Categoricity-like Properties in the First Order Realm

Ali Enayat and Mateusz Łełyk

Abstract: By classical results of Dedekind and Zermelo, second order logic imposes categoricity features on Peano Arithmetic and Zermelo-Fraenkel set theory. However, we have known since Skolem's anti-categoricity theorems that the first order formulations of Peano Arithmetic and Zermelo-Fraenkel set theory (i.e., PA and ZF) are not categorical. Here we investigate various categoricity-like properties (including tightness, solidity, and internal categoricity) that are exhibited by a distinguished class of first order theories that include PA and ZF, with the aim of understanding what is special about canonical foundational first order theories.

Keywords: Solidity, Tightness, Internal categoricity, Schemes, Biinterpretability, Retract, Nonstandard models, Peano Arithmetic, ZF

1. Introduction

By classical results of Dedekind and Zermelo, second order logic imposes categoricity features on Peano Arithmetic and Zermelo-Fraenkel set theory. More explicitly:

Dedekind Categoricity Theorem (1888). *There is a sentence σ in second order logic of the form $\forall X\, \varphi(X)$, where $\varphi(X)$ only has first order quantifiers, such that σ holds in a structure $\mathscr{M}$ iff $\mathscr{M} \cong (\mathbb{N}, S, 0)$, where S is the successor function.*

Zermelo Quasi-categoricity Theorem (1930). *There is a sentence θ in second order logic of the form $\forall X\, \psi(X)$, where $\psi(X)$ only has first order quantifiers, such that θ holds in a structure $\mathscr{M}$ iff $\mathscr{M} \cong (V_\kappa, \in)$, where κ is a strongly inaccessible cardinal.*

The above categoricity results have captured the imagination of several generations of philosophers in relation to the debate about the determinacy of the truth-value of arithmetical and set-theoretical statements. The debate is mediated by Skolem's ingenious theorems that indicate that categoricity fails dramatically for the first order formulation PA of Peano Arithmetic, and for the first order formulation ZF of Zermelo-Fraenkel set theory.[1]

Skolem Anti-categoricity Theorems.

(a) (1934) There is a structure $\mathscr{M}$ such that $\mathsf{Th}(\mathbb{N}, +, \cdot) = \mathsf{Th}(\mathscr{M})$, but $\mathscr{M}$ contains an 'infinite' element. Thus $\mathscr{M} \ncong (\mathbb{N}, +, \cdot)$.

(b) (1922) Every structure in a countable language has a countable elementary submodel. Thus, for any strongly inaccessible cardinal κ, there is a (countable) structure $\mathscr{M} \ncong (V_\kappa, \in)$ with $\mathsf{Th}(\mathscr{M}) = \mathsf{Th}(V_\kappa, \in)$.

Ali Enayat, University of Gothenburg, Sweden, ali.enayat@gu.se, ORCID: 0000-0003-0372-3354.
Mateusz Łełyk, University of Warsaw, Poland, mlelyk@uw.edu.pl.

Journal for the Philosophy of Mathematics, Vol. 1, 63-98, 2024 | ISSN 3035-1863 | DOI: 10.36253/jpm-2934

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.
Competing Interests: The Author(s) declare(s) no conflict of interest.

Firenze University Press | https://riviste.fupress.net/index.php/jpm

[1] Both results are seminal: the former led to the study of nonstandard models of arithmetic, and the latter is what is now commonly known as an instance of the Löwenheim-Skolem Theorem.

Here we bring together two distinct research threads in the foundation of mathematics that can be viewed as attempting to regain the 'lost' categoricity of PA and ZF by formulating 'categoricity-like' features that demonstrably hold for PA, ZF, and certain other canonical foundational theories.

One of these threads can be traced back to a potent result of Visser concerning PA that appears in his substantial paper (Visser, 2006) that builds a category-theoretic framework for the study of (relative) interpretability theory. Visser's result inspired Enayat (2016) to introduce the notion of solidity of an arbitrary first order theory, which allows Visser's result to be expressed as: PA *is a solid theory*. Solidity is defined model-theoretically (see Definition 7), but the completeness theorem of first order logic makes it clear that it implies a remarkable purely syntactic condition dubbed *tightness*: a theory T is tight if there is no pair of distinct deductively closed extensions of T (in the same language as T) that are bi-interpretable (for the definition of bi-interpretability, see Definition 1). Thus Visser's aforementioned theorem on the solidity of PA implies that if T_1 and T_2 are two consistent extensions of T in the same language as PA such that for some arithmetical sentence φ, we have:

$$T_1 \vdash \varphi \text{ and } T_2 \nvdash \varphi,$$

then T_1 and T_2 are not bi-interpretable. The solidity of certain canonical foundational theories, including ZF, was established in (Enayat, 2016) (see Theorem 14 of this paper). Enayat's work was further extended by Freire and Hamkins (2021) in the context of ZF and its fragments; their work shows that Z (Zermelo set theory) and ZF^- (ZF without the powerset axiom) fail to be tight (and therefore are not solid). More recently, Freire and Williams (2023) investigated tightness in the context of fragments of Kelley-Morse theory of classes and their arithmetical counterparts; their work shows that the commonly studied proper subtheories of the aforementioned theories fail to be tight.

The other research thread explored in this paper has a more complicated history, but for the purposes of this introduction it can be described as germinating in the introduction of the concept of internal categoricity of PA in full second order logic by Hellman and Parsons, which was followed by Väänänen's introduction of the notion of internal categoricity of Peano Arithmetic and Zermelo-Fraenkel set theory in the context of Henkin models of second order logic in (Väänänen, 2021) and in joint work with Wang (2015), and later in the context of first order logic, as in (Väänänen, 2019) and (Väänänen, 2021). Internal categoricity has been substantially explored and debated in the philosophical literature, as witnessed by Button and Walsh's monograph (2018), the recent monograph of Maddy and Väänänen (2023), and in the recent work of Fischer and Zicchetti (2023).

Both threads explore categoricity-like properties of first order foundational theories, but they differ in an important respect: the former thread is 'extensional' in the sense that its objects of study are first order theories viewed as a set of sentences, whereas the latter thread studies the 'intensional' *schematic representations* of PA and ZF. As we shall see in Theorem 52 of Section 4, this distinction is fundamental since despite the internal categoricity of the 'usual' axiomatizations of PA and ZF, there are other schematic representations of PA and ZF that fail to be internally categorical.

The paper is planned as follows. Section 2 is devoted to preliminary matters and a review of the pertinent results in the literature. Our novel technical results are presented in Sections 3 and 4. Section 3 presents a number of 'negative' results that demonstrate the failure of tightnness/soldity of commonly studied subtheories of PA and ZF, thus probing the optimality of the solidity proofs established in (Enayat, 2016). One of these negative results shows if T_0 is a subtheory of any of the known canonical solid theories T (such as $T = \mathsf{PA}$ or $T = \mathsf{ZF}$) that is axiomatized by a collection of sentences of bounded quantifier-complexity, then T_0 fails to be tight. Section 4 studies internal categoricity and the generalizations of the categoricity-like properties studied in (Enayat, 2016) to the context of schematic representations. This allows us to delineate the relationship between the aforementioned threads that explore categoricity-like properties of first order foundational theories. For example, we introduce a straightforward generalization of the concept of solidity (dubbed e-solidity) that is applicable to schematic presentations of theories and holds for the usual schematic axiomatization of PA (but not for ZF), and show that e-solidity implies internal categoricity. In Section 5 we reflect on the philosophical ramifications of the technical results of the paper; and Section 6 describes what remains to be done.

2. Preliminaries

2.1. Basic Definitions

Definition 1. We view a first order theory as a set of sentences, thus in the setting of this paper a theory need not be deductively closed. Suppose U and V are first order theories, formulated in relational languages[2] $\mathscr{L}_U$ and $\mathscr{L}_V$ (respectively); and let $\mathsf{Form}_{\mathscr{L}_U}$ and $\mathsf{Form}_{\mathscr{L}_V}$ be the set of first order $\mathscr{L}_U$-formulae and $\mathscr{L}_V$-formulae (respectively).

(a) We say that $\mathscr{I}$ *is an interpretation of* U *in* V, written $U \trianglelefteq^{\mathscr{I}} V$, if $\mathscr{I}$ specifies a *translation function*

$$\sigma : \mathsf{Form}_{\mathscr{L}_U} \to \mathsf{Form}_{\mathscr{L}_V}$$

such that for each $\varphi \in \mathscr{L}_U$,

$$U \vdash \varphi \Rightarrow V \vdash \sigma(\varphi),$$

and the translation $\varphi^\sigma := \sigma(\varphi)$ of φ satisfies the following three conditions:
(1) There is a designated $\mathscr{L}_V$-formula $\delta(x_1,...,x_k)$ referred to as the *domain formula* (and k is referred to as the *dimension* of the interpretation).
(2) There is a designated mapping $P \mapsto F_P$ that translates each n-ary $\mathscr{L}_U$-predicate P into some kn-ary $\mathscr{L}_V$-formula F_P (including the case when P is the equality relation).
(3) The translation function σ commutes with propositional connectives, and is subject to:

$$(\forall x \varphi)^\sigma = \forall x_1 \cdots x_n \, (\delta(x_1,\cdots,x_n) \to \varphi^\sigma).$$

The translation σ is called *direct* iff it is unrelativized, i.e.

$$\delta(x_1,\ldots,x_k) = (x_1 = x_1) \wedge \ldots \wedge (x_k = x_k),$$

and σ translates equality to equality on k-tuples.
Note that each translation $\sigma : \mathsf{Form}_{\mathscr{L}_U} \to \mathsf{Form}_{\mathscr{L}_V}$ gives rise to a uniform transformation of $\mathscr{L}_V$-structures into $\mathscr{L}_U$-structures. That is, given any $\mathscr{L}_V$-structure $\mathscr{M}$ we obtain an $\mathscr{L}_U$-structure $\sigma(\mathscr{M})$ whose domain is the set defined in $\mathscr{M}^k$ by δ and each predicate P (including the equality predicate) is interpreted as the set defined in $\mathscr{M}$ by F_P. Moreover, an interpretation $\mathscr{I}$ based on σ such that $U \trianglelefteq^{\mathscr{I}} V$ gives rise to an *internal model construction that* ***uniformly*** *builds a model* $\mathscr{M}^{\mathscr{I}} \models U$ *for any* $\mathscr{M} \models V$, where $\mathscr{M}^{\mathscr{I}} := \sigma(\mathscr{M})$.[3]

(b) U is *interpretable* in V, written $U \trianglelefteq V$, if $U \trianglelefteq^{\mathscr{I}} V$ for some interpretation $\mathscr{I}$. U and V are *mutually interpretable* when $U \trianglelefteq V$ and $V \trianglelefteq U$.

(c) We indicate the universe of each structure with the corresponding Roman letter, e.g., the universes of structures $\mathscr{M}$, $\mathscr{N}$, and $\mathscr{M}^*$ are respectively M, N, and M^*. Given an $\mathscr{L}$-structure $\mathscr{M}$ and $X \subseteq M^n$ (where n is a positive integer), we say that X is (*parametrically*) $\mathscr{M}$-*definable* iff there is an n-ary formula $\varphi(x_1,\ldots,x_n)$ in the language $\mathscr{L}_M$ (respectively $\mathscr{L}_M$, where $\mathscr{L}_M$ is the result of augmenting $\mathscr{L}$ with constant symbols c_m for each $m \in M$), such that $X = \varphi^{\mathscr{M}}$, where $\varphi^{\mathscr{M}} = \{(m_1,\cdots,m_k) \in M^n : (\mathscr{M},m)_{m\in M} \models \varphi(c_{m_1},\cdots,c_{m_k})\}$. We stress that in this paper $\mathscr{M}$-definability means "definability in $\mathscr{M}$ without parameters".

(d) Suppose $\mathscr{N}$ is an $\mathscr{L}_U$-structure and $\mathscr{M}$ is an $\mathscr{L}_V$-structure. We say that $\mathscr{M}$ (*parametrically*) *interprets* $\mathscr{N}$, written $\mathscr{M} \trianglerighteq \mathscr{N}$ ($\mathscr{M} \trianglerighteq_{\text{par}} \mathscr{N}$), iff there is a (parametric) translation σ of $\mathscr{L}_U$-formulae to $\mathscr{L}_V$-formulae such that $\mathscr{N} = \sigma(\mathscr{M})$. Unravelling the definition, this means that the universe of discourse of $\mathscr{N}$, as well as all the $\mathscr{N}$-interpretations of $\mathscr{L}_U$-predicates are (parametrically) $\mathscr{M}$-definable.[4] Note that both $\trianglelefteq$ and $\trianglelefteq_{\text{par}}$ are transitive relations.

(e) Suppose $\mathscr{M}$ is structure that (parametrically) interprets the structures $\mathscr{N}_0$ and $\mathscr{N}_1$. Let δ_0 and δ_1 be the domain formulae, and $\sim_0$, $\sim_1$ be translations of the equality relation in the respective translations. A (parametrically) $\mathscr{M}$-definable isomorphism between $\mathscr{N}_0$ and $\mathscr{N}_1$ is a (parametrically) $\mathscr{M}$-definable relation $R \subseteq \delta_0^{\mathscr{M}} \times \delta_1^{\mathscr{M}}$ such that the following hold:

2 This assumption is only for ease of exposition.
3 We often write $\mathscr{M}^{\mathscr{I}}$ instead of $\sigma(\mathscr{M})$ to emphasize that $\sigma(\mathscr{M}) \models U$ in the context of $U \trianglelefteq^{\mathscr{I}} V$. Note that translations relate formulae, and interpretations relate theories (and their models).
4 In the case of translations that are not equality-preserving, $\mathscr{N}$ is the quotient structure of a definable subset of M^n under an $\mathscr{M}$-definable equivalence relation, and the predicates and functions on $\mathscr{N}$ are treated accordingly.

1. The domain and the codomain of R are $\delta_0^{\mathscr{M}}$ and $\delta_1^{\mathscr{M}}$, respectively.

2. If $x_0 R x_1$ and $y_0 \sim_0 x_0$ and $y_1 \sim_1 x_1$, then $y_0 R y_1$.

3. The function $F : N_0 \to N_1$, defined: $F([x_0]_{\sim_0}) = [x_1]_{\sim_1}$ iff $x_0 R x_1$, is an isomorphism between $\mathscr{N}_0$ and $\mathscr{N}_1$.

(f) A structure $\mathscr{M}$ is *a (parametric) retract* of a structure $\mathscr{N}$ if there is an isomorphic copy $\mathscr{M}^*$ of $\mathscr{M}$ such that $\mathscr{M} \trianglerighteq \mathscr{N} \trianglerighteq \mathscr{M}^*$ ($\mathscr{M} \trianglerighteq_{\text{par}} \mathscr{N} \trianglerighteq_{\text{par}} \mathscr{M}^*$), and moreover there is a (parametrically) $\mathscr{M}$-definable isomorphism between $\mathscr{M}$ and $\mathscr{M}^*$.

(g) U is a *retract* of V iff there are interpretations $\mathscr{I}$ and $\mathscr{J}$ with $U \trianglelefteq^{\mathscr{I}} V$, and $V \trianglelefteq^{\mathscr{J}} U$ a binary U-formula F such that F is, U-verifiably, an isomorphism between id_U (the identity interpretation on U) and $\mathscr{I} \circ \mathscr{J}$ (where $\circ$ is the composition operation on interpretations). In model-theoretic terms, this translates to the requirement that every model of U is a retract of some model of V in *uniform and parameter-free* manner, i.e., that the following holds for every $\mathscr{M} \models U$:

$$F^{\mathscr{M}} : \mathscr{M} \xrightarrow{\cong} \mathscr{M}^* := \mathscr{M}^{\mathscr{I} \circ \mathscr{J}}.$$

(h) U and V are *bi-interpretable* iff there are interpretations $\mathscr{I}$ and $\mathscr{J}$ as above that witness that U is a retract of V, and additionally, there is a V-formula G, such that G is, V-verifiably, an isomorphism between id_V and $\mathscr{J} \circ \mathscr{I}$, where id_V is the identity interpretation. In particular, if U and V are bi-interpretable, then given $\mathscr{M} \models U$ and $\mathscr{N} \models V$, we have

$$F^{\mathscr{M}} : \mathscr{M} \xrightarrow{\cong} \mathscr{M}^* := \mathscr{M}^{\mathscr{I} \circ \mathscr{J}} \text{ and } G^{\mathscr{N}} : \mathscr{N} \xrightarrow{\cong} \mathscr{N}^* := \mathscr{N}^{\mathscr{J} \circ \mathscr{I}}.$$

(i) Suppose T is a theory formulated in a language $\mathscr{L}$, and T^+ is a theory formulated in a language $\mathscr{L}^+ \supseteq \mathscr{L}$, and assume without loss of generality that both $\mathscr{L}$ and $\mathscr{L}^+$ are relational languages. T^+ is said to be a *definitional extension* of T if for each n-ary relation $R \in \mathscr{L}^+ \setminus \mathscr{L}$ there is an n-ary $\mathscr{L}$-formula δ_R such that T^+ is logically equivalent to the $\mathscr{L}^+$-theory obtained by augmenting T with axioms of the form

$$\forall x_1 \cdots \forall x_n\, [R(x_1, \cdots, x_n) \leftrightarrow \delta_R(x_1, \cdots, x_n)].$$

With the above definition in mind, two theories T_1 and T_2 that are formulated in disjoint languages are said to be *definitionally equivalent* if they have a common definitional extension, i.e., there is a theory T such that T is a definitional extension of both T_1 and T_2. Definitional equivalence is also commonly referred to as *synonymy*, see (Lefever and Székel, 2019) for more detail.

Remark 2. As shown by Visser in Theorems 4.2 and 4.12 of his paper (2006), two theories U and V are definitionally equivalent iff they satisfy a stronger form of bi-interpretability, namely, in the definition of bi-interpretability, the condition that F is U-verifiably an isomorphism between id_U and $\mathscr{I} \circ \mathscr{J}$ is strengthened to the stronger condition that $F(x) = x$, and similarly, the condition that G is V-verifiably an isomorphism between id_V and $\mathscr{J} \circ \mathscr{I}$ is strengthened to the condition that $G(x) = x$. In model theoretic terms this translates to:

$$\mathscr{M}^{\mathscr{I} \circ \mathscr{J}} = \mathscr{M} \text{ for all } \mathscr{M} \models U, \text{ and } \mathscr{N}^{\mathscr{J} \circ \mathscr{I}} = \mathscr{N} \text{ for all } \mathscr{N} \models V.$$

Thus definitional equivalence is a stronger form of bi-interpretation; however, by a result of Friedman and Visser (2014), in many cases definitional equivalence is implied by bi-interpretability, namely, when the two theories involved are sequential[5], and the bi-interpretability between them is witnessed by a pair of one-dimensional, identity preserving interpretations. See also Theorem 16 below.

[5] At first approximation, a theory is sequential if it supports a modicum of coding machinery to handle finite sequences of all objects in the domain of discourse. Sequentiality is a modest demand for theories of arithmetic and set theory; however, by a theorem of Visser (2016), (Robinson's) Q is not sequential. However, by a theorem of Jeřábek (2012) the 'algebraic' fragment PA^- of PA is already sequential; for more on this theory, see the paragraph preceding Definition 30.

2.2. Examples

In what follows, $\mathbb{Q}$ is the set of rational numbers, ω is the set of finite ordinals (i.e., natural numbers), and $\mathsf{Th}(\mathscr{M})$ is the first order theory of the structure $\mathscr{M}$.

Theorem 3. *(J. Robinson, 1949)* $\mathsf{Th}(\mathbb{Q},+,\cdot)$ *and* $\mathsf{Th}(\omega,+,\cdot)$ *are bi-interpretable.*[6]

Theorem 4. *Let* $\mathsf{ZF_{fin}}$ *be the result of replacing the axiom of infinity in the usual axiomatization of* ZF *by its negation, and let* TC *denote the statement that every set has a transitive closure.*

(a) *(Ackermann, 1937; Mycielski, 1964; Kaye-Wong, 2007)* PA *and* $\mathsf{ZF_{fin}}+\mathsf{TC}$ *are definitionally equivalent. Moreover,* PA *is a retract of* $\mathsf{ZF_{fin}}$.

(b) *(Enayat-Schmerl-Visser, 2011)* $\mathsf{ZF_{fin}}$ *is not a retract of* PA.

Theorem 5 below arose from the work of Mostowski (based on earlier ideas going back to of Gödel).[7] In what follows,

$$\mathsf{ZF}^- := \mathsf{ZF}^{\mathsf{Sep+Coll}} \setminus \{\mathsf{Powerset}\},$$

where $\mathsf{ZF}^{\mathsf{Sep+Coll}}$ is the result of substituting the Replacement scheme in the usual axiomatization of ZF with the schemes of Separation and Collection.[8] Also, in part (b) $\mathsf{Inacc}(\kappa)$ expresses "κ is a strongly inaccessible cardinal greater than $\aleph_0$".

Theorem 5. *(Mostowski)*

(a) $\mathsf{Z}_2+\Pi^1_\infty$-$\mathsf{AC}$ *is bi-interpretable with* $\mathsf{ZF}^-+\forall x\ |x|\leq\aleph_0$.

(b) $\mathsf{KM}+\Pi^1_\infty$-AC *is bi-interpretable with* $\mathsf{ZF}^-+\exists\kappa(\mathsf{Inacc}(\kappa)\wedge\ \forall x\ |x|\leq\kappa)$.

Remark 6. As shown in Theorem 16 bi-interpretability cannot be strengthened to definitional equivalence in Theorem 5. Note that Friedman and Visser (2014) exhibited a pair of ad hoc theories to show that bi-interpretability does not imply definitional equivalence; thus Theorem 16 points to a counterexample involving canonical theories.

2.3. Categoricity-like properties

The notions of solidity, neatness, and tightness encapsulated in the following definition were formulated in (Enayat, 2016), but the notion of a minimalist theory in (a) below is new.

Definition 7. Suppose T is a first-order theory.

[6] As pointed out by Friedman and Visser, the main result of (2014) can be used to show that bi-interpretability can be improved to definitional equivalence here.

[7] For further bibliographical references for part (a) of Theorem 5, see Notes for section VII.3 of Simpson's encyclopedic exposition (2009), and for part (b) see Chapter 2 of Williams' doctoral dissertation (2018).

[8] Indeed this is how ZF is sometimes axiomatized, as in Chang and Keisler's textbook on model theory (1990), in contrast to set theory textbooks by Kunen (1980) and Jech (2003) that use the replacement scheme. Recall that the instances of the Separation scheme consist of universal generalizations of formula of the form

$$\forall b\exists a\forall x(x\in a\ \leftrightarrow\ \varphi(x)),$$

where the parameters of φ are suppressed, and instances of the Collection scheme consist of universal generalizations of formulae of the form

$$(\forall x\in a\ \exists y\ \psi(x,y))\rightarrow(\exists b\ \forall x\in a\ \exists y\in b\ \psi(x,y)),$$

where the parameters of ψ are suppressed. It is well-known that ZF and $\mathsf{ZF}^{\mathsf{Sep+Coll}}$ axiomatize the same theory; but in the absence of the powerset axiom, the latter theory is stronger; as demonstrated in (Gitman, Hamkins and Johnstone, 2016). By a classical theorem of Scott (1961), the same discrepancy between ZF and $\mathsf{ZF}^{\mathsf{Sep+Coll}}$ arises in the absence of the extensionality axiom.

- Each of the following definitions is in the form of an implication. Note that in parts (a) and (b), the parameters in the conclusion of each of the implications should be understood to be a subset of the parameters used in the premises. [9]

(a) T is *minimalist (maximalist)* iff for every model $\mathscr{M} \models T$ and every model $\mathscr{N} \models T$ such that $\mathscr{M} \trianglerighteq_{\mathrm{par}} \mathscr{N}$, there is a unique parametrically $\mathscr{M}$-definable embedding of $\mathscr{M}$ into $\mathscr{N}$ (resp. embedding of $\mathscr{N}$ into $\mathscr{M}$).

(b) T is *solid* iff the following holds for all models $\mathscr{M}$, $\mathscr{M}^*$, and $\mathscr{N}$ of T:

If $\mathscr{M} \trianglerighteq_{\mathrm{par}} \mathscr{N} \trianglerighteq_{\mathrm{par}} \mathscr{M}^*$ and there is a parametrically $\mathscr{M}$-definable isomorphism $i_0 : \mathscr{M} \to \mathscr{M}^*$, then there is a parametrically $\mathscr{M}$-definable isomorphism $i : \mathscr{M} \to \mathscr{N}$.

In other words, T is solid iff for every two models $\mathscr{M}$ and $\mathscr{N}$ of T, if $\mathscr{M}$ is a parametric retract of $\mathscr{N}$, then $\mathscr{M}$ and $\mathscr{N}$ are isomorphic via a parametrically $\mathscr{M}$-definable isomorphism.

(c) T is *neat* iff for any two deductively closed extensions U and V of T (both of which are formulated in the language of T), if U is a retract of V, then $V \subseteq U$.

(d) T is *tight* iff for any two deductively closed extensions U and V of T (both of which are formulated in the language of T), U and V are bi-interpretable iff $U = V$.

Remark 8. The following can be readily verified on the basis of the relevant definitions.

(a) If T is minimalist (maximalist), then T is solid.

(b) A solid theory is neat, and a neat theory is tight.

(c) Solidity, neatness and tightness are invariant under bi-interpretations.

Remark 9. Suppose we define a theory T to be *strongly solid* if, in the definition of solidity, the conclusion that there is a (parametrically) $\mathscr{M}$-definable isomorphism between $\mathscr{M}$ and $\mathscr{N}$, is changed to the existence of a (parametrically) $\mathscr{N}$-definable isomorphism between $\mathscr{N}$ and $\mathscr{M}^*$. It is easy to see that strong solidity implies solidity. Surprisingly, strong solidity is implied by solidity. We owe this observation to Leszek Kołodziejczyk. To see why, suppose that i is an isomorphism between $\mathscr{M}$ and $\mathscr{N}$, and j is an isomorphism between $\mathscr{M}$ and $\mathscr{M}^*$. Also let $h(p)$, where p denotes the parameters, be the translation that yields an interpretation of $\mathscr{M}^*$ in $\mathscr{N}$. Then $\mathscr{M}$ sees $h(i^{-1}(p))$ and sees an isomorphism between itself and the model given by $h(i^{-1}(p))$ (the isomorphism is $i^{-1} \circ j$). Hence by applying i to $\mathscr{M}$, $\mathscr{N}$ sees its isomorphism with the model given by $h(p)$, and hence with $\mathscr{M}^*$.

3. Categoricity of first-order theories

3.1. Positive results

In this section we present known results about categoricity-like properties. We begin with the following result of Albert Visser that inspired the first-named-author's paper (Enayat, 2016). Visser's result follows from Corollaries 9.4 and 9.6 of his paper (Visser, 2006). For a streamlined proof, see (Enayat, 2016).

Theorem 10. *(Visser)* PA *is solid.*

Another way of seeing that PA is solid is via part (a) of Remark 8 since PA can be readily shown to be a minimalist theory, as indicated in the proposition below.

Proposition 11. PA *is minimalist.*

[9] This restriction on parameters was only implicit in (Enayat, 2016); it is made explicit in light of two considerations: (1) the solidity proofs established in (Enayat, 2016) that inspired the definition abide by this restriction; and (2) The proof of Theorem 74 of this paper that establishes that internally categoricity is implied by e-solidity does not go through for the 'liberal' notion of solidity that imposes no restriction on the behaviour of the parameters (indeed we have an example that shows that the aforementioned implication provably fails if the restriction stipulated here on parameters is violated).

Proof. Suppose $\mathscr{M}$ and $\mathscr{N}$ are models of PA and $\mathscr{M} \trianglerighteq_{\text{par}} \mathscr{N}$. A recursion within $\mathscr{M}$ can be used to construct a parametrically definable embedding f of $\mathscr{M}$ onto an initial segment of $\mathscr{N}$, where $f(0^{\mathscr{M}}) = 0^{\mathscr{N}}$, and $f((x+1)^{\mathscr{M}}) = (f(x)+1)^{\mathscr{N}}$. Then, by using induction in $\mathscr{M}$, one can readily show that if g is a parametrically definable embedding of $\mathscr{M}$ onto an initial segment of $\mathscr{N}$, then $f(m) = g(m)$ for all m in $\mathscr{M}$.[10] □

In light of part (a) of Theorem 4 and part (b) of Remark 8, Theorem 10 yields the following corollary.

Corollary 12. $\mathsf{ZF}_{\mathsf{fin}} + \mathsf{TC}$ *is solid.*

Remark 13. It is easy to see that PA is not a maximalist theory, since for example a nonstandard model of PA is interpretable in the standard model of PA, but of course the latter cannot be embedded in the former. It is also noteworthy that ZF is neither minimalist nor maximalist. To see the former, note that if $\mathscr{M} \models \mathsf{ZF} + \exists\kappa \mathsf{Inacc}(\kappa)$, then $\mathscr{M}$ can interpret ZF via its definable submodel $(V_\kappa, \in)^{\mathscr{M}}$, however there can be no definable embedding of $\mathscr{M}$ into $(V_\kappa, \in)^{\mathscr{M}}$.[11] To see that ZF is not a maximalist theory, let $\mathscr{M}$ be a well-founded model of ZFC, and let $\mathscr{N}$ be an internal ultrapower of $\mathscr{M}$ modulo a nonprincipal ultrafilter over $\mathscr{P}(\omega)^{\mathscr{M}}$. Note that $\mathscr{N}$ is interpretable in $\mathscr{M}$ and $\mathscr{N}$ is ill-founded, and clearly there can be no embedding of $\mathscr{N}$ into $\mathscr{M}$.

The list of examples of solid theories was extended in the first-named-author's paper (Enayat, 2016) to other well-known foundational theories, as indicated in the next theorem. In what follows Z_n is n-th order arithmetic, and KM_n is n-th order Kelley-Morse theory of classes, where $\mathsf{Z}_1 := \mathsf{PA}$, and $\mathsf{KM}_1 := \mathsf{ZF}$. Also, Z_ω is the theory of types (with full comprehension) whose level-zero objects form a model of PA (equivalently $\mathsf{ZF}_{\mathsf{fin}} + \mathsf{TC}$), and KM_ω is the theory of types whose level-zero objects form a model of ZF are also solid theories.

Theorem 14. *Each of the theories in the family $\mathscr{S}$ of theories below is solid:*

$$\mathscr{S} = \{\mathsf{Z}_n : 1 \leq n \in \omega\} \cup \{\mathsf{KM}_n : 1 \leq n \in \omega\} \cup \{\mathsf{Z}_\omega, \mathsf{KM}_\omega\}.$$

As noted in (Enayat, 2016), the following corollary follows from putting part (c) of Remark 8 together with Theorem 14 and the bi-interpretability of the theories in (a) and (b) of Theorem 5 with appropriate extension of Z_2 and KM (respectively). Recall from the paragraph preceding Theorem 5 that ZF^- is the result of removing the Powerset axiom from $\mathsf{ZF}^{\mathsf{Sep+Coll}}$.

Theorem 15. *The following theories are solid:*

(a) $\mathsf{ZF}^- + \forall x\ |x| \leq \aleph_0$.

(b) $\mathsf{ZF}^- + \exists\kappa(\mathrm{Inacc}(\kappa) \wedge \forall x\ |x| \leq \kappa)$.

Theorem 16. *Bi-interpretability cannot be improved to definitional equivalence in parts (a) and (b) of Theorem 5 (assuming the consistency of* ZF *plus an inaccessible cardinal for part (a), and the consistency of* ZF *with two inaccessible cardinals for part (b)).*

Proof. We first deal with part (a). Note that by a classical argument going back to Mostowski (which inspired Theorem 5), within ZFC the standard model $(\mathscr{P}(\omega), \omega, +, \cdot, \in)$ of $\mathsf{Z}_2 + \Pi^1_\infty\text{-AC}$ is bi-interpretable with the standard model $(H(\omega_1), \in)$ of $\mathsf{ZF}^- + \forall x\ |x| \leq \aleph_0$. Thus, thanks to Theorem 5 and part (a) of Theorem 15 any model of $\mathsf{ZF}^- + \forall x\ |x| \leq \aleph_0$ that is bi-interpretable with $(\mathscr{P}(\omega), \omega, +, \cdot, \in)$ must be isomorphic to $(H(\omega_1), \in)$.

On the other hand, by a classical theorem, due to Solovay (1970, Theorem 2), assuming the existence of an inaccessible cardinal, there is a model $\mathscr{M}$ of ZFC (obtained by set forcing) in which CH holds (i.e., $2^{\aleph_0} = \aleph_1$) and all projective sets of reals are Lebesgue measurable. Note that (1) the submodel $H(\omega_1)^{\mathscr{M}}$ of hereditarily countable sets of such an $\mathscr{M}$ satisfies $\mathsf{ZF}^- + \forall x\ |x| \leq \aleph_0$, and the standard model $(\mathscr{P}(\omega), \omega, +, \cdot, \in)^{\mathscr{M}}$ of Z_2 of $\mathscr{M}$ satisfies Π^1_∞-AC. Within $\mathscr{M}$ evidently $H(\omega_1)$ has a parameter-free definable subset of order-type ω_1, but

[10] Note that the proof only uses the fact that $\mathscr{N} \models \mathsf{Q}$ (where Q is Robinson Arithmetic).

[11] One can bypass the appeal to the existence of an inaccessible cardinal in $\mathscr{M}$ with a more refined argument that takes advantage of the reflection theorem in $\mathscr{M}$.

$(\mathscr{P}(\omega), \omega, +, \cdot, \in)$ does not have a parametrically definable subset of order-type ω_1. More specifically, recall that the projective sets are precisely the subsets of $\mathscr{P}(\omega)$ that are parametrically definable in $(\mathscr{P}(\omega), \omega, +, \cdot, \in)$; and under CH, Fubini's theorem (in measure theory) implies that any linear order on $\mathscr{P}(\omega)$ that is of order type ω_1 fails to be measurable when viewed as a subset of the cartesian product of $\mathscr{P}(\omega)$ with itself as first observed by Sierpiñski. Here $\mathscr{P}(\omega)$ is identified with the coin-tossing product measure space 2^{ω}. In light of the first paragraph of the proof, this shows that the definitional equivalence of the theories $\mathsf{Z}_2 + \Pi^1_\infty$-AC and $\mathsf{ZF}^- + \forall x\ |x| \leq \aleph_0$ within $\mathscr{M}$ implies that, as viewed in $\mathscr{M}$, there is a projective nonmeasurable subset of the cartesian product of $\mathscr{P}(\omega)$ with itself, contradiction.

A similar argument can be carried out for part (b), using a generalization of Solovay's proof (which requires the existence of at least two inaccessible cardinals) in which ω is replaced by an inaccessible cardinal, and "projective" is replaced by "parametrically definable in the natural model of KM associated with V_κ in which classes are interpreted as members of $V_{\kappa+1}$". For a proof outline of this generalization, see Schultzenberg's MathOverflow answer (2023).

Note that the statement "U and V are definitionally equivalent", where both U and V are arithmetically definable theories (let alone computable theories) is an *arithmetical sentence*, and therefore its truth-value does not change in the passage to a Boolean-valued extension of the universe obtained by forcing; thus, we have shown (reasoning in ZFC plus "there is an inaccessible cardinal") that the two theories asserted to be bi-interpretable in part (a) of Theorem 5 not only fail to be definitionally equivalent in some model of ZFC, but are outright not definitionally equivalent assuming the existence of an inaccessible cardinal. A similar comment applies to part (b) of Theorem 5, assuming the existence of two inaccessible cardinals. □

Remark 17. In private communication, Vincenzo Dimonte has pointed out that the above generalization of Solovay's theorem (for the proof of part (b) of Theorem 16) can be derived from Theorem 2.19 and Lemma 2.21 of Schlicht' paper (2017) (that pertain to the same model considered by Schultzenberg). Moreover, one of the referees has pointed out the following two points. Firstly, Schultzenberg's aforementioned proof in (2023) contains a small gap that can be filled by consulting Kanamori's exposition (2009, p. 141) of Solovay's proof that the perfect set property holds. Secondly, the assumption for part (a) of Theorem 16 can probably be reduced to the consistency of ZF since by forcing with the partial order $\mathrm{Col}(\omega, <\mathrm{Ord})$ one obtains a model that sufficiently behaves like the Solovay model (and one can probably use the same idea to reduce the assumption in part (b) of Theorem 16 to the consistency of ZF plus an inaccessible).

The following adds another prominent theory to the list of solid theories.

Theorem 18. $\mathsf{ZF} \setminus \{\mathsf{Infinity}\} + \mathsf{TC}$ *is solid.*

Proof. As shown by Visser (see Corollaries 9.6 and 9.8 of Visser, 2006), if U is a consistent extension of PA and V is a consistent extension of ZF, then U is not a retract of V, and V is not a retract of U. An inspection of Visser's proofs of the aforementioned results allows us to conclude the following stronger statements:

(1) No model of PA is a parametric retract of a model of ZF.

(2) No model of ZF is a parametric retract of a model of PA.

Since TC is a theorem of ZF, it is routine to verify that $\mathsf{ZF} \setminus \{\mathsf{Infinity}\} + \mathsf{TC}$ is a solid theory with the help of (1) and (2) together with the bi-interpretability of $\mathsf{ZF}_{\mathsf{fin}} + \mathsf{TC}$ and PA (part (a) of Theorem 4), the solidity of PA and ZF (Theorem 14), and part (c) of Remark 8. □

3.2. Reflecting on the positive results

Let $\mathscr{S}$ be the list of theories whose solidity is asserted in Theorem 14. The following question is motivated by the fact that none of the theories $T \in \mathscr{S}$ is finitely axiomatizable. This is because each theory $T \in \mathscr{S}$ is *inductive*, i.e., T proves $\mathsf{Q}^{\mathscr{N}} + \mathsf{Ind}^{\mathscr{N}}(\mathscr{L}_T)$; here $\mathscr{N}$ is a designated interpretation $\mathscr{N}$ of arithmetic in T, Q is Robinson Arithmetic, and $\mathsf{Ind}^{N}(\mathscr{L}_T)$ is the scheme of induction over natural numbers for $\mathscr{L}_T$-formulae (in which the parameters are allowed to vary over the domain of discourse and are not limited to the 'numbers' of $\mathscr{N}$). Recall that by a classical theorem of Montague (1961), for a sequential[12] theory T, the T-provability of $\mathsf{Q}^{\mathscr{N}} + \mathsf{Ind}^{\mathscr{N}}(\mathscr{L}_T)$ implies that T is not finitely axiomatizable.

Question A. *Is there a consistent sequential finitely axiomatized solid theory*?

In Theorem 39 and 38 of this paper we give a partial negative answer to Question A by showing that finitely axiomatized subtheories of the theories PA, Z_2, ZF, and KM fail to be tight (and therefore they are not solid). Indeed our method can be used to show that no finitely axiomatized subtheory of any theory T in the list $\mathscr{S}$ of Theorem 14 is tight.

An inspection of the proofs of the different cases of Theorem 14 presented in (Enayat, 2016) makes it clear that the proofs of solidity of each of the $T \in \mathscr{S}$ uses the 'full power' of T. This prompts the next question.

Question B. *Is there an example T of one of the theories whose solidity is established in Theorem 14, and some solid deductively closed proper subtheory of T?*

Remark 19. Very recent joint work (in progress) of Piotr Gruza, Leszek Kołodziejczyk, and the second-named-author of this paper shows that Question B has a positive answer for $T = \mathsf{PA}$; indeed their work shows that the example in this case can be even required to contain any prescribed $\mathsf{I}\Sigma_n$. Also, note that by putting Corollary 12 and part (a) of Theorem 18 we obtain a positive answer to Question B for the subtheory $\mathsf{ZF} \setminus \{\mathsf{Infinity}\} + \mathsf{TC}$ of $T = \mathsf{ZF}$.

In the remaining subsections of this section, we first review published results concerning the failure of solidity/tightness established in (Enayat, 2016), (Freire and Hamkins, 2021), and (Freire and Williams, 2023) in Subsection 3.3, and then we present new results concerning the failure of solidity/tightness in Subsection 3.4.

3.3. Known negative results

As pointed out in (Enayat, 2016) inductive[13] sequential theories need not be tight since the theory $\mathsf{PA}(G)$, with no extra axioms for G, is not tight (where G is a fresh predicate). Here $\mathsf{PA}(G)$ is the result of augmenting PA with the scheme of induction for all formulae in the language $\mathscr{L}_{\mathsf{PA}} \cup \{G\}$. In the same paper, the proofs of the following results were outlined:

(1) $\mathsf{ZF}_{\mathsf{fin}}$ is not tight (the proof is based on a construction from Enayat, Schmerl and Visser, 2011). This makes it clear that $\mathsf{ZF} \setminus \{\mathsf{Infinity}\}$ is not tight, since tightness is inherited by theory extensions (in the same language).

(2) $\mathsf{ZF} \setminus \{\mathsf{Foundation}\}$ and $\mathsf{ZF} \setminus \{\mathsf{Extensionality}\}$ is not tight (the proofs employs classical techniques for building models in which Extensionality or Foundation fails).

The list of proper subtheories of ZF that fail to be tight/solid was further extended by Freire and Hamkins (2021), who showed that Z (Zermelo set theory) and $\mathsf{ZF} \setminus \{\mathsf{Powerset}\}$ also fail to be tight (even when ZF is formulated with the schemes of Separation and Collection).

More recently, Freire and Williams (2023) established the failure of tightness of well-known fragments of Z_2 and KM. Their work shows that for each $n \in \omega$, the fragments of Z_2 and KM in which the comprehension schema is limited to Π^1_n-formulae fails to be tight even when the full scheme of induction (for the case of fragments of Z_2) or the full scheme of $\in$-induction (for the case of fragments of KM) are included. In particular, their work shows that the subsystems Π^1_n-CA of Z_2 in which the comprehension scheme is limited to Π^1_n-formulae (but the full induction scheme is kept) fail to be tight.[14]

12 At first approximation, a theory is sequential if it supports a modicum of coding machinery to handle finite sequences of all objects in the domain of discourse. Sequentiality is a modest demand for theories of arithmetic and set theory; however, by a theorem of Visser (2016), (Robinson's) Q is not sequential. There are many equivalent definitions of sequentiality; the original definition due to Pudlák is as follows: A theory T is sequential if there is a formula $N(x)$, together with appropriate formulae providing interpretations of equality, and the operations of successor, addition, and multiplication for elements satisfying $N(x)$ such that T proves the translations of the axioms of Q (Robinson's arithmetic) when relativized to $N(x)$; and additionally, there is a formula $\beta(x,i,w)$ (whose intended meaning is that x is the i-th element of a sequence w) such that T proves that every sequence can be extended by any given element of the domain of discourse. For more detail and references see Visser's (2013).

13 Inductive theories are defined in the paragraph preceding Question A.

14 Independently and earlier, the first-named-author of this paper demonstrated the failure of tightness of the subsystem ACA (i.e., ACA_0 with the full scheme of induction) using a forcing argument similar to the one used by Freire and Williams. The proof was presented in an online seminar talk (Enayat, 2021).

3.4. New negative results

Definition 20. $\mathbb{N}$ is the standard model of PA, i.e., $(\omega, +, \cdot)$, and given a model of arithmetic $\mathcal{M}$,

$$\mathsf{Th}_{\Pi_n}(\mathcal{M}) = \{\varphi \in \Pi_n : \mathcal{M} \models \varphi\}.$$

Also $\mathsf{PA}_{\Pi_n} = \{\varphi \in \Pi_n : \mathsf{PA} \vdash \varphi\}$.[15] Here Π_n refers to the usual hierarchy of arithmetical formulae as in (Kaye, 1991) and (Hájek and Pudlák, 1998).

Definition 21. For $\mathcal{M} \models \mathsf{PA}$ and $n \in \omega$, $K_n(\mathcal{M})$ is the submodel of $\mathcal{M}$ whose universe consists of elements of $\mathcal{M}$ that are definable in $\mathcal{M}$ by a Σ_n-formula.

The following result is classical; it is due independently to the joint work of Kirby and Paris (1978) and to Lessan.[16] An exposition can be found in Kaye's monograph (1991).

In what follows the notation $\mathcal{A} \prec_{\Pi_n} \mathcal{B}$ means that Π_n-formulae are absolute in the passage between the arithmetical structures $\mathcal{A}$ and $\mathcal{B}$; $\mathsf{I}\Sigma_n$ is the fragment of PA in which the induction scheme is limited to Σ_n-formulae, and $\mathsf{B}\Sigma_n$ is the result of restricting the collection scheme to Σ_n-formulae.

Theorem 22. *(Kirby-Paris, Lessan) Suppose $n \in \omega$, $n \geq 1$, and $\mathcal{M}$ is a nonstandard model of* PA*, then:*

(a) $K_n(\mathcal{M}) \prec_{\Pi_n} \mathcal{M}$*, hence* $K_n(\mathcal{M}) \models \mathrm{Th}_{\Pi_{n+1}}(\mathcal{M})$.

(b) *If $K_n(\mathcal{M})$ is nonstandard, then* $K_n(\mathcal{M}) \models \mathsf{PA}_{\Pi_{n+1}} + \mathsf{I}\Sigma_{n-1} + \neg\mathsf{B}\Sigma_n$.

The following result also appears in Lessan's doctoral dissertation.

Theorem 23. *(Lessan) Suppose $n \in \omega$, $n \geq 1$, $\mathcal{M} \models$ PA, and $K_n(\mathcal{M})$ is nonstandard for some model $\mathcal{M} \models$ PA. Then the standard cut ω is first order definable in $K_n(\mathcal{M})$.*

Proof. (Outline) The key idea is that the complement of the standard cut can be defined in $K_n(\mathcal{M})$ via the formula that expresses "every element is definable by a Σ_n-formula whose code is below x". Recall that there is a Σ_k-definable satisfaction class of complexity Σ_k within models of $\mathrm{I}\Delta_0 + \mathsf{Exp}$ for each nonzero $k \in \omega$ (see Hájek and Pudlák, 1998). In particular since $K_n(\mathcal{M}) \models \mathsf{PA}_{\Pi_2}$ for $n \geq 1$, $K_n(\mathcal{M})$ carries a Σ_n-definable satisfaction class of complexity Σ_k for each nonzero $k \in \omega$. □

Theorem 24. *For each $n \in \omega$ the following statements hold:*

(a) *$\mathbb{N}$ is bi-interpretable with a nonstandard model of the form $K_{n+1}(\mathcal{M})$, where $\mathcal{M} \models$ PA, and* $K_{n+1}(\mathcal{M}) \models \mathrm{Th}_{\Pi_n}(\mathbb{N}) + \neg\mathsf{B}\Sigma_{n+1}$.

(b) $\mathrm{Th}(\mathbb{N})$ *is bi-interpretable with a complete extension of* $\mathsf{PA}_{\Pi_{n+2}} + \mathrm{Th}_{\Pi_n}(\mathbb{N}) + \neg\mathsf{B}\Sigma_{n+1}$.

Proof. To establish (a), we note that for a prescribed $n \in \omega$, $\mathrm{Th}_{\Pi_n}(\mathbb{N})$ is arithmetically definable, thanks to the existence of a definable Π_n-satisfaction class, i.e., an arithmetically definable predicate that satisfies Tarski's recursive clauses for a satisfaction predicate for all Π_n-formulae. Consider the arithmetically definable theories T_n and T_n^+ defined by:

$$T_n := \mathsf{PA} + \mathrm{Th}_{\Pi_n}(\mathbb{N}), \text{ and } T_n^+ := T_n + \neg\mathsf{Con}(T_n).$$

15 It is easy to see that $\mathrm{I}\Sigma_n$ is a proper subtheory of $\mathsf{PA}_{\Pi_{n+2}}$. The inclusion follows from the fact that each axiom of $\mathrm{I}\Sigma_n$ is of complexity at most Π_{n+2}; for the inequality, use Gödel's second incompleteness theorem and the fact that for all $n \in \omega$, $\mathsf{Con}_{\mathrm{I}\Sigma_n} \in \mathsf{PA}_{\Pi_1}$ (thanks to Mostowski's Reflection Theorem, which asserts that PA and all of its extensions prove the formal consistency of each of their finitely axiomatized subtheories).

16 The result appears in Lessan's doctoral dissertation (Manchester, 1978) that was reprinted in (Cégielski, Cornaros and Dimitracopoulos, 2013).

Recall that Gödel's second incompleteness theorem, when phrased for arbitrary consistent sound arithmetical theories U, states that $U + \neg\mathsf{Con}(U)$ is consistent (see, e.g., Chao and Seraji's (2018)). Here U is sound means that $\mathbb{N} \models U$. Therefore $\mathbb{N} \models \mathsf{Con}(T_n^+)$. Thanks to the arithmetized completeness theorem, this makes it clear that there is an arithmetical model $\mathscr{M} \models T_n^+$; thus $\mathbb{N} \trianglerighteq \mathscr{M}$, so we can fix a translation σ_1 such that $\mathscr{M} = \sigma_1(\mathbb{N})$.

Let $c \in M$ be the least proof of inconsistency of T_n in $\mathscr{M}$. Since True_{Π_n} is Π_n-definable (for $n \geq 1$), the predicate $P(x)$ expressing "x codes a proof of contradiction from $\mathsf{PA} + \mathsf{True}_{\Pi_n}$" is Π_n, and therefore the predicate $Q(x) := P(x) \wedge \forall y < x\, \neg P(x)$ is of the form $\Pi_n \wedge \Sigma_n$; hence c is Σ_{n+1}-definable in $\mathscr{M}$. This shows that $K_{n+1}(\mathscr{M})$ is nonstandard. Thanks to the availability of a truth-definition for Σ_{n+1}-formulae in $\mathscr{M}$, $K_{n+1}(\mathscr{M})$ is interpretable in $\mathscr{M}$, thus there is a translation σ_2 such that $K_{n+1}(\mathscr{M}) = \sigma_2(\mathscr{M})$. This makes it clear that by composing σ_1 and σ_2 we obtain a translation σ such that $K_{n+1}(\mathscr{M}) = \sigma(\mathbb{N})$.

By Theorem 22,

$$K_{n+1}(\mathscr{M}) \models \mathsf{PA}_{\Pi_{n+2}} + \neg \mathsf{B}\Sigma_{n+1},$$

and by Theorem 23, $\mathbb{N}$ is definable in $K_{n+1}(\mathscr{M})$. This makes it clear that $K_{n+1}(\mathscr{M}) \trianglerighteq \mathbb{N}$. It is routine now to verify that $\mathbb{N}$ is a retract of $K_{n+1}(\mathscr{M})$. Note that this shows that $\mathrm{Th}_{\Pi_n}(\mathbb{N})$ is not solid since $\mathbb{N}$ and $K_{n+1}(\mathscr{M})$ are non-isomorphic. Indeed, $K_{n+1}(\mathscr{M})$ is also a retract of $\mathbb{N}$ (with the same interpretations at work) since $K_{n+1}(\mathscr{M})$ has a truth-definition for Σ_{n+1}-formulae, so it can define an isomorphism between itself, and the model that $\mathbb{N}$ cooks up via the translation σ (recall that $K_{n+1}(\mathscr{M}) = \sigma(\mathbb{N})$). This concludes the proof of (a).

The proof of (b) is based on the observation that the proof of (a) works *uniformly* for all models of for $U = \mathrm{Th}(\mathbb{N})$ and $V = \mathrm{Th}(K_{n+1}(\mathscr{M}))$.

□

Part(b) of Theorem 24 immediately implies the following result.

Theorem 25. *The following statements hold for each $n \in \omega$.*

(a) $\mathsf{Th}_{\Pi_n}(\mathbb{N})$ *is not tight.*

(b) PA_{Π_n} *is not tight.*

The failure of tightness (and, a fortiori, the failure of solidity) for bounded fragments of PA (i.e., those axiomatized by a collection of sentences of bounded quantifier complexity) naturally prompts one to investigate the possibility of solidity/tightness of commonly studied fragments of PA that are unbounded. One such fragment is presented in the following definition.

Definition 26. Given a recursively enumerable theory T extending $\mathrm{I}\Delta_0 + \mathsf{Exp}$, let Ref_T consist of the collection of sentences of the form

$$\mathsf{Prov}_T(\ulcorner \varphi \urcorner) \to \varphi,$$

where φ ranges over arithmetical sentences, and $\mathsf{Prov}_T(x)$ is the arithmetical sentence that expresses "φ is provable in T ".

Proposition 27. *(Feferman, 1960).* $T + \mathsf{Th}_{\Pi_1}(\mathbb{N}) \vdash \mathsf{Ref}_T$.[17]

Proof. It suffices to show that if $\mathscr{M} \models T + \mathsf{Th}_{\Pi_1}(\mathbb{N})$, then $\mathscr{M} \models \mathsf{Prov}_T(\ulcorner \varphi \urcorner) \to \varphi$ for any given arithmetical sentence φ. Suppose not, then:

(1) $\mathscr{M} \models \mathsf{Prov}_T(\ulcorner \varphi \urcorner)$, and

(2) $\mathscr{M} \models \neg\varphi$.

The key observation is that since $\mathsf{Prov}_T(\ulcorner \varphi \urcorner)$ is a Σ_1-statement, $\mathbb{N} \models \mathsf{Prov}_T(\ulcorner \varphi \urcorner)$; otherwise $\mathbb{N} \models \neg\mathsf{Prov}_T(\ulcorner \varphi \urcorner)$, which by the assumption that $\mathscr{M} \models \mathsf{Th}_{\Pi_1}(\mathbb{N})$, contradicts (1). But if $\mathbb{N} \models \mathsf{Prov}_T(\ulcorner \varphi \urcorner)$, then φ is provable in T, which contradicts (2). □

17 In Feferman's formulation, T was specifically chosen as PA, but the same reasoning handles the general case. Also, as pointed by Lev Beklemishev (in private conversation) Feferman noticed this fact while thinking on Turing's (false) hope that iterated local reflection can prove all true Π_2-sentences.

Theorem 28. $\mathsf{PA}_{\Pi_n} + \mathsf{Ref}_{\mathsf{PA}_{\Pi_n}}$ *is not tight for each* $n \in \omega$.

Proof. This is readily established by putting Theorem 25 together with Proposition 27 (for $T = \mathsf{PA}_{\Pi_n}$). $\square$

Remark 29. It is important to bear in mind that the stronger scheme REF_T whose instances are *uniform versions* of the instances of Ref_T, i.e., implications of the form

$$\forall x\, \mathsf{Prov}_T(\ulcorner \varphi(\dot{x}) \urcorner) \to \varphi(x)),$$

behaves very differently, since by a classical result of Kreisel and Levy, (1968), for $T = \mathrm{I}\Delta_0 + \mathsf{Exp}$ the deductive closure of $T + \mathsf{REF}_T$ coincides with the deductive closure of PA.

Another unbounded subtheory of PA that we shall consider is $\mathsf{PA}^- + \mathsf{Collection}$. The following definition describes an important family of models of the well-known fragment PA^- of PA, whose axioms describe the non-negative substructure of discretely ordered rings (with no instance of the induction scheme, hence the minus superscript), as in Chapter 2 of Kaye's text (1991) on models of PA. As we shall see, $\mathsf{PA}^- + \mathsf{Collection}$ fails to be solid.

Definition 30. Let $\mathbb{Z}$ be the ring of integers and $(I, <_I)$ be a linear order.

(a) $\mathbb{Z}[X_i : i \in I]$ is the ring of polynomials with coefficients in $\mathbb{Z}$, whose indeterminates come from the I-indexed collection $\{X_i : i \in I\}$ of indeterminates.

(b) Each nonconstant $p \in \mathbb{Z}[X_i : i \in I]$ can be written as:

$$p(X_{i_1}, X_{i_2}, \cdots, X_{i_n}) \text{ where } i_1 <_I i_2 <_I \cdots <_I i_n,$$

for some finite subset $\{i_1, i_2, \cdots, i_n\}$ of I, with the understanding that $\{i_1, i_2, \cdots, i_n\}$ is the least (in the sense of inclusion) subset of $A \subseteq I$ such that $p \in \mathbb{Z}[X_i : i \in A]$. With this notation in mind, we refer to $\{i_1, i_2, \cdots, i_n\}$ as the *support* of p.

(c) The ordering on $\mathbb{Z}[X_i : i \in I]$ is defined by first declaring $p(X_{i_1}, X_{i_2}, \cdots, X_{i_n}) > 0$, provided the coefficient of X_{i_n} is positive; and then given p and q in $\mathbb{Z}[X_i : i \in I]$, we define $p > q$ iff $p - q > 0$.

The following proposition is well-known and easily verified.

Proposition 31. *For every linear order* $(I, <_I)$ *the substructure* $\mathbb{Z}[X_i : i \in I]^{\geq 0}$ *of non-negative elements of* $\mathbb{Z}[X_i : i \in I]$ *is a model of* PA^-.

Remark 32. The set ω of standard natural numbers (equivalently: constant non-negative polynomials) is defined in $\mathbb{Z}[X_i : i \in I]^{\geq 0}$ as the longest initial segment of elements with parity, i.e., for all $p \in \mathbb{Z}[X_i : i \in I]^{\geq 0}$, $p \in \omega$ iff $\forall q \leq p\, \exists r\, [(q = 2r) \vee (q = 2r+1)]$.

In the proposition below, $\mathbb{Q}$ is the set of rational numbers and $\mathbb{Q}^-$ is the set of negative rational numbers, both equipped with their natural ordering.

Proposition 33. $\mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0}$ *elementarily end extends* $\mathbb{Z}[X_i : i \in \mathbb{Q}^-]^{\geq 0}$.

Proof. It is routine to verify that $\mathbb{Z}[X_i : i \in \mathbb{Q}^-]^{\geq 0}$ is end extended by $\mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0}$, so we focus on verifying elementarity. Suppose $\varphi(p_1, \cdots, p_k)$ is an arithmetical sentence all of whose parameters $p_1, \cdots, p_k$ are from $\mathbb{Z}[X_i : i \in \mathbb{Q}^-]^{\geq 0}$. Since the support of each polynomial is finite, there is an initial segment $\mathbb{S}$ of $\mathbb{Q}^-$ such that:

(1) $\mathbb{S}$ has no last element;

(2) $\mathbb{Q}^- \setminus \mathbb{S}$ has no first element; and

(3) For $1 \leq i \leq k$, the support of p_i is a subset of $\mathbb{S}$.

Note that (1) implies that $\mathbb{S}$ is a DLO (dense linear order without end points), and (2) implies that both $\mathbb{Q}^-\backslash\mathbb{S}$ and $\mathbb{Q}\backslash\mathbb{S}$ are also DLOs. Thanks to $\aleph_0$-categoricity of DLOs, this shows that there is an order isomorphism $f : \mathbb{Q}^- \to \mathbb{Q}$ such that $f(s) = s$ for each $s \in \mathbb{S}$. This isomorphism naturally induces an isomorphism F with:

$$F : \mathbb{Z}\left[X_i : i \in \mathbb{Q}^-\right]^{\geq 0} \to \mathbb{Z}\left[X_i : i \in \mathbb{Q}\right]^{\geq 0},$$

such that F fixes all elements in $\mathbb{Z}\left[X_i : i \in \mathbb{S}\right]^{\geq 0}$. This makes it clear that $\varphi(p_1, \cdots, p_k)$ holds in $\mathbb{Z}\left[X_i : i \in \mathbb{Q}^-\right]^{\geq 0}$ iff $\varphi(p_1, \cdots, p_k)$ holds in $\mathbb{Z}\left[X_i : i \in \mathbb{Q}\right]^{\geq 0}$, since by an elementary fact of model theory if g is an isomorphism between two $\mathscr{L}$-structures $\mathscr{M}$ and $\mathscr{N}$, then for any n-ary $\mathscr{L}$-formula $\varphi(x_1, \cdots, x_k)$, and any k-tuple $(a_1, \cdots, a_k)$ from $\mathscr{M}$, $\mathscr{M} \models \varphi(a_1, \cdots, a_k)$ iff $\mathscr{N} \models \varphi(g(a_1), \cdots, g(a_k))$. □

In light of the well-known fact that the Collection scheme holds in linearly ordered structures that possess an elementary end extension[18], the following is an immediate consequence of Proposition 33.

Corollary 34. $\mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0} \models \mathsf{Collection}$.

Theorem 35. $\mathsf{PA}^- + \mathsf{Collection}$ *is not solid.*[19]

Proof. By Proposition 31 and Corollary 34 $\mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0}$ is a model of $\mathsf{PA}^- + \mathsf{Collection}$, and of course so is $\mathbb{N}$, and the two structures are nonisomorphic, so it suffices to verify that there are $\mathscr{I}$ and $\mathscr{J}$ such that:

$$\mathbb{N} \trianglerighteq^{\mathscr{I}} \mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0} \trianglerighteq^{\mathscr{J}} \mathbb{N},$$

and there is an $\mathbb{N}$-definable isomorphism $i : \mathbb{N} \to \mathbb{N}^{\mathscr{J} \circ \mathscr{I}}$. The description of $\mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0}$, and the fact that there is an arithmetically definable ordering $<_{\mathbb{Q}}$ on ω such that $(\omega, <_{\mathbb{Q}}) \cong \mathbb{Q}$ yields $\mathscr{I}$ such that $\mathbb{N} \trianglerighteq^{\mathscr{I}} \mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0}$. On the other hand, the definability of ω in $\mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0}$ (as indicated in Remark 32) provides $\mathscr{J}$ such that $\mathbb{Z}[X_i : i \in \mathbb{Q}]^{\geq 0} \trianglerighteq^{\mathscr{J}} \mathbb{N}$, and there is an $\mathbb{N}$-definable isomorphism $i : \mathbb{N} \to \mathbb{N}^{\mathscr{J} \circ \mathscr{I}}$. □

Remark 36. We conjecture that the method of proof of Theorem can be adapted to show the failure of solidity of the much stronger theory $\mathsf{IOpen} + \mathsf{Collection}$, where IOpen is the fragment of PA in which the induction scheme is limited to *open formulae* (i.e., quantifier-free formulae). Note that $\mathsf{IOpen} + \mathsf{Collection}$ is a proper subtheory of PA. Our conjecture is based on the following two well-known facts:

(1) There is a countable nonstandard model $\mathscr{M}_0$ of IOpen that has no nonstandard primes; this implies, thanks to Bertrand's 'postulate', that the standard cut is definable in $\mathscr{M}_0$ by the Δ_0 formula $\delta(x) :=$ "There is a prime number between x and $2x$". This result is due to Shepherdson (1964).

(2) As noted by Marker (1991) every model of IOpen has an end extension to a model of IOpen; the basic ingredients of Marker's construction are found in Shepherdson's paper (1964); the end extension is built using the technology of Puiseux series/polynomials.

More specifically, since the axioms of IOpen are of the form $\forall\exists$, they are preserved under unions of chains, so (2) can be used to build a continuous end-extension chain of countable model $\langle \mathscr{M}_\alpha : \alpha < \omega_1 \rangle$, where $\mathscr{M}_0$ is as in (1). This makes it clear that the union $\mathscr{M}_{\omega_1}$ of the chain satisfies $\mathsf{IOpen} + \mathsf{Coll}$ and end extends $\mathscr{M}_0$. It is evident that $\mathscr{M}_{\omega_1}$ is a model of IOpen in which the standard cut is definable, thus $\mathscr{M}_{\omega_1}$ is not a model of PA. By a variation of this argument, we can build an arithmetically definable model $\mathscr{M}(I)$ of IOpen for any arithmetically definable linear order $(I, <_I)$ such that $\mathscr{M}(\mathbb{Q}^-)$ is elementarily end extended by $\mathscr{M}(\mathbb{Q})$, and thus $\mathscr{M}(\mathbb{Q})$ satisfies the collection scheme. Then with an adaptation of the proof strategy of Theorem 35 one can verify the failure of solidity of $\mathsf{IOpen} + \mathsf{Collection}$. A detailed description of $\mathscr{M}(\mathbb{Q})$ is presented in (Enayat, Łełyk and Visser, to appear).

Remark 37. In contrast with Theorem 35 and Remark 36, $\mathrm{I}\Delta_0 + \mathsf{Collection}$ is a solid theory since it is well-known (see Kaye, 1991) that it is deductively equivalent to PA, which is solid by Theorem 10.

[18] Indeed, by a classical result of Keisler, for countable models, the converse also holds, i.e., if a countable linearly ordered model $\mathscr{M}$ (in a countable language) satisfies the collection scheme, then $\mathscr{M}$ has an elementary end extension.

[19] Perhaps the proof of this theorem can be fine-tuned so as to demonstrate the failure of tightness of $\mathsf{PA}^- + \mathsf{Collection}$, but we have not yet been able to formulate such a fine-tuning.

We close this subsection by showing that Theorem 25 can be extended to other theories whose solidity was asserted in Theorem 14.

Theorem 38. *No finitely axiomatizable subtheory of T is tight, where T is any of the theories in the list $\mathscr{S}$ of Theorem 14.*

Below in Theorem 39 we present a stronger version of the above result for $T = \mathsf{ZF}$; our method of proof together with Theorem 5 make it clear that similar results can be given for each of the theories T whose solidity was asserted in Theorem 14. What allows us to extend the argument is that the analogue of Theorem 5 holds for Z_α and KM_α for each $\alpha \leq \omega$, as shown in (Marek and Zbierski, 1978).

Our method of proof is similar to the proof of Theorem 24. In what follows we use the notation Δ_0, Σ_n, and Π_n in the context of the *Levy* hierarchy of set-theoretical formulae; and we use ZF_{Π_n} to denote $\{\varphi \in \Pi_n : \mathsf{ZF} \vdash \varphi\}$.

Theorem 39. *If ZF is consistent, then for each $n \in \omega$ the following hold:*

(a) ZF_{Π_n} *is not solid.*

(b) ZF_{Π_n} *is not tight.*

Proof. Assuming the consistency of ZF, by Gödel's second incompleteness theorem, and Gödel's relative consistency proof of $\mathsf{ZF} + \mathsf{V} = \mathsf{L}$, $\mathsf{ZF} + \mathsf{V} = \mathsf{L} + \neg\mathsf{Con}_{\mathsf{ZF}}$ is consistent, and therefore by the arithmetized completeness theorem there is a countable model $\mathscr{M}$ such that:

(1) $\mathscr{M} \models \mathsf{ZF} + \mathsf{V} = \mathsf{L} + \neg\mathsf{Con}_{\mathsf{ZF}}$, and

(2) The elementary diagram of $\mathscr{M}$ is definable in $\mathbb{N}$; in particular, $\mathbb{N} \trianglerighteq \mathscr{M}$.

As in the arithmetical case, given $n \in \omega$, let $K_n(\mathscr{M})$ be the submodel of $\mathscr{M}$ whose universe consists of elements of $\mathscr{M}$ that are definable in $\mathscr{M}$ by a Σ_n-formula.

The lemma below is the analogue of Theorem 22. Notice that there is a 'lag' in relation to the arithmetical case in the amount of theory-inheritance of $K_n(\mathscr{M})$ from the parent-structure $\mathscr{M}$; this is due to the higher complexity of the global well-ordering at work in set theory (under the assumption that $\mathsf{V} = \mathsf{L}$), as we shall see. Also notice that in part (b) there is no nonstandardness assumption. In part (b) of Lemma 40, $\mathsf{Coll}(\Sigma_n)$ is the collection scheme for Σ_n-formulae of set theory, whose instances are of the form:

$$(\forall x \in a\ \exists y\ \psi(x,y)) \to (\exists b\ \forall x \in a\ \exists y \in b\ \psi(x,y)),$$

where ψ is a Σ_n-formula whose parameters are suppressed.

Lemma 40. *Suppose $\mathscr{M} \models \mathsf{ZF} + \mathsf{V} = \mathsf{L}$, and $n \in \omega$ with $n \geq 2$.*

(a) $K_n(\mathscr{M}) \prec_{\Pi_{n-1}} \mathscr{M}$, *hence* $K_n(\mathscr{M}) \models \mathrm{Th}_{\Pi_n}(\mathscr{M})$.

(b) $K_n(\mathscr{M}) \models \mathsf{ZF}_{\Pi_n} + \neg\mathsf{Coll}(\Sigma_{n+1})$.

Proof of Lemma 40. The proof of part (a) is similar to the proof of part (a) of Theorem 22 except for one important difference: in the arithmetical case, the well-ordering $<$ of the universe has a Δ_0-graph, but within $\mathsf{ZF} + \mathsf{V} = \mathsf{L}$ the canonical well-order $<_{\mathsf{L}}$ of the universe has a Σ_1-graph (see Lemma 13.19 of Jech, 2003). We present the proof of (a) for $n = 2$, the general case is handled similarly.

By the Tarski-Vaught test, to show that $K_2(\mathscr{M}) \prec_{\Pi_1} \mathscr{M}$ it suffices to verify that if for some $a \in K_2(\mathscr{M})$ we have $\mathscr{M} \models \exists x \sigma(x,a)$ for some Σ_1-formula $\sigma(x,y)$, then there is some $b \in K_2(\mathscr{M})$ such that $\mathscr{M} \models \sigma(b,a)$. Towards this goal, assume that for some $a \in K_2(\mathscr{M})$ we have:

$$\mathscr{M} \models \exists x\ \exists v\ \delta(x,v,a),$$

where $\delta(x,v,a)$ is Δ_0. Thanks to the availability of a Δ_0-definable ordered pairing function $\langle\cdot,\cdot\rangle$ we can re-write the above as:

$$\mathscr{M} \models \exists w\ \delta((w)_0,(w)_1,a),$$

where $w = \langle (w)_0,(w)_1\rangle$. In the next step we minimize the w in (3) using the Σ_1-definable global well-ordering $<_{\mathsf{L}}$ by choosing c in $\mathscr{M}$ such that:

$$\mathscr{M} \models \delta((c)_0,(c)_1,a) \wedge \forall z\,(z <_{\mathsf{L}} c \to \neg\delta((c)_0,(c)_1,a)).$$

An inspection of the above makes it clear c has a Π_1-definition $\pi(y,a)$ (with parameter a) in $\mathscr{M}$. Since the element a is Σ_2-definable in $\mathscr{M}$ by assumption, let $\psi(y)$ be a Σ_2-formula defining a in $\mathscr{M}$. It suffices to show that by eliminating a from $\pi(y,a)$ with the help of $\psi(y)$ we reach a Σ_2-definition of c in $\mathscr{M}$. Towards this aim, consider the formula:

$$\theta(x) := \exists y(\pi(x,y) \wedge \psi(y)).$$

Clearly c is the unique element in $\mathscr{M}$ that satisfies $\theta(x)$. The formula $\theta(x)$ is of the form $\exists(\Sigma_2 \vee \Sigma_1)$, so it can be put in the form $\exists(\Sigma_2)$, which makes it clear that θ is Σ_2. This concludes the proof of part (a).

We now turn to the proof of (b). It is clear from part (a) that $K_n(\mathscr{M}) \models \mathsf{ZF}_{\Pi_{n-1}}$ since $\mathscr{M}$ is assumed to be a model of ZF. So we focus on the failure of Σ_n-Collection in $K_n(\mathscr{M}) \models \mathsf{ZF}_{\Pi_{n-1}}$. But first we need to address the question of definability of Σ_k-satisfaction predicate within *fragments* of ZF (it has been known since Levy's pioneering work that within full ZF there is a Σ_k-satisfaction class that is Σ_k-definable for each $k \geq 1$).

As pointed out in Theorem 2.9 of (Enayat and McKenzie, 2021) (the statement of which involves KP + the axiom of infinity, but the axiom of infinity is not used in the proof) within KP there is a Σ_k-satisfaction class that is Σ_k-definable for each $k \geq 1$.[20] Here KP is Kripke-Platek set theory; a frugal theory of sets whose axioms do not include the axiom of infinity. In our formulation (following recent practice, led by Mathias, 2021) the scheme of foundation is limited to Π_1-formulae (equivalently: the scheme of $\in$-induction for Σ_1-formulae). Thus in contrast to Barwise's KP in (1975), which includes the full scheme of foundation, our version of KP is *finitely axiomatizable*; and a straightforward calculation shows that it is axiomatizable by a Π_2-sentence.

With the above preliminaries in place, for $n \geq 2$, $K_n(\mathscr{M}) \models \mathsf{KP}$, thanks to Π_2-axiomatizability of KP and part (a) of Lemma 40. Therefore there is a Σ_n-satisfaction predicate in $K_n(\mathscr{M})$ for Σ_n-formulae. Consider the function f that maps (the code of) each Σ_n-formula in $K_n(\mathscr{M})$ to 0, if there is no element that satisfies $\sigma(x)$, and otherwise to the least $<_{\mathsf{L}}$-element satisfying $\sigma(x)$. Thanks to the availability of a Σ_n-definable satisfaction predicate in $K_n(\mathscr{M})$ for Σ_n-formula, the graph of f is readily seen to be defined by a Σ_{n+1}-formula. Note that the domain of f is a set in $K_n(\mathscr{M})$ but its range is the whole of $K_n(\mathscr{M})$. Thus $\mathsf{Coll}(\Sigma_{n+1})$-fails in $K_n(\mathscr{M})$. This concludes the proof of part (b) Lemma 40. □

Lemma 41. *The standard cut ω is definable in $K_n(\mathscr{M})$ for each $n \in \omega$ such that $n \geq 2$.*

Proof. Recall that $K_n(\mathscr{M})$ is not ω-standard since it satisfies $\neg\mathsf{Con}_{\mathsf{ZF}}$. So, similar to the proof of Theorem 23, i is a nonstandard member of $\omega^{K_n(\mathscr{M})}$ iff every element of $K_n(\mathscr{M})$ is definable by a Σ_n-formula below i. This is first order expressible, thanks to the definability of a Σ_n-satisfaction predicate in $K_n(\mathscr{M})$. □

By Lemma 40 given any $2 \leq n < k$, both in ω, $\mathbb{N}$ is bi-interpretable with a model of $\mathsf{ZF}_{\Pi_n} + \neg\mathsf{Coll}(\Sigma_{n+1})$, and also $\mathbb{N}$ is bi-interpretable with a model of ZF_{Π_k}. Since $\mathsf{Coll}(\Sigma_n)$ is of complexity at most Π_{n+3}, this shows that by choosing $k \geq n+3$, $\mathsf{Coll}(\Sigma_n) \in \mathsf{ZF}_{\Pi_k}$, which in turn makes it clear that ZF_{Π_n} is not solid.

To see that ZF_{Π_n} is not tight it suffices to note that the above argument works uniformly; thus there are distinct deductively closed extension of ZF_{Π_n} that are bi-interpretable. This concludes the proof of Theorem 39. □

[20] The existence of definable partial satisfaction classes in KP follows from two facts: (1) KP can prove that every set is contained in a transitive set; and (2) KP can define the satisfaction predicate for all of its internal set structures. The proofs of both of these facts can be found in Barwise's monograph (1975); the proofs therein make it clear that only Π_1-Foundation is invoked.

4. Categoricity of schemes

In this section we investigate various categoricity-like notions within the framework of first order logic that pertain specifically to *scheme-templates* (as opposed to arbitrary first order theories). These notions include adaptations of the notions introduced in Definition 7, but there is one which has no counterpart in Definition 7, namely, internal categoricity. Internal categoricity has been widely present in philosophical discussions in the context PA and ZF but to our knowledge it has not been previously probed in the context of general first-order schematic theories. In Subsection 4.1 we focus on internal categoricity, and in Subsection 4.2 we investigate various generalizations of the notions introduced in Definition 7.

Definition 42. Assume $\mathscr{L}$ is a finite language. In what follows all languages are allowed to be n-sorted for some $n \in \omega$.

(a) An *$\mathscr{L}$-template* (for a scheme) is given by an $\mathscr{L}$-sentence $\tau(P)$ formulated in the language obtained by augmenting $\mathscr{L}$ with an n-ary predicate $P(x_1,...,x_n)$ for some $n \in \omega$, where $x_1,\ldots x_n$ belong to the same sort.[21] Given a language $\mathscr{L}^+ \supseteq \mathscr{L}$, an $\mathscr{L}^+$-sentence ψ is said to be *an instance of* τ if ψ is of the form $\forall v\ \tau[\varphi(x_1,..,x_n,v)/P]$, where $\tau[\varphi(x_1,..,x_n,v)/P]$ is the result of substituting all subformulae of the form $P(t_1,...,t_n)$, where each t_i is a term, with $\varphi(t_1,t_2,\cdots,t_n,v)$ (and re-naming bound variables of φ to avoid unintended clashes).

By calling τ an $\mathscr{L}$-template we shall mean that $\mathscr{L}$ is the smallest language (w.r.t. to the inclusion) which augmented with P contains τ.

(b) Suppose $\mathscr{L}^+ \supseteq \mathscr{L}$. We use $\tau[\mathscr{L}^+]$ to denote the collection of all $\mathscr{L}^+$-formulae that are instances of τ. We say that an T is *axiomatized by a scheme* if $T = \tau[\mathscr{L}^+]$ for some scheme τ and some $\mathscr{L}^+$.

(c) For familiar schematic theories such as PA and ZF the notation $\mathsf{PA}(\mathscr{L}^+)$ and $\mathsf{ZF}(\mathscr{L}^+)$ is often used in the literature to respectively denote $\tau_{\mathsf{Ind}}[\mathscr{L}^+]$ and $\tau_{\mathsf{Repl}}[\mathscr{L}^+]$.[22] Here τ_{Ind} is the template that is conjunction of two sentences, the first conjunct of which in turn is the conjunction of the finitely many axioms of (Robinson's) Q, and the second conjunct of which is the sentence expressing "If P includes 0 and is closed under successors, then P is everything". Similarly, τ_{Repl} is the template that is conjunction of two sentences, the first of which asserts that Emptyset, Pairs, Union, Infinity, and Powerset hold, and the second conjunct of which is the sentence expressing "If P is a class function whose domain is a set, then its range is also a set".

Remark 43. Let $\mathscr{L}_{\mathsf{Arith}}$ denote the language of arithmetic. As shown in (Enayat and Łełyk, 2023), the complexity of the collection of (Gödel numbers) of $\mathscr{L}_{\mathsf{Arith}}$-templates τ such that $\tau[\mathscr{L}_{\mathsf{Arith}}] = \mathsf{PA}$ is a complete Π_2-set in the arithmetical hierarchy, and in particular it is not recursively enumerable.

4.1. Internal Categoricity

The notion of internal categoricity was initially explored in the context of second order logic (see the recent monograph by Maddy and Väänänen, 2023, for references). The definition below of internal categoricity generalizes Väänänen's formulation, who focused on the internal categoricity of the first order formulations of Peano Arithmetic and Zermelo-Fraenkel set theory, as summarized in (Väänänen, 2021). More specifically, the notion of internal categoricity encapsulated in the definition below coincides with the one formulated by Väänänen in the setting of the usual schematic formulations of PA and ZF, so we recommend the readers unfamiliar with the notion of internal categoricity to consult (Väänänen, 2021) for the motivation of the definition below.

21 The theories we study all have access to a pairing function and therefore by elementary coding we can limit ourselves to a single unary predicate P, but for the purposes of exposition we are not insisting on P being unary. Note that notion of a schematic axiomatization presented here is not affected in our context where a pairing function is available if the template τ is allowed to use finitely many predicate symbols $P_1,\cdots,P_k$ of various finite arities.

22 The notation PA^* is used in the model theory of arithmetic for $\mathsf{PA}(\mathscr{L})$, where $\mathscr{L}$ is clear from the context, or unimportant.

Definition 44. Suppose $\tau(P)$ is an $\mathscr{L}$-template.

(a) Let $\mathscr{L}^{\text{red}}$ and $\mathscr{L}^{\text{blue}}$ be two fresh disjoint copies of the language $\mathscr{L}$ and R and B be two fresh unary predicates. Let σ_r and σ_b be the trivial direct translations of $\mathscr{L}$ into $\mathscr{L}^{\text{red}}$ and $\mathscr{L}^{\text{blue}}$ respectively and let σ_r^{rel} (σ_b^{rel}) be the extension of σ_r (σ_b, resp.) which relativize the quantifiers to R (B, resp.). The result of applying σ_r (σ_b, σ_r^{rel}, σ_b^{rel}) to τ will be denoted τ^{red} (τ^{blue}, $\tau^{\text{red}}_{\text{rel}}$, $\tau^{\text{blue}}_{\text{rel}}$, respectively). We assume that each of the translations acts trivially on P.

(b) Let $\mathscr{L}^{\text{duo}} = \mathscr{L}^{\text{red}} \cup \mathscr{L}^{\text{blue}}$ and let

$$\tau^{\text{duo}} = \tau^{\text{red}}(P) \wedge \tau^{\text{blue}}(P).$$

$\tau(P)$ is said to be *internally categorical* if there is an $\mathscr{L}^{\text{duo}}$-binary formula $\alpha(x,y)$ such that, provably in $\tau^{\text{duo}}[\mathscr{L}^{\text{duo}}]$, α defines an isomorphism between the $\mathscr{L}^{\text{red}}$-reduct and the $\mathscr{L}^{\text{blue}}$-reduct of the universe, i.e., for all models $\mathscr{M}$ of $\tau^{\text{duo}}[\mathscr{L}^{\text{duo}}]$, α defines an isomorphism between $\mathscr{M}^{\text{red}}$ (the $\mathscr{L}^{\text{red}}$-reduct of $\mathscr{M}$) and $\mathscr{M}^{\text{blue}}$ (the $\mathscr{L}^{\text{blue}}$-reduct of $\mathscr{M}$).[23]

(c) Let $\mathscr{L}^{\text{duo}+} = \mathscr{L} \cup \mathscr{L}^{\text{blue}} \cup \mathscr{L}^{\text{red}} \cup \{R, B\}$ and let

$$\tau^{\text{duo}+} := \tau[\mathscr{L}] \cup \tau^{\text{red}}_{\text{rel}}[\mathscr{L}^{\text{duo}+}] \cup \tau^{\text{blue}}_{\text{rel}}[\mathscr{L}^{\text{duo}+}].$$

Thus in contrast with $\mathscr{L}^{\text{duo}}$-structures $\mathscr{M}$, the domain of discourse of the 'colored' submodels are allowed to be proper subsets of the domain of discourse of an $\mathscr{L}^{\text{duo}+}$ structure $\mathscr{M}$. τ is said to be *strongly internally categorical* iff there is an $\mathscr{L}^{\text{duo}+}$-binary formula $\alpha(x,y)$ such that, provably in $T^{\text{duo}+}$, α defines an isomorphism between the $\mathscr{L}^{\text{red}}$ model with the universe given by R (called *the red model*) and $\mathscr{L}^{\text{blue}}$ model with the universe given by B (*the blue model*). From the model-theoretic perspective, this is equivalent to α defining an isomorphism between $\sigma_r^{\text{rel}}(\mathscr{M})$ (the red model) and $\sigma_b^{\text{rel}}(\mathscr{M})$ (the blue model) for an arbitrary $\mathscr{L}^{\text{duo}+}$ model $\mathscr{M} \models T^{\text{duo}+}$.

Example 45. Let τ be τ_{Repl}, $\mathscr{L}$ be the language of set theory, and κ_1, κ_2, and κ_3 be strongly inaccessible cardinals, with $\kappa_1 > \kappa_2 > \kappa_3$. Then we can expand $(V_{\kappa_1}, \in)$ into a model $\tau^{\text{duo}+}$ by interpreting R as V_{κ_2}, B as V_{κ_3}, and interpreting $\in$, $\in^{\text{red}}$, and $\in^{\text{blue}}$ as the membership relation in the real world.

The following proposition motivates our choice of the terminology "strong internal categoricity":

Proposition 46. *If a scheme template τ is strongly internally categorical, then τ is internally categorical.*

Proof. Assume that τ is strongly internally categorical and let α be a $\mathscr{L}^{\text{duo}+}$ binary formula which witness the isomorphism. Let α_r result from α by

- substituting each occurrence of $R(t)$ and $B(t)$ with $t = t$ and
- colouring each symbol of $\mathscr{L}$ into red (i.e. applying σ_r to the formula obtained in the first step).

We claim that α_r provably in $\tau^{\text{duo}}[\mathscr{L}^{\text{duo}}]$ defines an isomorphism between the red and the blue part of the universe. Indeed, take any model $\mathscr{M} \models \tau^{\text{duo}}[\mathscr{L}^{\text{duo}}]$. By interpreting R and B as the whole universe and interpreting $\mathscr{L}$ symbols as their red counterparts, one can see that $\mathscr{M}$ definably expands to a model $\mathscr{M}^+ \models T^{\text{duo}+}$. Hence α defines in $\mathscr{M}^+$ an isomorphism between the red and the blue model. By definitions, α_r defines an isomorphism between $\mathscr{M}^{\text{red}}$ and $\mathscr{M}^{\text{blue}}$. □

The converse of the above proposition need not hold; this was implicitly observed in (Maddy and Väänänen, 2023). Part (a) of the theorem below appears without proof in (Maddy and Väänänen, 2023) as Theorem 9, p. 32, where the result is stated in a different notation.

[23] By Compactness, this is equivalent to: For all models $\mathscr{M}$ of $\tau^{\text{duo}}[\mathscr{L}^{\text{duo}}]$, there is an $\mathscr{M}$-definable isomorphism between $\mathscr{M}^{\text{red}}$ and $\mathscr{M}^{\text{blue}}$.

Theorem 47. *(Väänänen) Let τ_{Ind} and τ_{Repl} be the scheme templates axiomatizing* PA *and* ZF *(respectively) as in part (c) of Definition 42.*

(a) τ_{Ind} *is strongly internally categorical.*

(b) τ_{Repl} *is internally categorical, but not strongly internally categorical (assuming the consistency of* $\mathsf{ZF} + \exists\kappa\mathsf{Inacc}(\kappa)$*).*

Proof. To see that (a) holds, suppose $\mathscr{M}$ is a model of PA, and $\mathscr{N}_1$ and $\mathscr{N}_2$ are models of PA such that for $i \in \{1,2\}$ the domain of discourse N_i of $\mathscr{N}_i$ is a subset of the domain of discourse M of $\mathscr{M}$ (and therefore the graphs of addition and multiplication for $\mathscr{N}_i$ are subsets of M^2). Furthermore, assume:

(∇) For each subset D of M that is parametrically definable in $(\mathscr{M}, \mathscr{N}_1, \mathscr{N}_2)$, and for each $i \in \{1,2\}$, $(\mathscr{N}_i, D \cap N_i) \models \mathsf{PA}(X)$, where X is interpreted by $D \cap N_i$.

Let $\varphi(x, v_1, v_2)$ be the formula that expresses that $v_1 \in N_1$, $v_2 \in N_2$ and x codes an isomorphism between the initial segment of $\mathscr{N}_1$ determined by v_1 and the initial segment of $\mathscr{N}_2$ determined by v_2 (when addition and multiplication are viewed as ternary relations).

Let
$$I_1 = \{a \in N_1 : (\mathscr{M}, \mathscr{N}_1, \mathscr{N}_2) \models \exists x \exists v_2\ \varphi(x, a, v_2)\},$$
and
$$I_2 = \{b \in N_2 : (\mathscr{M}, \mathscr{N}_1, \mathscr{N}_2) \models \exists x \exists v_1\ \varphi(x, v_1, b)\}.$$
Using the assumptions on $\mathscr{N}_1$ and $\mathscr{N}_2$ it is easy to see that I_1 is an initial segment of N_1 that contains $0^{\mathscr{N}_1}$ and is closed under the successor relation of $\mathscr{N}_1$, and therefore by (∇) $I_1 = N_1$; and I_2 is an initial segment of N_2 that contains $0^{\mathscr{N}_2}$ and is closed under the successor relation of $\mathscr{N}_2$ and therefore $I_2 = N_2$. It can also be readily shown using (∇) that for every a in $\mathscr{N}_1$ there is a unique b in $\mathscr{N}_2$ such that $(\mathscr{M}, \mathscr{N}_1, \mathscr{N}_2) \models \exists x \varphi(x, a, b)$. This makes it clear that there is an $(\mathscr{M}, \mathscr{N}_1, \mathscr{N}_2)$-definable isomorphism between $\mathscr{N}_1$ and $\mathscr{N}_2$.[24]

For (b), the proof that τ_{Repl} is internally categorical appears in (Väänänen, 2019) (note the proof is stated for the schematic axiomatization of ZFC, but the axiom of choice can be dispensed with). Let $\mathscr{M}$ be a model of ZF and let $\kappa \in M$ be a strongly inaccessible cardinal (we do not assume that $\mathscr{M}$ is well-founded). We expand $\mathscr{M}$ to a model of $\mathscr{M}^+$ by interpreting

- $\in, \in^{\text{red}}, \in^{\text{blue}}$ as $\in^{\mathscr{M}}$,
- R as the whole universe and
- B as $(V_\kappa)^{\mathscr{M}}$.

Since in $\mathscr{M}^+$ the red model is just the whole universe, then clearly, the red part satisfies the replacement scheme. That for every parametrically definable subset D of $\mathscr{M}$ the expansion $(V_\kappa^{\mathscr{M}}, \in^{\mathscr{M}}, D \cap V_\kappa^{\mathscr{M}})$ of the blue model satisfies the replacement scheme in the extended language follows from basic properties of strongly inaccessible cardinals. □

Remark 48. The direct proof of tightness of ZF (as presented in Freire and Hamkins, 2021) is similar to proof of internal categoricity of τ_{Repl} (as presented in Väänänen, 2019), and they can be viewed as refinements of Zermelo's quasi-categoricity theorem for ZF. However, the proof of solidity of ZF in (Enayat, 2016) involves a step that is not present in the proofs of tightness/internal categoricity of ZF, namely, an appeal to Tarski's undefinability of truth theorem.

[24] It is worth pointing out that the proof makes it clear that the assumption $\mathscr{M} \models \mathsf{PA}$ can be substantially reduced to the assumption that the theory of $\mathscr{M}$ is sequential. Thus in light of the sequentiality of PA^- (established in Jeřábek, 2012) it is sufficient to assume that $\mathscr{M} \models \mathsf{PA}^-$.

Remark 49. Hamkins and Freire (2021) showed that two distinguished schematic subtheories of ZF fail to be internally categorical, namely Zermelo set theory (axiomatized by the separation scheme), and ZF^- (axiomatized by the schemes of separation and collection). In the same paper, Hamkins and Freire showed that internal categoricity of ZF implies the tightness of ZF using a soft argument (an easier similar soft argument gives the corresponding result for PA); see Theorem 16 of (Hamkins and Freire, 2021), and the paragraph following it. The proof of the implication employs a class version of the Schröder-Bernstein theorem, plus the fact that ZF has a property known in model theory as *eliminating imaginaries* (cf. Section 4.4 of Hodges, 1993), i.e., if E is a definable equivalence relation on the universe of discourse, then there is a definable function f that maps distinct E-equivalence classes to distinct objects (in ZF such a function can be readily described by what is commonly known as the 'Scott trick', which takes advantage of the stratification of the universe into the well-ordered family of sets of the form V_α, as α ranges over the ordinals; see Exercise 2 of Section 4.4 of Hodges, 1993). In light of the fact that the class version of Schröder-Bernstein is provable in all sequential theories as shown by Friedman and Visser (2014), the proof strategy of Hamkins and Freire can be used to show the following general proposition.

Proposition 50. *For sequential theories T that eliminate imaginaries, the internal categoricity of T implies the tightness of T.*

Remark 51. There is an interesting conceptual difference between internal categoricity and the notions from the previous section (tightness, solidity), namely: the formulation of the latter ones do not depend on T being a schematic theory, but internal categoricity applies primarily to schemes. More explicitly: given a scheme τ the theory $\tau^{\mathrm{duo}}[\mathscr{L}^{\mathrm{duo}}]$ is not simply the union of $\tau^{\mathrm{red}}[\mathscr{L}^{\mathrm{red}}]$ and $\tau^{\mathrm{blue}}[\mathscr{L}^{\mathrm{blue}}]$. Additionally, there seems to be no "natural" definition of such a doubling of a theory T, if T is not readily presented through a scheme. A way to bypass this issue would be to show that actually for any two schemes τ and τ' and a language $\mathscr{L}$, if $\tau[\mathscr{L}]$ and $\tau'[\mathscr{L}]$ are deductively equivalent (axiomatize the same theory), then τ is internally categorical iff τ' is internally categorical. The next theorem shows that, actually, the exact opposite of the above claim is true: every sufficiently interesting theory of arithmetic is axiomatizable via a scheme that is *not* internally categorical. In particular, even though the induction scheme template is internally categorical, there is a scheme template that can be used to axiomatize the same first-order theory (i.e. Peano Arithmetic) that is not internally categorical. Moreover, we cannot escape this conclusion by looking at "sufficiently strong" theories.

Theorem 52. *Let T be any consistent recursively enumerable extension of* PA *(in the same language) or of* ZF *(in the same language). Then there is a schematic axiomatization of T that is not internally categorical.*

Proof. We shall first explain the proof when T is an extension of PA. Fix T and let σ be a formula that strongly represents a primitive recursive axiomatization of T (such a σ exists by Craig's trick). Let $\mathsf{Sat}(P,x)$ be the usual formula that expresses that P is a satisfaction predicate for formulae of logical depth at most x. More explicitly, $\mathsf{Sat}(P,x)$ asserts:

(1) Members of P are ordered pairs $\langle\varphi,\alpha\rangle$, where φ is a formula of depth[25] at most x, and α is an assignment for φ; and

(2) P satisfies Tarski's clauses for a satisfaction predicate for all formulae of depth at most x.

Let $\mathsf{EA} = \mathsf{I}\Delta_0 + \mathsf{Exp}$; by a theorem of Wilkie, EA is finitely axiomatizable (see Hájek and Pudlák, 1998, for an exposition). Consider the following scheme template $\tau_{\mathsf{Vaught},\sigma}(P)$:

$$\mathsf{EA} + \forall x\big(\mathsf{Sat}(P,x) \to \forall\varphi\big(\mathsf{pdepth}(\varphi) \leq x \wedge \sigma(\varphi) \to P(\varphi,\varnothing)\big)\big).$$

In the above $\mathsf{pdepth}(\varphi) \leq x$ expresses that the pure depth[26] φ is at most x and $\varnothing$ is the empty assignment. As shown in Theorem 12 of (Enayat and Łełyk, 2023) $\tau_{\mathsf{Vaught},\sigma}$ axiomatizes T; the proof therein is a streamlined version of Vaught's original proof in (1967).

25 The depth of a formula φ is the length of the longest chain in the formation tree of φ (which allows for arbitrary atomic formulae as leaves).

26 The pure depth of a formula takes into account also the complexity of terms occurring in it. More precisely: the pure depth of an atomic formula is the maximal complexity of terms occurring in it. The pure depth of compound formulae is then defined recursively in the standard way. See (Enayat and Łełyk, 2023) for details.

To see that $\tau_{\mathsf{Vaught},\sigma}$ is not internally categorical take any two non-isomorphic nonstandard models $\mathscr{M}$ and $\mathscr{N}$ of T such that neither $\mathscr{M}$ nor $\mathscr{N}$ is recursively saturated. For example, using the assumptions on T, we can invoke the first and second incompleteness theorems to get hold of two distinct completions T_1 and T_2 of $T + \neg\mathsf{Con}_T$. Note that T_1 and T_2 have the property that all of their models are nonstandard. We can choose $\mathscr{M}$ to be a pointwise definable model of T_1, and choose $\mathscr{N}$ to be a pointwise definable model of T_2. Since $\mathscr{M}$ and $\mathscr{N}$ are both countable, without loss of generality assume that the universes of $\mathscr{M}$ and $\mathscr{N}$ are the same. Let $\mathscr{K}$ be the model with the same universe as that of $\mathscr{M}$ and which carries the interpretations from both $\mathscr{M}$ and $\mathscr{K}$; thus the red model of $\mathscr{K}$ is isomorphic to $\mathscr{M}$, and the blue model of $\mathscr{K}$ is isomorphic to $\mathscr{N}$. As previously, we shall refer to symbols of T as interpreted in $\mathscr{M}$ ($\mathscr{N}$) as the *red* (*blue*) ones. We claim that $\mathscr{K} \models \tau^{\mathrm{duo}}[\mathscr{L}^{\mathrm{duo}}]$. To see this is it is enough to argue that for any formula $\Psi(x,y) \in \mathscr{L}^{\mathrm{duo}}$ we have:

$$\mathscr{K} \models \forall x\big(\mathsf{Sat}(\Psi,x) \to \forall\varphi\big(\mathsf{pdepth}(\varphi) \leq x \wedge \sigma(\varphi) \to \Psi(\varphi,\varnothing)\big)\big),$$

where all basic symbols which appear outside of Ψ belong to $\mathscr{L}^{\mathrm{red}}$ (the same argument applies in the case of $\mathscr{L}^{\mathrm{blue}}$).

Fix Ψ and work in $\mathscr{K}$. Fix x, and assume that $\mathsf{Sat}(\Psi,x)$. We claim that x is a standard number (according to $\leq^{\mathrm{red}}$). If not, then the satisfaction class S defined by Ψ in $\mathscr{K}$ is a partial nonstandard satisfaction class on $\mathscr{M}$, and therefore by Lachlan's theorem (see 15.5 of Kaye, 1991) $\mathscr{M}$ must be recursively saturated, contradicting our choice of $\mathscr{M}$. So take an arbitrary φ of pure depth at most x. It follows that φ corresponds to a standard formula (of $\mathscr{L}^{\mathrm{red}}$). Assuming additionally that $\sigma(\varphi)$ holds, we conclude that φ is an axiom of T^{red}. So φ must hold in $\mathscr{M}$. Consequently $\Psi(\varphi,\varnothing)$ holds because Ψ satisfies the Tarski T-scheme for sentences of $\mathscr{L}^{\mathrm{red}}$.

The above proof strategy can be modified to work for extensions T of ZF. More specifically, the definition of the template $\tau_{\mathsf{Vaught},\sigma}(P)$ is modified by replacing EA with KP.[27] Since it is well-known that Lachlan's theorem also holds for ω-nonstandard models of ZF, to complete the proof we only need to explain how to get to hold of two countable nonisomorphic ω-nonstandard models of T that are not recursively saturated. As in the arithmetical case we can use the first and second incompleteness theorems to get hold of two distinct completions T_1 and T_2 of $T + \neg\mathsf{Con}_T$. Note that T_1 and T_2 have the property that all of their models are ω-nonstandard. Then we can choose $\mathscr{M}$ to be a Paris model of T_1 and $\mathscr{N}$ to be a Paris model of T_2 (A Paris model is a model all of whose ordinals are pointwise definable, therefore a Paris model is not recursively saturated). Since every consistent extension of ZF has a Paris model by a classical theorem of Paris (see Enayat, 2005) this concludes the proof. □

Remark 53. Since, as shown in (Jeřábek, 2012), the canonical theory of (the non-negative parts) of the discretely ordered rings, i.e. PA^- (see Kaye, 1991), is sequential, the above proof works also for an arguably more natural axiomatization of PA which results from $\tau_{\mathsf{Vaught},\sigma}(P)$ by replacing EA with PA^-.

Proposition below complements Theorem 52.

Proposition 54. *Every recursively enumerable extension T of* PA *or of* ZF *can be axiomatized by a scheme whose template is internally categorical.*

Proof. Use Craig's trick to get hold of σ that strongly represents a primitive recursive axiomatization of T. Thanks to Vaught's Theorem (1967) and Theorem 47, if T is an extension of PA, then the desired template is $\tau_{\mathsf{Vaught},\sigma}(P) \wedge \tau_{\mathsf{Ind}}(P)$; and if T is an extension of ZF, then the desired template is $\tau_{\mathsf{Vaught},\sigma}(P) \wedge \tau_{\mathsf{Repl}}(P)$. □

At first sight Theorem 52 and Proposition 54 might seem to diminish the importance of the notion of internal categoricity, because there is a strong intuition that a foundationally interesting property of theories shouldn't depend on how the theory is presented. However, in what follows we experiment with a different option: we check whether schemes can be viewed as a self-standing and, in a sense, primitive object of foundational analysis. In the rest of this section we scrutinize mathematical benefits from taking this point of view: we show that this perspective enables us to generalize and make explicit some patterns that can be observed in the hierarchy of solid/tight theories.

27 Recall that KP is Kripke-Platek set theory with Π_1-Foundation (as in Mathias, 2001); its well-known that KP is finitely axiomatizable.

Furthermore, there are natural adaptations of the notion of solidity to the context of schemes, which bring to light interesting distinctions between various natural foundational schemes, such as the scheme of induction and the replacement scheme. We shall discuss the philosophical consequences of taking this perspective in the next section.

Definition 55. The scheme template $\tau_{\mathsf{Coll}}^{\mathsf{Arith}}$ is the *arithmetical sentence* that is conjunction of two sentences; the first conjunct asserts that EA holds, and the second conjunct is the following sentence:

$$\forall x\big(\forall y < x\ \exists z\ P(y,z) \to \exists b\ \forall y < x\ \exists z < b\ P(y,z)\big).$$

Similarly, The scheme template $\tau_{\mathsf{Coll}}^{\mathsf{Set}}$ is the *set-theoretical sentence* that is conjunction of two sentences; the first conjunct asserts that KP holds, and the second conjunct is the following sentence:

$$\forall x\big(\forall y \in x\ \exists z\ P(y,z) \to \exists b\ \forall y \in x\ \exists z \in b\ P(y,z)\big).$$

Theorem 56. *The scheme templates $\tau_{\mathsf{Coll}}^{\mathsf{Arith}}$ and $\tau_{\mathsf{Coll}}^{\mathsf{Set}}$ are not internally categorical, even though they respectively axiomatize* PA *and* ZF.

Proof. We already commented in Remark 37 PA is axiomatized by all arithmetical instances of τ_{Coll}. It is also well-known that ZF is axiomatized by the set-theoretical instances of $\tau_{\mathsf{Coll}}^{\mathsf{Set}}$ (since an induction on the complexity of formulae shows that the full separation scheme is provable in the resulting theory, thanks to the availability of Δ_0-separation and full collection).

To demonstrate the failure of internal categoricity for $\tau_{\mathsf{Coll}}^{\mathsf{Arith}}$, we resort to the fact that there are (many) nonisomorphic ω_1-like models of PA (e.g., see p.237 of Kossak and Schmer, 2006). Recall that a model of PA is said to be ω_1-like if it is uncountable but every proper initial segment of it is countable.

Let $\mathscr{M}$ and $\mathscr{N}$ be any two non-isomorphic ω_1-like models of PA, let $\mathscr{K}$ be the disjoint union $\mathscr{M}, \mathscr{N}$ (as in the proof of Theorem 52). Obviously $\mathscr{K} \models \mathsf{EA}^{\mathsf{red}} \wedge \mathsf{EA}^{\mathsf{blue}}$. Moreover, since each component of $\mathscr{K}$ is ω_1-like, for every $\varphi(x,y) \in \mathscr{L}^{\mathsf{duo}}$,

$$\mathscr{K} \models \tau_{\mathsf{Coll}}^{\mathsf{Arith}}[\varphi(x,y)/P].$$

The set-theoretical case is handled similarly to the arithmetical case; thanks to the well-known fact that there are (many) nonisomorphic ω_1-like models of ZF; see (Enayat, 2004). Here, a model of ZF is said to be ω_1-like if it is uncountable but whose every proper rank-initial segment is countable. □

Remark 57. The Vaught schemes template $\tau_{\mathsf{Vaught},\sigma}(P)$ (introduced in the proof of Theorem 52, where σ is a formula that strongly represents a primitive recursive axiomatization of PA), and the Collection scheme template $\tau_{\mathsf{Coll}}^{\mathsf{Arith}}(P)$ (introduced in Definition 55) do not give rise to Π^1_1-statements $\forall X\, \tau(X)$ that characterize the standard model of arithmetic $\mathbb{N}$. Clearly the standard model of arithmetic satisfies the Π^1_1-statements corresponding to $\tau_{\mathsf{Vaught},\sigma}$ and to $\tau_{\mathsf{Coll}}^{\mathsf{Arith}}$. However, the Π^1_1-statement corresponding to $\tau_{\mathsf{Vaught},\sigma}$ is also satisfied in any countable model of PA that is not recursively saturated (by the proof of Theorem 52), and the Π^1_1-statement corresponding to $\tau_{\mathsf{Coll}}^{\mathsf{Arith}}$ is satisfied in any ω_1-like model of PA.

Remark 58. It is easy to see that if $\tau(P)$ is a template in the language of arithmetic that is internally categorical and the scheme it generates is true in $\mathbb{N}$, then the Π^1_1-statement $\forall X\, \tau(X)$ characterizes $\mathbb{N}$ in full second order logic. However, as shown in recent (to be published) joint work of the second-named-author with Piotr Gruza, there is a template $\tau(P)$ in the language of arithmetic that is not internally categorical but which has the property that the Π^1_1-statement $\forall X\, \tau(X)$ characterizes $\mathbb{N}$ in full second order logic.

Next, we show that internal categoricity is preserved upwards w.r.t. to the ordering determined by provability of Π^1_1 sentences in two-sorted logic.

Definition 59. For a theory T in an n-sorted language $\mathscr{L}$, let $\mathsf{PC}(T)$ be the extension of T in the $(n+1)$-sorted language by the predicative comprehension axioms, that is all sentences of the form

$$\exists X^{n+1}\forall X^{n}\big(X^{n}\in_{n+1}X^{n+1}\equiv\varphi(X^{n},\overline{Y})\big),$$

where $\varphi(X^{n},\overline{Y})$ is an $\mathscr{L}$-formula (see Pakhomov and Visser, 2009). The upper index of a variable denotes its sort.

Theorem 60. *Let τ and σ be two $\mathscr{L}$-scheme templates that $\tau[\mathscr{L}]$ and $\sigma[\mathscr{L}]$ axiomatize the same theory T. Suppose further that $\mathsf{PC}(T)\vdash\forall X\tau[X]\rightarrow\forall X\sigma[X]$. Then if σ is internally categorical, then so is τ.*

Proof. Take any $\mathscr{M}\models\tau^{\mathrm{red}}[\mathscr{L}^{\mathrm{duo}}]\wedge\tau^{\mathrm{blue}}[\mathscr{L}^{\mathrm{duo}}]$. Let $\mathfrak{X}$ consists of $\mathscr{L}^{\mathrm{duo}}$-definable subsets of $\mathscr{M}$. Then, it follows that for $\mathsf{col}\in\{\mathrm{red},\mathrm{blue}\}$,

$$(\mathscr{M},\mathfrak{X})\models\mathsf{PC}(T^{\mathsf{col}})\wedge\forall X\tau^{\mathsf{col}}[X].$$

In particular, by our assumption $(\mathscr{M},\mathfrak{X})\models\mathsf{PC}(T^{\mathsf{col}})\wedge\forall X\sigma^{\mathsf{col}}[X]$. It follows that

$$\mathscr{M}\models\sigma^{\mathsf{col}}[\mathscr{L}^{\mathsf{duo}}],$$

hence, by the internal categoricity of σ, it follows that there is a definable isomorphism between the red and blue parts of $\mathscr{M}$. □

4.2. Generalizations of solidity

In this subsection we adapt the categoricity-like properties of theories introduced in Subsection 2.3 to the context of schemes. As we shall see in Theorem 74, these generalizations can be viewed as providing a (short) hierarchy of ‘categoricity grades’ of scheme templates.

Definition 61. Below all the languages are arbitrary n-sorted languages. In the paper we shall deal only with translations which are sort-preserving, i.e. types of all the variables are preserved under the translation. Moreover we shall deal only with the translations which have the same dimension on every sort. Last but not least, we assume that the isomorphisms between sorted structures need to preserve sorts.

(a) Let $m\leq n$. A (k-dimensional) translation σ between an m-sorted language $\mathscr{L}_1$ and an n-sorted $\mathscr{L}_2$ is given by

- a tuple of formulae $(\delta_1(\bar{x}^1),\dots,\delta_m(\bar{x}^m))$, where each $\delta_i(\bar{x}^i)$ is a formula of the i-th order logic and all $\bar{x}^i$ are variables of the i-th sort.
- a mapping $P\mapsto A_P$, where P is an l-ary predicate of $\mathscr{L}_1$ and A_P is a $k\cdot l$-ary formula

As in the previous section we allow the equality to be redefined and our interpretations to be multidimensional. A translation σ is *direct* iff it is direct with respect to all sorts.

(b) Let $\tau(P)$ be an $\mathscr{L}_1$-template, where P is a predicate of arity l and σ be any k-dimensional translation $\mathscr{L}_1\rightarrow\mathscr{L}_2$. Let R be any fresh predicate of arity kl. Let $\sigma[P/R]$ be a translation $\mathscr{L}_1\cup\{P\}\rightarrow\mathscr{L}_2\cup\{R\}$ which translates P to R and behaves like σ otherwise. By τ^{σ} we denote the following scheme template:

$$(\text{"}R\text{ is a }\approx\text{-invariant }l\text{-ary relation on }k\text{-tuples"})\longrightarrow\tau^{\sigma[P/R]}.$$

In the above $\approx$ is the translation of equality according to σ and $\tau^{\sigma[P/R]}$ is the translation of τ according to $\sigma[P/R]$.

(c) We say that an $\mathscr{L}_2$-structure $\mathscr{M}$ *accommodates* a scheme template $\tau(P)$ iff there is a translation σ of $\mathscr{L}_1$ into $\mathscr{L}_2$ such that,

$$(\mathscr{M},\mathsf{Def}^{kn}(\mathscr{M}))\models\forall X\tau^{\sigma}(X/R),$$

where $\mathrm{Def}^{kn}(\mathscr{M})$ denotes the set of parametrically $\mathscr{M}$-definable subsets of M^{kn}. In other words, for each $X \in \mathrm{Def}^{kn}(\mathscr{M})$, $(\mathscr{M}, X) \models \tau^{\sigma}(X/R)$. Note that the translation σ is allowed to use parameters from $\mathscr{M}$.

(d) We say that an $\mathscr{L}_2$-structure $\mathscr{M}$ *directly accommodates* a scheme template τ iff $\mathscr{M}$ accommodates τ via a direct translation σ. In such a case, $\mathscr{M}$ is also called a *strong model of* τ.

(e) We say that a theory T *(directly) accommodates* a scheme τ iff there is a translation σ such that for every model $\mathscr{M} \models T$, $\mathscr{M}$ (directly) accommodates τ via σ.

Example 62. Since in this paper we are primarily interested in sequential theories, we can always assume that for our purposes interpretations are one-dimensional. Assuming additionally that τ is equality-preserving, we can easily show that $\mathscr{M}$ accommodates an $\mathscr{L}$-scheme τ via a one-dimensional and equality-preserving translation just in case the following condition holds:

There exists an $\mathscr{L}$-structure $\mathscr{N} \trianglelefteq_{\mathrm{par}} \mathscr{M}$ such that for every parametrically $\mathscr{M}$-definable set $P \subseteq M$, $(\mathscr{N}, P \cap N) \models \tau(P)$.

In particular, if $\mathscr{M}$ is a model of Z_2 (Second Order Arithmetic) or a model of Z (Zermelo set theory), then $\mathscr{M}$ accommodates the induction scheme τ_{Ind}. Similarly, every model of KM accommodates the replacement scheme τ_{Repl}.

Example 63. A structure $\mathscr{M}$ accommodates the induction scheme τ_{Ind}, just in case $\mathscr{M}$ parametrically interprets a model $\mathscr{N}$ of PA *with the additional property that* $\mathscr{N}$ *has no proper parametrically* $\mathscr{M}$*-definable proper cuts*. Note that under this scenario if the 'accommodating' structure $\mathscr{M} \models \mathsf{PA}$, then there is an $\mathscr{M}$-definable isomorphism between $\mathscr{M}$ and $\mathscr{N}$; this readily follows from the fact that PA is a minimalist theory (as indicated in part (a) of Remark 8).

Remark 64. We note that $\mathscr{M}$ is a strong model (in the sense of part (d) of Definition 61) if an $\mathscr{L}$-scheme τ just in case there is an extension (perhaps higher order) $\mathscr{L}^+ \supseteq \mathscr{L}$ and a definitional expansion $\mathscr{M}^+$ which is an $\mathscr{L}^+$ model, such that $\mathscr{M}^+ \models \tau[\mathscr{L}^+]$.

Definition 65. Suppose $\tau(P)$ is an $\mathscr{L}$-template.

(a) τ is *g-minimalist* (generalized minimalist) if the following holds for all strong models $\mathscr{M}$ and $\mathscr{N}$ of τ: If $\mathscr{M}$ parametrically accommodates τ via a translation σ_0, $\mathscr{N} \trianglelefteq_{\mathrm{par}} \mathscr{M}$, and $\mathscr{N}$ accommodates τ via a translation σ_1, **then** there is a **unique** $\mathscr{M}$-definable embedding

$$F : \sigma_0(\mathscr{M}) \to \sigma_1(\mathscr{N}).$$

(b) τ is *g-solid*[28](generalized solid) iff the following holds for all strong models $\mathscr{M}$, $\mathscr{N}$, and $\mathscr{M}^*$ of τ that respectively accommodate $\tau(P)$ via translations σ_l, σ_c and σ_r: If

$$\mathscr{M} \trianglerighteq_{\mathrm{par}} \mathscr{N} \trianglerighteq_{\mathrm{par}} \mathscr{M}^*,$$

and there is an $\mathscr{M}$-definable isomorphism $F : \sigma_l(\mathscr{M}) \to \sigma_r(\mathscr{M}^*)$, **then** there is an $\mathscr{N}$-definable isomorphism

$$G : \sigma_c(\mathscr{N}) \to \sigma_r(\mathscr{M}^*).$$

(c) τ is *e-solid* (expansion solid) iff τ satisfies the definition for g-solidity with the added requirement that σ_l, σ_c and σ_r are *direct* interpretations.

Remark 66. The e-solidity terminology is informed by the fact that e-solidity can be equivalently formulated in terms of the notion of *expansion* in model theory. Observe that an $\mathscr{L}$-template τ is e-solid just in case for all expansions $\mathscr{L}_1$, $\mathscr{L}_2$, $\mathscr{L}_3$ of $\mathscr{L}$, and all structures $\mathscr{M}_1$, $\mathscr{M}_2$, $\mathscr{M}_3$ such that $\mathscr{M}_i \models \tau[\mathscr{L}_i]$ for $i \in \{1,2,3\}$, the following condition holds:

[28] Note that general solidity is a generalization of the notion of strong solidity that was shown in Remark 9 to be equivalent to solidity. The proof strategy of the equivalence does not seem to generalize to the generalized setting. We have opted for this stronger notion because (1) it is satisfied by the usual schematic representations of PA and ZF, and (2) it comes handy in the proof of solidity of KF_μ (see Theorem 88).

if $\mathcal{M}_1 \trianglerighteq_{\text{par}} \mathcal{M}_2 \trianglerighteq_{\text{par}} \mathcal{M}_3$ and there is an $\mathcal{M}_1$-definable isomorphism between $\mathcal{M}_1\restriction_{\mathcal{L}}$ and $\mathcal{M}_3\restriction_{\mathcal{L}}$, then there is an $\mathcal{M}_2$-definable isomorphism between $\mathcal{M}_2\restriction_{\mathcal{L}}$ and $\mathcal{M}_3\restriction_{\mathcal{L}}$.

In the above $\mathcal{M}\restriction_{\mathcal{L}}$ denotes the reduct of a model $\mathcal{M}$ to the language $\mathcal{L}$.

Remark 67. Note that the categoricity-like notions defined in Definition 65 are attributes of *scheme templates*, whereas those defined in Definition 7 are attributes of *first order theories*.

Definition 68. The scheme template $\tau_{\mathsf{Repl+Tarski}}(P)$ is the result of adjoining Tarski's undefinability of truth theorem to the usual axiomatization of ZF using τ_{Repl}, i.e.,

$$\tau_{\mathsf{Repl+Tarski}}(P) := \exists x \in \omega \neg \mathrm{Sat}(P,x) \wedge \tau_{\mathsf{Repl}}(P).$$

In the above, $\mathrm{Sat}(P,x)$ is the formula obtained by adding the unary predicate P to the language $\{\in\}$ of set theory that expresses that P satisfies Tarski conditions for a satisfaction predicate for all set-theoretical formulae of depth at most x.

Remark 69. $\tau_{\mathsf{Repl+Tarski}}$ ensures that replacement holds for all formulae, and no formula defines a full satisfaction predicate for $\mathcal{L} = \{\in\}$. Thanks to Tarski's undefinability of truth theorem $\tau_{\mathsf{Rep+Tarski}}$ axiomatizes ZF. However, note that if κ is a strongly inaccessible cardinal, then the natural Kelley-Morse model associated with V_κ whose classes are elements of $V_{\kappa+1}$ does not satisfy the Π^1_1-sentence $\forall X\, \tau_{\mathsf{Repl+Tarski}}(X)$.

Theorem 70. *The scheme-templates* τ_{Ind} *and* $\tau_{\mathsf{Repl+Tarski}}$ *are e-solid.*[29]

Proof. (Outline). A close examination of the proof of the solidity of PA presented in (Enayat, 2016) shows that the same proof strategy succeeds in establishing the e-solidity of τ_{Ind}. The e-solidity of $\tau_{\mathsf{Repl+Tarski}}$ is established by using the proof strategy of the solidity proof of ZF presented in (Enayat, 2016) (or the one presented in Freire and Hamkins, 2021) with the difference that the Tarski clause comes to the rescue at the point in the proof of solidity of ZF where Tarski's undefinability of truth is invoked. More specifically, the Tarski clause of $\tau_{\mathsf{Repl+Tarski}}$ makes sure that none of the extra predicates included in the expansion of the first of the three models of ZF (in the set-up of the e-solidity proof) is a satisfaction predicate for the first model of ZF. □

Example 71. For $T = \mathsf{PA}$, Theorem 70 can be rephrased as follows: Suppose $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3 \models \mathsf{PA}$, and that $\mathcal{M}_i^+$ is an $\mathcal{L}_i$-structure that is an expansion of $\mathcal{M}_i$ and $\mathcal{M}_i^+ \models \mathsf{PA}(\mathcal{L}_i)$ for $i \in \{1,2,3\}$, where each $\mathcal{L}_i$ is an extension of the usual language of arithmetic. Suppose, furthermore, that $\mathcal{M}_1^+ \trianglerighteq \mathcal{M}_2^+ \trianglerighteq \mathcal{M}_3^+$ and there is an $\mathcal{M}_1^+$-definable isomorphism $i_0 : \mathcal{M}_1 \to \mathcal{M}_3$. **Then** there is an $\mathcal{M}_1^+$-definable isomorphism $i : \mathcal{M}_2 \to \mathcal{M}_3$.

Note that in the above, the existence of $i : \mathcal{M}_1 \to \mathcal{M}_2$ cannot in general be strengthened to the existence of $i : \mathcal{M}_1^+ \to \mathcal{M}_2^+$, e.g., let $\mathcal{M}_1^+ = (\mathcal{M}, D_1)$ and $\mathcal{M}_1^+ = (\mathcal{M}, D_2)$, where D_1 and D_2 are distinct $\mathcal{M}$-definable sets.

Example 72. τ_{Repl} is not e-solid. To see this, suppose κ and λ are strongly inaccessible cardinals with $\kappa < \lambda$. Let γ be an ordinal such that $\kappa \in \gamma \in \lambda$. By the Löwenheim-Skolem theorem, there is an elementary submodel $\mathcal{K} = (K, \in)$ of $(V_\lambda, \in)$ that $\gamma \in K$ and $V_\kappa \subseteq K$ and $|K| = |V_\kappa|$. Note that the order-type of the ordinals in K is higher than κ. Using a bijection g between K and V_κ there is a binary relation E on V_κ such that g is an isomorphism between (V_κ, E) and $(K, \in)$. Thus there is an embedding f of $(V_\kappa, \in)$ as a topped rank initial segment of (V_κ, E). By well-foundedness of (V_κ, E) there is some ordinal $\alpha \in V_\kappa$ such that $(V_\alpha, \in)$-as-calculated-in-(V_κ, E) is isomorphic to $(V_\kappa, \in)$; thus α is the first ordinal in (V_κ, E) that is not in the range of f (and the order-type of the E-predecessors of α in (V_κ, E) from an external point of view is precisely κ). Now let $\mathcal{M} = (V_\kappa, \in, E, f)$; this is obviously a model of $\mathsf{ZF}(E, f)$ since κ is strongly inaccessible. Also let $\mathcal{N} = (V_\kappa, E)$, and let $\mathcal{M}^* = (V_\alpha, \in)^{\mathcal{N}}$ (i.e., $(V_\alpha, \in)$-as-calculated-in-$\mathcal{N}$), where α is as define above. Note that:

1. $\mathcal{N}$ is a reduct of $\mathcal{M}$, so clearly $\mathcal{N}$ is parameter-free interpretable in $\mathcal{M}$.

[29] Recall from Definition 42 that τ_{Ind} axiomatizes PA and τ_{Repl} axiomatizes ZF.

2. $\mathscr{M}^*$ is interpretable in $\mathscr{N}$ with the use of the parameter α.[30]

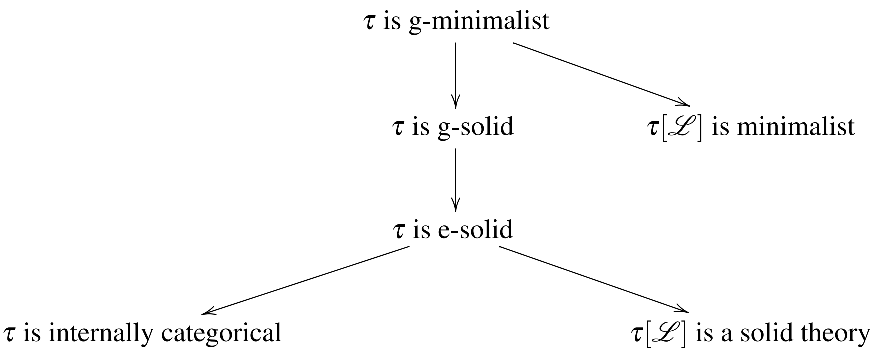
3. There is an $\mathscr{N}$-definable isomorphism between $(V_\kappa, \in)$ and $\mathscr{M}^*$, thanks to having f available as one of the "oracles" in $\mathscr{M}$.[31]

This shows that τ_{Repl} is not e-solid since $(V_\kappa, \in)$ is not isomorphic to (V_κ, E) since the latter is isomorphic to $(K, \in)$ whose ordinal height exceeds the ordinal height of $(V_\kappa, \in)$, i.e., κ.

Theorem 73. *The template τ_{Ind} for the induction scheme is g-minimalist.*

Proof. The proof is analogous to the proof of Proposition 11. □

Theorem 74. *The following implications hold for an arbitrary $\mathscr{L}$-scheme template τ:*



Proof. Note that based on the relevant definitions, three of the implications are trivial, namely that g-solidity implies e-solidity, e-solidity implies solidity and g-minimalist implies minimalist.

Recall that by Remark 8 a minimalist theory is a solid theory (on the basis of the relevant definitions). A similar argument shows that if τ is g-minimalist, then τ is g-solid.

It remains to show that e-solidity entails internal categoricity. In the interest of notational clarity, we will present the proof for the template τ_{Ind} of the usual schematic presentation of PA; the same proof strategy works in general. We want to show that if $\mathscr{M}^{\mathsf{duo}} := \left(M, +^{\mathsf{red}}, \cdot^{\mathsf{red}}, +^{\mathsf{blue}}, \cdot^{\mathsf{blue}}\right) \models \tau^{\mathsf{duo}}_{\mathsf{Ind}}[\mathscr{L}^{\mathsf{duo}}_{\mathsf{PA}}]$, then there is an $\mathscr{M}^{\mathsf{duo}}$-definable isomorphism $i : \mathscr{M}^{\mathsf{red}} \to \mathscr{M}^{\mathsf{blue}}$. Define:

$$\mathscr{M}_1 = \mathscr{M}_2 = \mathscr{M}_3 := \mathscr{M}^{\mathrm{duo}};$$
$$\sigma_1 = \sigma_3 := [+ \mapsto +^{\mathrm{red}}, \cdot \mapsto \cdot^{\mathrm{red}}], \sigma_2 := [+ \mapsto +^{\mathrm{blue}}, \cdot \mapsto \cdot^{\mathrm{blue}}].$$

Then clearly $\mathscr{M}_1 \trianglerighteq \mathscr{M}_2 \trianglerighteq \mathscr{M}_3$ and there is an $\mathscr{M}_1$-definable isomorphism $i_0 : \sigma_1(\mathscr{M}_1) \to \sigma_3(\mathscr{M}_3)$, so the desired $\mathscr{M}^{\mathsf{duo}}$ -definable isomorphism $i : \mathscr{M}^{\mathsf{red}} \to \mathscr{M}^{\mathsf{blue}}$ exists by e-solidity of τ_{Ind}. □

Remark 75. We suspect that the implication "g-solid implies e-solid" of Theorem 74 cannot be reversed; but we have not yet found a counterexample. The following show that none of the other implication arrows of Theorem 74 reverses.

1. Let $\tau_{\mathsf{CT[PA]}}$ be the template axiomatizing $\mathsf{CT[PA]}$ obtained conjuncting τ_{Ind} with finitely many axioms of compositional truth CT. Then by part (c) of Theorem 88 τ is g-solid, but not g-minimalist by Remark 89.

2. If τ is template for the Vaught schematic axiomatization of PA, and $\mathscr{L}$ is the language of arithmetic, then $\tau[\mathscr{L}]$ is solid by Theorem 10, but τ is not internally categorical by Theorem 52, and since e-solidity implies internal categoricity by Theorem 74, τ is not e-solid. A similar argument shows that the fact that $\tau[\mathscr{L}]$ is minimalist does not imply τ is g-minimalist.

[30] If κ is chosen as the first strongly inaccessible cardinal, then α is definable in $\mathscr{N}$ as the first strongly inaccessible cardinal and therefore the interpretation of $\mathscr{M}^*$ in $\mathscr{N}$ becomes parameter-free.

[31] Indeed, with a little extra work one can show that f is definable in $(V_\kappa, \in, E)$, but this flourish is not needed for our purposes.

3. As indicated in Example 72, τ_{Repl} is not e-solid, but by part (b) of Theorem 47 it is internally categorical.

Remark 76. In light of Remark 8 and Proposition 50 the bottom part of the diagram of implication arrows of Theorem 74 can be complemented with two more implications, one indicating that the solidity of $\tau[\mathscr{L}]$ implies the tightness of $\tau[\mathscr{L}]$; and the other indicating that the internal categoricity of $\tau[\mathscr{L}]$ implies the tightness of $\tau[\mathscr{L}]$ with the additional assumption that $\tau[\mathscr{L}]$ is a sequential theory that eliminates imaginaries.

Recall from part (b) of Theorem 47 that τ_{Repl} is not strongly internally categorical and not e-solid by Example 72. The result below shows that there is a schematic axiomatization of ZF that is both 'more internally categorical' and 'more solid' than the schematic axiomatization given by τ_{Repl}.

Theorem 77. *The scheme template $\tau_{\mathsf{Repl+Tarski}}$ (that axiomatizes ZF; see Definition 68 and the remark following it), is both strongly internally categorical and g-solid.*

Before presenting the proof of 77, we need some preliminary lemmas.

Lemma 78. *Suppose $\mathscr{K}$ and $\mathscr{M}$ are models of ZF, where $\mathscr{M}$ is a proper rank extension of $\mathscr{K}$ (i.e., the rank of every element in $M \setminus K$ is higher than the rank of every element of K). Assume furthermore that for every parametrically $\mathscr{M}$-definable subset D of M, $(\mathscr{K}, D\cap K) \models \mathsf{ZF}(P)$, where P is a fresh unary predicate interpreted by $D \cap K$.* ***Then*** *there is some $\gamma \in \mathrm{Ord}^{\mathscr{M}} \setminus \mathrm{Ord}^{\mathscr{K}}$ such that:*

$$\mathscr{K} \preceq (V_\gamma, \in)^{\mathscr{M}}.$$

In particular, there is a parametrically $\mathscr{M}$-definable set that is a full satisfaction class for $\mathscr{K}$.

Proof. This follows immediately from Theorem 4.4 of (Enayat, 2024) (whose proof is obtained by an adaptation of the proof of Theorem 3.3 of Enayat, 1986). □

Lemma 79. *Suppose $\mathscr{M}$ and $\mathscr{K}$ are models of ZF such that the universe K of $\mathscr{K}$ is a subset of the universe M of $\mathscr{M}$ (and therefore the membership relation $\in^{\mathscr{K}}$ of $\mathscr{K}$ is a subset of M^2). Moreover, assume for every parametrically $\mathscr{M}$-definable subset D of M, $(\mathscr{K}, D\cap K) \models \mathsf{ZF}(P)$, where P is a fresh unary predicate interpreted by $D \cap K$.* ***Then*** *precisely one of the following three conditions hold:*

(1) There is an $(\mathscr{M}, \mathscr{K})$-definable isomorphism between $\mathscr{K}$ and $(V_\kappa, \in)^{\mathscr{M}}$ for some strongly inaccessible cardinal κ of $\mathscr{M}$.

(2) There is an $(\mathscr{M}, \mathscr{K})$-definable isomorphism between $\mathscr{K}$ and $\mathscr{M}$.

(3) There is an $(\mathscr{M}, \mathscr{K})$-definable embedding of $\mathscr{M}$ as a proper rank initial segment of $\mathscr{K}$.

Proof. Using the assumptions on $\mathscr{K}$, it is easy to see that the membership relation $\in^{\mathscr{K}}$ of $\mathscr{K}$ is well-founded from the point of view of $\mathscr{M}$. The proof strategy of solidity of ZF (as in Enayat, 2016) can be used to show that (1) holds if K is a set in $\mathscr{M}$; (2) holds if K is a proper class in $\mathscr{M}$ such that $\in^{\mathscr{K}}$ is set-like[32] in the sense of $\mathscr{M}$; and (3) holds if K is a proper class in $\mathscr{M}$ such that $\in^{\mathscr{K}}$ is not set-like in the sense of $\mathscr{M}$. □

Proof of Theorem 77. We shall explain why $\tau_{\mathsf{Repl+Tarski}}$ is g-solid. A similar argument shows that it is also strongly internally categorical. Suppose $\mathscr{M}$, $\mathscr{M}_2$, and $\mathscr{M}_3$ are strong models of ZF, $\mathscr{M}_1 \trianglerighteq_{\mathsf{par}} \mathscr{M}_2 \trianglerighteq_{\mathsf{par}} \mathscr{M}_3$, and $\mathscr{M}_i$ accommodates $\tau_{\mathsf{Repl+Tarski}}$ for $i \in \{1,2,3\}$ via a model $\mathscr{K}_i$ such that for every parametrically $\mathscr{M}_i$-definable subset D of $\mathscr{M}_i$,

$$(\mathscr{K}_i, D \cap K_i) \models \tau_{\mathsf{Repl+Tarski}}(P).$$

Note that since ZF eliminates imaginaries, we can assume without loss of generality that the interpretation of each $\mathscr{K}_i$ in $\mathscr{M}_i$ is identity preserving. By Lemma 79, for each $i \in \{1,2,3\}$ either (1) there is an $\mathscr{M}_i$-definable isomorphism between $\mathscr{K}_i$ and $(V_\kappa, \in)^{\mathscr{M}_i}$ for some inaccessible cardinal κ of $\mathscr{M}_i$, or (2) there is an $\mathscr{M}_i$-definable

32 In other words, for each $x \in K$, the collection $\{y \in K : y \in^{\mathscr{K}} x\}$ forms a set in $\mathscr{M}$ as opposed to a proper class.

isomorphism between $\mathscr{M}_i$ and $\mathscr{K}_i$, or (3) there is an $\mathscr{M}_i$-definable embedding of $\mathscr{M}_i$ into a proper rank initial segment of $\mathscr{K}_i$.

Thanks to $\tau_{\mathsf{Repl+Tarski}}$ and Lemma 78, possibility (1) is ruled out, and (3) is ruled out by Tarski's theorem on undefinability of truth and Lemma 78. So for each $i \in \{1,2,3\}$ there is a $\mathscr{M}_i$-definable isomorphism between $\mathscr{M}_i$ and $\mathscr{K}_i$. This reduces the initial problem to the assumptions of e-solidity of ZF. Hence the desired conclusion follows from Theorem 70. □

Below we consider two natural operations that preserve g-solidity and e-solidity of schemes. The first one (the comprehension scheme) is connected to the process of enriching a given language $\mathscr{L}$ with a new sort of sets (that are interpreted as subsets of the domain of discourse of $\mathscr{L}$); and the second (extremal schemes) enriches $\mathscr{L}$ with a new predicate.

Definition 80 (The comprehension scheme)**.** If $\mathscr{L}$ is any n-sorted language, then $\mathscr{L}^{+1}$ is a natural extension of $\mathscr{L}$ with the $n+1$-st sort.

Suppose that $\mathscr{L}$ is an n-sorted language and P a fresh predicate for the n-th sort. The comprehension scheme $\tau_{\mathsf{Comp}}(P)$ is the following statement of $\mathscr{L}^{+1}$:

$$\exists X^{n+1} \forall X^n \big(X^n \in_{n+1} X^{n+1} \equiv P(X^n)\big).$$

Example 81. If $\mathscr{L}$ is the 1-sorted language of PA, then $\tau_{\mathsf{Comp}}(P)$ is the template for the comprehension scheme of Z_2 (Second Order Arithmetic); and if $\mathscr{L}$ is the 2-sorted language of Z_2, then $\tau_{\mathsf{Comp}}(P)$ is the template for the comprehension scheme of Z_3 (Third Order Arithmetic).

The following result generalizes the fact that Z_2 and KM and their higher older analogues mentioned in Theorem 14 are solid. Recall that, as indicated in Theorem 70, the scheme templates for usual axiomatizations of PA and ZF are e-solid.

Theorem 82. *Let $*$ be either g or e. Suppose $\tau(P)$ is an $\mathscr{L}$-template that is *-solid. Then $\tau \wedge \tau_{\mathsf{Compr}}$ is *-solid (as an $\mathscr{L}^{+1}$ template).*

Proof. We show the proof for g-solidity. Assume that $\mathscr{L}$ is an n-order language. Assume that $\mathscr{M}$, $\mathscr{N}$ and $\mathscr{M}^*$ are strong models of $\tau \wedge \tau_{\mathsf{Compr}}$ which accommodate $\tau \wedge \tau_{\mathsf{Compr}}$ via the translations σ_l, σ_c and σ_r respectively and that I is an $\mathscr{M}$-definable isomorphism between $\sigma_l(\mathscr{M})$ and $\sigma_r(\mathscr{M}^*)$.

For the ease of exposition we assume that the translations are one -dimensional and equality-preserving, however the proof below can clearly be adapted to the more general setting. Let $\widehat{\sigma}_l, \widehat{\sigma}_c, \widehat{\sigma}_r, \widehat{I}$ be the natural restrictions of σ's and I to the first n-sorts. Hence $\mathscr{M}$, $\mathscr{N}$ and $\mathscr{M}^*$ are strong models of τ which accommodate τ through $\widehat{\sigma}_l, \widehat{\sigma}_c, \widehat{\sigma}_r$ and $\widehat{I}$ is an isomorphism between $\widehat{\sigma}_l(\mathscr{M})$ and $\widehat{\sigma}_r(\mathscr{M}^*)$. Hence, by g-solidity of τ we have an $\mathscr{N}$-definable isomorphism $\widehat{J}$ between $\widehat{\sigma}_r(\mathscr{M}^*)$ and $\widehat{\sigma}_c(\mathscr{N})$. $\widehat{J}$ can be canonically extended to an $\mathscr{N}$-definable embedding $J : \sigma_r(\mathscr{M}^*) \to \sigma_c(\mathscr{N})$ by putting, for an arbitrary $(n+1)$-st order object X

$$J(X) := \Big\{\widehat{J}(x) \in \widehat{\sigma}_c(\mathscr{N}) \ \mid \ x \in X\Big\} \, (= \widehat{J}[X]).$$

The fact that J is a well defined function between the $(n+1)$-st sorts of $\sigma_r(\mathscr{M}^*)$ and $\sigma_c(\mathscr{N})$ follows from the fact that $\mathscr{N}$ accommodates the full comprehension scheme through σ_c.

We claim that J is onto the $(n+1)$-st sort of $\sigma_c(\mathscr{N})$. Assume not and consider $K := J \circ I$. Then K is an $\mathscr{M}$-definable embedding of $\sigma_l(\mathscr{M})$ into $\sigma_c(\mathscr{N})$, which is not onto. Let X be outside the image of this embedding. Let $\widehat{K}$ be the restriction of K to the first n-sorts, hence $\widehat{K} = \widehat{J} \circ \widehat{I}$. Hence $\widehat{K}$ is an isomorphism between $\widehat{\sigma}_l(\mathscr{M})$ and $\widehat{\sigma}_c(\mathscr{N})$. Consider the $\mathscr{M}$-definable subset Y of $\widehat{\sigma}_l(\mathscr{M})$ given by $\widehat{K}^{-1}[X]$. We claim that $K(Y) = X$ contradicting the choice of X. Indeed, fix an arbitrary $x \in \widehat{\sigma}_c(\mathscr{N})$ and $y \in \widehat{\sigma}_l(\mathscr{M})$ such that $x = \widehat{K}(y)$. Clearly we have:

$$x \in X \Leftrightarrow \widehat{K}(y) \in X \Leftrightarrow y \in Y \Leftrightarrow K(y) \in K(Y) \Leftrightarrow x \in K(Y).$$

By Extensionality, this shows that $J(Y) = X$, as desired. □

Definition 83 (Extremal schemes)**.** Let $\mathscr{L}$ be any language containing a predicate P. Let α be any $\mathscr{L}$-sentence and β be any $\mathscr{L}$-formula. The (α,β)-*minimality* scheme, denoted $\mu_\beta(\alpha)$ is the following $\mathscr{L}$-sentence

$$\alpha \to \forall x\big(\beta(x) \to P(x)\big).$$

The (α,β)-*maximality* scheme is defined analogously as

$$\alpha \to \forall x\big(P(x) \to \beta(x)\big).$$

Example 84. Both the induction scheme and the Burgess minimality scheme (Burge, 1996) are instances of the minimality schemes for sufficiently chosen α and β. Indeed, to see the former, take $\alpha := P(0) \wedge \forall x\big(P(x) \to P(x+1)\big)$ and $\beta := x = x$. In this way, $\mu_\beta(\alpha)$ says that the universe is its least subset closed under successor and containing 0. To see the latter, take α to be the conjunction of the compositional axioms of KF (with the predicate P substituted for T) and $\beta := T(x)$. By considering the extension of KF with the (α,β)-maximality scheme (for α,β defined as above) one obtains the dual version of Burgess extension of KF.

Theorem 85. *Let $*$ be either g or e. Suppose that $\tau(P)$ is a $*$-solid scheme for a language $\mathscr{L}$ and let Q be a fresh predicate. Let α be any sentence of $\mathscr{L}_P$. Then the schemes $\tau \wedge \alpha[Q/P] \wedge \mu_Q(\alpha)$ and $\tau \wedge \alpha[Q/P] \wedge \nu_Q(\alpha)$ are both $*$-solid (as $\mathscr{L}_Q$ templates).*

Proof. We run the proofs for g and e in parallel. Pick $\mathscr{M}_l, \mathscr{M}_c, \mathscr{M}_r$ as in the definition of $*$-solidity for $\tau \wedge \alpha[Q/P] \wedge \mu_Q(\alpha)$ (i.e. $\mathscr{M}_l$, $\mathscr{M}_c$ and $\mathscr{M}_r$ rename $\mathscr{M}$, $\mathscr{N}$ and $\mathscr{M}^*$, respectively, for notational convenience).

Let σ_l, σ_c and σ_r witness the accommodations of $\tau \wedge \alpha[Q/P] \wedge \mu_Q(\alpha)$, respectively. Finally let F be the $\mathscr{M}_l$-definable isomorphism between $\sigma_l(\mathscr{M}_l)$ and $\sigma_r(\mathscr{M}_r)$. Let $\sigma'_l, \sigma'_c, \sigma'_r$ be the restrictions of these translations to the language $\mathscr{L}$. By definition, each σ'_i witnesses that $\mathscr{M}_i$ accommodates τ.

By the $*$-solidity of τ we obtain an $\mathscr{M}_c$-definable isomorphism $G_c : \sigma'_c(\mathscr{M}_c) \to \sigma'_r(\mathscr{M}_r)$ and consequently an $\mathscr{M}_l$-definable isomorphism $G_c^{-1} \circ F =: G_l : \sigma'_l(\mathscr{M}_l) \to \sigma'_c(\mathscr{M}_c)$. For $i \in \{l,r,c\}$, let Q_i be the interpretation of Q in $\sigma_i(\mathscr{M}_i)$.

We shall argue that G_c maps Q_c to Q_r and hence is also an isomorphism of $\sigma_c(\mathscr{M}_c)$ with $\sigma_r(\mathscr{M}_r)$. Since $G_l^{-1}[Q_c]$ is an $\mathscr{M}_l$-definable subset of $\sigma_l(\mathscr{M}_l)$ such that

$$(\sigma'_l(\mathscr{M}_l), G_l^{-1}[Q_c]) \models \alpha[Q/P],$$

by the minimality scheme in $\mathscr{M}_l$ it follows that $Q_l \subseteq G_l^{-1}[Q_c]$. Hence $F[Q_l] \subseteq G_c[Q_c]$. However, since F is an isomorphism of $\mathscr{L}_Q$ structures, then $F[Q_l] = Q_r$. By applying G_c^{-1} we get $G_c^{-1}[Q_r] \subseteq Q_c$. However, since

$$(\sigma'_c(\mathscr{M}_c), G_c^{-1}[Q_r]) \models \alpha[Q/P],$$

and $G_c^{-1}[Q_r]$ is $\mathscr{M}_c$-definable, by the minimality scheme we obtain $Q_c \subseteq G_c^{-1}[Q_r]$, which concludes the proof. □

For completeness we note the following solidity variant of Theorem 60, which is proved with a similar argument.

Theorem 86. *Let τ and σ be two scheme-templates and $\mathscr{L}$ be a language such that $\tau[\mathscr{L}] \equiv \sigma[\mathscr{L}]$. Denote this theory with T. Suppose further that $\mathsf{PC}(T) \vdash \forall X \tau[X] \to \forall X \sigma[X]$. Then if σ is $*$-solid (for $* \in \{e,g\}$), then so is τ.*

We close this section by applying our results to canonical axiomatic theories of truth.

Definition 87. (Truth theories)

(a) CT[PA] is the theory obtained by augmenting PA with CT (finitely many axioms of compositional truth over arithmetic) with all instances of induction in the extended language (i.e., the language of PA augmented with a truth predicate).

(b) $\mathsf{KF}_\mu[\mathsf{PA}]$ is Burgess' extension (2014) of the Kripke-Feferman truth theory $\mathsf{KF}[\mathsf{PA}]$; the choice μ is informed by the fact that the truth theory corresponding to the *least* fixed point of the usual construction of a model of $\mathsf{KF}[\mathsf{PA}]$ yields a model of $\mathsf{KF}_\mu[\mathsf{PA}]$.

(c) Let $\mathsf{CT}_\sigma[\tau]$ be the theory in the language $\mathscr{L}_S$ with a fresh binary predicate S extending the schematic theory generated by the template $\tau \wedge \tau_{\mathsf{Ind}}[\mathscr{L}_S]$ with axioms that are the natural counterparts of compositional axioms of CT in the language with the satisfaction predicate, in which:

(1) the syntactical operations are translated by σ;

(2) the quantifiers over syntactical objects are relativized to the domain of σ;

(3) the quantifiers over assignments are unrelativized.

(d) $\mathsf{KF}_{\mu,\sigma}[\tau]$ is analogously defined as $\mathsf{CT}_\sigma[\tau]$, as an extension of the schematic theory generated by the template $\tau \wedge \tau_{\mathsf{Ind}} \wedge \mu_T(\bigwedge \mathsf{KF})$, where $\bigwedge \mathsf{KF}$ denotes the conjunction of the compositional axioms of KF.

Theorem 88. *The following axiomatic truth theories are solid:*

(a) $\mathsf{CT}[\mathsf{PA}]$*, and each of its finite iterates* $\mathsf{CT}[\mathsf{CT}[\mathsf{PA}]]$*, etc.*

(b) $\mathsf{KF}_\mu[\mathsf{PA}]$.

(c) *More generally, suppose that τ is a $*$-solid (where $*$ is either g or e) $\mathscr{L}$-scheme such that $\tau[\mathscr{L}]$ accommdates the induction scheme τ_{Ind} via a translation σ. **Then** $\mathsf{CT}_\sigma[\tau]$ and $\mathsf{KF}_{\mu,\sigma}[\tau]$ are solid theories.*

Proof. The solidity of $\mathsf{CT}[\mathsf{PA}]$ (as well as its finite iterates) can be readily obtained through the g-solidity of the scheme $\bigwedge \mathsf{CT} \wedge \tau_{\mathsf{Ind}}$, , where $\bigwedge \mathsf{CT}$ is the conjunction of the finitely many compositional axioms of CT. Indeed, observe that over PC($\mathsf{CT}[\mathsf{PA}]$), the sentence $\forall X \tau_{\mathsf{Ind}}(X)$ implies the minimality of T, i.e. $\forall X \mu_T(\bigwedge \mathsf{CT})$. Since $\tau_{\mathsf{Ind}} \wedge \bigwedge \mathsf{CT} \wedge \mu_T(\bigwedge \mathsf{CT})$ is g-solid (Theorem 85), so is $\bigwedge \mathsf{CT} \wedge \tau_{\mathsf{Ind}}$. Hence CT is indeed solid.

For the proof of (b) it is enough to observe that Burgess' KF_μ is axiomatized by the scheme $\sigma := \tau_{\mathsf{Ind}} \wedge \bigwedge \mathsf{KF} \wedge \mu_T(\bigwedge \mathsf{KF})$, where $\bigwedge \mathsf{KF}$ is the conjunction of the finitely many compositional axioms of KF. τ_{IND} is g-minimalist, hence by Theorem 85 σ is g-solid. Thus $\sigma[\mathscr{L}_T]$ is solid (where $\mathscr{L}_T$ is the language extending the language of arithmetic with a fresh unary predicate T).

The proof of (c) follows from Theorem 85, which generalizes parts (a) and (b). □

Remark 89. The minimalist property of PA is not shared by $\mathsf{CT}[\mathsf{PA}]$. To see this, it is sufficient to observe that if $\mathscr{M}$ and $\mathscr{N}$ are models of $\mathsf{CT}[\mathsf{PA}]$ and $\mathscr{M}$ is a submodel of $\mathscr{N}$, then the arithmetical theories of the two models are the same. Hence, in order to demonstrate that $\mathsf{CT}[\mathsf{PA}]$ is not minimalist, it is enough to consider $\mathscr{M} \models \mathsf{CT}[\mathsf{PA}] + \mathsf{Con}_{\mathsf{CT}[\mathsf{PA}]}$. Then $\mathscr{M}$, by the formalized second Gödel's incompleteness theorem, $\mathscr{M} \models \mathsf{Con}_{\mathsf{CT}[\mathsf{PA}]+\neg\mathsf{Con}_{\mathsf{CT}[\mathsf{PA}]}}$, hence $\mathscr{M}$ interprets a model $\mathscr{N} \models \mathsf{CT}[\mathsf{PA}] + \neg\mathsf{Con}_{\mathsf{CT}[\mathsf{PA}]}$ via the Arithmetized Completeness Theorem. However there can be no embedding between $\mathscr{M}$ and $\mathscr{N}$ as the arithmetical theories of the two models are different.

5. Philosophical discussion

In this section we outline the philosophical implications of the formal results in the preceding sections that we view as relevant for contemporary debates and discussions in philosophy of science and the foundations of mathematics. We focus mostly on two topics: (1) The use of bi-interpretability and synonymy (definitional equivalence) as good explications of sameness of theories, a view present in discussions in the philosophy of science; and (2) The use of categoricity-style arguments in the debate over the determinacy of some core mathematical concepts, such as the notion of a number, or the notion of a set. These two conceptual areas naturally correspond to Sections 3 and 4 of our paper.

5.1. The meaning of bi-interpretability, solidity and tightness

Synonymy is typically treated as a good explication of the notion of sameness of theories, which works nicely for theories in different signatures.[33] Intuitively, a definitional extension of a theory T "says no more" than T itself, hence two theories that have a common definitional extension "say the same thing", and thus they should have the same content. However, despite the conceptual attractiveness of synonymy, this approach needs to contend with some notorious examples of theories that are synonymous, yet seem to be clearly different. One of the most famous examples is provided by two extensions of $\mathsf{KF}[\mathsf{PA}]$ with axioms expressing (Cons) "No sentence is both true and false" and (Comp) "Every sentence is either true or false". By trivially disjoining the signatures of $\mathsf{KF}[\mathsf{PA}]+\mathsf{Comp}$ and $\mathsf{KF}[\mathsf{PA}]+\mathsf{Cons}$, one easily finds a common definitional extension of the two theories.[34] However, $\mathsf{KF}[\mathsf{PA}]\vdash \mathsf{Cons}\rightarrow\neg\mathsf{Comp}$, so the two theories are inconsistent.

Section 3.4 provides new examples of such unexpected failures of sameness between synonymous theories: each restricted fragment of any of the canonical theories from among $\{\mathsf{PA},\mathsf{ZF},\mathsf{Z}_2,\mathsf{KM}\}$ has two different complete bi-interpretable extensions. In particular for each $n, \mathrm{I}\Sigma_n$ has a model in which $\mathsf{I}\Sigma_{n+1}$ induction fails and yet the model is bi-interpretable with the standard model of arithmetic. By examining the proof, one notices that in fact the interpretations used in the bi-interpretability do not involve parameters and are equality preserving, hence by the Visser-Friedman theorem (Friedman and Visser, 2014) the theories of both models are synonymous (treated as theories in disjoint signatures). However, the theories disagree on whether the induction holds and it is hard to imagine a more striking example of a foundational disagreement.

One can counter the conclusions of the above paragraphs, by pointing out that we misinterpret the *content* of theories, as measured by synonymy. For if we agree that a definitional extension of a theory U says no more than U, then we implicitly assume that what we really care about is the structure of U-definable sets. We abstract from the complexity of the definitions and treat all the notions that can be expressed in U on a par. This explains the sameness of $\mathsf{KF}[\mathsf{PA}]+\mathsf{Cons}$ and $\mathsf{KF}[\mathsf{PA}]+\mathsf{Comp}$: using a consistent truth predicate we can indeed define a complete one and vice versa. Hence two theories are equally expressive, only they disagree on which truth predicate is "the basic one". We agree with this argument, however we think that for the philosophical use of axiomatic theories one would expect a more fine-grained notion. As we explain in the next paragraph, bi-interpretability over tight theories seems to preserve more information about what the theories are meant to axiomatize.

Positive results about solidity and tightness show that "counterexamples" of the above discussed type do not occur if we restrict attention to the extensions of appropriate theories: whenever U and V are theories in the same language which both extend any of the solid theories from the list including PA, ZF, Z_n, KM, $\mathsf{CT}[\mathsf{PA}]$, $\mathsf{KF}_\mu[\mathsf{PA}]\ldots$, then U is a retract of V implies $U\vdash V$. This leads to a conclusion that bi-interpretability and synonymy are more restrictive on "cones" originating from foundationally salient theories[35]

The above naturally leads to a question: What is so special about solid theories? Unlike internal categoricity, whose definition preserves many intuitions supporting the traditional notion of categoricity, solidity seems to lack such a ‘conceptual charm’. We shall try to explain the intuitions behind this notion more carefully. First of all, let us observe that solidity speaks about the *local* behaviour of a first-order U across all models of U, i.e., it considers what happens in each of the models of the theory separately. Speaking very informally, each model of a solid theory thinks of itself as a special or distinguished model from among the ones that it can talk about. More formally: consider a solid theory U and a model $\mathscr{M}\models U$ and let $U(\mathscr{M})$ be the set of all models of U that are parametrically definable in $\mathscr{M}$ (considered up to the $\mathscr{M}$-definable isomorphism). Of course, $\mathscr{M}$ itself is an element of $U(\mathscr{M})$. Elements of $U(\mathscr{M})$ are naturally pre-ordered by the relation of parametric definability. Intuitively speaking $\mathscr{N}$ is (strictly) greater in this ordering than a model $\mathscr{K}$, if $\mathscr{N}$ can see $\mathscr{K}$, but not vice versa. Solidity simply says that $\mathscr{M}$ is the greatest element in this pre-order: whenever $\mathscr{N}\in U_{\mathscr{M}}$ (hence $\mathscr{M}$ sees $\mathscr{N}$) and $\mathscr{N}$ is also greater or equal to $\mathscr{M}$ ($\mathscr{N}$ sees an isomorphic copy of $\mathscr{M}$), then actually $\mathscr{N}$ is equal to $\mathscr{M}$ (there is an $\mathscr{M}$-definable isomorphism between $\mathscr{N}$ and $\mathscr{M}$). We think that this explanation should suffice to

33 It is worth adding that sometimes synonymy is taken to be too restrictive and a weaker notion, called Morita equivalence is preferred; as shown in (McEldowney, 2020) Morita equivalence implies bi-interpretability under very mild assumptions.

34 One defines a complete/consistent truth predicate T_2 using a consistent/complete one, T_1 by putting $T_2(x) := \neg T_1(\dot\neg x)$.

35 Recent unpublished work due to Piotr Gruza, Leszek Kołodziejczyk and the second-named-author show that in some of these cases, and most probably in all of them, the full strength of these theories is not necessary to eliminate pathologies in the behaviour of bi-interpretability.

put solidity on the list of categoricity-like properties for first-order theories. Let us observe that, unlike internal categoricity, solidity is a categoricity-like notion which naturally applies to first-order theories as opposed to schemes.

Let us complete this subsection with a metaphor that guides our thinking about an important consequence of solidity, namely *neatness* (as in Definition 7), which links the notion of a retract with that of theory extension. One can "define" a natural or canonical theory as the one that "has a story to tell". If so, then one can see a neat theory to be akin to the first episode of a well-written series of stories, in which the characters, and the relationships and tensions between them, are well-developed. Such a successful first episode is self-standing in its own right, but also allows the next episodes to be developed on the basis of it, without being seen as its inevitable conclusion.

5.2. Internal categoricity: second order predicates, schemes vs. (first-order) theories

We shall start from a reconstruction of an extensively discussed philosophical puzzle: the structure of the natural numbers *seems to be* both (a) well determined (by the operation of the successor and basic recursive equations for addition and multiplication) and (b) first-order describable. Of course, a natural way of making the above claim mathematically precise demonstrates that in fact (a) contradicts (b): the full (complete) first-order theory of the natural numbers (the standard model) can be realized in various nonstandard models. One can react to this phenomenon by admitting that there is indeed an irreducible second-order factor in our understanding of the canonical structure of the natural numbers; or one might try to explicate the condition of "uniqueness" in some other way. The internal categoricity approach uncovers a path leading towards a philosophically intriguing explanation of the latter option. It starts with an observation that actually the categoricity of various prominent second-order systems is *essentially* a very concrete phenomenon (we shall refer to this loosely defined property as "concrete categoricity")[36]: the statement "Each two structures which satisfy the theory U are isomorphic" is provable in the canonical proof system for second-order logic (which is sound and complete w.r.t. Henkin semantics) for the cases when U is one of the canonical theories such as PA, $\mathsf{CT}[\mathsf{PA}]$, restricting to models of the same ordinal height in the case of ZF. Even though to obtain these results one makes use of the second-order logic with Henkin semantic (with full comprehension), the appeal to the full second-order semantics has been bypassed. Secondly, one can formulate a first-order version of this observation by further noticing that in each of the cases there is a procedure of finding the required isomorphism, which is uniform with respect to the definitions of the models for the given theory. In this way one arrives at the definitions of *internal categoricity* (for first-order theories) and what we here dubbed *strong internal categoricity* in Definition 44. In both the syntactic second-order and the first-order approaches one obtains *intolerance* results: If U is a *concretely* categorical theory, then for every sentence φ separately one can prove that either φ holds in all U-structures or $\neg\varphi$ holds in all U-structures.[37] An interesting variation of the syntactic second-order approach and the first-order approach was put forward by Fischer and Zicchetti (2023): the authors use an axiomatic theory of truth to obtain a first-order intolerance in the form of a single universal sentence (as opposed to a scheme).

Our results from Section 4 bring more information about the peculiarities of each of the above introduced approaches to formalising the idea of concrete categoricity, i.e., the syntactic second-order approach, the first-order approach and the truth-theoretic approach.[38] First of all, we developed a simple hierarchy of categoricity-like notions (as in Theorem 74) in the first-order realm that can serve as a simple indicator of a 'concrete categoricity degree' of a given first-order portion of mathematics. The internal categoricity is at the bottom of this hierarchy. Secondly, we claim that the properties defining each level in this categoricity scale (including internal categoricity) are best seen as properties of *schemes*, as opposed to first-order *theories*. This is supported by the following observations:

(a) A first order theory may be axiomatized by schemes with very different "degrees of concrete categoricity", ranging from those that are not internally categorical (the Vaught schemes, see Theorem 52) up to those that are g-minimalist (in the case of the usual schematic axiomatization of PA).

[36] This is because we decided to fix a meaning of "internal categoricity" that seems to be used in this context.
[37] For the precise explanation of these results in each scenario, consult (Maddy and Väänänen, 2023).
[38] Indeed, Fischer and Zicchetti consider also one more perspective based on the work of Feferman and Hellman.

(b) This perspective is fruitful: we can define a simple syntactic operations on schemes and show that some categoricity degrees are preserved by them thus explaining some regularities in the distribution of internally categorical theories.

What is more, there is a very clear philosophical idea explaining why schemes are better carriers of internal categoricity than first-order theories: as opposed to theories, schemes are potentially *open-ended*, which means that they are meaningfully applicable to any first-order language. Continuing along these lines, mathematical agents can accept a given scheme τ as an open-ended entity, which means that they commit to extending τ to any language they find comprehensible. Mathematically speaking, schemes are functions which, when applied to a first-order language, return a first-order theory. Seen from this perspective, Theorem 52 shows that two such functions may coincide on a given language, however behave very differently on another ones. Further degrees on our scale, e-solidity, g-solidity and g-minimalist, are underpinned by the same intuition: how competent a given scheme is when presented with sets definable in various other languages. However, in designing the formal definitions we were admittedly led by analogies with the categoricity notions of theories and interpretability theory (the distinction between g- and e-solidity reflects the difference between the direct and relative interpretation).

The above highlights the core difference between schemes and first-order theories. How about schemes and second order formulations of theories, present in the syntactic second-order approach? Here we point to one interesting phenomenon linked to the Tarski component added to various schemes. Consider the replacement scheme conjuncted with the scheme expressing "no formula defines the truth for the set-theoretic universe" (as in $\tau_{\mathsf{Repl+Tarski}}$ of Definition 68). As shown in Theorem 77 this scheme is, on our scale, at least one level 'more categorical' in relation to the usual schematic axiomatization of ZF.[39] Moreover, there is a clear philosophical idea why mathematical agents might like to accept such a scheme in an open-ended fashion. Let's think about a set-theorist claiming that there is one all-embracing universe of sets, V, and all of mathematics takes place within it. It would be natural to think that there is no external-to-V point of view. By the Tarski undefinability of truth theorem, the truth for V should be absolutely undefinable, a claim which is reflected in the acceptance of the Tarski component of our scheme.[40]

What can be said about the categoricity of $\tau_{\mathsf{Repl+Tarski}}$ in the context of second-order logic with Henkin semantics and full comprehension? Indeed, in this second order context $\tau_{\mathsf{Repl+Tarski}}$ is categorical (in the sense that any two models of the second order theory are isomorphic), provably in the second order logic. However, this is because of the fact that provably in second order logic there is no structure that satisfies the second order version of $\tau_{\mathsf{Repl+Tarski}}$ (for a scheme τ in the signature with one binary relational symbol and for two second order variables X, E, $\tau^{X,E}$ denotes the result of relativizing the quantifiers in τ to X and substituting E for the relational symbol):

Theorem 90. *The following sentence is provable in pure second order logic (with full comprehension):*

$$\forall X\ \forall E\ \neg\forall Y \subseteq E\ \ \tau^{X,E}_{\mathsf{Repl+Tarski}}(Y).$$

Proof. Suppose otherwise and consider a Henkin model $\mathscr{M}$ such that some pair of sets (X, E) satisfies $\tau_{\mathsf{Repl+Tarski}}$ with respect to all subsets of X which exists in $\mathscr{M}$. However, the Tarskian satisfaction class for (X, E) can be seen (via coding) as a subset of X, which is Σ^1_1-definable in (X, E), so by comprehension it exists as an element of $\mathscr{M}$. This contradicts the fact that (X, E) satisfies the Tarski component of $\tau_{\mathsf{Repl+Tarski}}$. □

We think that the above observation speaks in favour of the schematic approach to the phenomenon of concrete categoricity. Unlike in the syntactic second order approach, the schematic approach can do justice to the coherence of the standpoint motivating the acceptance of schemes with the Tarski component. The fact that this cannot be accomplished in the syntactic second-order framework is, in our opinion, due to the fact that the full second-order comprehension introduces a perspective that is (at least partially) external to second order theories under consideration. In our proof, the Tarskian satisfaction class, although undefinable in the structure

39 And two levels "more categorical" if g-solidity is strictly stronger than e-solidity, which we conjecture.

40 The philosophical ramifications of the open-ended feature of schemes has been explored by Charles Parsons, see, e.g., (Parsons, 2008).

(X,E) can be shown to exists by the use of Σ^1_1-comprehension, available in the full second-order logic. Let us observe that these arguments are in line with the known objections against making use of impredicative (Π^1_1) comprehension to explain Parson's problem on how can mathematicians know that they are talking about the same structure of natural numbers. We refer to (Fischer and Zicchetti, 2023) for a discussion.

Last but not least, we stress how our results can be used to extend the truth-theoretic approach, put forward in (Fischer and Zicchetti, 2023). An important step in this approach was to show that the the usual scheme axiomatizing the theory $\mathsf{CT}[\mathsf{PA}]$ is strongly internally categorical. We show (Theorem 85) that also an untyped theory of truth, $\mathsf{KF}_\mu[\mathsf{PA}]$ is g-solid, and a similar argument can be applied to show that it is strongly internally categorical. Hence one can devise a truth-theoretic approach based on an untyped truth predicate.

6. Future work

There is still much to be learned about categoricity-like properties of first order theories. The following list of questions and conjectures provides a glimpse of the unexplored part of the territory.

1. **Question.** *Is there a consistent sequential theory that is maximalist?* The notion of a maximalist theory was introduced in Definition 7.
2. **Conjecture.** *There is no finitely axiomatizable sequential tight theory.* Note that by Theorem 38 the conjecture is confirmed for finitely axiomatized subtheories of the canonical theories in the list $\mathscr{S}$ of Theorem 14.
3. **Question.** *Is there a solid deductively closed proper subtheory of* ZF *that includes* KP *(Kripke-Platek set theory) and the axiom of infinity?* This question is informed by the solidity of ZF.
4. **Question.** *Is the* $\mathsf{PA}^- + \mathsf{Collection}$ *a tight theory?* Recall from Theorem 35 that $\mathsf{PA}^- + \mathsf{Collection}$ is not a solid theory.
5. **Conjecture.** *For every sequential r.e. theory T in a finite language $\mathscr{L}$ there is an $\mathscr{L}$-scheme template τ such that $\tau[\mathscr{L}] = T$ and τ is not internally categorical.* This conjecture is informed by Theorem 52.
6. **Question.** *What is the relationship between the notion of internal categoricity in the context of Henkin models of pure second order logic (as in Button and Walsh, 2018) and the categoricity-like concepts of scheme templates studied in this paper?* This question is motivated by juxtaposing the (proof of) Theorem 77 with Theorem 90.
7. **Question.** *What are the categoricity-like properties of the so-called restricted schemes (in the sense of Wilkie, 1987)?*
8. **Conjecture.** *There is an e-solid scheme template τ that is not g-solid and τ axiomatizes a sequential theory.*
9. **Question.** *Is there an arithmetical template that is e-solid, the theory T it generates includes* PA^-*, and the deductive closure of T is a proper subtheory of* PA*?*
10. **Question.** *Do the results in this paper shed any light on the Parsons' dilemma about the determinacy of the structure of the natural numbers?*

Acknowledgements

We have benefitted from conversations (some via email) with several colleagues concerning the topics explored in this paper, including (in reverse alphabetical order of last names) Matteo Zicchetti, Kameryn Williams, Albert Visser, Jouko Väänänen, Stephen Simpson, Philipp Schlicht, Michael Rathjen, Fedor Pakhomov, Toby Meadows, David Marker, Roman Kossak, Leszek Kołodziejczyk, Joel David Hamkins, Piotr Gruza, Alfredo Freire, Thomas Forster, Vincenzo Dimonte, Tim Button, and Lev Beklemishev. A special note of thanks goes to

Daniel Isaacson and other participants of the Oxford Philosophy of Mathematics Seminar for their probing questions and perceptive comments following a seminar presentation in January 2024 by the first-named-author. Finally, hats off to the two anonymous referees whose reports were instrumental in the preparation of the final draft. First author is supported by the National Science Centre, Poland (NCN), grant number 2019/34/A/HS1/00399. Second author is supported by the National Science Centre, Poland (NCN), grant number 2022/46/E/HS1/00452.

References

Ackermann, W. (1937). Die Widerspruchsfreiheit der allgemeinen Mengenlehre. *Math. Ann.*, 114(1), 305–315.

Barwise, J. (1975). *Admissible sets and structures*. Perspectives in Mathematical Logic. Springer-Verlag.

Burge, T. (1996). Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society*, 96(1), 91–116.

Burgess, J. P. (2014). Friedman and the axiomatization of Kripke's theory of truth. In *Foundational adventures* (Vol. 22, Tributes, pp. 125–148). Coll. Publ.

Button, T. & Walsh, S. (2018). *Philosophy and model theory*. Oxford University Press.

Cégielski, P, Cornaros, C. & Dimitracopoulos, C. (Eds.). (2013). *New studies in weak arithmetics* (Vol. 211, CSLI Lecture Notes). CSLI Publications; Presses Universitaires du Pôle de Recherche et d'Enseignement Supérieur Paris-Est.

Chang, C. C. & Keisler, H. J. (1990). *Model theory* (3rd ed., Vol. 73, Studies in Logic and the Foundations of Mathematics). North-Holland Publishing Co.

Chao, C. & Seraji, P. (2018). Gödel's second incompleteness theorem for Σ_n-definable theories. *Log. J. IGPL*, 26(2), 255–257.

Enayat, A. (1986). Conservative extensions of models of set theory and generalizations. *J. Symbolic Logic*, 51(4), 1005–1021.

Enayat, A. (2004). Leibnizian models of set theory. *J. Symbolic Logic*, 69(3), 775–789.

Enayat, A. (2005). Models of set theory with definable ordinals. *Arch. Math. Logic*, 44(3), 363–385.

Enayat, A. (2016). Variations on a Visserian theme. In *A tribute to Albert Visser* (Vol. 30, Tributes, pp. 99–110). Coll. Publ.

Enayat, A. (2021). Tight theories. Online Workshop on Gödel Incompleteness Theorems. http://yongcheng.whu.edu.cn/webPageContent/Goedel2021/Goedel-Enayat.pdf

Enayat, A. (2024). Models of set theory: Extensions and dead-ends. arXiv:1804.09526 [math.LO].

Enayat, A. & Łełyk, M. (2023). Axiomatizations of Peano Arithmetic: A truth-theoretic view. *J. Symb. Log.*, 88(4), 1526–1555.

Enayat, A., Łełyk, M. & Visser, A. (2024). Completions of restricted complexity. To appear.

Enayat, Ali & McKenzie, Z. (2021). Initial self-embeddings of models of set theory. *J. Symb. Log.*, 86(4), 1584–1611.

Enayat, A., Schmerl, J. H., & Visser, A. (2011). ω-models of finite set theory. In *Set theory, arithmetic, and foundations of mathematics: theorems, philosophies* (Vol. 36, Lect. Notes Log., pp. 43–65). Assoc. Symbol. Logic.

Feferman, S. (1960). Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*.

Fischer, M. & Zicchetti, M. (2023). Internal categoricity, truth and determinacy. *J. Philos. Logic*, 52(5), 1295–1325.

Freire, A. & Hamkins, J. D. (2021). Bi-interpretation in weak set theories. *J. Symb. Log.*, 86(2), 609–634.

Freire, A. & Williams, K. J. (2023). Non-tightness in class theories and second order arithmetic. *J. Symb. Log.*, 1–28.

Friedman, H. & Visser, A. (2014). When bi-interpretation implies synonymy. *Logic Group Preprint Series*, 9(3), 1–19.

Gitman, V., Hamkins, J. D. & Johnstone, T. A. (2016). What is the theory ZFC without power set? *MLQ Math. Log. Q.*, 62(4-5), 391–406.

Hájek, P. & Pudlák, P. (1998). *Metamathematics of first-order arithmetic*. Springer.

Hodges, W. (1993). *Model theory* (Vol. 42, Encyclopedia of Mathematics and its Applications). Cambridge University Press.

Jech, T. (2003). *Set theory*. Springer Monographs in Mathematics. Springer-Verlag.

Jeřábek, E. (2012). Sequence encoding without induction. *MLQ Math. Log. Q.*, 58(3), 244–248.

Kanamori, A. (2009). *The higher infinite* (2nd ed.). Springer-Verlag.

Kaye, R. (1991). *Models of Peano Arithmetic*. Clarendon Press.

Kaye, R. & Wong, T. L. (2007). On interpretations of arithmetic and set theory. *Notre Dame J. Formal Logic*, 48(4), 497–510.

Kossak, R. & Schmerl, J. (2006). *The Structure of Models of Peano Arithmetic*. Clarendon Press.

Kreisel, G. & Lévy, A. (1968). Reflection principles and their use for establishing the complexity of axiomatic systems. *Z. Math. Logik Grundlagen Math.*, 14, 97–142.

Kunen, K. (1980). *Set theory* (Vol. 102, Studies in Logic and the Foundations of Mathematics). North-Holland Publishing Co.

Lefever, K. & Székely, G. (2019). On generalization of definitional equivalence to non-disjoint languages. *J. Philos. Logic*, 48(4), 709–729.

Maddy, P. & Väänänen, J. (2023). *Philosophical Uses of Categoricity Arguments*. Cambridge University Press.

Marek, W. & Zbierski, P. (1978). On a class of models of the nth order arithmetic. In *Higher set theory* (Vol. 669, Lecture Notes in Math., pp. 361–374). Springer.

Marker, D. (1991). End extensions of normal models of open induction. *Notre Dame J. Formal Logic*, 32(3), 426–431.

Mathias, A. R. D. (2001). The strength of Mac Lane set theory. *Ann. Pure Appl. Logic*, 110(1-3), 107–234.

McEldowney, P. A. (2020). On morita equivalence and interpretability. *The Review of Symbolic Logic*, 13, 388–415.

Montague, R. (1961). Semantical closure and non-finite axiomatizability. I. In *Infinitistic Methods* (pp. 45–69). Pergamon.

Mycielski, J. (1964). Definition der arithmetischen Operationen im Ackermannschen Modell. *Algebra Logika*, 3(5-6), 64–65.

Pakhomov, F. & Visser, A. (2019). On a question of Krajewski's. *J. Symbolic Logic*, 84, 343–358.

Paris, J. B., & Kirby, L. A. S. (1978). Σ_n-collection schemas in arithmetic. In *Logic Colloquium '77* (Vol. 96, Stud. Logic Found. Math., pp. 199–209). North-Holland Publishing Co.

Parsons, C. (2008). *Mathematical thought and its objects*. Cambridge University Press.

Robinson, J. (1949). Definability and decision problems in arithmetic. *J. Symbolic Logic*, 14, 98–114.

Schlicht, P. (2017). Perfect subsets of generalized Baire spaces and long games. *J. Symb. Log.*, 82(4), 1317–1355.

Schultzenberg, F. (2023). Projective well-ordered sets, higher up. *MathOverflow*. Retrieved from https://mathoverflow.net/q/460467

Scott, D. (1961). More on the axiom of extensionality. In *Essays on the foundations of mathematics* (pp. 115–131). Magnes Press, The Hebrew University.

Shepherdson, J. C. (1964). A non-standard model for a free variable fragment of number theory. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 12, 79–86.

Simpson, S. G. (2009). *Subsystems of second order arithmetic* (2nd ed.). Cambridge University Press.

Solovay, R. M. (1970). A model of set-theory in which every set of reals is Lebesgue measurable. *Ann. of Math.*, 92, 1–56.

Väänänen, J. (2012). Second order logic or set theory? *Bull. Symbolic Logic*, 18(1), 91–121.

Väänänen, J. (2019). An extension of a theorem of Zermelo. *Bull. Symbolic Logic*, 25(2), 208–212.

Väänänen, J. (2021). Tracing internal categoricity. *Theoria*, 87(4), 986–1000.

Väänänen, J. & Wang, T. (2015). Internal categoricity in arithmetic and set theory. *Notre Dame J. Formal Logic*, 56(1), 121–134.

Vaught, R. L. (1967). Axiomatizability by a schema. *J. Symbolic Logic*, 32, 473–479.

Visser, A. (2006). Categories of theories and interpretations. In *Logic in Tehran* (Vol. 26, Lect. Notes Log., pp. 284–341). Assoc. Symbol. Logic.

Visser, A. (2013). What is sequentiality? In *New studies in weak arithmetics* (Vol. 211, CSLI Lecture Notes, pp. 229–269). CSLI Publ.

Visser, A. (2016). On Q. *Soft Computing*, 21(3), 39–56.

Wilkie, A. J. (1987). On schemes axiomatizing arithmetic. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2* (pp. 331–337). Amer. Math. Soc.

Williams, K. (2018). The structure of models of second order set theories. *arXiv:1804.09526* [math.LO].