variable"; $x \in \mathsf{VarSeq}$ expresses "x is (the code of) a sequence of variables"; $x \in \mathsf{Form}^{\leq n}_{\mathcal{L}_{\mathsf{PA}}}$ expresses "x is a (the code of) formula of $\mathcal{L}_{\mathsf{PA}}$ with at most n distinct free variables"(n is not a variable in $\mathsf{Form}^{\leq n}_{\mathcal{L}_{\mathsf{PA}}}$); $\underline{x}$ is (the code of) the numeral representing x; $\varphi[\underline{x}/v]$ is (the code of) the formula obtained by substituting the variable v with the numeral representing x and $\varphi[\bar{s}/\bar{v}]$ has an analogous meaning for simultaneous substitution of terms from the sequence $\bar{s}$ for variables from the sequence $\bar{v}$; s° denotes the value of the term s, and $\bar{s}^\circ$ denotes the sequence of numbers that correspond to values of terms from the sequence of terms $\bar{s}$.

Finally, for better readability we will sometimes skip formulae denoting syntactic operations and write the effect of the operations instead. Thus, for example, we will write $T(\neg\varphi)$ to denote "There exists ψ which is the negation of the sentence φ and $T(\psi)$.". For similar reasons, we shall often identify formulae with their Gödel codes. Where it is helpful to distinguish between the two, $\ulcorner\varphi\urcorner$ will denote the Gödel number of φ or the numeral naming this number (depending on the context).

CT1 $\forall s, t \in \mathsf{ClTerm}_{\mathcal{L}_{\mathsf{PA}}} \ \ T(s = t) \leftrightarrow s^\circ = t^\circ$.

CT2 $\forall \varphi, \psi \in \mathsf{Sent}_{\mathcal{L}_{\mathsf{PA}}} \ \ T(\varphi \vee \psi) \leftrightarrow T(\varphi) \vee T(\psi)$.

CT3 $\forall \varphi \in \mathsf{Sent}_{\mathcal{L}_{\mathsf{PA}}} \ \ T(\neg\varphi) \leftrightarrow \neg T(\varphi)$.

CT4 $\forall v \in \mathsf{Var} \forall \varphi \in \mathsf{Form}^{\leq 1}_{\mathcal{L}_{\mathsf{PA}}} \ \ T(\exists v \varphi) \leftrightarrow \exists x\, T(\varphi[\underline{x}/v])$.

CT5 $\forall \varphi(\bar{v}) \in \mathsf{Form}_{\mathcal{L}_{\mathsf{PA}}} \forall \bar{s}, \bar{t} \in \mathsf{ClTermSeq}_{\mathcal{L}_{\mathsf{PA}}} \ \ \big(\bar{s^\circ} = \bar{t^\circ} \to T(\varphi[\bar{s}/\bar{v}]) \leftrightarrow T(\varphi[\bar{t}/\bar{v}])\big)$.

The axiom CT5 is sometimes called *generalized regularity*, or *generalized term-extensionality*, and is not included in the accounts of CT^- provided in the monographs of [3, 10]. The conservativity of this particular version of $\mathsf{CT}^-[\mathsf{PA}]$ can be verified by a refinement of the model-theoretic method introduced in [7], as presented both in [5, 12]. Moreover, the following strengthening of the conservativity result in [5].

THEOREM 4. *There is a polynomial-time computable function f such that for every* $\mathsf{CT}^-[\mathsf{PA}]$*-proof π of an arithmetical sentence φ, $f(\pi)$ is a* PA*-proof of φ. Moreover the correctness of f is verifiable in* PA.

The above result shows that $\mathsf{CT}^-[\mathsf{PA}]$ is *feasibly reducible* to PA. In particular, the basic truth theory $\mathsf{CT}^-[\mathsf{PA}]$ admits at most a polynomial speed-up over PA.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

Moreover, as shown in [5], PA proves the consistency of every finitely axiomatizable subtheory of $\mathsf{CT}^-[\mathsf{PA}]$, which together with the arithmetized completeness theorem and Orey's compactness theorem shows that $\mathsf{CT}^-[\mathsf{PA}]$ is interpretable in PA.

Theorem 4 witnesses the "flatness" of $\mathsf{CT}^-[\mathsf{PA}]$ over its base theory PA. The so-called Tarski Boundary project, seeks to map out the extent of this phenomenon. More concretely, given a metamathematical property of theories P which is exhibited by $\mathsf{CT}^-[\mathsf{PA}]$ we are interested in determining which extensions of $\mathsf{CT}^-[\mathsf{PA}]$ also exhibit P. In particular $P(x)$ can stand for any of the properties below:

- x is conservative over PA.
- x is relatively interpretable in PA.
- x admits at most a polynomial speed-up over PA.

There is an obvious way of obtaining a natural strengthening of $\mathsf{CT}^-[\mathsf{PA}]$ which fails to have any of the above properties. To describe this strengthening, given a theory $\mathcal{T}$ let $\mathsf{Pr}_{\mathcal{T}}(\varphi)$ be the arithmetical formula that expresses "φ is provable from $\mathcal{T}$," where the axioms of $\mathcal{T}$ are given by some arithmetical formula. The *Global Reflection* for $\mathcal{T}$ is the following truth principle:

$$\forall \varphi \in \mathsf{Sent}_{\mathcal{L}_{\mathcal{T}}}\big(\mathsf{Pr}_{\mathcal{T}}(\varphi) \rightarrow T(\varphi)\big). \qquad (\mathrm{GRP}(\mathcal{T}))$$

We stress that $\mathsf{GRP}(\mathcal{T})$ depends not only on $\mathcal{T}$ but also on the particularly chosen formula, which represents the axiom set of $\mathcal{T}$. Below PA denotes the canonical formula which naturally represents the set of axioms of PA (as in Definition 1). Note that $\mathsf{CT}^-[\mathsf{EA}] + \mathsf{GRP}(\mathsf{PA})$ is non-conservative over PA since $\mathsf{Con}_{\mathsf{PA}}$ is provable in $\mathsf{CT}^-[\mathsf{EA}] + \mathsf{GRP}(\mathsf{PA})$. However, $\mathsf{CT}^-[\mathsf{EA}] + \mathsf{GRP}(\mathsf{PA})$ is much stronger, as indicated by the following result.

THEOREM 5 (Kotlarski [13]–Smoryński [29], Łełyk [16]). *The arithmetical consequences of* $\mathsf{CT}^-[\mathsf{EA}] + \mathsf{GRP}(\mathsf{PA})$ *coincides with* $\mathsf{REF}^{<\omega}(\mathsf{PA})$.

In the above $\mathsf{REF}^0(\mathcal{T}) := \mathcal{T}$, $\mathsf{REF}^{n+1}(\mathcal{T}) := \mathsf{REF}(\mathsf{REF}^n(\mathcal{T}))$, $\mathsf{REF}^{<\omega}(\mathcal{T}) := \bigcup_{n\in\omega} \mathsf{REF}^n(\mathcal{T})$, where $\mathsf{REF}(\mathcal{T})$ denotes the extension of $\mathcal{T}$ with all instances of the Uniform Reflection Scheme for $\mathcal{T}$, i.e., $\mathsf{REF}(\mathcal{T})$ consists of all sentences of the following form, where φ ranges over $\mathcal{L}_{\mathcal{T}}$-formulae with at most one free variable:

$$\forall x\big(\mathsf{Pr}_{\mathcal{T}}(\varphi(\underline{x})) \rightarrow \varphi(x)\big).$$

Interestingly enough, over $\mathsf{CT}^-[\mathsf{EA}]$, $\mathsf{GRP}(\mathsf{PA})$ lends itself to many different characterisations, some of which express very basic properties of the truth predicate.

THEOREM 6. *Over* $\mathsf{CT}^-[\mathsf{EA}]$ *the following are all equivalent to* $\mathsf{GRP}(\mathsf{PA})$:

1. Δ_0-*induction scheme for* $\mathcal{L}_T$ (*see* [16, 17]).
2. $\mathsf{GRP}(\emptyset)$, *i.e.*, $\forall \varphi\big(\mathsf{Pr}_{\emptyset}(\varphi) \rightarrow T(\varphi)\big)$ (*see* [2])
3. $\forall c\big($"*c codes a set of sentences*" $\wedge\, T(\bigvee_{\varphi\in c}\varphi) \rightarrow \exists \varphi \in c\ T(\varphi)\big)$ (*see* [4]).

Theorem 6 reveals the surprising robustness of the theory $\mathsf{CT}^-[\mathsf{EA}] + \mathsf{GRP}(\mathsf{PA})$. Out of the three above principles, the third one looks especially modest, being only one direction of a straightforward generalisation (often dubbed *disjunctive correctness*) of the compositional axiom CT2 of CT^- for disjunctions.[6]

[6] The last part of Theorem 6 refines the main result of [6], which shows that CT_0 can be axiomatized by simply adding the disjunctive correctness axiom to $\mathsf{CT}^-[\mathsf{EA}]$.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

This shows that conceptually $\mathsf{CT}^{-}[\mathsf{PA}]$ is closer to the Tarski Boundary than previously conceived. One of the achievements of the current research is the discovery of the remarkable fact that this "conceptually small" area is populated by very different natural theories of truth, each of which "merely" expresses that PA is true.

- Note that by part (1) of Theorem 6, $\mathsf{CT}^{-}[\mathsf{EA}] + \mathsf{GRP}(\mathsf{PA})$ is also axiomatizable by the theory $\mathsf{CT}_0[\mathsf{EA}]$, which is obtained by augmenting $\mathsf{CT}^{-}[\mathsf{EA}]$ with Δ_0-induction scheme for $\mathcal{L}_T$. Since this theory plays a very important role in our paper, for the sake of convenience we omit the reference to the base theory in $\mathsf{CT}_0[\mathsf{EA}]$ and refer to it as CT_0. This is additionally justified by the fact that $\mathsf{CT}_0[\mathsf{EA}] = \mathsf{CT}_0[\mathsf{B}]$ for any base theory B (i.e., any subtheory of PA that extends EA).

As mentioned already in Section 1, our main focus in the current paper is on finite extensions of $\mathsf{CT}^{-}[\mathsf{EA}]$ that expresses "PA is true." As shown in Theorem 58, if we admit all elementary presentations of PA, then each true Π_2-statement can be proved in a theory of this form. Hence, it is natural to look for some intuitive restrictions on "admissible" presentations of PA. We investigate two such admissible families of axiomatizations: *schematic* axiomatizations (introduced in Section 2.2) and *prudent* axiomatizations (introduced in Section 2.3). The former family is well-known; the latter family is defined in this paper as consisting of axiomatizations whose deductive equivalence to PA is verifiable in the weak, finitistically justified metatheory Primitive Recursive Arithemtic (PRA).

2.2. Schematic axiomatizations.

DEFINITION 7. A *template* (for a scheme) is given by a sentence $\tau[P]$ formulated in the language obtained by augmenting $\mathcal{L}_{\mathsf{PA}}$ with a predicate P, where P is unary.[7] An $\mathcal{L}_{\mathsf{PA}}$-sentence ψ is said to be *an instance of* τ if ψ is of the form $\forall v\, \tau[\varphi(x,v)/P]$, where $\tau[\varphi(x,v)/P]$ is the result of substituting all subformulae of the form $P(t)$, where t is a term, with $\varphi(t,v)$ (and re-naming bound variables of φ to avoid unintended clashes). We use S_τ to denote the collection of all instances of τ, and we refer to S_τ as *the scheme generated by* τ.

- We will use $T[\tau]$ to refer to the $\mathcal{L}_T$-sentence that says that each instance of S_τ is true; more formally:

$$T[\tau] := \forall v, w \in \mathsf{Var}\, \forall \varphi(v,w) \in \mathsf{Form}^{\leq 2}_{\mathcal{L}_{\mathsf{PA}}}\ \forall z\ T(\tau[\varphi(v,\underline{z}/w)/P]).$$

In the above, the quantification $\forall \varphi(v,w) \in \mathsf{Form}^{\leq 2}_{\mathcal{L}_{\mathsf{PA}}}$ expresses "for all formulae with at most two free variables v, w." ($\forall \varphi(v) \in \mathsf{Form}^{\leq 1}_{\mathcal{L}_{\mathsf{PA}}}$ below has an analogous meaning.) We note that, over $\mathsf{CT}^{-}[\mathsf{EA}]$, $T[\tau]$ is equivalent to the assertion

$$\forall v \in \mathsf{Var}\, \forall \varphi(v) \in \mathsf{Form}^{\leq 1}_{\mathcal{L}_{\mathsf{PA}}}\ T(\tau[\varphi(v)/P]).$$

[7]Thanks to the coding apparatus available in arithmetic, we can limit ourselves to a single unary predicate P. In other words, the notion of a schematic axiomatization presented here is not affected in our context if the template τ is allowed to use finitely many predicate symbols $P_1, \ldots, P_n$ of various finite articles.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

We sometimes write "T is τ-correct" instead of $T[\tau]$.

As mentioned in Section 1, the special case of the following theorem was established for $\mathsf{B} = \mathsf{PA}$ by Kotlarski, Krajewski, and Lachlan [14], and in full generality by Enayat and Visser [7], and Leigh [15].

THEOREM 8. $\mathsf{CT}^{-}[\mathsf{B}] + T[\tau]$ *is conservative over* B *for every base theory* B *and every scheme template* τ *such that* $\mathsf{B} \vdash S_{\tau}$.

We will need the following definition and classical result about partial truth definitions in the proof of Theorem 12 below.

DEFINITION 9. The *depth* of a formula φ is defined recursively by setting the depth of an arbitrary atomic formula to be zero, putting the depth of $\neg\varphi$ and $\forall x\varphi$ to be one plus the depth of φ, the depth of $\varphi \vee \psi$ to be one plus the maximum of depths of φ and ψ. The depth of a term is defined similarly: the depth of a variable or a constant is zero and the depth of $t + s$ and $t \cdot s$ is one plus the maximum of depths of t, s. The *pure depth* of the formula φ is the defined analogously to depth of φ, except for the condition for atomic formulae: the pure depth of a formula $s = t$ is one plus the maximum of the depths of s, t. The depth of a formula φ will be denoted with $\mathsf{depth}(\varphi)$, whereas its pure depth by $\mathsf{pdepth}(\varphi)$. Observe that the depth of φ is always bounded above by its pure depth. We will write

$$\mathsf{True}(y, P),$$

where P is a unary predicate and y is a variable, for the formula obtained from the conjunction of CT1 through CT4 of Definition 3 in which (1) the predicate T is replaced by P, and (2) the universal quantifiers on φ and ψ are limited to formulae of depth at most y. Intuitively speaking, $\mathsf{True}(y, P)$ says that P satisfies the Tarskian compositional clauses for formulae of depth at most y.

EXAMPLE 10. The depth of an atomic formula is 0, whereas its pure depth can be arbitrarily large. The depth of $\exists x\big(x = S(S(0)) \vee \neg x = x\big)$ is 3, whereas its pure depth is 6.

The following theorem is classical; see [9] for a proof.

THEOREM 11 (Partial Truth Definitions). *For each* $n \in \omega$ *there is a unary* $\mathcal{L}_{\mathsf{PA}}$-*formula* $\mathsf{True}_n(x)$ *such that the formula obtained by replacing* y *with* $\underline{n}$ *and* P *with* $\mathsf{True}_n(x)$ *in the formula* $\mathsf{True}(y, P)$ *is provable in* EA.

THEOREM 12 (Vaught [31], Visser [32]). *Let* $\mathcal{T}$ *be an r.e. theory with enough coding*[8], *and let* $\mathcal{L}_{\mathcal{T}}$ *be the language of* $\mathcal{T}$. *There is a primitive recursive function* f *(indeed* f *is elementary) such that given any unary* Σ_1-*formula* σ *that defines a set of* $\mathcal{L}_{\mathcal{T}}$-*sentences* Φ *in* $\mathbb{N}$, $f(\sigma)$ *is a scheme template such that the deductive closures of* $\mathcal{T} + S_{f(\sigma)}$ *and* $\mathcal{T} + \Phi$ *coincide.*

[8]Visser [32] showed that supporting a pairing function is "enough coding" in this context; this improved the main result of Vaught's paper [31], in which "enough coding" meant being able to interpret an $\in$-relation for which the statement: For all objects $x_0, \dots, x_{n-1}$ there is an object y such that for all objects t, $t \in y$ iff $(t = x_0$ or ...or $t = x_{n-1})$" holds for each $n \in \omega$ (sequential theories support such an $\in$-relation).

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

PROOF OUTLINE FOR $\mathcal{T} = \mathrm{EA}$. Suppose $\sigma(x)$ is a Σ_1-formula that defines a set Φ of sentences of $\mathcal{L}_{\mathsf{PA}}$ in the standard model of arithmetic. (By Craig's trick, σ can be chosen to be an elementary formula, this does not play a role in this proof, but it will come handy in the proof of Proposition 21, which is based on this one.) Let $\mathsf{True}(y, P)$ be as in Definition 9. The desired scheme template τ is

$$\forall y \left[\mathsf{True}(y, P) \to \forall x \left[\left(\sigma(x) \wedge \mathsf{pdepth}(x) \leq y\right) \to P(x)\right]\right].$$

We note that:

(1) $\mathsf{EA} + S_\tau \vdash \Phi$, because for each $n \in \omega$ the truth predicate for formulae of depth at most n is definable by Theorem 11; and

(2) $\mathsf{EA} + \Phi \vdash S_\tau$. To see this, suppose to the contrary that some instance ψ of S_τ is not provable in $\mathsf{EA} + \Phi$. Then by the completeness theorem of first order logic there is a model $\mathcal{M}$ of $\mathsf{EA} + \Phi + \neg\psi$. Since by the definition of S_τ there is a formula $\varphi(x, v)$ such that ψ is a sentence of the form $\forall v\ \tau[\varphi(x, v)/P]$, we have

$$\mathcal{M} \models \mathsf{EA} + \Phi + \neg(\forall v\ \tau[\varphi(x, v)/P]).$$

Thus $\mathcal{M}$ is a model of $\mathsf{EA} + \Phi$ in which the sentence $\exists v\ \neg\tau[\varphi(x, v)/P])$ holds, i.e., $\mathcal{M} \models \exists v \exists y \theta(v, y)$, where

$$\theta(v, y) := \left[\mathsf{True}(y, \varphi(x, v)/P) \wedge \exists x \left[\left(\sigma(x) \wedge \mathsf{pdepth}(x) \leq y\right) \wedge \neg\varphi(x, v)\right]\right].$$

Let a and b be elements in $\mathcal{M}$ such that $\mathcal{M} \models \theta(a, b)$. The key observation at this point is that b cannot be a standard element since $\mathcal{M} \models \Phi$. (It is precisely at this step that the argument would have broken down if we had used depth instead instead of pure depth in our formulation of the scheme template τ.) Together with the fact that $\mathcal{M} \models \mathsf{True}(b, \varphi(x, a)/P)$, this implies that the formula $\varphi(x, a)$ defines a subset of $\mathcal{M}$ that satisfies Tarski's compositional clauses for all standard formulae, thus contradicting Tarski's undefinability of truth theorem. ⊣

REMARK 13. The proof of the above theorem would not go through, if in the definition of τ, pdepth was changed to depth. Indeed, assume τ is modified accordingly. It is enough to take $\Phi := \{\mathsf{Con}_{\mathsf{EA}}(\underline{n}) \mid n \in \omega\}$, where $\mathsf{Con}_{\mathsf{EA}}(x)$ expresses "there is no proof of inconsistency of EA whose code is below x." Let σ be the natural elementary definition of Φ, i.e.,

$$\sigma(x) := \exists y < x\left(x = \ulcorner \mathsf{Con}_{\mathsf{EA}}(\underline{y}) \urcorner\right).$$

Observe that each sentence in Φ has the same, standard depth, call it k. Assume that θ is a truth predicate for formulae of depth k. Then the sentence

$$\forall y \left[\mathsf{True}(y, \theta) \to \forall z \left[\left(\sigma(z) \wedge \mathsf{depth}(z) \leq y\right) \to \theta(z)\right]\right]$$

clearly implies $\mathsf{Con}_{\mathsf{EA}}$, hence S_τ is, over EA, properly stronger than Φ.

The above is the main reason for introducing both depth and pure depth of a formula into the picture. On the one hand, the natural definition of partial truth predicates involves the notion of depth. On the other, we need pure depth to make the proof of Theorem 12 work. The crucial difference between the two notions of depth is that in an arbitrary model $\mathcal{M} \models \mathsf{EA}$ and for an arbitrary standard number n, if $\mathcal{M} \models \mathsf{Sent}_{\mathcal{L}_{\mathsf{PA}}}(\varphi) \wedge \mathsf{pdepth}(\varphi) \leq n$, then φ is "almost" a standard sentence: there is a standard sentence ψ of the same pure depth as φ (ψ differs with φ only

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

w.r.t. the indices of bounded variables) such that for every formula $P(x)$ such that $\mathcal{M} \models \mathsf{True}(n, P(x))$ we have $\mathcal{M} \models P(\varphi) \leftrightarrow P(\psi)$.

REMARK 14. Note that by coupling Theorem 12 with the KKL Theorem we can readily obtain the so called Kleene–Vaught Theorem for extensions of EA that asserts that every r.e. extension of EA can be finitely axiomatized in an extended language. For another line of reasoning, see the proof of Proposition 47.

REMARK 15. Let $\mathsf{Con}_{\mathsf{ZF}}$ be the arithmetical statement asserting the consistency of ZF, and for each $n \in \omega$ let $\mathsf{Con}_{\mathsf{ZF}}(\underline{n})$ be the restricted consistency statement for ZF (that expresses "there is no proof of inconsistency of ZF whose code is below n"). Consider the following extension PA^+ of PA:

$$\mathsf{PA}^+ := \mathsf{PA} + \{\mathsf{Con}_{\mathsf{ZF}}(\underline{n}) \mid n \in \omega\}.$$

Then provably in ZF:

$$\text{"}\mathsf{PA}^+ \text{ is conservative over PA" iff } \mathsf{Con}_{\mathsf{ZF}}.$$

To see that the above holds, we reason in ZF. Suppose PA^+ is conservative over PA. Then for all $n \in \omega$, PA proves $\mathsf{Con}_{\mathsf{ZF}}(\underline{n})$. On the other hand, ZF "knows" that PA holds in the standard model of arithmetic, so for all $n \in \omega$, n is really not a proof of inconsistency of ZF, i.e., $\mathsf{Con}_{\mathsf{ZF}}$ holds. On the other hand, if $\mathsf{Con}_{\mathsf{ZF}}$ holds, then by Σ_1-completeness of PA, PA^+ is conservative over PA.

Moreover, by invoking Theorem 12, there is a *scheme* whose instances are provable inPA (assuming $\mathsf{Con}_{\mathsf{ZF}}$), but ZF cannot verify this. Moreover, coupled with Theorem 8, and using part(c) of Definition 22, this also shows that there is a scheme template τ such that

$$\mathsf{ZF} \vdash \left[\mathsf{Con}_{\mathsf{ZF}} \leftrightarrow \tau \in \mathsf{Sch}^{\mathsf{T}}_{\mathsf{PA}}\right].$$

2.3. Prudent axiomatizations. In Section 4 we will investigate another intuitive restriction on "admissible" axiomatizations of PA, namely axiomatizations that are *prudent* in the sense that their correctness can be verified in a *finitistic* metatheory. To formalize this intuition we use the well-entrenched notion of *proof-theoretic reducibility*.

DEFINITION 16. Let δ, δ' range over elementary formulae with one free variable. We say that δ is *proof-theoretically reducible* to δ' ($\delta \leq_{pt} \delta'$) if

$$\mathrm{I}\Sigma_1 \vdash \forall\varphi\big(\mathsf{Pr}_\delta(\varphi) \to \mathsf{Pr}_{\delta'}(\varphi)\big).$$

In the above $\mathsf{Pr}_\delta(x)$ is the canonical provability predicate that expresses "There is a proof of x in First-Order Logic using the sentences from the set of axioms described by δ as additional assumptions." We write δ_{PA} for the elementary formula representing the usual axiomatization of PA (as in Definition 1), i.e., $\delta_{\mathsf{PA}}(x)$ expresses: x is either (the code of) an axiom of Q or (the code of) an instance of the induction scheme. We say that δ is *proof-theoretically equivalent* to δ_{PA} (written as $\delta \sim_{pt} \delta_{\mathsf{PA}}$) if

$$\mathrm{I}\Sigma_1 \vdash \forall\varphi\big(\mathsf{Pr}_\delta(\varphi) \leftrightarrow \mathsf{Pr}_{\delta_{\mathsf{PA}}}(\varphi)\big).$$

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

It is a classical fact due to Parsons [24, 25] that $\mathrm{I}\Sigma_1$ and the system of Primitive Recursive Arithmetic, known as PRA, have the same Π_2-consequences. In particular it follows that whenever $\delta \sim_{p.t.} \delta'$, then in fact δ and δ' are deductively equivalent *provably in* PRA. As a consequence there are primitive recursive proof transformations mapping proofs in δ to proofs with the same conclusions in δ' and *vice-versa*.

- For the purposes of the results obtained in this paper, we do not need the full power of the proof-theoretic equivalence of δ and δ' to be verifiable in $\mathrm{I}\Sigma_1$ since a theory as weak as Buss's S^1_2 would be sufficient. (Thus we can require that there are *polynomial-time computable* proof transformations mapping proofs in δ to proofs with the same conclusions in δ' and *vice-versa*.) However, we decided to stick to the more well-known notion of proof-theoretic reducibility rather than feasible reducibility, especially since the former notion is philosophically well-motivated by Hilbert's finitism, as argued forcefully by Tait [30].
- We focus primarily on elementary presentations, rather than on, possibly more natural, r.e. axiomatizations for two reasons. First of all, for most of our main results, the simpler the axiomatizations, the better. (The results concerning them, mostly, have the form "For every x there is a prudent axiomatization δ such that ....") Secondly, from the philosophical perspective, the elementary formulae, being decidable in multi-exponential time and hence absolute between models of EA, guarantee (or at least come closer to guaranteeing) that the notion of an axiom of the given theory is determinate. From these perspectives, feasible axiomatizations, i.e., P-Time decidable, axiomatizations would be even better, but we leave the investigation of such axiomatizations for further research.

DEFINITION 17. We use Δ^* to denote the collection of unary elementary formulae $\delta(x)$ such that $\delta^{\mathbb{N}} := \{n \in \omega \mid \mathbb{N} \models \delta(\underline{n})\}$ codes an $\mathcal{L}_{\mathsf{PA}}$-theory that is deductively equivalent to PA. We sometimes refer to the members of Δ^* as *elementary presentations* of PA.

- Given any arithmetical formula φ with exactly one free variable,

  $$T[\varphi(x)] := \forall x\big(\varphi(x) \to T(x)\big),$$

  where x is the unique free variable of φ. So $T[\varphi]$ is the $\mathcal{L}_T$-sentence expressing that the theory described by φ is true. Moreover, we put

  $$\mathsf{CT}^-[\![\varphi]\!] := \mathsf{CT}^-[\mathsf{EA}] + T[\varphi].$$

- We use Δ to denote the subset of Δ^* consisting of formulae $\delta \in \Delta^*$ such that δ is proof-theoretically equivalent to δ_{PA}. *Thus* Δ *is the collection of (defining formulae of) prudent axiomatizations of* PA. Occasionally we also need the extension of Δ, denoted Δ^-, defined

  $$\Delta^- := \left\{\delta \in \Delta^* \mid \delta \leq_{pt} \mathsf{PA}\right\}.$$

  On Δ^- and Δ we shall consider the relation $\leq_{\mathsf{CT}^-}$ given by

  $$\delta \leq_{\mathsf{CT}^-} \delta' \iff \mathsf{CT}^-[\mathsf{EA}] \vdash T[\delta] \to T[\delta'].$$

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

Convention 18. *Simplifying things a little bit, when talking about the structures* $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$ *and* $\langle \Delta^-, \leq_{\mathsf{CT}^-}\rangle$, *we shall assume that* Δ *is replaced by the quotient set* $\Delta/\sim$, *where* $\sim$ *is the least equivalence relation that makes* $\leq_{\mathsf{CT}^-}$ *antisymmetric, to wit*:

$$\delta \sim \delta' \text{ iff } \delta \leq_{\mathsf{CT}^-} \delta' \text{ and } \delta' \leq_{\mathsf{CT}^-} \delta.$$

- Let us stress an important difference between $\mathsf{CT}^-[\mathsf{PA}]$ and $\mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$: the latter but not the former includes the sentence "All induction axioms are true." In particular, the latter is finitely axiomatizable, while the former is known to be reflexive and therefore not finitely axiomatizable.
- Note that the meaning of $T[\cdots]$ depends on whether the object within the brackets is a scheme template, in which case $T[\cdots]$ is interpreted as in Definition 7, or an arithmetical formula, in which case $T[\cdots]$ has the meaning given in Definition 17. We shall try to reserve the use of variables τ, σ, etc. in this context for schematic templates and φ, ψ, δ, etc. for formulae.

Proposition 19. *Both* $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$ *and* $\langle \Delta^-, \leq_{\mathsf{CT}^-}\rangle$ *are distributive lattices.*

Proof. We only present the proof for the case of Δ as it is $(1+\varepsilon)$-times harder. For showing that both structures are distributive lattices, it is enough to show that given $\delta, \delta' \in \Delta$, one can find elements $\delta \oplus \delta'$ and $\delta \otimes \delta'$ of Δ such that over $\mathsf{CT}^-[\mathsf{PA}]$ we have

$$T[\delta] \wedge T[\delta'] \leftrightarrow T[\delta \oplus \delta'], \tag{1}$$

$$T[\delta] \vee T[\delta'] \leftrightarrow T[\delta \otimes \delta']. \tag{2}$$

Indeed, this is because the Lindenbaum algebra of CT^- is a distributive lattice. It can be readily seen that if we define

$$\delta \oplus \delta'(x) := \delta(x) \vee \delta'(x),$$

then $\delta \oplus \delta' \in \Delta$ and (1) is satisfied. For (2) it is sufficient to define

$$\delta \otimes \delta'(x) := \exists y, z < x\big(\delta(y) \wedge \delta'(z) \wedge x = y \vee z\big),$$

where $x = y \vee z$ expresses that x is a disjunction of y and z. To see that (2) holds and $\delta_{\mathsf{PA}} \leq_{pt} \delta \otimes \delta'$ one simply applies reasoning by cases; the proof of $\delta \otimes \delta' \leq_{pt} \delta_{\mathsf{PA}}$ is trivial. ⊣

Remark 20. If $\delta \in \Delta$ corresponds to a *schematic axiomatization* of PA (i.e., for some template $\tau[P]$, $\delta(x)$ says that x is the result of substituting P with some unary arithmetical formula), then $\mathsf{CT}^-[\![\delta]\!]$ is a conservative extension of PA by Theorem 26. In contrast, even for very natural $\delta \in \Delta$, $\mathsf{CT}^-[\![\delta]\!]$ may be a highly non-conservative extension of PA. For example, consider

$$\mathsf{REF}_{\mathsf{EA}} = \left\{\forall x\big(\mathsf{Pr}_{\mathsf{EA}}(\varphi(\underline{x})) \rightarrow \varphi(x)\big) \;\middle|\; \varphi(x) \in \mathcal{L}_{\mathsf{PA}}\right\}.$$

By a classical theorem of Kreisel, the union of EA and $\mathsf{REF}_{\mathsf{EA}}$ is deductively equivalent to PA (see, e.g., [1, p. 39]). Let $\delta(x)$ be a natural elementary definition of $\mathsf{EA} \cup \mathsf{REF}_{\mathsf{EA}}$. Then, in fact $\delta \in \Delta$. An easy argument shows that

$$\mathsf{CT}^-[\![\delta]\!] \vdash \forall \varphi\big(\mathsf{Pr}_{\mathsf{EA}}(\varphi) \rightarrow T(\varphi)\big).$$

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

However, by a theorem of [2], over $\mathsf{CT}^-[\mathsf{EA}]$, the above consequence of $\mathsf{CT}^-[\![\delta]\!]$ implies the Global Reflection Principle for PA.

PROPOSITION 21. *Every $\mathcal{L}_{\mathsf{PA}}$-theory $\mathcal{T}$ extending* EA *whose axioms are described by an elementary formula δ (in the standard model of arithmetic) has a proof-theoretically equivalent presentation δ' such that $\mathsf{CT}^-[\![\delta']\!]$ is a conservative extension of $\mathcal{T}$.*

PROOF. Fix $\mathcal{T}$ and δ as in the assumptions. Let δ' be the natural elementary definition of the set S_τ, where τ is the template defined as in the proof of Theorem 12. This works, since in the proof of Theorem 12, the verification of the fact that the deductive closure of $\mathsf{EA} + S_\tau$ and $\mathsf{EA} + \Phi$ coincide formalizes smoothly in the subsystem WKL_0^* of second order arithmetic, which is well-known to be a conservative extension of EA, as first shown in [28]. More explicitly, the verification of $\mathsf{EA} + S_\tau \vdash \Phi$ requires only the existence of well-behaved partial truth predicates (that can be developed within EA, as demonstrated, e.g., in [1, Proposition 2.6]). On the other hand, the verification of $\mathsf{EA} + \Phi \vdash S_\tau$ requires the completeness theorem of first order logic (which is readily available in WKL_0^*) together with Tarski's undefinability of truth theorem. Although Tarski's theorem presupposes the consistency of Φ, this can be assumed, because if Φ is inconsistent, so is S_τ by the proof of the first implication, and in such a scenario the two theories clearly coincide. Hence δ' is indeed proof-theoretically reducible to δ. However, in this case $\mathsf{CT}^-[\![\delta']\!]$ is trivially equivalent to $\mathsf{CT}^-[\mathsf{PA}] + T[\tau]$, hence is a conservative extension of $\mathcal{T}$, due to Theorem 8. ⊣

§3. Schematically correct axiomatizations.

3.1. Complexity.

DEFINITION 22. In the following definitions τ ranges over scheme templates and S_τ is the corresponding scheme (in the sense of Definition 7) generated by τ.

(a) $\mathsf{Sch}^-_{\mathsf{PA}} := \{\tau : \mathsf{PA} \vdash S_\tau\}$, i.e., $\mathsf{Sch}^-_{\mathsf{PA}}$ is the collection of templates whose corresponding scheme is PA-provable.
(b) $\mathsf{Sch}_{\mathsf{PA}} := \{\tau \in \mathsf{Sch}^-_{\mathsf{PA}} : S_\tau \vdash \mathsf{PA}\}$, i.e., $\mathsf{Sch}_{\mathsf{PA}}$ is the collection of templates whose corresponding scheme is an axiomatization of PA.
(c) $\mathsf{Sch}^{\mathsf{T}}_{\mathsf{PA}}$ is the collection of templates τ such that the arithmetical consequences of $\mathsf{CT}^-[\mathsf{EA}] + T[\tau]$ coincides with PA (recall that $T[\tau]$ says that T is τ-correct, as in Definition 7).
(d) $\mathsf{Cons} := \{\varphi \in \mathcal{L}_T : \mathsf{CT}^-[\mathsf{PA}] + \varphi$ is conservative over PA$\}$.

Recall that in Section 1 we defined $\leq_{\mathsf{CT}^-}$ on $\mathsf{Sch}^-_{\mathsf{PA}}$ as follows:

$$\tau \leq_{\mathsf{CT}^-} \tau' \iff \mathsf{CT}^- \vdash T[\tau] \to T[\tau'].$$

Note that at this point $\leq_{\mathsf{CT}^-}$ denotes both the ordering on scheme templates and the ordering on prudent axiomatizations. As the notation "$\leq_{\mathsf{CT}^-}$" will never be used in isolation this shouldn't lead to serious confusion. When talking about the structural properties of $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ we shall tacitly assume that $\mathsf{Sch}_{\mathsf{PA}}$ is factored out by an appropriate equivalence relation, so as to make $\leq_{\mathsf{CT}^-}$ a partial order (as in Convention 18).

PROPOSITION 23. *$\langle \mathsf{Sch}^-_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ and $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ are distributive lattices.*

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

PROOF. As previously we do the case of a smaller structure, with $\mathsf{Sch}_{\mathsf{PA}}$ as the universe. Arguing as previously in Proposition 19, it is enough to define $\oplus$ and $\otimes$ such that $\mathsf{CT}^{-}[\mathsf{PA}]$ proves the following for all $\tau,\ \tau' \in \mathsf{Sch}_{\mathsf{PA}}$:

$$T[\tau] \wedge T[\tau'] \leftrightarrow T[\tau \oplus \tau'], \tag{3}$$

$$T[\tau] \vee T[\tau'] \leftrightarrow T[\tau \otimes \tau']. \tag{4}$$

The case of $\oplus$ is trivial. We put

$$\tau \oplus \tau' := \tau \wedge \tau'.$$

The case of $\otimes$ is (a little bit) harder. We put

$$\tau \otimes \tau' := \tau \vee (\tau'[Q/P]),$$

where Q is a fresh unary predicate. As remarked earlier (compare footnote 4) thanks to the coding apparatus, $\tau \otimes \tau'$ can be expressed as a scheme with a single unary predicate P. Then we obtain

$$\mathsf{CT}^{-}[\mathsf{EA}] \vdash T[\tau \otimes \tau'] \equiv \forall\varphi\, \forall\psi\ T\big(\tau[\varphi/P] \vee \tau'[\psi/Q]\big).$$

It is very easy now to check that (4) is satisfied. ⊣

We note that the above proof adapts to the case of $\langle \mathsf{Sch}^{\mathsf{T}}_{\mathsf{PA}}, \leq_{\mathsf{CT}^{-}} \rangle$. Quite in the opposite direction, it can be shown that $\langle \mathsf{Cons}, \leq_{\mathsf{CT}^{-}} \rangle$ is not even a lattice as there are two sentences $\varphi, \psi \in \mathsf{Cons}$ such that $\mathsf{CT}^{\lceil}\mathsf{PA}] + \varphi + \psi$ is a non-conservative extension of PA. First examples of such sentences were discovered by Bartosz Wcisło (unpublished). We plan to present a family of such examples in the forthcoming sequel [18] to the current paper.

THEOREM 24 (KKL-Theorem, first formulation). $\mathsf{CT}^{-}[\mathsf{PA}] + T[\tau]$ *is conservative over* PA *for each* $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}}$.

Let Θ be the union of sentences of the form $T[\tau]$ (expressing that T is τ-correct) as τ ranges in $\mathsf{Sch}^{-}_{\mathsf{PA}}$. Since the union of two schemes is axiomatizable by a single scheme, the KKL-theorem can be reformulated as:

THEOREM 25 (KKL-Theorem, second formulation). $\mathsf{CT}^{-}[\mathsf{PA}] + \Theta$ *is conservative over* PA.

The above formulation naturally suggests the question: *How complicated is* Θ *(viewed as a subset of* ω*)? Is it recursively enumerable?* The result below shows that Θ is Π_2-complete, since Θ is readily seen to be recursively isomorphic to $\mathsf{Sch}^{-}_{\mathsf{PA}}$ (indeed the isomorphism is witnessed by an elementary function). Therefore, Θ is pretty far from being recursively enumerable.

THEOREM 26. *The sets* $\mathsf{Sch}^{-}_{\mathsf{PA}}$, $\mathsf{Sch}_{\mathsf{PA}}$, $\mathsf{Sch}^{\mathsf{T}}_{\mathsf{PA}}$, *and* Cons *are all* Π_2*-complete.*

PROOF. Each of the four sets is readily seen to be definable by a Π_2-formula, so it suffices to show that each is Π_2-hard, i.e., the complete Π_2-set $\mathsf{True}^{\mathbb{N}}_{\Pi_2}$ consisting of (Gödel numbers of) Π_2-sentences that are true in the standard model $\mathbb{N}$ of PA is many-one reducible (denoted $\leq_{\mathsf{m}}$) to each of them. Recall that $\leq_{\mathsf{m}}$ is defined among subsets of ω via:

$A \leq_{\mathsf{m}} B$ iff there is a total recursive function f such that: $\forall n \in \omega\ (n \in A \Leftrightarrow f(n) \in B)$.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

The proof will be complete once we demonstrate the following four assertions:

(i) $\mathsf{True}^{\mathbb{N}}_{\Pi_2} \leq_{\mathsf{m}} \mathsf{Sch}^{-}_{\mathsf{PA}}$.[9]
(ii) $\mathsf{Sch}^{-}_{\mathsf{PA}} \leq_{\mathsf{m}} \mathsf{Sch}_{\mathsf{PA}}$.
(iii) $\mathsf{Sch}^{-}_{\mathsf{PA}} \leq_{\mathsf{m}} \mathsf{Sch}^{T}_{\mathsf{PA}}$.
(iv) $\mathsf{True}^{\mathbb{N}}_{\Pi_2} \leq_{\mathsf{m}} \mathsf{Cons}$.

To prove (i), suppose $\pi = \forall x\ \exists y\ \delta(x, y)$ is a Π_2-statement, where $\delta(x, y)$ is Δ_0. We first observe that by Σ_1-completeness of PA:

$(*)$ $\pi \in \mathrm{True}^{\mathbb{N}}_{\Pi_2}$ iff $\forall n \in \omega$ PA $\vdash \exists y\ \delta(\underline{n}, y)$.

On the other hand, $R = \{\exists y\ \delta(\underline{n}, y) : n \in \omega\}$ is a recursive set of sentences, so by Theorem 12 there is τ such that $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}}$ iff PA $\vdash R$. To finish the proof, it remains to observe that the transition from π to the Σ_1-formula σ that defines R in $\mathbb{N}$ is given by a recursive function g, therefore if f is the total recursive function as in Theorem 12 then we have

$$\pi \in \mathsf{True}^{\mathbb{N}}_{\Pi_2} \text{ iff } f(g(\pi)) \in \mathsf{Sch}^{-}_{\mathsf{PA}}.$$

The proof of (ii) is based on the observation that $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}}$ iff $h(\tau) \in \mathsf{Sch}_{\mathsf{PA}}$, where $h(\tau) := \tau \wedge \tau_{\mathsf{PA}}$, and τ_{PA} is defined as follows:

$$\tau_{\mathsf{PA}} := \mathsf{Q} \wedge [P(0) \wedge \forall x(P(x) \to P(S(x))) \to \forall x P(x)].$$

To verify (iii), we claim that $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}}$ iff $(\tau \wedge \tau_{\mathsf{PA}}) \in \mathsf{Sch}^{T}_{\mathsf{PA}}$. The implication $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}} \Rightarrow (\tau \wedge \tau_{\mathsf{PA}}) \in \mathsf{Sch}^{T}_{\mathsf{PA}}$ follows directly from Theorem 3 (since PA proves $S_{\tau\wedge\tau_{\mathsf{PA}}}$ if $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}}$). On the other hand, if $(\tau \wedge \tau_{\mathsf{PA}}) \in \mathsf{Sch}^{T}_{\mathsf{PA}}$, then by the definition of $\mathsf{Sch}^{T}_{\mathsf{PA}}$, PA proves S_τ, so $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}}$.

Finally, to establish (iv) suppose $\pi = \forall x\ \exists y\ \delta(x, y)$ is a Π_2-statement, where $\delta(x, y)$ is Δ_0. In the proof of part (i) we showed that there are recursive functions f and g such that

$$\pi \in \mathsf{True}^{\mathbb{N}}_{\Pi_2} \Longleftrightarrow f(g(\pi)) \in \mathsf{Sch}^{-}_{\mathsf{PA}}.$$

Let h be the function that takes a template τ as input, and outputs the sentence $T[\tau] \in \mathcal{L}_T$ expressing "T is τ-correct." Clearly h is a recursive function. Also, it is evident that $\tau \in \mathsf{Sch}^{-}_{\mathsf{PA}}$ iff $T[\tau] \in \mathsf{Cons}$ (the direction $\Rightarrow$ follows from Theorem 8; and the direction $\Leftarrow$ follows from the relevant definitions). Therefore,

$$\pi \in \mathsf{True}^{\mathbb{N}}_{\Pi_2} \Longleftrightarrow h\left(f(g(\pi))\right) \in \mathsf{Cons}. \qquad \dashv$$

Proposition 27. *Let φ_s be the single $\mathcal{L}_T$-sentence that expresses "every* PA*-provable scheme is true." Then* CT_0 *can be axiomatized by* $\mathsf{CT}^{-}[\mathsf{EA}] + \varphi_s$.

Proof. By Theorem 6, CT_0 can be axiomatized by $\mathsf{CT}^{-}[\mathsf{EA}] + \mathsf{GRP}(\mathsf{PA})$. This makes it clear that φ_s is provable in CT_0. For the other direction, working in $\mathsf{CT}^{-}[\mathsf{EA}] + \varphi_s$, suppose ψ is PA-provable. Then the scheme given by $\forall x(\psi \vee P(x))$ is PA-provable, so the instance of this scheme in which P is replaced with $x \neq x$

[9]The proof of (i) shows that Sch^{-}_{T} is Π_2-complete for any extension T of Robinson's Q that is Σ_1-sound, and which also supports a pairing function.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

is true, but since $T(\forall x(x = x))$, we have $T(\psi)$. Thus, since ψ was arbitrary, $\mathsf{CT}^-[\mathsf{EA}] + \varphi_s \vdash \mathsf{GRP}(\mathsf{PA})$. ⊣

3.2. Structure of schematically correct extensions. In this subsection we take a closer look at the structure of $\mathsf{Sch}_{\mathsf{PA}}$. In particular, we look at interpretability properties of its elements, where by "interpretability" we always mean relative interpretability, as described in [9]. The most basic tool we shall use is a modification of the Vaught operation from the proof of Theorem 12. Let us introduce the relevant definition:

Definition 28. For arithmetical formulae $\varphi(x), \delta(x)$ with at most one free variable let the *φ-restricted Vaught schematization of δ* be the scheme template.

$$V_{(\varphi,\delta)}[P] := \forall y\big[\big(\varphi(y) \wedge \mathrm{True}(y, P)\big) \rightarrow \forall x\big((\delta(x) \wedge \mathsf{pdepth}(x) \leq y) \rightarrow P(x)\big)\big].$$

For a single formula δ, $V_\delta[P]$ abbreviates $V_{(x=x,\delta)}[P]$ and we often omit the reference to P. Similarly $V_{\varphi,\delta}[\theta(x)]$ abbreviates $V_{\varphi,\delta}[\theta(x)/P(x)]$.

Convention 29. *Working in* $\mathsf{CT}^-[\mathsf{EA}]$ *and having fixed an (possibly nonstandard) arithmetical formula with one free variable* $\theta(v)$, $T{*}\theta(x)$ *will abbreviate the formula* $T(\theta[\underline{x}/v])$. *Hence* $T{*}\theta(x)$ *says that* x *satisfies* θ. *This notation was first introduced in* [19] *and is very successful in decreasing the number of brackets and improving readability.*

Below, we shall borrow a notation used in the context of prudent axiomatizations: $\mathsf{CT}^-[\![\tau]\!]$ is the theory $\mathsf{CT}^-[\mathsf{EA}] + T[\tau]$, i.e., $\mathsf{CT}^-[\mathsf{EA}]$ together with the assertion that T is τ-correct. In such contexts the variables such as σ, τ, or $V_{\_}$ below will always denote scheme templates.

Theorem 30. *If* $\psi \in \mathcal{L}_T$ *is such that for every* $\tau \in \mathsf{Sch}_{\mathsf{PA}}$, ψ *is interpretable in* $\mathsf{CT}^-[\![\tau]\!]$, *then* ψ *is interpretable in* $\mathsf{CT}^-[\mathsf{PA}]$.

Proof. We prove the contrapositive. Fix ψ which is not interpretable in $\mathsf{CT}^-[\mathsf{PA}]$. We modify the Pakhomov–Visser diagonalization from [22, Theorem 4.1]. Observe that for two finite theories α, β, the condition "α interprets β" is Σ_1. Let $\alpha \rhd \beta$ denote the formalization of this relation. Consider a Σ_1-sentence $\varphi = \exists x \varphi'(x)$, where $\varphi'(x) \in \Delta_0$ such that the following equivalence is provable in $\mathsf{CT}^-[\mathsf{PA}]$:

$$\varphi \leftrightarrow \left[\mathsf{CT}^-[\![V_{(\forall z \leq y \neg\varphi'(z), \delta_{\mathsf{PA}})}]\!] \rhd \psi\right].$$

Similarly to the Pakhomov–Visser argument, we argue that φ is false. Suppose not and take the least $n \in \omega$ such that $\varphi'(n)$ holds. Then, in Q, $\forall z \leq x \neg\varphi'(z)$ is equivalent to $x < \underline{n}$, hence the following is provable in $\mathsf{CT}^-[\mathsf{PA}]$:

$$\forall\theta(x)\big(T\big(V_{(\forall z \leq y \neg\varphi'(z), \delta_{\mathsf{PA}})}[\theta]\big) \leftrightarrow T\big(V_{(y<\underline{n}, \delta_{\mathsf{PA}})}[\theta]\big)\big).$$

We claim that

$$\mathsf{CT}^-[\mathsf{PA}] \vdash \forall\theta(x)\ T\big(V_{(y<\underline{n}, \delta_{\mathsf{PA}})}[\theta]\big). \qquad (*)$$

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

Indeed, working in $\mathsf{CT}^-[\mathsf{PA}]$ fix $\theta \in \mathsf{Form}^{\leq 1}_{\mathcal{L}_{\mathsf{PA}}}$. By compositional conditions $T\big(V_{(y<\underline{n},\delta_{\mathsf{PA}})}[\theta]\big)$ is equivalent to

$$\bigwedge_{i<n}\big[\big(T{*}\mathrm{True}(\underline{i},\theta)\big) \rightarrow \forall x\big((\delta_{\mathsf{PA}}(x) \wedge \mathsf{pdepth}(x) \leq \underline{i}) \rightarrow T{*}\theta(x)\big)\big].$$

However, once again by compositional conditions imposed on T, $T{*}\mathrm{True}(\underline{i},\theta)$ is equivalent to $\mathrm{True}(\underline{i}, T{*}\theta(x))$, hence to the assertion that $T{*}\theta(x)$ is a compositional truth predicate for formulae of depth at most i. Assuming that this is the case, since i is standard, every induction axiom of pure depth at most i is true in the sense of $T{*}\theta(x)$. This concludes our proof of $(*)$.

Now, since φ is true, it follows that

$$\mathsf{CT}^-[\mathsf{PA}] + \forall\theta(x)\ T\big(V_{(y<\underline{n},\delta_{\mathsf{PA}})}[\theta]\big) \text{ interprets } \psi.$$

However, by the above argument it would mean that $\mathsf{CT}^-[\mathsf{PA}]$ interprets ψ, contrary to the assumption.

Since φ is false, $V_{(\forall z \leq y \neg\varphi'(z),\delta_{\mathsf{PA}})}[P]$ is a scheme template, such that the scheme associated with it axiomatizes PA. Moreover, $\mathsf{CT}^-[\mathsf{PA}] + T\big[V_{(\forall z \leq y \neg\varphi'(z),\delta_{\mathsf{PA}})}\big]$ does not interpret ψ. ⊣

Since $\mathsf{CT}^-[\mathsf{PA}]$ is interpretable in PA (see [7, 15]), we obtain the following corollary.

Corollary 31. *For every $\psi \in \mathcal{L}_T$ such that* PA *does not interpret ψ there is a scheme template $\tau \in \mathsf{Sch}_{\mathsf{PA}}$ such that $\mathsf{CT}^-[\![\tau]\!]$ does not relatively interpret ψ.*

Since $\mathsf{PA} \not\rhd \mathsf{Q} + \mathsf{Con}_{\mathsf{PA}}$ (see [26]) we obtain the following corollary. It is of interest because it gives an example of a natural theory that is not interpretable in PA (because it is finite) but this is not due to the reason that the theory interprets the consistency of PA (like most known finite extensions of PA).

Corollary 32. *There is a scheme template $\tau \in \mathsf{Sch}_{\mathsf{PA}}$ such that $\mathsf{CT}^-[\![\tau]\!]$ does not interpret $\mathsf{Q} + \mathsf{Con}_{\mathsf{PA}}$.*

Corollary 33. *For every scheme template $\tau \in \mathsf{Sch}_{\mathsf{PA}}$ there is a scheme template $\tau' \in \mathsf{Sch}_{\mathsf{PA}}$ such that $\mathsf{CT}^-[\![\tau]\!]$ interprets $\mathsf{CT}^-[\![\tau']\!]$, but not vice versa.*

Proof. Fix τ and apply Corollary 31 to $\psi := \mathsf{CT}^-[\![\tau]\!]$. This is legal, since the latter theory is a finitely axiomatizable extension of PA, hence it is not interpretable in PA.[10] So there is a scheme $\tau'' \in \mathsf{Sch}_{\mathsf{PA}}$ such that $\mathsf{CT}^-[\![\tau'']\!]$ does not relatively interpret $\mathsf{CT}^-[\![\tau]\!]$. Now it is sufficient to take $\tau' := \tau \otimes \tau''$, as in the proof of Proposition 23. ⊣

Next we will consider more structural properties of $\mathsf{Sch}_{\mathsf{PA}}$. These properties will be shown to be transferable to the Lindenbaum Algebra of CT_0.

- For the rest of this section δ and δ' are arbitrary elementary formulae that, provably in EA, specify arithmetical theories, i.e., possibly infinite sets of arithmetical sentences. We will write $\delta \subseteq \delta'$ as an abbreviation of $\forall x(\delta(x) \rightarrow$

[10] This is because otherwise $\mathsf{CT}^-[\![\tau]\!]$, being a finite theory, would be interpretable in a finite fragment of PA, call it $\mathcal{T}$. But then, since $\mathsf{CT}^-[\![\tau]\!]$ extends PA and PA is reflexive, $\mathsf{CT}^-[\![\tau]\!] \vdash \mathsf{Con}_{\mathcal{T}}$. Hence $\mathcal{T}$ would interpret $Q + \mathsf{Con}_{\mathcal{T}}$, which is impossible by the interpretability version of the Second Incompleteness Theorem, see [9] (we owe this argument to Albert Visser).

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

$\delta'(x)$). Recall (from Definition 17) that $T[\delta]$ is the following sentence expressing "T is δ correct":

$$\forall x\big(\delta(x) \to T(x)\big).$$

Note the difference between $T[V_\delta]$ and $T[\delta]$.

The first result is immediate:

Proposition 34. *For every δ and δ',* $\mathsf{CT}^-[\mathsf{PA}] \vdash \forall x\big(\delta(x) \to \delta'(x)\big) \to \big(T[V_{\delta'}] \to T[V_\delta]\big)$.

For many applications, the condition $\delta \subseteq \delta'$ from the antecedent is too restrictive. One would like to relax it to $\delta \vdash \delta'$, however, this one is too weak to guarantee (over $\mathsf{CT}^-[\mathsf{PA}]$) that the implication $T[V_{\delta'}] \to T[V_\delta]$ holds. This is because the truth predicate axiomatized by pure $\mathsf{CT}^-[\mathsf{PA}]$ is far from being closed under logic (compare with Theorem 6). The next proposition is a fair compromise between the two solutions.

- Given a unary arithmetical formula $\varphi(x)$, in the proposition below we use the convention of using $\delta_\varphi(x)$ to refer to the formula that defines the set of (codes of) sentences of the form $\varphi(\underline{n})$ (in the standard model $\mathbb{N}$ of arithmetic).

Proposition 35. *For arbitrary arithmetical formulae $\varphi(x)$ and $\psi(x)$,*

$$\mathsf{CT}^-[\mathsf{PA}] \vdash \forall x\big(\varphi(x) \to \psi(x)\big) \to \big(T[V_{\delta_\varphi}] \to T[V_{\delta_\psi}]\big).$$

Proof. Fix arbitrary arithmetical formulae ψ and φ with exactly one free variable. Let δ_ψ and δ_φ be defined in the bullet point above the current proposition. Without loss of generality, assume that the variable x occurs in ψ. Working in $\mathsf{CT}^-[\mathsf{PA}]$ assume that $\forall x\big(\varphi(x) \to \psi(x)\big)$ and $T[V_{\delta_\varphi}]$ hold. We argue that $T[V_{\delta_\psi}]$ holds as well. Fix arbitrary a, θ, b such that $\mathsf{True}(a, T{*}\theta)$ and $\mathsf{pdepth}(\psi(\underline{b})) \leq a$. It follows that for some standard n, $\mathsf{pdepth}(\varphi(\underline{b})) \leq a + n$, hence there exists a formula $\theta'(x)$ such that

$$\mathsf{True}(a + n, T{*}\theta').$$

By $T[V_{\delta_\varphi}]$ we conclude $T{*}\theta'(\varphi(\underline{b}))$. However, since $\varphi(x)$ is of standard depth, it follows that $\varphi(b)$ holds. Hence $\psi(b)$ holds as well. Since $\psi(b)$ is also of standard depth, we conclude that $T{*}\theta(\psi(\underline{b}))$, which ends the proof. ⊣

The proposition below is an important tool for discovering various patterns in $\mathsf{Sch}_{\mathsf{PA}}$. It enables us to switch from somewhat less readable Vaught schematizations of elementary presentations of theories to more workable presentations themselves. It says that over CT_0, δ-correctness is equivalent to V_δ-correctness.

Proposition 36. *For every δ,* $\mathsf{CT}_0 \vdash T[\delta] \leftrightarrow T[V_\delta]$.

Proof. We start by showing that provably in CT_0 all arithmetical partial truth predicates are coextensive, i.e., the following is provable in CT_0:

$$\forall x\ \forall\theta \in \mathsf{Form}^{\leq 1}_{\mathcal{L}_{\mathsf{PA}}}\ \forall\varphi \in \mathsf{Sent}_{\mathcal{L}_{\mathsf{PA}}}\big[\big(\mathsf{True}(x, T{*}\theta) \wedge \mathsf{depth}(\varphi) \leq x\big) \to \big(T{*}\theta(\varphi) \leftrightarrow T(\varphi)\big)\big]. \tag{$*$}$$

Fix an arbitrary $(\mathcal{M}, T) \models \mathsf{CT}_0$. For an arbitrary $c \in M$, let T_c denote the $((\mathcal{M}, T)$-definable) restriction of T to all sentences of depth at most c. Then

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

$(\mathcal{M}, T_c) \models \mathsf{True}(c, T)$. However, as proved in [20, Fact 32], $(\mathcal{M}, T_c)$ satisfies full induction scheme for $\mathcal{L}_T$. Hence T_c is a fully inductive truth predicate for formulae of depth at most c. Using this we argue that $(*)$ holds in $(\mathcal{M}, T)$. Working in the model, fix an arbitrary a and an arbitrary $\theta \in \mathsf{Form}^{\leq 1}_{\mathcal{L}_{\mathsf{PA}}}$. Assume that the depth of θ is b and let $c = \max\{a, b\}$. Assume $\mathsf{True}(a, T{*}\theta)$, i.e. the formula $T{*}\theta$ is a partial truth predicate for formulae of depth $\leq a$. Since for every formula φ of depth at most c, $T_c(\varphi)$ is equivalent to $T(\varphi)$, we conclude that $\mathsf{True}(a, T_c{*}\theta)$ holds. Moreover, it is sufficient to show that

$$\forall x \big(T_c{*}\theta(x) \leftrightarrow T_c(x)\big).$$

In other words, it is sufficient to prove that

$$(\mathcal{M}, T_c) \models \forall x \big(T{*}\theta(x) \leftrightarrow T(x)\big).$$

The above can be demonstrated by a routine induction on the build-up of formulae. More precisely, let

$$\Xi(y) := \forall \varphi \in \mathsf{depth}(y)\big(T{*}\theta(\varphi) \leftrightarrow T(\varphi)\big).$$

Then $\Xi(0)$ and $\forall x < a\big(\Xi(x) \rightarrow \Xi(x+1)\big)$ hold (in $(\mathcal{M}, T_c)$), because both $T_c{*}\theta$ and T_c are partial truth predicates for formulae of depth at most a. Since $\Xi(y)$ is a formula of $\mathcal{L}_T$, in $(\mathcal{M}, T_c)$ we have an induction axiom for it, and we can conclude

$$(\mathcal{M}, T_c) \models \forall y \leq a\ \Xi(y).$$

This completes the proof of $(*)$.

We show that over $\mathsf{CT}^{-}[\mathsf{PA}]$, $T[\delta]$ implies $T[V_\delta]$. We fix an arbitrary δ and working in CT_0 assume that $\forall x\big(\delta(x) \rightarrow T(x)\big)$. We show that T is V_δ-correct, i.e., for every arithmetical formula θ (possibly nonstandard) $T(V_\delta[\theta])$ holds. By the compositional conditions, $T(V_\delta[\theta])$ is equivalent to

$$\forall x \left(\mathsf{True}(x, T{*}\theta) \rightarrow \left[\forall y\big(\delta(y) \wedge \mathsf{pdepth}(y) \leq x\big) \rightarrow T{*}\theta(y)\big)\right]\right).$$

Fix x, assume $\mathsf{True}(x, T{*}\theta)$ and fix an arbitrary y such that $\mathsf{pdepth}(y) \leq x$ and $\delta(y)$. By δ-correctness $T(y)$ holds, hence y is a formula and since $\mathsf{pdepth}(y) \leq x$, y is a formula of depth at most x. Then, by the previous claim $(*)$ we know that for every φ whose depth is at most x, $T{*}\theta(\varphi)$ is equivalent to $T(\varphi)$. Hence $T{*}\theta(y)$ holds as well.

For the converse direction, we assume T is V_δ-correct. Fix an arbitrary x and assume that $\delta(x)$ holds. In particular x is a formula. Let y be the depth of x and let θ be any arithmetical truth predicate such that $\mathsf{Pr}_{\mathsf{PA}}(\mathsf{True}(\underline{y}, \theta))$ holds. By the Global Reflection in CT_0, $T(\mathsf{True}(\underline{y}, \theta))$ holds as well, and this in turn implies, by compositional conditions, $\mathsf{True}(y, \overline{T}{*}\theta)$. Consequently, by V_δ-correctness, $T{*}\theta(x)$ holds. Finally, it follows that $T(x)$ holds by our claim $(*)$. This concludes the proof of δ-correctness and the whole proof. ⊣

COROLLARY 37. *For every* δ, δ', *if* $\mathsf{CT}^{-}[\mathsf{EA}] \vdash T[V_\delta] \rightarrow T[V_{\delta'}]$, *then* $\mathsf{CT}_0 \vdash T[\delta] \rightarrow T[\delta']$.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

The above corollary yields a versatile tool for studying the structure of $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$, where $\leq_{\mathsf{CT}^-}$ is defined by: $\tau_1 \leq_{\mathsf{CT}^-} \tau_2$ iff $\mathsf{CT}^-[\mathsf{EA}] \vdash T[\tau_1] \rightarrow T[\tau_2]$. We show the crucial application:

THEOREM 38. *$\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ is a countably universal partial order.*

We recall that a partial order $\langle P, \leq_P\rangle$ is said to be countably universal if every countable partial order $\langle Q, \leq_Q\rangle$ can be embedded into $\langle P, \leq_P\rangle$, i.e., there is an injection $f : Q \rightarrow P$ such that for every $a, b \in P$, $f(a) \leq_P f(b) \iff a \leq_Q b$.

The above theorem reduces immediately to the one below.[11] This is thanks to the work of Hubička and Nešetřil [11, Corollary 2.6], where a particular countably universal partial order is defined. It is clear from the presentation that the order $\langle \mathcal{W}, \leq_{\mathcal{W}}\rangle$ is decidable and provably a partial order in PA.

THEOREM 39. *Suppose that $\preceq$ is a decidable partial order on ω such that* PA *proves that $\preceq$ is a partial order. Then there is an embedding $\langle \omega, \preceq\rangle \hookrightarrow \langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$.*

PROOF. Suppose that $\preceq$ satisfies the assumptions. Firstly, we build a family of consistent theories $\{\sigma_n\}_{n\in\omega}$ such that the following hold for all $m, n \in \omega$:

1. If $m \preceq n$, then $\mathsf{PA} \vdash \sigma_n \subseteq \sigma_m$.
2. $\mathsf{CT}_0 \nvdash \mathsf{Con}_{\sigma_m}$.
3. If $m \npreceq n$, then $\mathsf{CT}_0 + \mathsf{Con}_{\sigma_m} \nvdash \mathsf{Con}_{\sigma_n}$.

As shown in [21, Section 2.3, Theorem 11], there is a Π_1-formula $\pi(x)$ that is flexible over $\mathsf{REF}^{<\omega}(\mathsf{PA})$, i.e., for every Π_1-formula $\theta(x)$, the following theory is consistent:

$$\mathsf{REF}^{<\omega}(\mathsf{PA}) + \forall x\big(\pi(x) \leftrightarrow \theta(x)\big).$$

For each $n \in \omega$ let σ_n be the natural Σ_1-definition of the following set of sentences[12]:

$$\mathsf{PA} + \{\pi(\underline{k}) \mid n \preceq k\}.$$

Now, condition (1) easily follows from the (PA-provable) transitivity of $\preceq$. Condition (2) easily reduces to Condition (3), so let us now show the latter. Aiming at a contradiction assume $m \npreceq n$ and $\mathsf{CT}_0 \vdash \mathsf{Con}_{\sigma_m} \rightarrow \mathsf{Con}_{\sigma_n}$. Let $\theta(x) := m \preceq x$. By flexibility there exists model $\mathcal{M}$ such that

$$\mathcal{M} \models \mathsf{REF}^{<\omega}(\mathsf{PA}) + \forall x\big(\pi(x) \leftrightarrow \theta(x)\big).$$

By the choice of $\mathcal{M}$ it follows that $\mathcal{M} \models \neg\pi(n)$. As a consequence, by provable Σ_1-completeness of PA, $\mathcal{M} \models \mathsf{Pr}_{\mathsf{PA}}(\neg\pi(\underline{n}))$, and $\mathcal{M} \models \neg\mathsf{Con}_{\sigma_n}$. However, since $\mathcal{M} \models \mathsf{REF}(\mathsf{PA})$, as viewed in $\mathcal{M}$, PA is consistent with Π_1-truth (of $\mathcal{M}$). Consequently, since $\mathcal{M} \models \forall x\big(m \preceq x \rightarrow \pi(x)\big)$, it follows that $\mathcal{M} \models \mathsf{Con}_{\sigma_m}$. Hence $\mathcal{M} \models \mathsf{Con}_{\sigma_n}$ as well, which contradicts our previous conclusions.

We are ready to construct the promised embedding. Fix the family $\{\sigma_n\}_{n\in\omega}$ as above and for each $m \in \omega$ choose $\delta_m \in \Delta^*$ to be the natural elementary definition of the following set of sentences:

$$\mathsf{PA} + \{\mathsf{Con}_{\sigma_m}(\underline{n}) \mid n \in \omega\},$$

[11] We are grateful to Fedor Pakhomov for pointing our this more general result.

[12] Observe that since $\preceq$ need not be elementary; also σ_n need not be elementary either. However, σ is not our final axiomatization.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

where $\mathsf{Con}_{\sigma_m}(\underline{n})$ asserts that there is no proof of contradiction of σ_m with Gödel code $\leq n$. Since for every $m \in \omega$, σ_m is consistent, δ_m is really an axiomatization of PA, hence $V_{\delta_m} \in \mathsf{Sch}_{\mathsf{PA}}$. We check that the map

$$m \mapsto V_{\delta_m}$$

is an embedding of $\langle \omega, \preceq \rangle$ into $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-} \rangle$. Fix $m, n \in \omega$ and assume $m \preceq n$. Then clearly $\mathsf{PA} \vdash \forall x \big(\mathsf{Con}_{\sigma_m}(x) \to \mathsf{Con}_{\sigma_n}(x)\big)$. Consequently, applying Proposition 35 to $\varphi(x) := \mathsf{Con}_{\sigma_m}(x)$ and $\psi(x) := \mathsf{Con}_{\sigma_n}(x)$, we obtain

$$\mathsf{CT}^-[\mathsf{PA}] \vdash T[V_{\delta_m}] \to T[V_{\delta_n}].$$

Suppose now $m \not\preceq n$ and aiming at a contradiction, assume that $\mathsf{CT}^-[\mathsf{PA}] \vdash T[V_{\delta_m}] \to T[V_{\delta_n}]$. Then, by Corollary 37, $\mathsf{CT}_0 \vdash T[\delta_m] \to T[\delta_n]$. However, since $\mathsf{CT}_0 \vdash T[\delta_{\mathsf{PA}}]$, $\mathsf{CT}_0 \vdash T[\delta_i] \leftrightarrow \mathsf{Con}_{\sigma_i}$ for every $i \in \omega$. Hence $\mathsf{CT}_0 \vdash \mathsf{Con}_{\sigma_m} \to \mathsf{Con}_{\sigma_n}$, which is impossible by our previous considerations, since $m \not\preceq n$. ⊣

Corollary 40. *The following partial orders are countably universal* (*we take the ordering* $\leq_{\mathsf{CT}^-}$ *on* Cons *to be inherited from the Lindenbaum Algebra of* CT^-):

- $\langle \mathsf{Sch}^-_{\mathsf{PA}}, \leq_{\mathsf{CT}^-} \rangle$.
- $\langle \mathsf{Sch}^T_{\mathsf{PA}}, \leq_{\mathsf{CT}^-} \rangle$.
- $\langle \mathsf{Cons}, \leq_{\mathsf{CT}^-} \rangle$.

Proof. This follows since $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-} \rangle$ can be easily embedded into each of the above partial orders. ⊣

§4. Prudently correct axiomatizations. Recall (from Definition 17) that Δ is the collection of prudent axiomatizations of PA. In the first subsection we classify the extensions of PA that can be axiomatized by theories of the form $\mathsf{CT}^-[\![\delta]\!]$ and measure the complexity of the Tarski Boundary problem for such theories.

4.1. Universality and complexity. As indicated by the proposition below, theories of the form $\mathsf{CT}^-[\![\delta]\!]$ for $\delta \in \Delta$ are never too strong.

Proposition 41. *For every* $\delta \in \Delta$, $\mathsf{CT}_0 \vdash \mathsf{CT}^-[\![\delta]\!]$.

Proof. This follows immediately from Theorem 6 that $\mathsf{CT}_0 \vdash \forall \varphi \big(\mathsf{Pr}_{\mathsf{PA}}(\varphi) \to T(\varphi)\big)$. ⊣

Therefore, the theory CT_0 provides an upper-bound for the strength of theories in question. The following theorem is this section's main result.

Theorem 42. *For any r.e.* $\mathcal{L}_{\mathsf{PA}}$*-theory* $\mathcal{T}$ *extending* PA *such that* $\mathsf{CT}_0 \vdash \mathcal{T}$ *there exists a* $\delta \in \Delta$ *such that* $\mathcal{T}$ *and* $\mathsf{CT}^-[\![\delta]\!]$ *have the same arithmetical theorems.*

Proposition 41 and Theorem 42 when put together, yield the following characterization of arithmetical theories provable in $\mathrm{REF}^{<\omega}(\mathsf{PA})$.

Corollary 43. *For every arithmetical recursively enumerable theory* $\mathcal{T}$ *extending* PA *the following are equivalent*:

1. $\mathsf{REF}^{<\omega}(\mathsf{PA}) \vdash \mathcal{T}$.
2. *There exists a* $\delta \in \Delta$ *such that* $\mathcal{T}$ *and* $\mathsf{CT}^-[\![\delta]\!]$ *coincide on arithmetical theorems.*

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

To prove Theorem 42 we need to arrange δ such that

- $\delta \in \Delta$.
- $\mathsf{CT}^-[\![\delta]\!]$ does not overgenerate, i.e., its arithmetical consequences do not transcend those of $\mathcal{T}$.

To satisfy the first condition we recall that by (Cieśliński's) Theorem 6, uniform reflection over logic is an example of a principle which is provable in PA and whose "globalized" version is equivalent to CT_0. We shall often use the notation described in the following definition. We recall that, for heuristic reasons, we sometimes write $\ulcorner\varphi\urcorner$ to denote either the Gödel number of φ or the numeral naming this number (depending on the context).

Definition 44. For two sentences θ and φ, $\sigma_\varphi[\theta]$ abbreviates the sentence

$$(\mathsf{Pr}_\emptyset(\ulcorner\theta\urcorner) \wedge \neg\theta) \rightarrow \varphi.$$

The map $\langle\varphi,\theta\rangle \mapsto \sigma_\varphi[\theta]$ is clearly elementary and we shall identify it with its elementary definition.

To satisfy the second condition we could use Vaught's theorem on axiomatizability by a scheme, as we did earlier (see Remark 14). However, we prefer to introduce an original method of finding "deductively weak" axiomatizations of arithmetical theories. The very essence of our method was noted already in the original KKL-paper [14]: there are models of $\mathsf{CT}^-[\mathsf{PA}]$ in which nonstandard pleonastic disjunctions of obviously false statements are deemed true by the truth predicate. For example, if $\mathcal{M}$ is a countable recursively saturated model of PA and a is any nonstandard element, then there is a truth class $T \subseteq M$ such that $(\mathcal{M}, T) \models \mathsf{CT}^-[\mathsf{PA}]$ and the sentence

$$\underbrace{0 \neq 0 \vee (0 \neq 0 \vee (\ldots \vee 0 \neq 0)\ldots)}_{a \text{ many disjuncts}}$$

is deemed true by T. This phenomenon was quite recently pushed to the extreme by the following result of Bartosz Wcisło that appears in [4].

Theorem 45. *If $\mathcal{M} \models \mathsf{EA}$, then there is an elementary extension $\mathcal{N}$ of $\mathcal{M}$ that has an expansion $(\mathcal{N}, T) \models \mathsf{CT}^-[\mathsf{EA}]$, which has the property that every disjunction of nonstandard length in $\mathcal{N}$ is deemed true by T. Moreover, if $\mathcal{M} \models \mathsf{PA}$, then $(\mathcal{N}, T)$ can be taken to be a model of $\mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$.*

The above theorem provides us with a new method of finding finite conservative axiomatizations of arithmetical theories extending EA.

Definition 46. Given an arithmetical sentence φ, the *pleonastic disjunction* of φ is the sentence

$$\underbrace{\varphi \vee (\varphi \vee (\ldots \vee \varphi)\ldots)}_{\ulcorner\varphi\urcorner \text{ times}}.$$

The pleonastic disjunction of φ will be denoted with $\bigvee\varphi$.

Note that the above definition formalizes smoothly in EA (in which case φ is identified with $\ulcorner\varphi\urcorner$ and treated both as a number and as a formula) and that in a

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

nonstandard model of this theory $\bigvee \varphi$ has standard length if and only if φ is (coded by) a standard number.

PROPOSITION 47. *Every r.e.* $\mathcal{T} \supseteq \mathsf{EA}$ *can be finitely axiomatized by a theory of the form* $\mathsf{CT}^-[\mathsf{EA}] + T[\varphi]$, *for some elementary formula* $\varphi(x)$.

We recall that, by our conventions, $\mathsf{CT}^-[\mathsf{EA}] + T[\varphi]$ and $\mathsf{CT}^-[\![\varphi]\!]$ mean the same thing.

PROOF. Let $\varphi'(x)$ formalize an elementary axiomatization of $\mathcal{T}$ (which exists by Craig's trick). Define

$$\varphi(x) := \exists \psi < x\big(\varphi'(\psi) \wedge x = \bigvee \psi\big).$$

That is to say that x satisfies φ if it is a pleonastic disjunction of a formula from an elementary axiomatization of $\mathcal{T}$. Observe first that $\mathsf{CT}^-[\![\varphi]\!] \vdash \mathcal{T}$. Indeed, it is sufficient to show that for every sentence ψ we have

$$\mathsf{CT}^-[\![\varphi]\!] \vdash \varphi'(\psi) \rightarrow T(\psi).$$

Observe that, over EA, $\varphi'(\psi)$ implies $\varphi(\bigvee \psi)$, which in turn, over $\mathsf{CT}^-[\![\varphi]\!]$ implies $T(\bigvee \psi)$. However, over pure $\mathsf{CT}^-[\mathsf{EA}]$ the last sentence implies $T(\psi)$ by compositional conditions, since $\bigvee \psi$ is a disjunction of length $\ulcorner \psi \urcorner$ and hence is standard.

We show conservativity: pick any model $\mathcal{M} \models \mathcal{T}$. By Theorem 45 there is a $(\mathcal{N}, T) \models \mathsf{CT}^-[\mathsf{EA}]$ such that $\mathcal{N}$ is an elementary extension of $\mathcal{M}$ and every disjunction of nonstandard length is made true by T. It follows that $(\mathcal{N}, T) \models \mathsf{CT}^-[\![\varphi]\!]$. Indeed, firstly observe that if $\mathcal{N} \models \varphi(a)$ then there exists ψ such that $\mathcal{N} \models \varphi'(\psi) \wedge a = \bigvee \psi$. Now the argument splits into two cases:

1. ψ is a standard sentence. In this case $\bigvee \psi$ is standard and $\mathcal{N} \models \bigvee \psi$, by elementarity. Consequently $(\mathcal{N}, T) \models T(\bigvee \psi)$ by compositional clauses; or
2. ψ is not a standard sentence. In this case $\bigvee \psi$ is a disjunction of nonstandard length, hence is made true in $(\mathcal{N}, T)$. ⊣

We shall recycle the above conservativity argument in the proof of Theorem 42, which we now turn to.

PROOF OF THEOREM 42. Fix $\mathcal{T}$ such that

$$\mathsf{CT}_0 \vdash \mathcal{T}.$$

Let ϱ be an arbitrary elementary axiomatization of $\mathcal{T}$. Let $\sigma_\varphi[\theta]$ denote the map from Definition 44. We now observe that the proof of the reflexive property of PA is formalizable in $\mathrm{I}\Sigma_1$. This follows from two well-known facts: firstly, the cut-elimination theorem formalizes in $\mathrm{I}\Sigma_1$ (thus provably in $\mathrm{I}\Sigma_1$, every provable sentence has a proof that has the subformula property) and secondly, $\mathrm{I}\Sigma_1$ is enough for the formalization of the proof of existence of partial truth predicates in PA. As a consequence, we obtain

$$\mathrm{I}\Sigma_1 \vdash \forall \theta\ \mathsf{Pr}_{\mathsf{PA}}(\neg(\mathsf{Pr}_\emptyset(\ulcorner \theta \urcorner) \wedge \neg \theta)).$$

Hence $\delta(x) \in \Delta$, where δ is defined as follows:

$$\delta(x) := \delta_{\mathsf{PA}}(x) \vee \exists \theta, \varphi < x\big(\varrho(\varphi) \wedge x = \sigma_{\bigvee \varphi}[\theta]\big).$$

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

Observe that δ naturally defines the following set of sentences:

$$\mathsf{Q} \cup \{\mathsf{Ind}(\varphi) \mid \varphi \in \mathcal{L}_{\mathsf{PA}}\} \cup \left\{(\mathsf{Pr}_\emptyset(\ulcorner\theta\urcorner) \wedge \neg\theta) \rightarrow \bigvee\varphi \mid \theta \in \mathcal{L}_{\mathsf{PA}}, \varphi \in \mathcal{T}\right\}.$$

We argue first that $\mathsf{CT}^-[\![\delta]\!]$ is conservative over $\mathcal{T}$. To see this, fix an arbitrary model $\mathcal{M} \models \mathcal{T}$ and let $(\mathcal{N}, T) \models \mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$ be a model from Theorem 45. Then $(\mathcal{N}, T) \models \mathsf{CT}^-[\![\delta]\!]$ since, reasoning by cases as in the proof of Proposition 47, for every $\varphi \in N$ such that $\mathcal{N} \models \varrho(\varphi)$ we have

$$(\mathcal{N}, T) \models T\left(\bigvee\varphi\right).$$

Now, we argue that $\mathsf{CT}^-[\![\delta]\!] \vdash \mathcal{T}$. Let φ be an arbitrary ϱ–axiom of $\mathcal{T}$. We claim

$$\mathsf{CT}^-[\![\delta]\!] \vdash \varphi.$$

To see why the last claim holds, reason in $\mathsf{CT}^-[\![\delta]\!]$. We have

$$\forall\theta\ T(\sigma_{\bigvee\varphi}[\theta]).$$

By the axioms of CT^- the above is equivalent to

$$\big(\exists\theta\big(\mathsf{Pr}_\emptyset(\ulcorner\theta\urcorner) \wedge \neg T(\theta)\big)\big) \rightarrow T\left(\bigvee\varphi\right). \tag{$*$}$$

Now we reason by cases: either $\forall\theta\ \big(\mathsf{Pr}_\emptyset(\ulcorner\theta\urcorner) \rightarrow T(\theta)\big)$ or not. If the latter holds, we have $T(\bigvee\varphi)$ by Modus Ponens applied to $(*)$. Hence φ holds by compositional conditions, because $\bigvee\varphi$ is a disjunction of standard length and φ is a standard sentence. If the former holds, we have CT_0 by Theorem 6 and φ holds, because we assumed that $\mathsf{CT}_0 \vdash \mathcal{T}$. ⊣

We conclude this subsection with complexity results that complement Theorem 26.

Proposition 48. *The set Δ^* is Π_2-complete.*

Proof. Clearly Δ^* is Π_2-definable. Consider the map f that takes a scheme template τ as input and outputs the formula $\delta_\tau(x)$ that expresses "x is an instance of τ." f is clearly recursive (indeed elementary) and satisfies

$$\tau \in \mathsf{Sch}_{\mathsf{PA}} \text{ iff } \delta_\tau \in \Delta^*.$$

Therefore $\mathsf{Sch}_{\mathsf{PA}}$ is many-one reducible to Δ^*, which in light of the Π_2-completeness of $\mathsf{Sch}_{\mathsf{PA}}$ (established in Theorem 26), completes the verification of Π_2-completeness of Δ^*. ⊣

Proposition 49. *The sets Δ and Δ^- are both Σ_1-complete.*

Proof. Both sets are clearly r.e., because both definitions require just the provability of a particular sentence in $\mathrm{I}\Sigma_1$. To verify completeness we sketch the reduction f of the set of true Σ_1-sentences to Δ (an analogous reduction works for Δ^-). Given a Σ_1-sentence φ we define a formula $f(\varphi) := \delta_{\mathsf{PA}}(x) \vee x = \varphi$. It follows that

$$\mathbb{N} \models \varphi \iff f(\varphi) \sim_{pt} \delta_{\mathsf{PA}}.$$

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

The left-to-right direction follows by $\mathrm{I}\Sigma_1$-provable Σ_1-completeness of PA. The right-to-left direction follows by the soundness of $\mathrm{I}\Sigma_1$ and PA. ⊣

THEOREM 50. *The set* $\{\delta \in \Delta \mid T[\delta] \in \mathsf{Cons}\}$ *is* Π_2*-complete.*

In what follows, Π_2-REF(PA) denotes the extension of EA with all sentences of the form

$$\forall x\big(\mathsf{Pr}_{\mathsf{PA}}(\ulcorner \varphi(\underline{x})\urcorner) \to \varphi(x)\big)$$

for $\varphi(x) \in \Pi_2$. It is a folklore result [1] that this theory is finitely axiomatizable. We need the following folklore lemma, proved, e.g., in [23]:

LEMMA 51. $\mathsf{PA} + \neg\Pi_2\text{-REF}(\mathsf{PA})$ *is* Π_2*-sound.*

PROOF OF THEOREM 50. Fix a Π_2-sentence $\pi := \forall x\varphi(x)$, where $\varphi(x)$ is Σ_1. Let δ^π be the formula in Δ that describes the union of (the canonical axiomatization of) PA with the following set of sentences:

$$\left\{\mathsf{Pr}_{\emptyset}(\ulcorner\chi\urcorner) \wedge \neg\chi \to \bigvee \varphi(\underline{n}) \;\mid\; \chi \in \mathcal{L}_{\mathsf{PA}}, n \in \omega\right\}.$$

The function $\pi \mapsto \delta^\pi$ is clearly recursive, and $\delta^\pi \in \Delta$. Let $\theta(x) := \Pi_2\text{-REF}(\mathsf{PA}) \vee \varphi(x)$ and observe that for every n, $\mathsf{CT}^-[\![\delta^\pi]\!] \vdash \theta(\underline{n})$. Indeed, work in $\mathsf{CT}^-[\![\delta^\pi]\!]$ and assume $\neg\Pi_2$-REF(PA). Then clearly $\neg\mathsf{CT}_0$ and consequently, as in the proof of Theorem 42 we get $T\left(\bigvee \varphi(\underline{n})\right)$. Finally, the latter implies $\varphi(\underline{n})$, since it is a standard sentence.

Let $\mathsf{True}^{\mathbb{N}}_{\Pi_2}$ be the set of Π_2-statements that are true in $\mathbb{N}$. We will prove

$$\pi \in \mathsf{True}^{\mathbb{N}}_{\Pi_2} \iff \mathsf{CT}^-[\![\delta^\pi]\!] \text{ is conservative over PA.}$$

Assume first that $\pi \in \mathsf{True}^{\mathbb{N}}_{\Pi_2}$ and $\pi = \forall x\, \varphi(x)$, for some $\varphi(x) \in \Sigma_1$. In particular $\varphi(\underline{n})$ is a true Σ_1 sentence for every $n \in \omega$, hence,

$$\mathsf{PA} \vdash \varphi(\underline{n}) \text{ for every } n \in \omega.$$

As usual, fix any model $\mathcal{M} \models \mathsf{PA}$ and take its elementary extension $(\mathcal{N}, T) \models \mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$ in which every disjunction of nonstandard length is true. As previously, it follows that $(\mathcal{N}, T) \models \mathsf{CT}^-[\![\delta^\pi]\!]$.

Conversely, assume that $\mathsf{CT}^-[\![\delta^\pi]\!]$ is conservative over PA. Then for every $n \in \omega$, $\mathsf{PA} \vdash \theta(\underline{n})$. In particular, for every $n \in \omega$, $\mathsf{PA} + \neg\Pi_2\text{-REF}(\mathsf{PA}) \vdash \varphi(\underline{n})$. By the soundness of this theory we conclude that π is true. ⊣

4.2. Structure of prudent axiomatizations. Theorem 42 allows us to transfer results about the fragment of the Lindenbaum algebra of PA consisting of sentences provable in CT_0 to results about the structure of Tarski Boundary. Let us isolate the former structure: put

$$\mathsf{CT}_0/\mathsf{PA} := \{[\varphi]_{\mathsf{PA}} \mid \varphi \in \mathcal{L}_{\mathsf{PA}} \wedge \mathsf{CT}_0 \vdash \varphi\},$$

where $[\varphi]_{\mathsf{PA}}$ denotes φ-equivalence class modulo PA-provable equivalence, i.e., the element of the Lindenbaum algebra of PA that contains φ. Then, it is fairly easy to see that the following holds:

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

OBSERVATION 52. $\mathsf{CT}_0/\mathsf{PA}$ *with the operations inherited from the Lindenbaum algebra of* PA *is a lattice with a greatest but not a least element (obviously assuming the consistency of* CT_0*). The lack of the least element follows from the fact that the arithmetical consequences of* CT_0 *is a theory in* $\mathcal{L}_{\mathsf{PA}}$ *which extends* PA. *In particular this theory is not finitely axiomatizable by essential reflexivity of* PA. *Moreover the greatest element has no immediate predecessors. This follows by the classical fact that the Lindenbaum Algebra of* PA *is dense (see* [27]*) for a proof.*

The following is an easy corollary to Theorem 42.

PROPOSITION 53. *There exists a lattice embedding* $\mathsf{CT}_0/\mathsf{PA} \hookrightarrow \langle \Delta, \leq_{\mathsf{CT}^-} \rangle$.

PROOF. To each $[\varphi]_{\mathsf{PA}}$ we assign $\delta^\varphi \in \Delta$ as in the proof of Theorem 42. $\varrho(x)$ is now simply $x = \underline{\ulcorner \varphi \urcorner}$. Hence by compositional axioms, and the fact that φ is standard we have

$$\mathsf{CT}^-[\mathsf{EA}] \vdash \forall \theta \big(T(\sigma_{\bigvee \varphi}[\theta]) \leftrightarrow T(\sigma_\varphi[\theta]) \big).$$

Consequently, δ^φ can be taken to axiomatize the (natural definition of the) following set of sentences:

$$\mathsf{PA} \cup \{ (\mathsf{Pr}_\emptyset(\ulcorner \theta \urcorner) \wedge \neg \theta) \to \varphi \mid \theta \in \mathcal{L}_{\mathsf{PA}} \}.$$

We claim that for an arbitrary $\varphi \in \mathcal{L}_{\mathsf{PA}}$, over $\mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$, $\mathsf{CT}^-[\![\delta^\varphi]\!]$ is equivalent to φ. Working in $\mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$ assume first that φ holds. Then for every θ we have

$$T(\sigma_\varphi[\theta]),$$

since $T(\sigma_\varphi[\theta])$ is equivalent to an implication with a true conclusion. Hence every sentence satisfying δ^φ is true. For the converse implication, working over $\mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$, assume $\mathsf{CT}^-[\![\delta^\varphi]\!]$. We argue by cases:

- If CT_0 holds, then φ holds, by assumption.
- If CT_0 fails, then, as in the proof of Theorem 42, φ holds.

We show that the mapping $\varphi \mapsto \delta^\varphi$ is a lattice embedding. Firstly, we show that the mapping preserves the partial ordering. To this end, we prove that the following are equivalent for arbitrary arithmetical formulae φ, ψ that are provable in CT_0:

- $\mathsf{PA} \vdash \varphi \to \psi$.
- $\mathsf{CT}^-[\mathsf{PA}] \vdash T[\delta^\varphi] \to T[\delta^\psi]$.

Indeed the top-to-bottom direction follows easily, since for an arbitrary $\varphi \in \mathcal{L}_{\mathsf{PA}}$, $\mathsf{CT}^-[\![\delta^\varphi]\!]$ is equivalent to φ and $\mathsf{CT}^-[\![\delta^\varphi]\!]$ proves $\mathsf{CT}^-[\![\delta_{\mathsf{PA}}]\!]$. The bottom-to-up direction uses the same observations plus additionally the conservativity of $\mathsf{CT}^-[\mathsf{PA}]$ over PA. Finally, we show that the mapping preserves infima and suprema. It is enough to observe that for φ and ψ as above, $\mathsf{CT}^-[\![\delta^{\varphi\vee\psi}]\!]$ is equivalent to $\mathsf{CT}^-[\![\delta^\varphi]\!] \vee \mathsf{CT}^-[\![\delta^\psi]\!]$ and the same with $\wedge$. This concludes the proof. ⊣

The next proposition slightly lies on the margins of our considerations as it does not concern *axiomatizations* of PA, but rather concerns *the set of theorems* of PA. However, we include it, since it reveals an interesting feature of the Tarski Boundary.

PROPOSITION 54. *There is an embedding* $\iota : \mathsf{CT}_0/\mathsf{PA} \hookrightarrow \langle \Delta^-, \leq_{\mathsf{CT}^-} \rangle$ *that is cofinal in the region below (i.e., the nonconservative side of) the Tarski Boundary. More*

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

precisely, for every $\alpha \in \mathcal{L}_T$ *such that* $\mathsf{CT}^-[\mathsf{PA}] + \alpha$ *is non-conservative over* PA, *there is an* $a \in \mathsf{CT}_0/\mathsf{PA}$ *such that* $T[\imath(a)]$ *is strictly above* α (*i.e., is logically weaker*) *and* $\mathsf{CT}^-[\mathsf{PA}] + T[\imath(a)]$ *is non-conservative over* PA.

PROOF. The embedding $\imath$ is defined as in the proof of the previous proposition with the only exception that we do not add PA to δ^{ψ}. More concretely, if $[\varphi]_{\mathsf{PA}} \in \mathsf{CT}_0/\mathsf{PA}$, then we put $\imath([\varphi]_{\mathsf{PA}})$ to be the natural elementary definition of the following set of sentences:

$$\{(\mathsf{Pr}_{\emptyset}(\ulcorner\theta\urcorner) \wedge \neg\theta) \rightarrow \varphi \ \mid \ \theta \in \mathcal{L}_{\mathsf{PA}}\}.$$

Denote the canonical elementary definition of this set with δ^{φ}. As in the proof of the previous proposition, we obtain that for every $[\varphi]_{\mathsf{PA}} \in \mathsf{CT}_0/\mathsf{PA}$, provably in $\mathsf{CT}^-[\mathsf{PA}]$, φ is equivalent to $T[\delta^{\varphi}]$. Consequently, $\imath$ is a lattice embedding. Now we claim that $\imath$ is cofinal with the Tarski Boundary in the sense explained. Pick any $\alpha \in \mathcal{L}_T$ such that $\mathsf{CT}^-[\mathsf{PA}] + \alpha$ is non-conservative over PA (but consistent). By definition, $\mathsf{CT}^-[\mathsf{PA}] + \alpha \vdash \varphi$ for some PA - unprovable sentence $\varphi \in \mathcal{L}_{\mathsf{PA}}$. Then, since the Lindenbaum algebra of PA is atomless there is a sentence $\psi \in \mathcal{L}_{\mathsf{PA}}$, which is logically strictly weaker than φ. Then there is a sentence θ such that $[\theta]_{\mathsf{PA}} \in \mathsf{CT}_0/\mathsf{PA}$ and $\psi \vee \theta$ is unprovable in PA. This holds, since it is known that over PA, REF(PA) (which is a consequence of CT_0) does not follow from any finite, consistent, set of sentences. Hence $[\psi \vee \theta]_{\mathsf{PA}} \in \mathsf{CT}_0/\mathsf{PA}$ is not the greatest element. Consequently, $T[\imath(\psi \vee \theta)] = T[\delta^{\psi\vee\theta}]$ is below the Tarski Boundary. However, since ψ does not prove φ (over PA), *a fortiori* $\psi \vee \theta$ does not prove φ. Hence $\mathsf{CT}^-[\![\delta^{\psi\vee\theta}]\!]$ does not prove $\mathsf{CT}^-[\mathsf{PA}] + \alpha$. Additionally, $\mathsf{CT}^-[\mathsf{PA}] + \alpha \vdash \mathsf{CT}^-[\![\delta^{\psi\vee\theta}]\!]$, since $\psi \vee \theta$ follows from α. ⊣

PROPOSITION 55. *There are recursive infinite antichains in* $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$.

PROOF. We shall make use of a Π_1-formula that is PA-independent, i.e., for every binary sequence s of length $n \in \omega$ the following sentence is unprovable in PA:

$$\left(\pi(\underline{0})^{s(0)} \wedge \pi(\underline{1})^{s(1)} \wedge \ldots \wedge \pi(\underline{n-1})^{s(n-1)}\right),$$

where for any formula φ, $\varphi^0 := \varphi$, and $\varphi^1 := \neg\varphi$. We will use the construction of such a Π_1-formula described in [21, Theorem 9, Chapter 2]. Let $\pi(x)$ be such a formula. Assuming that each $\pi(\underline{k})$ is provable in CT_0, $\{\pi(\underline{k})\}_{k\in\omega}$ is an infinite antichain in $\mathsf{CT}_0/\mathsf{PA}$. By Proposition 53 this implies that $\{\delta^{\pi(\underline{k})}\}_{k\in\omega}$ is an infinite antichain in Δ. These considerations show that it suffices to verify

$$\mathsf{CT}_0 \vdash \pi(\underline{k}), \text{ for each } k \in \omega. \qquad (*)$$

The verification of $(*)$ is a straightforward formalization of the reasoning in [21, Theorem 9, Chapter 2], so it is delegated to the Appendix. ⊣

PROPOSITION 56. *There is an embedding* $(\mathbb{Q}, <) \hookrightarrow \langle \Delta, \leq_{\mathsf{CT}^-}\rangle$.

PROOF. This is an immediate consequence of the existence of an embedding $(\mathbb{Q}, <) \hookrightarrow \mathsf{CT}_0/\mathsf{PA}$, which in turn follows from the well-known fact that the Lindenbaum Algebra of PA is dense (see [27] for a proof). ⊣

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

PROPOSITION 57. *There are* $\delta, \delta' \in \Delta$ *such that* $\mathsf{CT}^-[\![\delta]\!]$ *and* $\mathsf{CT}^-[\![\delta']\!]$ *are non-conservative extensions of* PA, *but* $\mathsf{CT}^-[\![\delta]\!] \vee \mathsf{CT}^-[\![\delta']\!]$ *is a conservative extension of* PA.

PROOF. Consider $\varphi := \mathsf{Con}_{\mathsf{PA}+\neg\mathsf{Con}_{\mathsf{PA}}}$ and $\psi := \mathsf{Con}_{\mathsf{PA}} \to \mathsf{Con}_{\mathsf{PA}+\mathsf{Con}_{\mathsf{PA}}}$. Both φ and ψ generate different non-zero elements in $\mathsf{CT}_0/\mathsf{PA}$ but it is easy to see that

$$\mathsf{PA} \vdash \varphi \vee \psi.$$

Hence the desired $\delta, \delta' \in \Delta$ can be chosen as $\delta := \delta^\varphi$ and $\delta' := \delta^\psi$ (defined as in the proof of Proposition 53). ⊣

§5. Coda: The arithmetical reach of $\mathsf{CT}^-[\![\delta]\!]$ for $\delta \in \Delta^*$. Recall from Definition 17 that Δ^* is the collection of elementary presentations of PA, i.e., elementary formulae that define (in $\mathbb{N}$) a theory that is deductively equivalent to PA. We are now in a position to fulfill our promise given in the introduction and characterize the set denoted sup PA of arithmetical sentences that are provable in some theory of the form $\mathsf{CT}^-[\![\delta]\!]$, where $\delta \in \Delta^*$.

THEOREM 58. sup PA *is deductively equivalent to* $\mathsf{True}^{\mathbb{N}}_{\Pi_2} + \mathsf{REF}^{<\omega}(\mathsf{PA})$.

PROOF. First note that $\mathsf{REF}^{<\omega}(\mathsf{PA}) \subseteq \sup \mathsf{PA}$ is an immediate corollary to Theorem 42. Also, the proof of $\mathsf{True}^{\mathbb{N}}_{\Pi_2} \subseteq \sup \mathsf{PA}$ is morally contained in the proof of Theorem 26: for every true Π_2-sentence $\pi := \forall x \exists y \varphi(x, y)$, the theory

$$\mathsf{PA} \cup \{\exists y \varphi(\underline{n}, y) \mid n \in \omega\}$$

is deductively equivalent to PA, hence the natural arithmetical definition of the above set witnesses that $\sup \mathsf{PA} \vdash \pi$. To prove the converse inclusion[13], assume that for some $\delta \in \Delta$, $\mathsf{CT}^-[\![\delta]\!] \vdash \varphi$. Let π be the true Π_2-sentence

$$\forall x \big(\mathsf{Pr}_\delta(x) \to \mathsf{Pr}_{\mathsf{PA}}(x)\big),$$

expressing that every theorem of δ is provable already in PA. Then it is easy to observe that

$$\mathsf{CT}^-[\mathsf{PA}] + \pi + \mathsf{GRP}(\mathsf{PA}) \vdash \varphi.$$

However, by any of the proofs of Theorem 5, the theory $\mathsf{CT}^-[\mathsf{PA}] + \pi + \mathsf{GRP}(\mathsf{PA})$ is arithmetically conservative over $\mathsf{EA} + \mathsf{REF}^{<\omega}(\mathsf{PA}) + \pi$.[14] Hence $\mathsf{EA} + \mathsf{REF}^{<\omega}(\mathsf{PA}) + \pi \vdash \varphi$. Since $\mathsf{EA} + \pi$ is a true Π_2-sentence the proof is complete. ⊣

§6. Open problems.

(I) Are the lattices $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ and $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$ dense? Does $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$ have maximal or minimal elements? Does $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ have minimal elements (by the proof of Theorem 30 no $\leq_{\mathsf{CT}^-}$-maximal element exists)?

---

[13]This proof is due to Fedor Pakhomov and appears here with his kind permission.

[14]The crucial lemma in all the known proofs states that for every model $\mathcal{M} \models \mathsf{REF}^{<\omega}(\mathsf{PA})$ there is a model $\mathcal{N}$ which is elementarily equivalent to $\mathcal{M}$ and $T \subseteq N$ such that $(\mathcal{N}, T) \models \mathsf{CT}^-[\mathsf{PA}] + \mathsf{GRP}(\mathsf{PA})$.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

(II) Are the lattices $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ and $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$ universal for countable distributive lattices?[15]

(III) How do $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ and $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$ fit in the Lindenbaum algebra of $\mathsf{CT}^-[\mathsf{EA}]$?

(IV) Is the Lindenbaum algebra of Cons dense?

(V) Do $\langle \mathsf{Sch}_{\mathsf{PA}}, \leq_{\mathsf{CT}^-}\rangle$ and $\langle \Delta, \leq_{\mathsf{CT}^-}\rangle$ have decidable copies? If not, how undecidable are they?

(VI) How close can we get to the Tarski Boundary *from below* using theories $\mathsf{CT}^-[\![\delta]\!]$, where $\delta \in \Delta$? In other words, if $\mathsf{CT}^-[\mathsf{PA}] + \alpha$ is nonconservative over PA, is there some $\delta \in \Delta$ such that $\mathsf{CT}^-[\![\delta]\!]$ is nonconservative over PA, and $\mathsf{CT}^-[\mathsf{PA}] + \alpha \vdash T[\delta]$?

(VII) How close can we get to the Tarski Boundary *from above* using theories $\mathsf{CT}^-[\![\delta]\!]$, where $\delta \in \Delta$? In other words, if $\mathsf{CT}^-[\mathsf{PA}] + \alpha$ is conservative over PA, is there some $\delta \in \Delta$ such that $\mathsf{CT}^-[\![\delta]\!]$ is conservative over PA, and $\mathsf{CT}^-[\mathsf{PA}] + T[\delta] \vdash \alpha$?

(VIII) Do the answers to Questions (VI) and (VII) change if $\mathsf{CT}^-[\![\delta]\!]$ is required to be a subtheory of CT_0?

§7. Appendix.

Proof Verification of $(*)$ of the proof of Proposition 55. To lighten the notation, we will identify numerals with their denotations, and formulae with their codes. We wish to show that if $\pi(x)$ is the Π_1-formula $\pi(x)$ constructed in [21, Theorem 9, Chapter 2], then for every $k \in \omega$, $\mathsf{CT}_0 \vdash \pi(k)$. Let us revisit the construction of $\pi(x)$. Given a finite binary sequence s of length n, and a unary arithmetical formula $\varphi(x)$, let φ^s abbreviate the following sentence:

$$\big(\varphi(0)^{s(0)} \wedge \varphi(1)^{s(1)} \wedge \cdots \wedge \varphi(n-1)^{s(n-1)}\big).$$

For a unary formula φ, let $\varrho(x, i, \varphi, p)$ express:

there is a binary sequence s of length $x+1$ such that $s(x) = i$ an p is a proof in PA of $\neg\varphi^s$.

Finally, let $\pi(x)$ be a formula assured to exist by the diagonal lemma such that the following is provable in PA:

$$\pi(x) \leftrightarrow \forall p\big(\varrho(x, 1, \pi, p) \rightarrow \exists q \leq p\ \varrho(x, 0, \pi, q)\big).$$

By metainduction on $n \in \omega$, we show that for every $n \in \omega$, $\mathsf{CT}_0 \vdash \big(\mathsf{len}(s) = n+1 \rightarrow \neg\mathsf{Pr}_{\mathsf{PA}}(\neg\pi^s)\big)$. Observe that this implies that for every $n \in \omega$, $\pi(n)$ is provable in CT_0. We first show that $\pi(0)$ is provable in CT_0. Working in CT_0, assume that $\neg\pi(0)$ holds. It follows that for some p, $\varrho(0, 1, \pi, p)$ holds, hence in particular, $\mathsf{Pr}_{\mathsf{PA}}(\pi(0))$ holds. However, in CT_0 the theorems of PA are true, so $\pi(0)$ holds, contrary to the assumption. Hence $\mathsf{CT}_0 \vdash \neg\mathsf{Pr}_{\mathsf{PA}}(\neg\pi(0))$. Moreover, since $\pi(0)$ holds, for every PA-proof of $\pi(0)$ there exists a smaller PA-proof of $\neg\pi(0)$. Consequently, since CT_0 proves the consistency of PA, for $n = 0$, $\mathsf{CT}_0 \vdash \forall s\big(\mathsf{len}(s) = n+1 \rightarrow \neg\mathsf{Pr}_{\mathsf{PA}}(\neg\pi^s)\big)$.

Now, assume $n = k+1$, $\mathsf{CT}_0 \vdash \forall s\big(\mathsf{len}(s) = n \rightarrow \neg\mathsf{Pr}_{\mathsf{PA}}(\neg\pi^s)\big)$. Working in CT_0 assume for some s of length $n+1$, $\mathsf{Pr}_{\mathsf{PA}}(\neg\pi^s)$. Fix s such that the proof of π^s in PA is

[15]This question was communicated to us by Fedor Pakhomov.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

the least possible (among s's of length $n+1$). Denote (the code of) this proof with p. We distinguish two cases:

1. $s(n) = 0$. Then, by the definition of π^s, we have $\mathsf{Pr}_{\mathsf{PA}}(\pi^{s\restriction n} \rightarrow \neg\pi(n))$. Moreover, both $\varrho(n, 0, \pi, p)$ and $\forall q \leq p\ \neg\varrho(n, 1, \pi, q)$ hold. Since ϱ is a Δ_0-formula, we have

$$\mathsf{Pr}_{\mathsf{PA}}\big(\varrho(n, 0, \pi, p) \wedge \forall q \leq p\ \neg\varrho(n, 1, \pi, q)\big).$$

   In particular, $\mathsf{Pr}_{\mathsf{PA}}(\pi(n))$. Hence $\mathsf{Pr}_{\mathsf{PA}}(\neg\pi^{s\restriction n})$, which is impossible by the induction step, since $s\restriction_n$ has length n.
2. $s(n) = 1$. Then, as before, $\mathsf{Pr}_{\mathsf{PA}}(\pi^{s\restriction n} \rightarrow \pi(n))$. Moreover, by minimality of p, we have $\varrho(n, 1, \pi, p)$ and $\forall q < p\ \neg\varrho(n, 0, \pi, q)$. Hence, as before we obtain $\mathsf{Pr}_{\mathsf{PA}}(\neg\pi(n))$, which contradicts the induction assumption.

This concludes the proof of the induction step and the whole proof. ⊣

§8. Acknowledgements. We have both directly and indirectly benefitted from conversations with several colleagues concerning the topics explored in this paper, including (in reverse alphabetical order of last names) Bartosz Wcisło, Albert Visser, Fedor Pakhomov, Carlo Nicolai, Roman Kossak, Cezary Cieśliński, Lev Beklemishev, and Athar Abdul-Quader. The research presented in this paper was supported by the National Science Centre, Poland (NCN; Grant Number 2019/34/A/HS1/00399).

REFERENCES

[1] L. Beklemishev, *Reflection principles and provability algebras in formal arithmetic*. ***Russian Mathematical Surveys***, vol. 60 (2005), no. 2, pp. 197–268.

[2] C. Cieśliński, *Deflationary truth and pathologies*. ***Journal of Philosophical Logic***, vol. 39 (2010), no. 3, pp. 325–337.

[3] ———, ***The Epistemic Lightness of Truth: Deflationism and its Logic***, Cambridge University Press, Cambridge, 2018.

[4] C. Cieśliński, M. Łełyk, and B. Wcisło, *The two halves of disjunctive correctness*. ***Journal of Mathematical Logic*** (2022), https://doi.org/10.1142/S021906132250026X.

[5] A. Enayat, M. Łełyk, and B. Wcisło, *Truth and feasible reducibility*, ***Journal of Symbolic Logic***, vol. 85 (2020), pp. 367–421.

[6] A. Enayat and F. Pakhomov, *Truth, disjunction, and induction*. ***Archive for Mathematical Logic***, vol. 58 (2019), nos. 5–6, pp. 753–766.

[7] A. Enayat and A. Visser, *New constructions of satisfaction classes*, ***Unifying the Philosophy of Truth*** (T. Achourioti, H. Galinon, J. M. Fernández, and K. Fujimoto, editors), Springer, Dordecht, 2015, pp. 321–325.

[8] S. Feferman, *Reflecting on incompleteness*, ***Journal of Symbolic Logic***, vol. 56 (1991), no. 1, pp. 1–49.

[9] P. Hájek and P. Pudlák, ***Metamathematics of First-Order Arithmetic***, Springer, Berlin, 1993.

[10] V. Halbach, ***Axiomatic Theories of Truth***, Cambridge University Press, Cambridge, 2011.

[11] J. Hubička and J. Nešetřil, *Some examples of universal and generic partial orders*, ***Model Theoretic Methods in Finite Combinatorics***, Contemporary Mathematics, vol. 558, American Mathematical Society, Providence, 2011, pp. 293–317.

[12] R. Kossak and B. Wcisło, *Disjunctions with stopping conditions*. ***Bulletin of Symbolic Logic***, vol. 27 (2021), no. 3, pp. 231–253.

[13] H. Kotlarski, *Bounded induction and satisfaction classes*. ***Zeitschrift für matematische Logik und Grundlagen der Mathematik***, vol. 32 (1986), pp. 531–544.

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press

[14] H. Kotlarski, S. Krajewski, and A. Lachlan, *Construction of satisfaction classes for nonstandard models*. ***Canadian Mathematical Bulletin***, vol. 24 (1981), pp. 283–293.

[15] G. Leigh, *Conservativity for theories of compositional truth via cut elimination*, ***Journal of Symbolic Logic***, vol. 80 (2015), no. 3, pp. 845–865.

[16] M. Łełyk, ***Axiomatic theories of truth, bounded induction and reflection principles***, Ph.D. thesis, University of Warsaw, 2017.

[17] ———, *Model theory and proof theory of the global reflection principle*. ***The Journal of Symbolic Logic,*** First View, (2022), pp. 1–42, https://doi.org/10.1017/jsl.2022.39.

[18] ———, *Axiomatizations of Peano arithmetic: A truth-theoretic perspective*, pt. 2, in preparation, forthcoming.

[19] M. Łełyk and B. Wcisło, *Models of positive truth*. ***Review of Symbolic Logic***, vol. 12 (2018), pp. 144–172.

[20] ———, *Local collection and end-extensions of models of compositional truth*. ***Annals of Pure and Applied Logic***, vol. 172 (2021), no. 6, Article no. 102941.

[21] P. Lindström, ***Aspects of Incompleteness***, Lecture Notes in Logic, Cambridge University Press, Cambridge, 2017.

[22] F. Pakhomov and A. Visser, *On a question of Krajewski's*, ***Journal of Symbolic Logic***, vol. 84 (2019), no. 1, pp. 343–358.

[23] F. Pakhomov and J. Walsh, *Reflection ranks and ordinal analysis*, ***Journal of Symbolic Logic***, vol. 86 (2020), no. 4, pp. 1350–1384.

[24] C. Parsons, *On a number theoretic choice schema and its relation to induction*, ***Intuitionism and Proof Theory: Proceedings of the Summer Conference at Buffalo N.Y. 1968*** (A. Kino, J. Myhill, and R.E. Vesley, editors), Studies in Logic and the Foundations of Mathematics, vol. 60, Elsevier, Amsterdam, 1970, pp. 459–473.

[25] ———, *On n-quantifier induction*, ***Journal of Symbolic Logic***, vol. 37 (1972), no. 3, pp. 466–482.

[26] P. Pudlák, *Cuts, consistency statements and interpretations,* ***Journal of Symbolic Logic***, vol. 50 (1985), pp. 423–441.

[27] V. Y. Shavrukov and A. Visser, *Uniform density in Lindenbaum algebras*. ***Notre Dame Journal of Formal Logic***, vol. 55 (2014), no. 4, pp. 569–582.

[28] S. G. Simpson and R. L. Smith, *Factorization of polynomials and Σ^0_1 induction*. ***Annals of Pure and Applied Logic***, vol. 31 (1986), nos. 2–3, pp. 289–306.

[29] C. Smoryński, *ω-consistency and reflection*, ***Colloque International de Logique (Colloq. Int. CNRS)***, CNRS Inst. B. Pascal, Paris, 1977, pp. 167–181.

[30] W. W. Tait, *Finitism*. ***Journal of Philosophy***, vol. 78 (1981), no. 9, pp. 524–546.

[31] R. Vaught, *Axiomatizability by a schema*, ***Journal of Symbolic Logic***, vol. 32 (1967), pp. 473–479.

[32] A. Visser, *Vaught's theorem on axiomatizability by a scheme*. ***Bulletin of Symbolic Logic***, vol. 18 (2012), no. 3, pp. 382–402.

DEPARTMENT OF PHILOSOPHY, LINGUISTICS
AND THEORY OF SCIENCE
UNIVERSITY OF GOTHENBURG
GOTHENBURG, SWEDEN
E-mail: ali.enayat@gu.se

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF WARSAW
WARSAW, POLAND
E-mail: mlelyk@uw.edu.pl

https://doi.org/10.1017/jsl.2022.83 Published online by Cambridge University Press