The Review of Symbolic Logic
Volume 17, Number 1, March 2024

DISJUNCTION AND EXISTENCE PROPERTIES IN MODAL ARITHMETIC

TAISHI KURAHASHI
Kobe University
and
MOTOKI OKUDA
Independent Scholar

Abstract. We systematically study several versions of the disjunction and the existence properties in modal arithmetic. First, we newly introduce three classes B, $\Delta(\mathrm{B})$, and $\Sigma(\mathrm{B})$ of formulas of modal arithmetic and study basic properties of them. Then, we prove several implications between the properties. In particular, among other things, we prove that for any consistent recursively enumerable extension T of $\mathbf{PA}(\mathbf{K})$ with $T \nvdash \Box\bot$, the $\Sigma(\mathrm{B})$-disjunction property, the $\Sigma(\mathrm{B})$-existence property, and the B-existence property are pairwise equivalent. Moreover, we introduce the notion of the $\Sigma(\mathrm{B})$-soundness of theories and prove that for any consistent recursively enumerable extension of $\mathbf{PA}(\mathbf{K4})$, the modal disjunction property is equivalent to the $\Sigma(\mathrm{B})$-soundness.

§1. Introduction. A theory or a logic T is said to have the *disjunction property* (DP) if for any sentences φ and ψ in the language of T, if $T \vdash \varphi \vee \psi$, then $T \vdash \varphi$ or $T \vdash \psi$. This is a property that may be considered to represent the constructivity of intuitionistic logic. Gödel [8] noted that the intuitionistic propositional logic has DP. Gentzen [7] and Kleene [14] proved that the intuitionistic quantified logic and Heyting arithmetic **HA** have DP, respectively. A property in arithmetic that is related to DP is the (numerical) existence property. We say that a theory T of arithmetic has the *existence property* (EP) if for any formula $\varphi(x)$ that has no free variables except x, if $T \vdash \exists x \varphi(x)$, then $T \vdash \varphi(\overline{n})$ for some natural number n. Here $\overline{n}$ is the numeral for n. Kleene [14] also proved that **HA** has EP. Moreover, Friedman [5] proved that for any recursively enumerable (r.e.) extension T of **HA**, T has DP if and only if T has EP.

A similar situation has been shown to be true for modal arithmetic. Modal arithmetic is a framework of arithmetic equipped with the unary modal operator $\Box$. Let $\mathcal{L}_A$ and $\mathcal{L}_A(\Box)$ be the languages of arithmetic and modal arithmetic, respectively. A prominent $\mathcal{L}_A(\Box)$-theory of modal arithmetic is **EA** (epistemic arithmetic) which is obtained by adding **S4** into Peano arithmetic **PA**. The theory **EA** was independently introduced by Shapiro [21] and Reinhardt [19, 20]. In this framework, $\Box$ is intended to represent knowability or informal provability, and the language $\mathcal{L}_A(\Box)$ has the expressive power to make analyses about these concepts. Moreover, it was shown that **HA** is faithfully

Received: October 6, 2021.
2020 *Mathematics Subject Classification*: Primary 03B45, 03F30, 03F40.
Key words and phrases: modal arithmetic, disjunction property, existence property.

doi:10.1017/S1755020322000363

embeddable into **EA** via Gödel's translation (cf. [4, 9, 21]). This result verifies Shapiro's suggestion that **EA** is a system about both classical and intuitionistic mathematics. From his suggestion, **EA** may possess some constructive properties. A theory or a logic T is said to have the *modal disjunction property* (MDP) if for any $\mathcal{L}_A(\Box)$-sentences φ and ψ, if $T \vdash \Box\varphi \vee \Box\psi$, then $T \vdash \varphi$ or $T \vdash \psi$. Also, T is said to have the *modal existence property* (MEP) if for any $\mathcal{L}_A(\Box)$-formula $\varphi(x)$ that has no free variables except x, if $T \vdash \exists x \Box\varphi(x)$, then $T \vdash \varphi(\overline{n})$ for some natural number n. Then, Shapiro [21] proved that **EA** has both MDP and MEP. Moreover, Friedman and Sheard [6] proved that for any r.e. $\mathcal{L}_A(\Box)$-theory T extending **EA**, T has MDP if and only if T has MEP.[1]

In the case of classical logic, DP is related to the completeness of theories. Indeed, it is easy to see that a consistent theory T based on classical logic has DP if and only if T is complete. Hence, Gödel–Rosser's first incompleteness theorem is restated as follows: For any consistent r.e. extension T of **PA**, T does not have DP. In this context, Gödel–Rosser's first incompleteness theorem can be strengthened. For a class Γ of formulas, we say that a theory T has the Γ-*disjunction property* (Γ-DP) if for any Γ sentences φ and ψ, if $T \vdash \varphi \vee \psi$, then $T \vdash \varphi$ or $T \vdash \psi$. Also, T is said to have the Γ-*existence property* (Γ-EP) if for any Γ formula $\varphi(x)$ that has no free variables except x, if $T \vdash \exists x \varphi(x)$, then $T \vdash \varphi(\overline{n})$ for some natural number n. Then, it is shown that for any consistent r.e. extension T of **PA**, T does not have Π_1-DP (see [13]). On the other hand, for extensions of **PA**, a similar situation to that of DP and EP in intuitionistic logic has been shown to hold. That is, it is known that **PA** has both Σ_1-DP and Σ_1-EP. Moreover, Guaspari [10] proved that Σ_1-DP, Σ_1-EP, and the Σ_1-soundness are pairwise equivalent for any consistent r.e. extension of **PA**.

In the usual proof of the incompleteness theorems, a provability predicate $\mathrm{Pr}_T(x)$, that is, a Σ_1 formula weakly representing the provability relation of a theory T plays an important role. Besides the context in which $\Box$ is intended as informal provability, a modal logical study of the notion of formalized provability has been developed by interpreting $\Box$ in terms of $\mathrm{Pr}_T(x)$. One of the important results of this study is Solovay's arithmetical completeness theorem which states that if T is Σ_1-sound, then the propositional modal logic **GL** is exactly the logic of all T-verifiable principles [22]. In this framework, MDP also makes sense. It is known that **GL** enjoys MDP. Rather than corresponding to some constructive property, this fact corresponds to the fact that if T is Σ_1-sound, then $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \vee \mathrm{Pr}_T(\ulcorner \psi \urcorner)$ implies $T \vdash \varphi$ or $T \vdash \psi$.

Our motivation for the research in the present paper is to provide a unified viewpoint on MDP and Γ-DP, which have been discussed in different contexts and frameworks. In particular, we would like to unify the arguments on $\Box$ as an informal provability and $\Box$ as a provability predicate. For this purpose, instead of fixing a modal logic such as **S4** or **GL**, we discuss the theory **PA**(L) obtained by adding an arbitrary normal modal logic L to **PA**. In particular, **K4** is a common sublogic of **S4** and **GL**, and thus an investigation for extensions of **PA**(**K4**) would be applicable to both of the two different interpretations of $\Box$. For example, we prove that for any r.e. $\mathcal{L}_A(\Box)$-theory T extending **PA**(**K4**), T has MDP if and only if T has MEP. This is a strengthening of the above mentioned form of Friedman and Sheard's result.

[1] Actually, Friedman and Sheard proved this theorem for a wider class of $\mathcal{L}_A(\Box)$-theories. This will be discussed in Remark 5.41.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

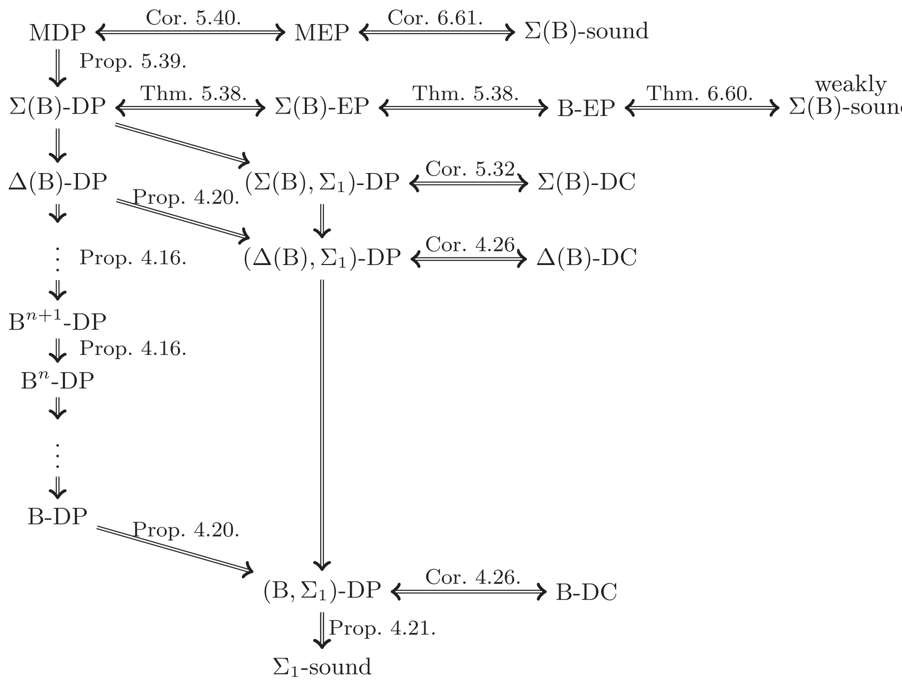

Fig. 1. Implications for consistent r.e. extensions T of $\mathbf{PA}(\mathbf{K4})$ with $T \nvdash \Box\bot$.

We would also like to analyze the possibility of applying existing methods for studying properties such as Γ-DP to modal arithmetic. In particular, as suggested by Guaspari's result, MDP and MEP may be characterized by soundness with respect to some class of $\mathcal{L}_A(\Box)$-formulas. For this reason, in the present paper, we introduce three new classes B, $\Delta(\mathrm{B})$, and $\Sigma(\mathrm{B})$ of $\mathcal{L}_A(\Box)$-formulas. Then, we prove that for any consistent r.e. $\mathcal{L}_A(\Box)$-theory T extending $\mathbf{PA}(\mathbf{K4})$, T has MDP if and only if T is $\Sigma(\mathrm{B})$-sound. We also provide a systematic analysis of the disjunction and the existence properties in modal arithmetic, including investigations of DP and EP concerning these new classes of formulas.

The present paper is organized as follows. In Section 2, we introduce several theories of modal arithmetic and show that each of them is a conservative extension of **PA**. In Section 3, we introduce three new classes B, $\Delta(\mathrm{B})$, and $\Sigma(\mathrm{B})$ of $\mathcal{L}_A(\Box)$-formulas and show some basic properties of these classes. Section 4 is devoted to the study of B-DP, $\Delta(\mathrm{B})$-DP, and related properties. In Section 5, we study $\Sigma(\mathrm{B})$-DP and related properties. In particular, we prove that for any r.e. extension T of the theory $\mathbf{PA}(\mathbf{K})$, if $T \nvdash \Box\bot$, then $\Sigma(\mathrm{B})$-DP, $\Sigma(\mathrm{B})$-EP, and B-EP are pairwise equivalent. From this result, the equivalence of MDP and MEP for any consistent r.e. $\mathcal{L}_A(\Box)$-theory extending $\mathbf{PA}(\mathbf{K4})$ is obtained. In Section 6, as generalizations of the notions of the soundness and the Σ_1-soundness of $\mathcal{L}_A$-theories, we introduce the notions of the $\mathcal{L}_A(\Box)$-soundness and the $\Sigma(\mathrm{B})$-soundness of $\mathcal{L}_A(\Box)$-theories. We study these notions precisely, and then, we prove that for any consistent r.e. extension T of $\mathbf{PA}(\mathbf{K4})$, T has MDP if and only T is $\Sigma(\mathrm{B})$-sound. This is a modal arithmetical analogue of Guaspari's theorem. Figure 1 summarizes our results obtained in Sections 4–6. We also show some non-implications between the properties: Σ_1-soundness does not imply (B, Σ_1)-DP (Proposition 4.28), $\Sigma(\mathrm{B})$-EP does not imply MDP (Proposition 6.65), and $\Sigma(\mathrm{B})$-DC does not imply B-DP (Proposition 6.66). Finally, in the last section, we list several unsolved problems.

§2. Theories of modal arithmetic. We work within the framework of modal arithmetic. The language $\mathcal{L}_A(\Box)$ of modal arithmetic consists of logical connectives $\bot, \land, \lor, \to, \neg$, quantifiers $\forall, \exists$, elements of the language $\mathcal{L}_A = \{0, S, +, \times, \leq, =\}$ of first-order arithmetic, and modal operator $\Box$. The formulas $\varphi \leftrightarrow \psi$ and $\Diamond\varphi$ are abbreviations for $(\varphi \to \psi) \land (\psi \to \varphi)$ and $\neg\Box\neg\varphi$, respectively. A set of sentences is called a *theory*. In the present paper, we always assume that the inference rules of every $\mathcal{L}_A(\Box)$-theory are modus ponens (MP) $\dfrac{\varphi \to \psi \quad \varphi}{\psi}$, generalization (Gen) $\dfrac{\varphi}{\forall x\varphi}$, and necessitation (Nec) $\dfrac{\varphi}{\Box\varphi}$. Since we will study modal arithmetic from a broader perspective than just **EA**, we also deal with $\mathcal{L}_A(\Box)$-theories obtained by adding normal modal propositional logics other than **S4** into **PA**. Let $\mathbf{PA}_\Box$ be the $\mathcal{L}_A(\Box)$-theory obtained by adding the logical axioms of first-order logic for $\mathcal{L}_A(\Box)$-formulas and the induction axioms for $\mathcal{L}_A(\Box)$-formulas into **PA**. Notice that the value of each $\mathcal{L}_A$-term $t(\vec{x})$ can be effectively computed from the input $\vec{x}$, and thus universal instantiation $\forall x\varphi(x) \to \varphi(t)$ (where t is an $\mathcal{L}_A$-term substitutable for x in φ), which is problematic in modal predicate logic, is not a problem in our framework. As in [3, 6, 19, 21], we adopt universal instantiation as an axiom scheme of $\mathbf{PA}_\Box$. Of course, this is not the case in general framework (see [21, sec. 7]).

For each normal modal propositional logic L, let $\mathbf{PA}(L)$ denote the $\mathcal{L}_A(\Box)$-theory obtained by adding universal closures of formulas corresponding to modal axioms of L into $\mathbf{PA}_\Box$. We deal with the following $\mathcal{L}_A(\Box)$-theories.

- $\mathbf{PA}(\mathbf{K}) = \mathbf{PA}_\Box + \{\forall\vec{x}(\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)) \mid \varphi, \psi \text{ are } \mathcal{L}_A(\Box)\text{-formulas}\}$;
- $\mathbf{PA}(\mathbf{K4}) = \mathbf{PA}(\mathbf{K}) + \{\forall\vec{x}(\Box\varphi \to \Box\Box\varphi) \mid \varphi \text{ is an } \mathcal{L}_A(\Box)\text{-formula}\}$;
- $\mathbf{PA}(\mathbf{KT}) = \mathbf{PA}(\mathbf{K}) + \{\forall\vec{x}(\Box\varphi \to \varphi) \mid \varphi \text{ is an } \mathcal{L}_A(\Box)\text{-formula}\}$;
- $\mathbf{PA}(\mathbf{S4}) = \mathbf{EA} = \mathbf{PA}(\mathbf{KT}) + \{\forall\vec{x}(\Box\varphi \to \Box\Box\varphi) \mid \varphi \text{ is an } \mathcal{L}_A(\Box)\text{-formula}\}$;
- $\mathbf{PA}(\mathbf{S5}) = \mathbf{PA}(\mathbf{S4}) + \{\forall\vec{x}(\Diamond\varphi \to \Box\Diamond\varphi) \mid \varphi \text{ is an } \mathcal{L}_A(\Box)\text{-formula}\}$;
- $\mathbf{PA}(\mathbf{Triv}) = \mathbf{PA}(\mathbf{K}) + \{\forall\vec{x}(\Box\varphi \leftrightarrow \varphi) \mid \varphi \text{ is an } \mathcal{L}_A(\Box)\text{-formula}\}$;
- $\mathbf{PA}(\mathbf{GL}) = \mathbf{PA}(\mathbf{K4}) + \{\forall\vec{x}(\Box(\Box\varphi \to \varphi) \to \Box\varphi) \mid \varphi \text{ is an } \mathcal{L}_A(\Box)\text{-formula}\}$;
- $\mathbf{PA}(\mathbf{Verum}) = \mathbf{PA}(\mathbf{K}) + \{\Box\bot\}$.

Interestingly, Došen [3, Lemma 7] proved that $\mathbf{PA}(\mathbf{S5})$ and $\mathbf{PA}(\mathbf{Triv})$ are deductively equivalent.

Here we discuss the principle $x = y \to (\varphi(x) \to \varphi(y))$ of identity. Our system has this principle only for atomic formulas $\varphi(x)$ as identity axioms as in the case of classical first-order logic. On the other hand, this principle for all $\mathcal{L}_A(\Box)$-formulas is not generally valid in our framework because our language has the symbol $\Box$. Shapiro [21] states that the following proposition holds for $\mathbf{PA}(\mathbf{S4})$.

Proposition 2.1.

1. $\mathbf{PA}(\mathbf{K}) \vdash x = y \to \Box x = y$.
2. *For any $\mathcal{L}_A(\Box)$-formula $\varphi(x)$,* $\mathbf{PA}(\mathbf{K}) \vdash x = y \to (\varphi(x) \to \varphi(y))$.

Proof. 1. Let $\varphi(x, y)$ be the formula $x = y \to \Box x = y$. Firstly, we prove $\mathbf{PA}_\Box \vdash \forall y\varphi(0, y)$. Since $\mathbf{PA} \vdash 0 = 0$, we have $\mathbf{PA}_\Box \vdash \Box 0 = 0$, and hence $\mathbf{PA}_\Box \vdash \varphi(0, 0)$.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

Since $\mathbf{PA} \vdash 0 \neq S(y)$, we also have $\mathbf{PA}_\Box \vdash \varphi(0, S(y))$, and thus $\mathbf{PA}_\Box \vdash \forall y(\varphi(0, y) \to \varphi(0, S(y)))$. By the induction axiom for $\varphi(0, y)$, we obtain $\mathbf{PA}_\Box \vdash \forall y\varphi(0, y)$.

Secondly, we prove $\mathbf{PA}(\mathbf{K}) \vdash \forall y\varphi(x, y) \to \forall y\varphi(S(x), y)$. Since $\mathbf{PA} \vdash S(x) \neq 0$, we have $\mathbf{PA}_\Box \vdash \varphi(S(x), 0)$. It follows from $\mathbf{PA} \vdash S(x) = S(y) \to x = y$ that $\mathbf{PA}_\Box \vdash \varphi(x, y) \wedge S(x) = S(y) \to \Box x = y$. Since $\mathbf{PA} \vdash x = y \to S(x) = S(y)$, we have $\mathbf{PA}(\mathbf{K}) \vdash \Box x = y \to \Box S(x) = S(y)$. Thus, we get

$$\mathbf{PA}(\mathbf{K}) \vdash \varphi(x, y) \wedge S(x) = S(y) \to \Box S(x) = S(y).$$

This means $\mathbf{PA}(\mathbf{K}) \vdash \varphi(x, y) \to \varphi(S(x), S(y))$. By the universal instantiation, we have $\mathbf{PA}(\mathbf{K}) \vdash \forall y\varphi(x, y) \to \varphi(S(x), S(y))$, and hence

$$\mathbf{PA}(\mathbf{K}) \vdash \forall y\varphi(x, y) \to \forall y(\varphi(S(x), y) \to \varphi(S(x), S(y))).$$

From this with $\mathbf{PA}_\Box \vdash \varphi(S(x), 0)$, we obtain $\mathbf{PA}(\mathbf{K}) \vdash \forall y\varphi(x, y) \to \forall y\varphi(S(x), y)$ by the induction axiom for $\varphi(S(x), y)$.

Finally, by the induction axiom for $\forall y\varphi(x, y)$, we conclude $\mathbf{PA}(\mathbf{K}) \vdash \forall x\forall y\varphi(x, y)$.

2. This is proved by induction on the construction of $\varphi(x)$. We only prove the case that $\varphi(x)$ is of the form $\Box\psi(x)$ and the statement holds for $\psi(x)$. By the induction hypothesis, $\mathbf{PA}(\mathbf{K}) \vdash x = y \to (\psi(x) \to \psi(y))$. Then, $\mathbf{PA}(\mathbf{K})$ proves $\Box x = y \to (\Box\psi(x) \to \Box\psi(y))$. By combining this with Clause 1, we conclude $\mathbf{PA}(\mathbf{K}) \vdash x = y \to (\Box\psi(x) \to \Box\psi(y))$. □

We say that a theory T is a *subtheory* of a theory U, $U \vdash T$, if every axiom of T is provable in U. Makinson's theorem [17] states that every consistent normal modal propositional logic L is a sublogic of **Triv** or **Verum** (see also [12]). Hence, every $\mathcal{L}_A(\Box)$-theory of the form $\mathbf{PA}(L)$ for some consistent normal propositional modal logic L is a subtheory of $\mathbf{PA}(\mathbf{Triv})$ or $\mathbf{PA}(\mathbf{Verum})$. We prove that every such logic is a conservative extension of **PA**.

First, we prove that $\mathbf{PA}(\mathbf{Triv})$ is a conservative extension of **PA**. In order to prove this, we introduce a translation α of $\mathcal{L}_A(\Box)$-formulas into $\mathcal{L}_A$-formulas.

Definition 2.2 (α-translation). *We define a translation α of $\mathcal{L}_A(\Box)$-formulas into $\mathcal{L}_A$-formulas inductively as follows*:

1. *If φ is an $\mathcal{L}_A$-formula, then $\alpha(\varphi) :\equiv \varphi$.*
2. *α preserves logical connectives and quantifiers.*
3. $\alpha(\Box\varphi) :\equiv \alpha(\varphi)$.

It is obvious that for any $\mathcal{L}_A(\Box)$-formula φ, $\mathbf{PA}(\mathbf{Triv}) \vdash \varphi \leftrightarrow \alpha(\varphi)$. Moreover:

Proposition 2.3. *For any $\mathcal{L}_A(\Box)$-formula φ, if $\mathbf{PA}(\mathbf{Triv}) \vdash \varphi$, then $\mathbf{PA} \vdash \alpha(\varphi)$.*

Proof. We prove the proposition by induction on the length of proofs of φ in $\mathbf{PA}(\mathbf{Triv})$.

- If φ is an axiom of **PA**, then $\alpha(\varphi) \equiv \varphi$ and $\mathbf{PA} \vdash \alpha(\varphi)$.
- If φ is a logical axiom, then so is $\alpha(\varphi)$, and it is **PA**-provable.
- If φ is an induction axiom in the language $\mathcal{L}_A(\Box)$, then $\alpha(\varphi)$ is also an induction axiom in $\mathcal{L}_A$, and so $\mathbf{PA} \vdash \alpha(\varphi)$.
- If φ is $\forall\vec{x}(\Box(\psi \to \sigma) \to (\Box\psi \to \Box\sigma))$, then $\alpha(\varphi)$ is the **PA**-provable sentence $\forall\vec{x}\big((\alpha(\psi) \to \alpha(\sigma)) \to (\alpha(\psi) \to \alpha(\sigma))\big)$.
- If φ is $\forall\vec{x}(\psi \leftrightarrow \Box\psi)$, then $\alpha(\varphi)$ is $\forall\vec{x}(\alpha(\psi) \leftrightarrow \alpha(\psi))$. This is provable in **PA**.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

- If φ is derived from ψ and $\psi \to \varphi$ by MP, then by the induction hypothesis, $\mathbf{PA} \vdash \alpha(\psi)$ and $\mathbf{PA} \vdash \alpha(\psi) \to \alpha(\varphi)$, and hence $\mathbf{PA} \vdash \alpha(\varphi)$.
- If φ is derived from $\psi(x)$ by Gen, then $\varphi \equiv \forall x \psi(x)$. By the induction hypothesis, $\mathbf{PA} \vdash \alpha(\psi(x))$ and hence $\mathbf{PA} \vdash \forall x \alpha(\psi(x))$. Therefore, $\mathbf{PA} \vdash \alpha(\varphi)$.
- If φ is derived from ψ by Nec, then $\varphi \equiv \Box\psi$. By the induction hypothesis, $\mathbf{PA} \vdash \alpha(\psi)$. Since $\alpha(\varphi) \equiv \alpha(\psi)$, we have $\mathbf{PA} \vdash \alpha(\varphi)$. □

Let $\mathbb{N}$ be the standard model of arithmetic in the language $\mathcal{L}_A$. We say that an $\mathcal{L}_A(\Box)$-theory T is $\mathcal{L}_A$-*sound* if for any $\mathcal{L}_A$-sentence φ, $\mathbb{N} \models \varphi$ whenever $T \vdash \varphi$.

Corollary 2.4. $\mathbf{PA}(\mathbf{Triv})$ *is a conservative extension of* $\mathbf{PA}$. *In particular,* $\mathbf{PA}(\mathbf{Triv})$ *is* $\mathcal{L}_A$*-sound.*

Proof. Let φ be any $\mathcal{L}_A$-sentence such that $\mathbf{PA}(\mathbf{Triv}) \vdash \varphi$. By Proposition 2.3, $\mathbf{PA} \vdash \alpha(\varphi)$. Since $\alpha(\varphi) \equiv \varphi$, $\mathbf{PA} \vdash \varphi$. Furthermore, by the $\mathcal{L}_A$-soundness of $\mathbf{PA}$, $\mathbf{PA}(\mathbf{Triv})$ is also $\mathcal{L}_A$-sound. □

Next, we prove that $\mathbf{PA}(\mathbf{Verum})$ is a conservative extension of $\mathbf{PA}$. We also introduce another translation β.

Definition 2.5 (β-translation). *We define a translation β of $\mathcal{L}_A(\Box)$-formulas into $\mathcal{L}_A$-formulas inductively as follows*:

1. *If φ is an $\mathcal{L}_A$-formula, then* $\beta(\varphi) :\equiv \varphi$.
2. *β preserves logical connectives and quantifiers.*
3. $\beta(\Box\varphi) :\equiv 0 = 0$.

As in the case of α, for any $\mathcal{L}_A(\Box)$-formula φ, $\mathbf{PA}(\mathbf{Verum}) \vdash \varphi \leftrightarrow \beta(\varphi)$. Moreover:

Proposition 2.6. *For any $\mathcal{L}_A(\Box)$-formula φ, if* $\mathbf{PA}(\mathbf{Verum}) \vdash \varphi$, *then* $\mathbf{PA} \vdash \beta(\varphi)$.

Proof. As in the proof of Proposition 2.3, this proposition is proved by induction on the length of proofs of φ in $\mathbf{PA}(\mathbf{Verum})$. We only give proofs of the following three cases:

- If φ is $\forall \vec{x}(\Box(\psi \to \sigma) \to (\Box\psi \to \Box\sigma))$, then $\beta(\varphi)$ is the $\mathbf{PA}$-provable sentence $\forall \vec{x}(0 = 0 \to (0 = 0 \to 0 = 0))$.
- If φ is $\Box\bot$, then $\beta(\varphi)$ is the $\mathbf{PA}$-provable sentence $0 = 0$.
- If φ is derived from ψ by Nec, then $\varphi \equiv \Box\psi$. Since $\beta(\varphi) \equiv 0 = 0$, this is $\mathbf{PA}$-provable. □

Corollary 2.7. $\mathbf{PA}(\mathbf{Verum})$ *is a conservative extension of* $\mathbf{PA}$. *In particular,* $\mathbf{PA}(\mathbf{Verum})$ *is* $\mathcal{L}_A$*-sound.*

We close this section by showing that the notion of Σ_1 formulas has a high affinity with modal arithmetic. The following theorem is proved by applying a schematic proof of formalized Σ_1-completeness theorem (see [2, 16, 18]). This is also implicitly stated in [6].

Theorem 2.8 (Formalized Σ_1-completeness theorem). *For any Σ_1 formula φ, we have* $\mathbf{PA}(\mathbf{K}) \vdash \varphi \to \Box\varphi$.

Proof. Before proving the theorem, we show that for any Σ_1 formula ψ, there exists a Σ_1 formula ψ' such that $\mathbf{PA} \vdash \psi \leftrightarrow \psi'$, ψ' does not contain the connectives $\neg$ and $\to$, and every atomic formula contained in ψ' is of the form $t_1 = t_2$ for some $\mathcal{L}_A$-terms

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

t_1 and t_2. First, we easily find a Σ_1 formula ψ_0 without the connective $\to$ such that ψ_0 is logically equivalent to ψ and every negation symbol $\neg$ in ψ_0 is applied to an atomic formula. Then, by replacing every negated atomic formula $\neg(t_1 = t_2)$ or $\neg(t_1 < t_2)$ of ψ_0 by $t_1 < t_2 \vee t_2 < t_1$ or $t_1 = t_2 \vee t_2 < t_1$ respectively, we obtain a **PA**-equivalent Σ_1 formula ψ_1 without having $\neg$. Finally, by replacing every atomic formula $t_1 < t_2$ of ψ_1 by $\exists y(t_1 + S(y) = t_2)$, we obtain a required equivalent Σ_1 formula ψ'. Since $\mathbf{PA}(\mathbf{K}) \vdash \Box\psi \leftrightarrow \Box\psi'$, to prove the theorem, it suffices to show that $\mathbf{PA}(\mathbf{K}) \vdash \varphi \to \Box\varphi$ for any Σ_1 formula φ such that it does not contain the connectives $\neg$ and $\to$, and that every atomic formula contained in φ is of the form $t_1 = t_2$ for some $\mathcal{L}_A$-terms t_1 and t_2.

This theorem is proved by induction on the construction of φ. If φ is $t_1 = t_2$, then $\mathbf{PA}(\mathbf{K}) \vdash t_1 = t_2 \to \Box t_1 = t_2$ follows from $\mathbf{PA}(\mathbf{K}) \vdash x = y \to \Box x = y$ (Proposition 2.1) by substituting t_1 and t_2 into x and y, respectively. The cases for $\wedge$, $\vee$, $\forall x < t$ and $\exists$ are proved as in the proof of Theorem 3.13. ☐

§3. Classes of $\mathcal{L}_A(\Box)$-formulas. In first-order arithmetic, it is important to classify $\mathcal{L}_A$-formulas according to the arithmetic hierarchy. In this section, we introduce three classes B, $\Delta(\mathrm{B})$, and $\Sigma(\mathrm{B})$ of $\mathcal{L}_A(\Box)$-formulas, and investigate basic properties of formulas in these classes. Our classes $\Delta(\mathrm{B})$ and $\Sigma(\mathrm{B})$ are modal arithmetical counterparts of Δ_0 and Σ_1, respectively.

Definition 3.9 (B, $\Delta(\mathrm{B})$, and $\Sigma(\mathrm{B})$).

- *Let* B *be the class of all $\mathcal{L}_A(\Box)$-formulas of the form $\Box\varphi$.*
- *Let $\Delta(\mathrm{B})$ be the smallest class of $\mathcal{L}_A(\Box)$-formulas satisfying the following conditions*:
    1. $\Delta_0 \cup \mathrm{B} \subseteq \Delta(\mathrm{B})$.
    2. *If φ and ψ are in $\Delta(\mathrm{B})$, then so are $\varphi \wedge \psi$, $\varphi \vee \psi$, $\forall x < t\, \varphi$ and $\exists x < t\, \varphi$, where t is an $\mathcal{L}_A$-term in which x does not occur.*
- *Let $\Sigma(\mathrm{B})$ be the smallest class of $\mathcal{L}_A(\Box)$-formulas satisfying the following conditions*:
    1. $\Sigma_1 \cup \mathrm{B} \subseteq \Sigma(\mathrm{B})$.
    2. *If φ and ψ are in $\Sigma(\mathrm{B})$, then so are $\varphi \wedge \psi$, $\varphi \vee \psi$, $\exists x\varphi$ and $\forall x < t\, \varphi$, where t is an $\mathcal{L}_A$-term in which x does not occur.*

We emphasize here that in some sense the class $\Sigma(\mathrm{B})$ is a natural extension of the class Σ_1. For each r.e. theory T, let $\mathrm{Pr}_T(x)$ be a fixed Σ_1 provability predicate of T. In the context of interpreting $\Box$ by $\mathrm{Pr}_T(x)$, each $\mathcal{L}_A(\Box)$-formula of the form $\Box\varphi$ is interpreted by a Σ_1 formula, and hence every $\Sigma(\mathrm{B})$ formula is also recognized as a Σ_1 formula. From this perspective, we will attempt to extend the properties possessed by Σ_1 formulas in first-order arithmetic to $\Sigma(\mathrm{B})$ formulas in modal arithmetic. Note, however, that $\Delta(\mathrm{B})$, unlike Δ_0, is not closed under taking negation and implication. For example, it can be shown that there is no $\Delta(\mathrm{B})$ sentence φ such that $\mathbf{PA}(\mathbf{K}) \vdash \neg\Box\bot \leftrightarrow \varphi$ (see Corollary 3.12).

The following proposition states that the relationship between $\Sigma(\mathrm{B})$ and $\Delta(\mathrm{B})$ is similar to the relationship between Σ_1 and Δ_0.

Proposition 3.10. *For any $\Sigma(\mathrm{B})$ formula φ, there exist a variable v and a $\Delta(\mathrm{B})$ formula ψ such that $\mathbf{PA}_\Box \vdash \varphi \leftrightarrow \exists v\psi$.*

Proof. We prove the proposition by induction on the construction of φ.

- If φ is Σ_1, then there exists a Δ_0 formula ψ such that $\mathbf{PA} \vdash \varphi \leftrightarrow \exists v\psi$.
- If φ is of the form $\Box\varphi_0$, then $\varphi \in \Delta(\mathbf{B})$ and $\mathbf{PA}_\Box \vdash \varphi \leftrightarrow \exists v\varphi$ for a variable v not contained in φ.
- Let $\circ \in \{\wedge, \vee\}$. If φ is of the form $\varphi_0 \circ \varphi_1$, then by the induction hypothesis, there exist distinct variables v_0 and v_1 and $\Delta(\mathbf{B})$ formulas ψ_0 and ψ_1 such that $\mathbf{PA}_\Box$ proves $\varphi_0 \leftrightarrow \exists v_0\psi_0$ and $\varphi_1 \leftrightarrow \exists v_1\psi_1$. Let v be any variable that does not occur in ψ_0 or ψ_1, and is not v_0 or v_1. Then, $\mathbf{PA}_\Box$ proves the equivalence $\varphi \leftrightarrow \exists v\, \exists v_0 < v\, \exists v_1 < v\, (\psi_0 \circ \psi_1)$.
- If φ is of the form $\exists x\varphi_0$, then by the induction hypothesis, there exist a variable v_0 and a $\Delta(\mathbf{B})$ formula ψ_0 such that $\mathbf{PA} \vdash \varphi_0 \leftrightarrow \exists v_0\psi_0$. Let v be any variable not contained in ψ_0 and is not v_0 or x. Then, $\mathbf{PA}_\Box$ proves the equivalence $\varphi \leftrightarrow \exists v\, \exists x < v\, \exists v_0 < v\, \psi_0$.
- The case that φ is of the form $\exists x < t\, \varphi_0$, where t is an $\mathcal{L}_A$-term in which x does not occur is proved as in the proof of the case of $\exists$.
- Suppose φ is of the form $\forall x < t\, \varphi_0$, where t is an $\mathcal{L}_A$-term in which x does not occur. By the induction hypothesis, there exists a variable v_0 and a $\Delta(\mathbf{B})$ formula ψ_0 such that $\mathbf{PA}_\Box \vdash \varphi_0 \leftrightarrow \exists v_0\psi_0$. Then, $\mathbf{PA}_\Box \vdash \varphi \leftrightarrow \forall x < t\, \exists v_0\psi_0$. By the collection principle for $\mathcal{L}_A(\Box)$-formulas derived from the induction axioms for $\mathcal{L}_A(\Box)$-formulas, we obtain

$$\mathbf{PA}_\Box \vdash \varphi \leftrightarrow \exists v\, \forall x < t\, \exists v_0 < v\, \psi_0$$

for some appropriate variable v. $\Box$

PROPOSITION 3.11. *For any $\Delta(\mathbf{B})$ sentence φ, there exist a natural number k and sentences $\psi_0, \dots, \psi_{k-1}$ such that $\mathbf{PA}(\mathbf{K}) \vdash \varphi \leftrightarrow \bigvee_{i<k} \Box\psi_i$. Here $\bigvee_{i<0} \Box\psi_i$ denotes $\bot$.*

Proof. We prove the proposition by induction on the construction of φ.

- If φ is a Δ_0 sentence, then either $\mathbf{PA} \vdash \varphi$ or $\mathbf{PA} \vdash \neg\varphi$. Thus, $\mathbf{PA}(\mathbf{K}) \vdash \varphi \leftrightarrow \Box 0 = 0$ or $\mathbf{PA}(\mathbf{K}) \vdash \varphi \leftrightarrow \bot$.
- If φ is of the form $\Box\psi$, then the statement is trivial.
- If φ is of the form $\psi \wedge \sigma$, then there exist sentences $\xi_0, \dots, \xi_{k-1}, \eta_0, \dots, \eta_{l-1}$ such that $\mathbf{PA}(\mathbf{K}) \vdash \psi \leftrightarrow \bigvee_{i<k} \Box\xi_i$ and $\mathbf{PA}(\mathbf{K}) \vdash \sigma \leftrightarrow \bigvee_{j<l} \Box\eta_j$. Then, $\mathbf{PA}(\mathbf{K}) \vdash \varphi \leftrightarrow \bigvee_{i<k} \bigvee_{j<l} \Box(\xi_i \wedge \eta_j)$.
- If φ is of the form $\psi \vee \sigma$, then the statement is obvious by the induction hypothesis.
- If φ is of the form $\exists y < t\, \psi(y)$ for some $\mathcal{L}_A$-term t, then t is a closed term because φ is a sentence. Let m be the value of t, then $\mathbf{PA}(\mathbf{K}) \vdash \varphi \leftrightarrow \bigvee_{i<m} \psi(\bar{i})$. Then, the statement holds by the induction hypothesis.
- If φ is of the form $\forall y < t\, \psi(y)$ for some closed $\mathcal{L}_A$-term t, then for the value m of the term t, $\mathbf{PA}(\mathbf{K}) \vdash \varphi \leftrightarrow \bigwedge_{i<m} \psi(\bar{i})$. We can prove the statement by the induction hypothesis as in the proof of the case $\wedge$. $\Box$

COROLLARY 3.12. *Let T be a theory extending $\mathbf{PA}(\mathbf{K})$ such that $T \nvdash \Box\bot$ and $T \nvdash \neg\Box\bot$. Then, there is no $\Delta(\mathbf{B})$ sentence φ such that $T \vdash \neg\Box\bot \leftrightarrow \varphi$.*

Proof. Suppose, towards a contradiction, that φ is a $\Delta(\mathbf{B})$ sentence such that $T \vdash \neg\Box\bot \leftrightarrow \varphi$. By Proposition 3.11, there exist k and $\psi_0, \dots, \psi_{k-1}$ such that $\mathbf{PA}(\mathbf{K}) \vdash \varphi \leftrightarrow \bigvee_{i<k} \Box\psi_i$. Then, $T \vdash \neg\Box\bot \leftrightarrow \bigvee_{i<k} \Box\psi_i$. Since $T \nvdash \Box\bot$, we get $k > 0$. Then,

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

$T \vdash \Box\psi_0 \to \neg\Box\bot$. On the other hand, $T \vdash \neg\Box\psi_0 \to \neg\Box\bot$ because T is an extension of $\mathbf{PA}(\mathbf{K})$. Therefore, we obtain $T \vdash \neg\Box\bot$. This is a contradiction. $\Box$

We naturally extend Theorem 2.8 into the framework of modal arithmetic.

Theorem 3.13 (Formalized $\Sigma(\mathrm{B})$-completeness theorem). *For any* $\varphi \in \Sigma(\mathrm{B})$, $\mathbf{PA}(\mathbf{K4}) \vdash \varphi \to \Box\varphi$.

Proof. We prove the theorem by induction on the construction of φ.

- If φ is a Σ_1 formula, then $\mathbf{PA}(\mathbf{K}) \vdash \varphi \to \Box\varphi$ by Theorem 2.8.
- If φ is of the form $\Box\psi$, then $\mathbf{PA}(\mathbf{K4}) \vdash \varphi \to \Box\varphi$.
- If φ is $\psi \wedge \sigma$, then by the induction hypothesis, $\mathbf{PA}(\mathbf{K4}) \vdash \varphi \to \Box\psi \wedge \Box\sigma$. We have $\mathbf{PA}(\mathbf{K4}) \vdash \varphi \to \Box\varphi$.
- If φ is $\psi \vee \sigma$, then by the induction hypothesis, $\mathbf{PA}(\mathbf{K4})$ proves $\psi \to \Box(\psi \vee \sigma)$ and $\sigma \to \Box(\psi \vee \sigma)$. Hence, $\mathbf{PA}(\mathbf{K4}) \vdash \varphi \to \Box\varphi$.
- Suppose that φ is $\exists x\psi$. Since $\mathbf{PA}(\mathbf{K}) \vdash \psi \to \exists x\psi$, we have $\mathbf{PA}(\mathbf{K}) \vdash \Box\psi \to \Box\exists x\psi$. By the induction hypothesis, $\mathbf{PA}(\mathbf{K4}) \vdash \psi \to \Box\psi$. Thus, $\mathbf{PA}(\mathbf{K4}) \vdash \psi \to \Box\exists x\psi$. Then, we obtain $\mathbf{PA}(\mathbf{K4}) \vdash \varphi \to \Box\varphi$.
- Before proving the case that φ is of the form $\forall x < t\, \psi(x)$ generally, we prove the restricted case that t is some variable y not occurring in ψ. Suppose that $\varphi(y)$ is of the form $\forall x < y\, \psi(x)$ for some variable y not occurring in $\psi(x)$. Let $\xi(y)$ be the formula $\varphi(y) \to \Box\varphi(y)$, and then we prove $\mathbf{PA}(\mathbf{K}) \vdash \forall y\xi(y)$ by using the induction axiom.

  For the base step, since trivially $\mathbf{PA}(\mathbf{K}) \vdash \forall x < 0\, \psi(x)$, we have $\mathbf{PA}(\mathbf{K}) \vdash \Box\varphi(0)$ and hence $\mathbf{PA}(\mathbf{K}) \vdash \xi(0)$.

  For the induction step, since

  $$\mathbf{PA}(\mathbf{K}) \vdash [\forall x < S(y)\, \psi(x)] \leftrightarrow [(\forall x < y\, \psi(x)) \wedge \psi(y)],$$

  we have

  $$\mathbf{PA}(\mathbf{K}) \vdash \varphi(S(y)) \leftrightarrow [\varphi(y) \wedge \psi(y)]. \tag{1}$$

  By the induction hypothesis, $\mathbf{PA}(\mathbf{K4}) \vdash \psi(y) \to \Box\psi(y)$. By combining this with (1) and the definition of $\xi(y)$,

  $$\mathbf{PA}(\mathbf{K4}) \vdash \xi(y) \wedge \varphi(S(y)) \to \Box\varphi(y) \wedge \Box\psi(y).$$

  Then, by (1) again,

  $$\mathbf{PA}(\mathbf{K4}) \vdash \xi(y) \wedge \varphi(S(y)) \to \Box\varphi(S(y)).$$

  Equivalently, $\mathbf{PA}(\mathbf{K4}) \vdash \xi(y) \to \xi(S(y))$.

  Therefore, by the induction axiom, we conclude $\mathbf{PA}(\mathbf{K4}) \vdash \forall y\xi(y)$.

  Finally, suppose that φ is of the form $\forall x < t\, \psi$ for some $\mathcal{L}_A$-term t. We have already proved that $\mathbf{PA}(\mathbf{K4})$ proves $\forall x < y\, \psi \to \Box\forall x < y\, \psi$ for some variable y not occurring in ψ. By substituting t for y in this formula, we obtain $\mathbf{PA}(\mathbf{K4}) \vdash \varphi \to \Box\varphi$. $\Box$

Corollary 3.14 ($\Sigma(\mathrm{B})$-deduction theorem). *Let T be any extension of* $\mathbf{PA}(\mathbf{K4})$ *and let X be any set of* $\Sigma(\mathrm{B})$ *sentences. Then, for any* $\mathcal{L}_A(\Box)$*-formula φ, if $T + X \vdash \varphi$, then there exist $\sigma_0, \ldots, \sigma_{k-1} \in X$ such that $T \vdash \sigma_0 \wedge \cdots \wedge \sigma_{k-1} \to \varphi$.*

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

Proof. This is proved by induction on the length of a proof of φ in $T + X$. We only give a proof of the case that φ is derived from ψ by the rule NEC. Then, φ is of the form $\Box\psi$. By the induction hypothesis, $T \vdash \sigma_0 \wedge \cdots \wedge \sigma_{k-1} \to \psi$ for some $\sigma_0, \dots, \sigma_{k-1} \in X$. Then, $T \vdash \Box\sigma_0 \wedge \cdots \wedge \Box\sigma_{k-1} \to \Box\psi$. By Theorem 3.13, $T \vdash \sigma_0 \wedge \cdots \wedge \sigma_{k-1} \to \Box\psi$. $\Box$

§4. B-DP, $\Delta(\mathrm{B})$-DP and related properties.

We introduce several versions of the partial disjunction property.

DEFINITION 4.15. *Let T be a theory and let Γ and Θ be classes of formulas.*

- *T is said to have the* modal disjunction property (MDP) *if for any sentences φ and ψ, if $T \vdash \Box\varphi \vee \Box\psi$, then $T \vdash \varphi$ or $T \vdash \psi$.*
- *T is said to have the* modal existence property (MEP) *if for any formula $\varphi(x)$ that has no free variables except x, if $T \vdash \exists x \Box\varphi(x)$, then for some natural number n, $T \vdash \varphi(\overline{n})$.*
- *T is said to have the* Γ-disjunction property (Γ-DP) *if for any Γ sentences φ and ψ, if $T \vdash \varphi \vee \psi$, then $T \vdash \varphi$ or $T \vdash \psi$.*
- *T is said to have the* Γ-existence property (Γ-EP) *if for any Γ formula $\varphi(x)$ that has no free variables except x, if $T \vdash \exists x \varphi(x)$, then for some natural number n, $T \vdash \varphi(\overline{n})$.*
- *T is said to have the* (Γ, Θ)-disjunction property ((Γ, Θ)-DP) *if for any Γ sentence φ and any Θ sentence ψ, if $T \vdash \varphi \vee \psi$, then $T \vdash \varphi$ or $T \vdash \psi$.*
- *For $n \geq 2$, T is said to have the* n-fold B-disjunction property (B^n-DP) *if for any $\mathcal{L}_A(\Box)$-sentences $\varphi_1, \dots, \varphi_n$, if $T \vdash \Box\varphi_1 \vee \cdots \vee \Box\varphi_n$, then $T \vdash \Box\varphi_i$ for some i $(1 \leq i \leq n)$.*
- *If T is r.e., then T is said to be* Γ-disjunctively correct (Γ-DC) *if for any Γ sentence φ, if $T \vdash \varphi \vee \mathrm{Pr}_T(\ulcorner\varphi\urcorner)$, then $T \vdash \varphi$.*
- *We say that T is closed under the* box elimination rule *if for any sentence φ, if $T \vdash \Box\varphi$, then $T \vdash \varphi$.*

Here $\mathrm{Pr}_T(x)$ is a fixed natural Σ_1 provability predicate of T. We also fix a primitive recursive proof predicate $\mathrm{Prf}_T(x, y)$ of T saying that y encodes a T-proof of x, whose existence is guaranteed by Craig's trick.

Of course, (Γ, Γ)-DP and B^2-DP are exactly Γ-DP and B-DP, respectively. The notion of Γ-DC was introduced in [15]. It is known that for any consistent r.e. extension T of **PA**, T is Σ_1-DC if and only if T is Σ_1-sound (cf. [15]).

PROPOSITION 4.16. *Let T be any extension of* $\mathbf{PA}(\mathbf{K})$.

1. *For any $n \geq 2$, if T has B^{n+1}-DP, then T also has B^n-DP.*
2. *T has B^n-DP for all $n \geq 2$ if and only if T has $\Delta(\mathrm{B})$-DP.*

Proof. 1. Let $\varphi_1, \dots, \varphi_n$ be any sentences such that $T \vdash \Box\varphi_1 \vee \cdots \vee \Box\varphi_n$. Then, $T \vdash \Box\varphi_1 \vee \cdots \vee \Box\varphi_n \vee \Box\varphi_n$. By B^{n+1}-DP, for some i $(1 \leq i \leq n)$, we have $T \vdash \Box\varphi_i$.

2. $(\Rightarrow)$: Let φ and ψ be any $\Delta(\mathrm{B})$ sentences such that $T \vdash \varphi \vee \psi$. By Proposition 3.11, there exist sentences $\varphi_0, \dots, \varphi_{k-1}$ and $\psi_0, \dots, \psi_{l-1}$ such that $T \vdash \varphi \leftrightarrow \bigvee_{i<k} \Box\varphi_i$ and $T \vdash \psi \leftrightarrow \bigvee_{j<l} \Box\psi_j$. Then, $T \vdash \bigvee_{i<k} \Box\varphi_i \vee \bigvee_{j<l} \Box\psi_j$. If $k = 0$ or $l = 0$, then we easily obtain $T \vdash \varphi$ or $T \vdash \psi$. Thus, we may assume both k and l are larger than 0. Then, $k + l \geq 2$. By B^{k+l}-DP, there exists $i < k$ or $j < l$ such that $T \vdash \Box\varphi_i$ or $T \vdash \Box\psi_j$. Then, we obtain that $T \vdash \varphi$ or $T \vdash \psi$.

($\Leftarrow$): We prove this implication by induction on $n \geq 2$. Since $\mathrm{B} \subseteq \Delta(\mathrm{B})$, T has B^2-DP. Suppose that T has B^n-DP and we would like to prove that T also has B^{n+1}-DP. Let $\varphi_1, \dots, \varphi_n, \varphi_{n+1}$ be any sentences such that $T \vdash \Box\varphi_1 \vee \cdots \vee \Box\varphi_n \vee \Box\varphi_{n+1}$. Since both $\Box\varphi_1 \vee \cdots \vee \Box\varphi_n$ and $\Box\varphi_{n+1}$ are $\Delta(\mathrm{B})$ sentences, we have $T \vdash \Box\varphi_1 \vee \cdots \vee \Box\varphi_n$ or $T \vdash \Box\varphi_{n+1}$ by $\Delta(\mathrm{B})$-DP. In the former case, we obtain $T \vdash \Box\varphi_i$ for some i $(1 \leq i \leq n)$ by the induction hypothesis. We have proved that T has B^{n+1}-DP. $\Box$

The following proposition is immediate from the definitions.

PROPOSITION 4.17. *Let T be any $\mathcal{L}_A(\Box)$-theory.*

1. *T has* MDP *if and only if T has* B-DP *and is closed under the box elimination rule.*
2. *T has* MEP *if and only if T has* B-EP *and is closed under the box elimination rule.*

We show that each existence property yields the corresponding disjunction property.

PROPOSITION 4.18. *Let T be any $\mathcal{L}_A(\Box)$-theory.*

1. *If T is an extension of* $\mathbf{PA}(\mathbf{K})$ *and T has* MEP (*resp.* B-EP), *then T has* MDP (*resp.* B-DP).
2. *If T has* $\Delta(\mathrm{B})$-EP (*resp.* $\Sigma(\mathrm{B})$-EP), *then T has* $\Delta(\mathrm{B})$-DP (*resp.* $\Sigma(\mathrm{B})$-DP).

Proof. We only give a proof of Clause 1 for MEP and MDP. Let φ and ψ be any sentences such that $T \vdash \Box\varphi \vee \Box\psi$. Then, $T \vdash \exists x \Box\big((x = 0 \wedge \varphi) \vee (x \neq 0 \wedge \psi)\big)$. By MEP, there exists a natural number n such that $T \vdash (\overline{n} = 0 \wedge \varphi) \vee (\overline{n} \neq 0 \wedge \psi)$. If $n = 0$, $T \vdash \varphi$; if $n \neq 0$, $T \vdash \psi$. Therefore, T has MDP. $\Box$

In the literature so far, modal disjunction and existence properties in modal arithmetic have been considered only for theories which are closed under the box elimination rule. As shown in Proposition 4.17, if T is closed under the box elimination rule, then MDP and B-DP are equivalent. Hence, MDP and B-DP have often been identified in the literature. Since the present paper also deals with theories that are not necessarily closed under the box elimination rule, we distinguish between MDP and B-DP. In fact, as Figure 1 shows, there seems to be a large gap between the strength of these properties.

We explore nontrivial implications between $\Delta(\mathrm{B})$-DP, $(\Delta(\mathrm{B}), \Sigma_1)$-DP, $\Delta(\mathrm{B})$-DC, B-DP, (B, Σ_1)-DP, and B-DC.

LEMMA 4.19. *Let T be any r.e. extension of* $\mathbf{PA}(\mathbf{K})$ *having* B^{n+1}-DP. *Then, for any $\mathcal{L}_A(\Box)$-sentences $\varphi_1, \dots, \varphi_n$ and Σ_1 sentence σ, if $T \vdash \Box\varphi_1 \vee \cdots \vee \Box\varphi_n \vee \sigma$, then $T \vdash \Box\varphi_i$ for some i $(1 \leq i \leq n)$ or $T \vdash \sigma$.*

Proof. Suppose $T \vdash \Box\varphi_1 \vee \cdots \vee \Box\varphi_n \vee \sigma$ and $T \nvdash \sigma$, and we would like to show $T \vdash \Box\varphi_i$ for some i. We may assume that σ is of the form $\exists x \delta(x)$ for some Δ_0 formula $\delta(x)$. Then, $\mathbb{N} \models \forall x \neg\delta(x)$ because $T \nvdash \sigma$. By the Fixed Point Lemma, for each i with $1 \leq i \leq n$, let ψ_i^0 and ψ_i^1 be Σ_1 sentences satisfying the following equivalences:

- $\mathbf{PA} \vdash \psi_i^0 \leftrightarrow \exists x\Big(\big(\delta(x) \vee \mathrm{Prf}_T(\ulcorner \Box\psi_i^1 \urcorner, x)\big) \wedge \forall y < x\, \neg\mathrm{Prf}_T(\ulcorner \Box(\varphi_i \vee \psi_i^0) \urcorner, y)\Big)$,
- $\mathbf{PA} \vdash \psi_i^1$
$\leftrightarrow \exists y\Big(\mathrm{Prf}_T(\ulcorner \Box(\varphi_i \vee \psi_i^0) \urcorner, y) \wedge \forall x \leq y\, \big(\neg\delta(x) \wedge \neg\mathrm{Prf}_T(\ulcorner \Box\psi_i^1 \urcorner, x)\big)\Big)$.

Then, for each i, we get $\mathbf{PA} \vdash \sigma \to \psi_i^0 \vee \psi_i^1$. Hence, we have

$$\mathbf{PA} \vdash \sigma \to \psi_1^0 \vee \cdots \vee \psi_n^0 \vee (\psi_1^1 \wedge \cdots \wedge \psi_n^1).$$

By Theorem 2.8,

$$\mathbf{PA}(\mathbf{K}) \vdash \sigma \to \Box\psi_1^0 \vee \cdots \vee \Box\psi_n^0 \vee \Box(\psi_1^1 \wedge \cdots \wedge \psi_n^1).$$

Hence,

$$\mathbf{PA}(\mathbf{K}) \vdash \sigma \to \Box(\varphi_1 \vee \psi_1^0) \vee \cdots \vee \Box(\varphi_n \vee \psi_n^0) \vee \Box(\psi_1^1 \wedge \cdots \wedge \psi_n^1). \qquad (2)$$

On the other hand, for each i, we have $\mathbf{PA}(\mathbf{K}) \vdash \Box\varphi_i \to \Box(\varphi_i \vee \psi_i^0)$. From our supposition, we obtain

$$T \vdash \Box(\varphi_1 \vee \psi_1^0) \vee \cdots \vee \Box(\varphi_n \vee \psi_n^0) \vee \sigma.$$

By combining this with (2),

$$T \vdash \Box(\varphi_1 \vee \psi_1^0) \vee \cdots \vee \Box(\varphi_n \vee \psi_n^0) \vee \Box(\psi_1^1 \wedge \cdots \wedge \psi_n^1).$$

By B^{n+1}-DP, we have $T \vdash \Box(\varphi_i \vee \psi_i^0)$ for some i or $T \vdash \Box(\psi_1^1 \wedge \cdots \wedge \psi_n^1)$. If $T \vdash \Box(\psi_1^1 \wedge \cdots \wedge \psi_n^1)$, then $T \vdash \Box\psi_i^1$ for each i.

- If $T \vdash \Box(\varphi_i \vee \psi_i^0)$ and $T \nvdash \Box\psi_i^1$, then $\mathbb{N} \models \psi_i^1$ by the choice of ψ_i^1 because $\mathbb{N} \models \forall x \neg\delta(x)$. Thus, $T \vdash \psi_i^1$ by Σ_1-completeness, and hence $T \vdash \Box\psi_i^1$. This is a contradiction.
- If $T \vdash \Box\psi_i^1$ and $T \nvdash \Box(\varphi_i \vee \psi_i^0)$, then $\mathbb{N} \models \psi_i^0$, and hence $T \vdash \psi_i^0$. Thus, $T \vdash \varphi_i \vee \psi_i^0$ and hence $T \vdash \Box(\varphi_i \vee \psi_i^0)$, a contradiction.

We have shown that in either case, for some i, both $\Box(\varphi_i \vee \psi_i^0)$ and $\Box\psi_i^1$ are provable in T. Since $\mathbf{PA} \vdash \psi_i^1 \to \neg\psi_i^0$, we have $T \vdash \Box\neg\psi_i^0$ for such an i. Therefore, we conclude $T \vdash \Box\varphi_i$. ☐

From Propositions 3.11 and 4.16 and Lemma 4.19, we obtain the following proposition.

PROPOSITION 4.20. *Let T be any r.e. extension of* $\mathbf{PA}(\mathbf{K})$.

1. *If T has* $\Delta(\mathrm{B})$-DP, *then T has* $(\Delta(\mathrm{B}), \Sigma_1)$-DP.
2. *If T has* B-DP, *then T has* (B, Σ_1)-DP.

PROPOSITION 4.21. *Let T be any r.e. extension of* $\mathbf{PA}(\mathbf{K})$ *with* $T \nvdash \Box\bot$. *If T has* (B, Σ_1)-DP, *then T is Σ_1-sound.*

Proof. We prove the contrapositive. Suppose that $T \nvdash \Box\bot$ and T is not Σ_1-sound. Then, there exists a Δ_0 formula $\delta(x)$ such that $T \vdash \exists x\delta(x)$ and $\mathbb{N} \models \forall x \neg\delta(x)$. Let σ_0 and σ_1 be Σ_1 sentences satisfying the following equivalences:

- $\mathbf{PA} \vdash \sigma_0 \leftrightarrow \exists x\Big(\big(\delta(x) \vee \mathrm{Prf}_T(\ulcorner\sigma_1\urcorner, x)\big) \wedge \forall y < x\, \neg\mathrm{Prf}_T(\ulcorner\Box\sigma_0\urcorner, y)\Big).$
- $\mathbf{PA} \vdash \sigma_1 \leftrightarrow \exists y\Big(\mathrm{Prf}_T(\ulcorner\Box\sigma_0\urcorner, y) \wedge \forall x \leq y\, \big(\neg\delta(x) \wedge \neg\mathrm{Prf}_T(\ulcorner\sigma_1\urcorner, x)\big)\Big).$

Since $T \vdash \exists x\delta(x)$, we have $T \vdash \sigma_0 \vee \sigma_1$. Therefore, $T \vdash (\Box\sigma_0) \vee \sigma_1$ by Theorem 2.8.

Suppose, towards a contradiction, that $T \vdash \Box\sigma_0$ or $T \vdash \sigma_1$. Let p be the smallest T-proof of $\Box\sigma_0$ or σ_1. If p is a proof of $\Box\sigma_0$, then $\mathbb{N} \models \sigma_1$ by the choice of σ_1. Hence, $T \vdash \sigma_1$ and thus $T \vdash \Box\sigma_1$. Since $T \vdash \sigma_0 \wedge \sigma_1 \to \bot$, we have $T \vdash \Box\bot$ because $T \vdash \Box\sigma_0 \wedge \Box\sigma_1$. This is a contradiction. If p is a proof of σ_1, then it is shown $T \vdash \sigma_0$. This contradicts the consistency of T. Thus, we have shown that $T \nvdash \Box\sigma_0$ and $T \nvdash \sigma_1$. This means that T does not have (B, Σ_1)-DP. ☐

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

PROPOSITION 4.22. *Let T be any consistent r.e. extension of* $\mathbf{PA}(\mathbf{K})$ *with* $T \nvdash \Box\bot$ *and let* Γ *be a class of formulas with* $\mathrm{B} \subseteq \Gamma$. *If T has* (Γ, Σ_1)-DP, *then T is* Γ-DC.

Proof. Suppose that T has (Γ, Σ_1)-DP. Let φ be any Γ sentence such that $T \vdash \varphi \lor \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$. By (Γ, Σ_1)-DP, $T \vdash \varphi$ or $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$. Since $\mathrm{B} \subseteq \Gamma$, by Proposition 4.21, T is Σ_1-sound. Thus, in either case, we obtain $T \vdash \varphi$. □

The converse implication also holds when Γ is B or $\Delta(\mathrm{B})$. In order to prove this, we generalize the Fixed Point Lemma to modal arithmetic. It is proved by repeating a well-known proof, and so we omit it (see [1]).

LEMMA 4.23 (The Fixed Point Lemma). *For any* $\mathcal{L}_A(\Box)$-*formulas* $\varphi_0(x_0, \dots, x_{k-1}), \dots, \varphi_{k-1}(x_0, \dots, x_{k-1})$ *with only the free variables* $x_0, \dots, x_{k-1}$, *we can effectively find* $\mathcal{L}_A(\Box)$-*sentences* $\psi_0, \dots, \psi_{k-1}$ *such that for each* $i < k$,

$$\mathbf{PA}_\Box \vdash \psi_i \leftrightarrow \varphi_i(\ulcorner \psi_0 \urcorner, \dots, \ulcorner \psi_{k-1} \urcorner).$$

Moreover, for each $i < k$, *if* $\varphi_i(x_0, \dots, x_{k-1})$ *is a* $\Sigma(\mathrm{B})$ *formula, then such a* ψ_i *can be found as a* $\Sigma(\mathrm{B})$ *sentence.*

PROPOSITION 4.24. *Let T be any r.e. extension of* $\mathbf{PA}(\mathbf{K})$.

1. *If T is* $\Delta(\mathrm{B})$-DC, *then T has* $(\Delta(\mathrm{B}), \Sigma_1)$-DP.
2. *If T is* B-DC, *then T has* (B, Σ_1)-DP.

Proof. We prove only Clause 1. Clause 2 is proved similarly. Suppose that T is $\Delta(\mathrm{B})$-DC. Let φ be any $\Delta(\mathrm{B})$ sentence and let $\delta(x)$ be any Δ_0 formula such that $T \vdash \varphi \lor \exists x \delta(x)$ and $T \nvdash \exists x \delta(x)$. We would like to show $T \vdash \varphi$. In this case, $\mathbb{N} \models \forall x \neg \delta(x)$. By Proposition 3.11, we may assume that φ is of the form $\Box\psi_0 \lor \cdots \lor \Box\psi_{k-1}$. By the Fixed Point Lemma, let $\xi_0, \dots, \xi_{k-1}$ be $\mathcal{L}_A(\Box)$-sentences satisfying the following equivalences for all $i < k$:

$$\mathbf{PA}_\Box \vdash \xi_i \leftrightarrow \Big[\psi_i \lor \exists x \big(\delta(x) \land \forall y < x\, \neg \mathrm{Prf}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner, y)\big)\Big].$$

Since $\mathbf{PA}_\Box \vdash \psi_i \to \xi_i$, we have $\mathbf{PA}(\mathbf{K}) \vdash \Box\psi_i \to \Box\xi_i$. Also,

$$\begin{aligned}\mathbf{PA}(\mathbf{K}) \vdash \exists x \delta(x) \land& \neg \mathrm{Pr}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner) \\ &\to \exists x \big(\delta(x) \land \forall y < x\, \neg \mathrm{Prf}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner, y)\big), \\ &\to \mathrm{Pr}_T\Big(\ulcorner \exists x \big(\delta(x) \land \forall y < x\, \neg \mathrm{Prf}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner, y)\big) \urcorner\Big), \\ &\to \mathrm{Pr}_T(\ulcorner \xi_i \urcorner).\end{aligned}$$

Since $\mathbf{PA}(\mathbf{K}) \vdash \mathrm{Pr}_T(\ulcorner \xi_i \urcorner) \to \mathrm{Pr}_T(\ulcorner \Box\xi_i \urcorner)$ because **PA** can prove that the consequences of T are closed under the rule NEC, we obtain $\mathbf{PA}(\mathbf{K}) \vdash \mathrm{Pr}_T(\ulcorner \xi_i \urcorner) \to \mathrm{Pr}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner)$. Thus, we have

$$\mathbf{PA}(\mathbf{K}) \vdash \big[\exists x \delta(x) \land \neg \mathrm{Pr}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner)\big] \to \mathrm{Pr}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner),$$

and hence

$$\mathbf{PA}(\mathbf{K}) \vdash \exists x \delta(x) \to \mathrm{Pr}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner).$$

Then, by combining this with our assumption that $T \vdash \Box\psi_0 \lor \cdots \lor \Box\psi_{k-1} \lor \exists x \delta(x)$, we obtain

$$T \vdash \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \lor \mathrm{Pr}_T(\ulcorner \Box\xi_0 \lor \cdots \lor \Box\xi_{k-1} \urcorner).$$

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

By $\Delta(\mathrm{B})$-DC, we have

$$T \vdash \Box\xi_0 \vee \cdots \vee \Box\xi_{k-1}. \tag{3}$$

Since

$$\mathbb{N} \models \exists y\big(\mathrm{Prf}_T(\ulcorner \Box\xi_0 \vee \cdots \vee \Box\xi_{k-1} \urcorner, y) \wedge \forall x \leq y\, \neg\delta(x)\big),$$

this sentence is provable in $\mathbf{PA}(\mathbf{K})$. Thus,

$$\mathbf{PA}(\mathbf{K}) \vdash \neg\exists x\big(\delta(x) \wedge \forall y < x\, \neg\mathrm{Prf}_T(\ulcorner \Box\xi_0 \vee \cdots \vee \Box\xi_{k-1} \urcorner, y)\big).$$

Then, by the choice of ξ_i, $\mathbf{PA}(\mathbf{K}) \vdash \xi_i \to \psi_i$ for each i. From (3), we conclude $T \vdash \Box\psi_0 \vee \cdots \vee \Box\psi_{k-1}$ and hence $T \vdash \varphi$. □

In the statements of Propositions 4.21 and 4.22, the condition "$T \nvdash \Box\bot$" is assumed. On the other hand, for consistent theories T with $T \vdash \Box\bot$, the situation changes. Indeed, every B formula is provable in such a theory T. Thus, T does not have MDP and MEP. Also, every $\Delta(\mathrm{B})$ formula is T-provably equivalent to some Δ_0 formula. Moreover, every $\Delta(\mathrm{B})$ sentence φ is either provable or refutable in T. Therefore, we obtain the following proposition.

Proposition 4.25. *Let T be any extension of* $\mathbf{PA}(\mathbf{Verum})$. *Then, T has* $\Delta(\mathrm{B})$-DP, $(\Delta(\mathrm{B}), \Sigma_1)$-DP, *and* B-EP. *Also, T is* B-DC.

From Propositions 4.22, 4.24, and 4.25, we have:

Corollary 4.26. *Let T be any consistent r.e. extension of* $\mathbf{PA}(\mathbf{K})$.

1. *If $T \nvdash \Box\bot$, then T has* $(\Delta(\mathrm{B}), \Sigma_1)$-DP *if and only if T is* $\Delta(\mathrm{B})$-DC.
2. *T has* (B, Σ_1)-DP *if and only if T is* B-DC.

For consistent r.e. extensions of $\mathbf{PA}(\mathbf{Verum})$, $\Delta(\mathrm{B})$-DC is strictly weaker than Σ_1-soundness.

Proposition 4.27. *Let T be any consistent r.e. extension of* $\mathbf{PA}(\mathbf{Verum})$. *Then, the following are equivalent*:

1. *T is* $\Delta(\mathrm{B})$-DC.
2. $T \nvdash \neg\mathrm{Con}_T$.

Proof. ($1 \Rightarrow 2$): Suppose that T is $\Delta(\mathrm{B})$-DC. Since $T \nvdash \bot$, we obtain $T \nvdash \mathrm{Pr}_T(\ulcorner \bot \urcorner) \vee \bot$. Hence, $T \nvdash \neg\mathrm{Con}_T$.

($2 \Rightarrow 1$): Suppose $T \nvdash \neg\mathrm{Con}_T$. Let φ be any $\Delta(\mathrm{B})$ sentence with $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \vee \varphi$. If $T \vdash \neg\varphi$, then φ is T-equivalent to $\bot$. We have $T \vdash \mathrm{Pr}_T(\ulcorner \bot \urcorner) \vee \bot$, and so $T \vdash \neg\mathrm{Con}_T$. This is a contradiction. Therefore, $T \vdash \varphi$. □

There are $\mathcal{L}_A$-sound theories that do not have even (B, Σ_1)-DP.

Proposition 4.28.
1. $\mathbf{PA}(\mathbf{Triv})$ *does not have* (B, Σ_1)-DP.
2. *Let T be any r.e. theory such that* $\mathbf{PA}(\mathbf{Verum}) \vdash T \vdash \mathbf{PA}(\mathbf{K4})$ *and let* $U := T + \{\mathrm{Pr}_T(\ulcorner \Box\bot \urcorner) \vee \Box\bot\}$. *If $T \nvdash \Box\bot$, then U is $\mathcal{L}_A$-sound but is not* B-DC.

Proof. 1. Let φ be a Π_1 Gödel sentence of $\mathbf{PA}$. Since $\mathbf{PA} \vdash \varphi \vee \neg\varphi$, we have $\mathbf{PA}(\mathbf{Triv}) \vdash \Box\varphi \vee \neg\varphi$. Since $\mathbf{PA}(\mathbf{Triv})$ is a conservative extension of $\mathbf{PA}$ (Corollary 2.4), $\mathbf{PA}(\mathbf{Triv}) \nvdash \varphi$ and $\mathbf{PA}(\mathbf{Triv}) \nvdash \neg\varphi$. Then, $\mathbf{PA}(\mathbf{Triv}) \nvdash \Box\varphi$ and $\mathbf{PA}(\mathbf{Triv}) \nvdash \neg\varphi$.

2. Since U is a subtheory of $\mathbf{PA}(\mathbf{Verum})$, U is $\mathcal{L}_A$-sound by Corollary 2.7. Since $U \vdash \mathrm{Pr}_T(\ulcorner \Box\bot \urcorner) \vee \Box\bot$, we have $U \vdash \mathrm{Pr}_U(\ulcorner \Box\bot \urcorner) \vee \Box\bot$. Suppose, towards a contradiction, $U \vdash \Box\bot$. Since $\mathrm{Pr}_T(\ulcorner \Box\bot \urcorner) \vee \Box\bot$ is a $\Sigma(\mathrm{B})$ sentence, $T \vdash \mathrm{Pr}_T(\ulcorner \Box\bot \urcorner) \vee \Box\bot \to \Box\bot$ by the $\Sigma(\mathrm{B})$-deduction theorem (Corollary 3.14). In particular, $T \vdash \mathrm{Pr}_T(\ulcorner \Box\bot \urcorner) \to \Box\bot$. By Löb's theorem, $T \vdash \Box\bot$. This is a contradiction. Therefore, $U \nvdash \Box\bot$. Hence, U is not B-DC. □

§5. $\Sigma(\mathrm{B})$-DP and related properties.

First of all, we consider the case that T proves $\Box\bot$.

Proposition 5.29. *Let T be any consistent r.e. extension of* $\mathbf{PA}(\mathbf{Verum})$. *Then, the following are equivalent*:

1. *T is Σ_1-sound.*
2. *T has $\Sigma(\mathrm{B})$-DP.*
3. *T has $(\Sigma(\mathrm{B}), \Sigma_1)$-DP.*
4. *T has $\Sigma(\mathrm{B})$-EP.*
5. *T has $\Delta(\mathrm{B})$-EP.*
6. *T is $\Sigma(\mathrm{B})$-DC.*

Proof. Since every $\Sigma(\mathrm{B})$ (resp. $\Delta(\mathrm{B})$) formula is T-provably equivalent to some Σ_1 (resp. Δ_0) formula, we have $(2 \Leftrightarrow 3)$. Also by Guaspari's theorem [10] on the equivalence of the Σ_1-soundness and Σ_1-DP, the equivalence $(1 \Leftrightarrow 2)$ holds. Moreover, since the implications "Σ_1-sound $\Rightarrow \Sigma_1$-EP," "Σ_1-EP $\Rightarrow \Delta_0$-EP," and "Δ_0-EP $\Rightarrow \Sigma_1$-sound" are easily verified, we obtain that Clauses 1, 4, and 5 are pairwise equivalent. Finally, since the equivalence of the Σ_1-soundness and Σ_1-DC is shown in [15], we get $(1 \Leftrightarrow 6)$. □

Corollary 5.30. $\mathbf{PA}(\mathbf{Verum})$ *has* $\Sigma(\mathrm{B})$-EP *but does not have* MDP.

Proof. Since $\mathbf{PA}(\mathbf{Verum})$ is Σ_1-sound by Corollary 2.7, $\mathbf{PA}(\mathbf{Verum})$ has $\Sigma(\mathrm{B})$-EP. On the other hand, $\mathbf{PA}(\mathbf{Verum}) \vdash \Box\bot$ and $\mathbf{PA}(\mathbf{Verum}) \nvdash \bot$, and thus $\mathbf{PA}(\mathbf{Verum})$ does not have MDP. □

We then discuss theories in which $\Box\bot$ is not necessarily provable. Unlike the cases of $\Delta(\mathrm{B})$ and B (Proposition 4.20), $(\Sigma(\mathrm{B}), \Sigma_1)$-DP directly follows from $\Sigma(\mathrm{B})$-DP because $\Sigma_1 \subseteq \Sigma(\mathrm{B})$. Also, as in the cases of $\Delta(\mathrm{B})$ and B (Proposition 4.24), we obtain the following proposition.

Proposition 5.31. *Let T be any r.e. $\mathcal{L}_A(\Box)$-theory extending* $\mathbf{PA}$. *If T is* $\Sigma(\mathrm{B})$-DC, *then T has* $(\Sigma(\mathrm{B}), \Sigma_1)$-DP.

Proof. Suppose that T is $\Sigma(\mathrm{B})$-DC. Let φ be any $\Sigma(\mathrm{B})$ sentence and let $\delta(x)$ be any Δ_0 formula such that $T \vdash \varphi \vee \exists x \delta(x)$ and $T \nvdash \exists x \delta(x)$. We would like to show $T \vdash \varphi$. In this case, $\mathbb{N} \models \forall x \neg \delta(x)$. Let σ be a Σ_1 sentence satisfying

$$\mathbf{PA} \vdash \sigma \leftrightarrow \exists x \big(\delta(x) \wedge \forall y < x \, \neg \mathrm{Prf}_T(\ulcorner \varphi \vee \sigma \urcorner, y)\big). \tag{4}$$

Since T is an extension of $\mathbf{PA}$ and σ is Σ_1, we have $\mathbf{PA} \vdash \sigma \to \mathrm{Pr}_T(\ulcorner \sigma \urcorner)$, and hence $\mathbf{PA} \vdash \sigma \to \mathrm{Pr}_T(\ulcorner \varphi \vee \sigma \urcorner)$. By the equivalence (4), we obtain

$$\mathbf{PA} \vdash \exists x \delta(x) \wedge \neg \mathrm{Prf}_T(\ulcorner \varphi \vee \sigma \urcorner) \to \sigma.$$

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

It follows $\mathbf{PA} \vdash \exists x \delta(x) \land \neg \mathrm{Prf}_T(\ulcorner \varphi \lor \sigma \urcorner) \to \mathrm{Pr}_T(\ulcorner \varphi \lor \sigma \urcorner)$, and hence

$$\mathbf{PA} \vdash \exists x \delta(x) \to \mathrm{Prf}_T(\ulcorner \varphi \lor \sigma \urcorner).$$

Since $T \vdash \varphi \lor \exists x \delta(x)$, we obtain $T \vdash (\varphi \lor \sigma) \lor \mathrm{Pr}_T(\ulcorner \varphi \lor \sigma \urcorner)$. Since $\varphi \lor \sigma$ is a $\Sigma(\mathrm{B})$ sentence, by $\Sigma(\mathrm{B})$-DC, we have $T \vdash \varphi \lor \sigma$. Since $\mathbb{N} \models \forall x \neg \delta(x)$, $\mathbb{N} \models \exists y \big(\mathrm{Prf}_T(\ulcorner \varphi \lor \sigma \urcorner, y) \land \forall x \leq y \, \neg \delta(x)\big)$ and this is provable in T. Then, $T \vdash \neg \sigma$, and thus $T \vdash \varphi$. $\square$

From Propositions 4.22, 5.29, and 5.31, we obtain the following corollary.

COROLLARY 5.32. *For any r.e. extension T of* $\mathbf{PA}(\mathbf{K})$, *T has* $(\Sigma(\mathrm{B}), \Sigma_1)$-DP *if and only if T is* $\Sigma(\mathrm{B})$-DC.

Before proving our main theorem of this section, we prepare some notations and lemmas.

DEFINITION 5.33. *For each $\Sigma(\mathrm{B})$ formula φ, we define the $\mathcal{L}_A(\Box)$-formula φ^- inductively as follows*:

1. *If φ is Σ_1, then $\varphi^- :\equiv \varphi$.*
2. *If φ is of the form $\Box \psi$, then $\varphi^- :\equiv \psi$.*
3. *Otherwise if φ is of the form $\psi \land \sigma$, $\psi \lor \sigma$, $\exists x \psi$ or $\forall x < t \, \psi$, then φ^- is respectively $\psi^- \land \sigma^-$, $\psi^- \lor \sigma^-$, $\exists x \psi^-$ or $\forall x < t \, \psi^-$.*

The operation $(\cdot)^-$ removes the outermost $\Box$ of nested occurrences of $\Box$'s in the formula. For example, $(\Box\Box\varphi \lor \Box\psi)^-$ is $\Box\varphi \lor \psi$. The following lemma is a strengthening of Theorem 2.8.

LEMMA 5.34. *For any $\Sigma(\mathrm{B})$ formula φ,* $\mathbf{PA}(\mathbf{K}) \vdash \varphi \to \Box(\varphi^-)$.

Proof. This lemma is proved by induction on the construction of φ as in the proof of Theorem 3.13. Notice that if φ is of the form $\Box\psi$, then $\varphi^- \equiv \psi$ and thus $\mathbf{PA}(\mathbf{K}) \vdash \varphi \to \Box(\varphi^-)$ holds. $\square$

DEFINITION 5.35. *For each $\Delta(\mathrm{B})$ formula $\varphi(\vec{x})$, we define the $\Delta(\mathrm{B})$ formula $\varphi^*(\vec{x}, \vec{v})$ with zero or more additional free variables $\vec{v}$ which do not occur in $\varphi(\vec{x})$ inductively as follows*:

1. *If $\varphi(\vec{x})$ is either Δ_0 or of the form $\Box\psi(\vec{x})$, then $\varphi^*(\vec{x}) :\equiv \varphi(\vec{x})$.*
2. *Otherwise if $\varphi(\vec{x})$ is of the form $\psi(\vec{x}) \land \sigma(\vec{x})$, then $\varphi^*(\vec{x}, \vec{u}, \vec{v}) :\equiv \psi^*(\vec{x}, \vec{u}) \land \sigma^*(\vec{x}, \vec{v})$ where $\vec{u}$ and $\vec{v}$ are pairwise disjoint.*
3. *Otherwise if $\varphi(\vec{x})$ is of the form $\psi(\vec{x}) \lor \sigma(\vec{x})$, then*
$$\varphi^*(\vec{x}, \vec{u}, \vec{v}, w) :\equiv \Big[\big(w = 0 \land \psi^*(\vec{x}, \vec{u})\big) \lor \big(w \neq 0 \land \sigma^*(\vec{x}, \vec{v})\big)\Big],$$
where $\vec{u}$, $\vec{v}$, and w *are pairwise disjoint.*
4. *Otherwise if $\varphi(\vec{x})$ is of the form $\forall y < t \, \psi(\vec{x}, y)$, then $\varphi^*(\vec{x}, \vec{v}) :\equiv \forall y < t \, \psi^*(\vec{x}, y, \vec{v})$.*
5. *Otherwise if $\varphi(\vec{x})$ is of the form $\exists y < t \, \psi(\vec{x}, y)$, then $\varphi^*(\vec{x}, \vec{v}, w)$ is the formula $\exists y < t \, (y = w \land \psi^*(\vec{x}, y, \vec{v}))$.*

From the definition, we can easily prove the following lemma by induction on the construction of $\varphi(\vec{x}) \in \Delta(\mathrm{B})$.

LEMMA 5.36. *For any $\Delta(\mathrm{B})$ formula $\varphi(\vec{x})$,* $\mathbf{PA}_\Box \vdash \varphi(\vec{x}) \leftrightarrow \exists \vec{v} \varphi^*(\vec{x}, \vec{v})$.

The following lemma is an important feature of our two transformations – and $*$.

LEMMA 5.37. *Let T be any r.e. extension of* $\mathbf{PA}(\mathbf{K})$ *such that* $T \nvdash \Box\bot$. *For any* $\Delta(\mathrm{B})$ *sentence* φ, *if there exist numbers* $\vec{p}$ *such that* $T \vdash \Box(\varphi^*)^-(\vec{\overline{p}})$, *then* $T \vdash \varphi$.

Proof. We prove the lemma by induction on the construction of φ.

- If φ is a Δ_0 sentence, then $(\varphi^*)^- \equiv \varphi^- \equiv \varphi$. Suppose $T \vdash \Box(\varphi^*)^-$, i.e., $T \vdash \Box\varphi$. If $\mathbb{N} \models \neg\varphi$, then $T \vdash \neg\varphi$ and $T \vdash \Box\neg\varphi$. We have $T \vdash \Box\bot$, a contradiction. Hence, $\mathbb{N} \models \varphi$. We conclude $T \vdash \varphi$.
- If φ is of the form $\Box\psi$, then $(\varphi^*)^- \equiv (\Box\psi)^- \equiv \psi$. Suppose $T \vdash \Box(\varphi^*)^-$. Then, $T \vdash \varphi$.
- If φ is of the form $\psi \wedge \sigma$, then $(\varphi^*)^-(\vec{u}, \vec{v}) \equiv (\psi^*)^-(\vec{u}) \wedge (\sigma^*)^-(\vec{v})$. Suppose $T \vdash \Box(\varphi^*)^-(\vec{\overline{p}}, \vec{\overline{q}})$. Then, $T \vdash \Box(\psi^*)^-(\vec{\overline{p}})$ and $T \vdash \Box(\sigma^*)^-(\vec{\overline{q}})$. By the induction hypothesis, $T \vdash \psi$ and $T \vdash \sigma$. We conclude $T \vdash \varphi$.
- If φ is of the form $\psi \vee \sigma$, then

$$(\varphi^*)^-(\vec{u}, \vec{v}, w) \equiv \Big[\big(w = 0 \wedge (\psi^*)^-(\vec{u})\big) \vee \big(w \neq 0 \wedge (\sigma^*)^-(\vec{v})\big)\Big].$$

  Suppose $T \vdash \Box(\varphi^*)^-(\vec{\overline{p}}, \vec{\overline{q}}, \overline{r})$. Then,

$$T \vdash \Box\Big[\big(\overline{r} = 0 \wedge (\psi^*)^-(\vec{\overline{p}})\big) \vee \big(\overline{r} \neq 0 \wedge (\sigma^*)^-(\vec{\overline{q}})\big)\Big].$$

  If $r = 0$, then $T \vdash \Box(\psi^*)^-(\vec{\overline{p}})$. By the induction hypothesis, $T \vdash \psi$. If $r \neq 0$, then $T \vdash \Box(\sigma^*)^-(\vec{\overline{q}})$. By the induction hypothesis, $T \vdash \sigma$. In either case, $T \vdash \varphi$.
- If φ is of the form $\forall x < t\, \psi(x)$ for some $\mathcal{L}_A$-term t, then $(\varphi^*)^-(\vec{v})$ is the formula $\forall x < t\, (\psi^*)^-(\vec{v}, x)$. Since φ is a sentence, t is a closed term. Let k be the value of the term t and suppose $T \vdash \Box(\varphi^*)^-(\vec{\overline{p}})$. Then, for all $n < k$, $T \vdash \Box(\psi^*)^-(\vec{\overline{p}}, \overline{n})$. By the induction hypothesis, $T \vdash \psi(\overline{n})$. We obtain $T \vdash \varphi$.
- If φ is of the form $\exists x < t\, \psi(x)$ for some closed term t, then

$$(\varphi^*)^-(\vec{v}, w) \equiv \exists x < t\big(x = w \wedge (\psi^*)^-(\vec{v}, x)\big).$$

  Suppose $T \vdash \Box(\varphi^*)^-(\vec{\overline{p}}, \overline{q})$. Then, $T \vdash \Box\exists x < t\big(x = \overline{q} \wedge (\psi^*)^-(\vec{\overline{p}}, x)\big)$. Since $T \nvdash \Box\bot$, the value of t is larger than q. Since $T \vdash \Box(\psi^*)^-(\vec{\overline{p}}, \overline{q})$, by the induction hypothesis, $T \vdash \psi(\overline{q})$. Then, $T \vdash \exists x < t\, \psi(x)$, that is, $T \vdash \varphi$. □

We are ready to prove our main theorem of this section.

THEOREM 5.38. *Let T be any r.e. extension of* $\mathbf{PA}(\mathbf{K})$ *such that* $T \nvdash \Box\bot$. *Then, the following are equivalent*:

1. *T has* $\Delta(\mathrm{B})$-DP *and* $(\Sigma(\mathrm{B}), \Sigma_1)$-DP.
2. *T has* B-EP.
3. *T has* $\Sigma(\mathrm{B})$-EP.
4. *T has* $\Sigma(\mathrm{B})$-DP.

Proof. ($1 \Rightarrow 2$): Let $\varphi(x)$ be any $\mathcal{L}_A(\Box)$-formula with no free variables except possibly x, such that $T \vdash \exists x \Box\varphi(x)$. By the Fixed Point Lemma, let ψ be a $\Sigma(\mathrm{B})$ sentence satisfying

$$\mathbf{PA}_{\Box} \vdash \psi \leftrightarrow \exists x\big(\Box\varphi(x) \wedge \forall y < x\, \neg \mathrm{Prf}_T(\ulcorner \psi \urcorner, y)\big).$$

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

Since $\mathbf{PA}_\Box \vdash \exists x \Box \varphi(x) \wedge \neg \mathrm{Pr}_T(\ulcorner \psi \urcorner) \to \psi$, we have $T \vdash \mathrm{Pr}_T(\ulcorner \psi \urcorner) \vee \psi$. Since T is $\Sigma(\mathrm{B})$-DC by Corollary 5.32, we obtain $T \vdash \psi$. By the choice of ψ,

$$T \vdash \exists x \big(\Box \varphi(x) \wedge \forall y < x\, \neg \mathrm{Prf}_T(\ulcorner \psi \urcorner, y) \big). \tag{5}$$

Let p be a proof of ψ in T, then $T \vdash \mathrm{Prf}_T(\ulcorner \psi \urcorner, \overline{p})$ and thus $T \vdash \exists x \leq \overline{p}\, \Box \varphi(x)$ by (5). Then, $T \vdash \Box \varphi(\overline{0}) \vee \cdots \vee \Box \varphi(\overline{p})$. Since T has B^{p+1}-DP by Proposition 4.16.2, there exists $k \leq p$ such that $T \vdash \Box \varphi(\overline{k})$. Therefore, T has B-EP.

$(2 \Rightarrow 3)$: Let $\varphi(x)$ be any $\Sigma(\mathrm{B})$ formula without having free variables except x such that $T \vdash \exists x \varphi(x)$. By Proposition 3.10, there exists a $\Delta(\mathrm{B})$ formula $\psi(x, y)$ such that $\mathbf{PA}_\Box \vdash \varphi(x) \leftrightarrow \exists y \psi(x, y)$. Also, by Lemma 5.36, $\mathbf{PA}_\Box \vdash \psi(x, y) \leftrightarrow \exists \vec{v} \psi^*(x, y, \vec{v})$. Then, $T \vdash \exists x \exists y \exists \vec{v} \psi^*(x, y, \vec{v})$ and

$$T \vdash \exists w\, \exists x \leq w\, \exists y \leq w\, \exists \vec{v} \leq w\, \big(w = \langle x, y, \vec{v} \rangle \wedge \psi^*(x, y, \vec{v}) \big).$$

Here $\langle x, y, \vec{v} \rangle$ is an appropriate iteration of usual Δ_0 representable bijective pairing function $\langle \cdot, \cdot \rangle$. We may assume that $\mathbf{PA}$ proves $x \leq \langle x, y \rangle$ and $y \leq \langle x, y \rangle$. By Lemma 5.34,

$$T \vdash \exists w\, \Box \exists x \leq w\, \exists y \leq w\, \exists \vec{v} \leq w\, \big(w = \langle x, y, \vec{v} \rangle \wedge (\psi^*)^-(x, y, \vec{v}) \big).$$

By B-EP, there exists a natural number k such that

$$T \vdash \Box \exists x \leq \overline{k}\, \exists y \leq \overline{k}\, \exists \vec{v} \leq \overline{k}\, \big(\overline{k} = \langle x, y, \vec{v} \rangle \wedge (\psi^*)^-(x, y, \vec{v}) \big).$$

For the unique p, q, and $\vec{r}$ such that $k = \langle p, q, \vec{r} \rangle$,

$$T \vdash \Box (\psi^*)^-(\overline{p}, \overline{q}, \vec{\overline{r}}).$$

By Lemma 5.37, we obtain $T \vdash \psi(\overline{p}, \overline{q})$. Then, $T \vdash \varphi(\overline{p})$. Therefore, T has $\Sigma(\mathrm{B})$-EP.

$(3 \Rightarrow 4)$: By Proposition 4.18.2.

$(4 \Rightarrow 1)$: This is trivial. □

In order to derive the equivalence of MDP and MEP from Theorem 5.38, we prove a proposition that connects MDP (resp. MEP) and $\Sigma(\mathrm{B})$-DP (resp. $\Sigma(\mathrm{B})$-EP).

PROPOSITION 5.39. *Let T be any r.e. extension of* $\mathbf{PA}(\mathbf{K4})$.

1. *If T has* MDP, *then T also has* $\Sigma(\mathrm{B})$-DP.
2. *If T has* MEP, *then T also has* $\Sigma(\mathrm{B})$-EP.

Proof. 1. Let φ and ψ be any $\Sigma(\mathrm{B})$ sentences such that $T \vdash \varphi \vee \psi$. By Theorem 3.13, $T \vdash \Box \varphi \vee \Box \psi$. By MDP, we obtain $T \vdash \varphi$ or $T \vdash \psi$. Therefore, T has $\Sigma(\mathrm{B})$-DP.

Clause 2 is proved similarly. □

COROLLARY 5.40. *For any r.e. extension T of* $\mathbf{PA}(\mathbf{K4})$, *T has* MDP *if and only if T has* MEP.

Proof. Since MEP implies MDP by Proposition 4.18.1, it suffices to show that MDP implies MEP. We may assume that T is consistent. If T has MDP, then T has $\Sigma(\mathrm{B})$-DP by Proposition 5.39. Also, T is closed under the box elimination rule by Proposition 4.17. Then, $T \nvdash \Box \bot$ by the consistency of T. By Theorem 5.38, T has B-EP. By Proposition 4.17 again, we conclude that T has MEP. □

REMARK 5.41. *In the introduction, we imprecisely mentioned the result of Friedman and Sheard* [6]. *Firstly, Friedman and Sheard actually proved their theorem in the setting*

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

where the use of the rule NEC *and the axiom* $\forall\vec{x}(\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi))$ *is restricted, that is, in the non-normal setting. In our normal setting, the following result follows from their theorem: For any r.e. extension T of* $\mathbf{PA}(\mathbf{K4}) + \{\forall\vec{x}\Box(\Box\varphi \to \varphi) \mid \varphi \in \Delta_0\}$, *if T is closed under the box elimination rule, then T has* B-DP *if and only if T has* B-EP. *Then, in the light of Proposition 4.17, this statement can be rewritten as follows: For any r.e. extension T of* $\mathbf{PA}(\mathbf{K4}) + \{\forall\vec{x}\Box(\Box\varphi \to \varphi) \mid \varphi \in \Delta_0\}$, *$T$ has* MDP *if and only if T has* MEP. *Therefore, our Corollary 5.40 shows that the same consequence is obtained without using the axiom scheme* $\{\forall\vec{x}\Box(\Box\varphi \to \varphi) \mid \varphi \in \Delta_0\}$. *Notice that by Theorem 2.8, over* $\mathbf{PA}(\mathbf{K})$, $\{\forall\vec{x}\Box(\Box\varphi \to \varphi) \mid \varphi \in \Delta_0\}$ *is equivalent to a single sentence* $\Box\neg\Box\bot$.

§6. Generalizations of the notions of soundness and Σ_1-soundness. In this section, we introduce several notions related to the soundness of theories of modal arithmetic with respect to $\mathcal{L}_A(\Box)$-sentences. This section consists of three subsections. In the first subsection, we introduce the notion of the $\mathcal{L}_A(\Box)$-soundness and prove that several $\mathcal{L}_A(\Box)$-theories are actually $\mathcal{L}_A(\Box)$-sound. In the second subsection, we introduce the notions of the $\Sigma(\mathrm{B})$-soundness and the weak $\Sigma(\mathrm{B})$-soundness. Then, we prove that over appropriate theories, the $\Sigma(\mathrm{B})$-soundness and the weak $\Sigma(\mathrm{B})$-soundness characterize MDP and $\Sigma(\mathrm{B})$-DP, respectively. In the last subsection, we prove two non-implications between the properties as applications of the results we have obtained so far.

6.1. $\mathcal{L}_A(\Box)$***-soundness.*** We formulate the notion of the $\mathcal{L}_A(\Box)$-soundness under the interpretation that boxed formulas represent the provability of some formula in the standard model $\mathbb{N}$ of arithmetic. To do so, we once translate each $\mathcal{L}_A(\Box)$-sentence into an $\mathcal{L}_A$-sentence using a provability predicate, and then consider the truth of the translated sentence in $\mathbb{N}$. First, we introduce two types of translations π_T and π'_T.

DEFINITION 6.42 (π-translations). *Let T be any r.e. $\mathcal{L}_A(\Box)$-theory. We define a translation π_T of $\mathcal{L}_A(\Box)$-formulas into $\mathcal{L}_A$-formulas inductively as follows:*

1. *If φ is an $\mathcal{L}_A$-formula, then* $\pi_T(\varphi) :\equiv \varphi$.
2. *π_T preserves logical connectives and quantifiers.*
3. $\pi_T(\Box\varphi(\vec{x})) :\equiv \mathrm{Pr}_T(\ulcorner\varphi(\dot{\vec{x}})\urcorner)$.

Here $\ulcorner\varphi(\dot{\vec{x}})\urcorner$ is a primitive recursive term corresponding to a primitive recursive function calculating the Gödel number of $\varphi(\overline{\vec{n}})$ from $\vec{n}$. Note that $\vec{x}$ are free variables in the formula $\mathrm{Pr}_T(\ulcorner\varphi(\dot{\vec{x}})\urcorner)$.

DEFINITION 6.43 (π'-translations). *Let T be any r.e. $\mathcal{L}_A(\Box)$-theory. We define a translation π'_T of $\mathcal{L}_A(\Box)$-formulas into $\mathcal{L}_A$-formulas inductively as follows:*

1. *If φ is an $\mathcal{L}_A$-formula, then* $\pi'_T(\varphi) :\equiv \varphi$.
2. *π'_T preserves logical connectives and quantifiers.*
3. $\pi'_T(\Box\varphi(\vec{x})) :\equiv \mathrm{Pr}_T(\ulcorner\varphi(\dot{\vec{x}})\urcorner) \land \pi'_T(\varphi(\vec{x}))$.

The translation π'_T is a formalization of Shapiro's slash interpretation [21], introduced in [11] under the name σ_T.

DEFINITION 6.44. *Let T be any r.e. $\mathcal{L}_A(\Box)$-theory.*

- *T is said to be $\mathcal{L}_A(\Box)$*-sound *if for any $\mathcal{L}_A(\Box)$-sentence φ, if $T \vdash \varphi$, then* $\mathbb{N} \models \pi_T(\varphi)$.

- *T is said to be* alternatively $\mathcal{L}_A(\Box)$-sound *if for any $\mathcal{L}_A(\Box)$-sentence φ, if $T \vdash \varphi$, then $\mathbb{N} \models \pi'_T(\varphi)$.*

Actually, these two notions are equivalent.

PROPOSITION 6.45. *For any r.e. $\mathcal{L}_A(\Box)$-theory T, the following are equivalent:*

1. *T is $\mathcal{L}_A(\Box)$-sound.*
2. *T is alternatively $\mathcal{L}_A(\Box)$-sound.*

Proof. ($1 \Rightarrow 2$): Suppose that T is $\mathcal{L}_A(\Box)$-sound. We prove by induction on the construction of φ that for all $\mathcal{L}_A(\Box)$-sentences φ, $\mathbb{N} \models \pi_T(\varphi) \leftrightarrow \pi'_T(\varphi)$. If φ is an atomic $\mathcal{L}_A$-sentence, then $\pi_T(\varphi)$ coincides with $\pi'_T(\varphi)$. The cases for Boolean connectives are easy.

If φ is of the form $\forall x \psi(x)$, then for any natural number n, $\mathbb{N} \models \pi_T(\psi(\overline{n})) \leftrightarrow \pi'_T(\psi(\overline{n}))$ by the induction hypothesis. Then, $\mathbb{N} \models \forall x\big(\pi_T(\psi(x)) \leftrightarrow \pi'_T(\psi(x))\big)$ and hence $\mathbb{N} \models \forall x \pi_T(\psi(x)) \leftrightarrow \forall x \pi'_T(\psi(x))$. This means $\mathbb{N} \models \pi_T(\varphi) \leftrightarrow \pi'_T(\varphi)$.

Suppose that φ is of the form $\Box\psi$. Since $\pi_T(\Box\psi)$ is $\mathrm{Pr}_T(\ulcorner\psi\urcorner)$, by the $\mathcal{L}_A(\Box)$-soundness of T, $\mathbb{N} \models \pi_T(\Box\psi)$ if and only if $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\psi\urcorner) \land \pi_T(\psi)$. By the induction hypothesis, $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\psi\urcorner) \land \pi_T(\psi)$ if and only if $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\psi\urcorner) \land \pi'_T(\psi)$. Thus, $\mathbb{N} \models \pi_T(\Box\psi) \leftrightarrow \pi'_T(\Box\psi)$.

($2 \Rightarrow 1$): Suppose that T is alternatively $\mathcal{L}_A(\Box)$-sound. Similarly, we only prove that for all $\mathcal{L}_A(\Box)$-sentences ψ, $\mathbb{N} \models \pi'_T(\Box\psi) \leftrightarrow \pi_T(\Box\psi)$. $\mathbb{N} \models \pi'_T(\Box\psi)$ is equivalent to $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\psi\urcorner) \land \pi'_T(\psi)$. Then, by the alternative $\mathcal{L}_A(\Box)$-soundness of T, this is equivalent to $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\psi\urcorner)$. This is exactly $\mathbb{N} \models \pi_T(\Box\psi)$. $\Box$

Here we show some propositions that help to prove the $\mathcal{L}_A(\Box)$-soundness of each $\mathcal{L}_A(\Box)$-theory.

PROPOSITION 6.46. *Let T be an $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* **PA**(**GL**) *into* $\mathbf{PA}_\Box$. *For any $\mathcal{L}_A(\Box)$-formula φ, if $T \vdash \varphi$, then for any r.e. extension U of T,* **PA** $\vdash \pi_U(\varphi)$.

Proof. Let U be any r.e. extension of T. As in the proof of Proposition 2.3, by induction on the length of proofs of φ in T, we prove that for any $\mathcal{L}_A(\Box)$-formula φ, if $T \vdash \varphi$, then **PA** $\vdash \pi_U(\varphi)$. We only give proofs of the following four cases.

- If φ is $\forall \vec{x}\big(\Box(\psi(\vec{x}) \to \sigma(\vec{x})) \to (\Box\psi(\vec{x}) \to \Box\sigma(\vec{x}))\big)$, then $\pi_U(\varphi)$ is
$$\forall \vec{x}\big(\mathrm{Pr}_U(\ulcorner\psi(\dot{\vec{x}}) \to \sigma(\dot{\vec{x}})\urcorner) \to (\mathrm{Pr}_U(\ulcorner\psi(\dot{\vec{x}})\urcorner) \to \mathrm{Pr}_U(\ulcorner\sigma(\dot{\vec{x}})\urcorner))\big),$$
and this is provable in **PA**.
- If φ is $\forall \vec{x}\big(\Box\psi(\vec{x}) \to \Box\Box\psi(\vec{x})\big)$, then $\pi_U(\varphi)$ is
$$\forall \vec{x}\big(\mathrm{Pr}_U(\ulcorner\psi(\dot{\vec{x}})\urcorner) \to \mathrm{Pr}_U(\ulcorner\Box\psi(\dot{\vec{x}})\urcorner)\big).$$
Since **PA** proves the fact that the consequences of U are closed under the rule NEC, this sentence is provable in **PA**.
- If φ is $\forall \vec{x}\big(\Box(\Box\psi(\vec{x}) \to \psi(\vec{x})) \to \Box\psi(\vec{x})\big)$, then we reason as follows: By invoking NEC,
$$\mathbf{PA} \vdash \mathrm{Pr}_U(\ulcorner\Box\psi(\dot{\vec{x}}) \to \psi(\dot{\vec{x}})\urcorner) \to \mathrm{Pr}_U(\ulcorner\Box(\Box\psi(\dot{\vec{x}}) \to \psi(\dot{\vec{x}}))\urcorner).$$
Since U is an extension of T, we have **PA** $\vdash \mathrm{Pr}_U(\ulcorner\varphi\urcorner)$ and hence
$$\mathbf{PA} \vdash \mathrm{Pr}_U(\ulcorner\Box(\Box\psi(\dot{\vec{x}}) \to \psi(\dot{\vec{x}}))\urcorner) \to \mathrm{Pr}_U(\ulcorner\Box\psi(\dot{\vec{x}})\urcorner).$$

Then,

$$\mathbf{PA} \vdash \mathrm{Pr}_U(\ulcorner \Box\psi(\dot{\vec{x}}) \to \psi(\dot{\vec{x}})\urcorner) \to \mathrm{Pr}_U(\ulcorner \Box\psi(\dot{\vec{x}})\urcorner),$$

and thus

$$\mathbf{PA} \vdash \mathrm{Pr}_U(\ulcorner \Box\psi(\dot{\vec{x}}) \to \psi(\dot{\vec{x}})\urcorner) \to \mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}})\urcorner).$$

This means $\mathbf{PA} \vdash \pi_U(\varphi)$.

- If $\varphi(\vec{x})$ is derived from $\psi(\vec{x})$ by Nec, then $\varphi(\vec{x}) \equiv \Box\psi(\vec{x})$. Since $U \vdash \psi(\vec{x})$, $\mathbf{PA} \vdash \mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}})\urcorner)$. Thus, $\mathbf{PA} \vdash \pi_U(\varphi(\vec{x}))$. □

Proposition 6.47. *Let T be an $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* $\mathbf{PA}(\mathbf{GL})$ *into* $\mathbf{PA}_\Box$*, and let U be any r.e. extension of T. If $\mathbb{N} \models \pi_U(\varphi)$ for all $\varphi \in U \setminus T$, then U is $\mathcal{L}_A(\Box)$-sound.*

Proof. Suppose that $\mathbb{N} \models \pi_U(\varphi)$ for all $\varphi \in U \setminus T$. We prove by induction on the length of a proof of φ in U that for all $\mathcal{L}_A$-formulas φ, if $U \vdash \varphi$, then $\mathbb{N} \models \pi_U(\forall\vec{x}\varphi)$.

- If φ is an axiom of T or a logical axiom, then $\mathbf{PA} \vdash \pi_U(\forall\vec{x}\varphi)$ by Proposition 6.46. Thus, $\mathbb{N} \models \pi_U(\forall\vec{x}\varphi)$.
- If φ is in $U \setminus T$, then $\mathbb{N} \models \pi_U(\varphi)$ by the supposition.
- If φ is derived from ψ and $\psi \to \varphi$ by MP, then by the induction hypothesis, $\mathbb{N} \models \pi_U(\forall\vec{x}\psi)$ and $\mathbb{N} \models \pi_U(\forall\vec{x}(\psi \to \varphi))$. Then, $\mathbb{N} \models \pi_U(\forall\vec{x}\varphi)$.
- If φ is derived from $\psi(y)$ by Gen, then $\varphi \equiv \forall y\psi(y)$. By the induction hypothesis, $\mathbb{N} \models \pi_U(\forall\vec{x}\forall y\psi(y))$. Hence, $\mathbb{N} \models \pi_U(\forall\vec{x}\varphi)$.
- If φ is derived from ψ by Nec, then $\varphi \equiv \Box\psi$ and $U \vdash \psi$. We have $\mathbb{N} \models \forall\vec{x}\,\mathrm{Pr}_U(\ulcorner \psi\urcorner)$, and equivalently $\mathbb{N} \models \pi_U(\forall\vec{x}\varphi)$. □

Corollary 6.48. *The theories* $\mathbf{PA}_\Box$, $\mathbf{PA}(\mathbf{K})$, $\mathbf{PA}(\mathbf{K4})$, *and* $\mathbf{PA}(\mathbf{GL})$ *are $\mathcal{L}_A(\Box)$-sound.*

Here, we give some more examples of $\mathcal{L}_A(\Box)$-sound theories. Let $x \in W_y$ be a Σ_1 formula saying that x is in the y-th r.e. set. Reinhardt's Weak Mechanistic Thesis (WMT) is the following schema:

- $\exists y\forall x\big(\Box\varphi(x) \leftrightarrow x \in W_y\big)$, where $\varphi(x)$ is an $\mathcal{L}_A(\Box)$-formula having lone free variable x.

When $\Box$ is interpreted as knowledge, WMT can be thought as a formalization of "Knowledge is mechanical." Concerning WMT, we obtain the following corollary to Proposition 6.47.

Corollary 6.49. *Let T be an r.e. $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* $\mathbf{PA}(\mathbf{GL})$ *into* $\mathbf{PA}_\Box$*. Then, the theory $U := T + \mathrm{WMT}$ is $\mathcal{L}_A(\Box)$-sound.*

Proof. Since $\mathrm{Pr}_U(\ulcorner \varphi(\dot{x})\urcorner)$ is a Σ_1 formula, there exists a natural number e such that

$$\mathbb{N} \models \forall x\big(\mathrm{Pr}_U(\ulcorner \varphi(\dot{x})\urcorner) \leftrightarrow x \in W_{\overline{e}}\big).$$

Then, we have $\mathbb{N} \models \pi_U\Big(\exists y\forall x\big(\Box\varphi(x) \leftrightarrow x \in W_y\big)\Big)$. By Proposition 6.47, the theory U is $\mathcal{L}_A(\Box)$-sound. □

We prove an analogue of Proposition 6.46 with respect to π'-translations.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

Proposition 6.50. *Let T be an $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* $\mathbf{PA}(\mathbf{S4})$ *into* $\mathbf{PA}_\Box$*. For any $\mathcal{L}_A(\Box)$-formula φ, if $T \vdash \varphi$, then for any r.e. extension U of T,* $\mathbf{PA} \vdash \pi'_U(\varphi)$.

Proof. Let U be any r.e. extension of T. As in the proof of Proposition 2.3, we prove by induction on the length of proofs of φ in T that for any $\mathcal{L}_A(\Box)$-formula φ, if $T \vdash \varphi$, then $\mathbf{PA} \vdash \pi'_U(\varphi)$. We only give proofs of the following four cases.

- If φ is $\forall \vec{x}\big(\Box(\psi(\vec{x}) \to \sigma(\vec{x})) \to (\Box\psi(\vec{x}) \to \Box\sigma(\vec{x}))\big)$, then $\pi'_U(\varphi)$ is
$$\forall \vec{x}\Big(\mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}}) \to \sigma(\dot{\vec{x}})\urcorner) \wedge \pi'_U(\psi(\vec{x}) \to \sigma(\vec{x})) \\ \to \big([\mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}})\urcorner) \wedge \pi'_U(\psi(\vec{x}))] \to [\mathrm{Pr}_U(\ulcorner \sigma(\dot{\vec{x}})\urcorner) \wedge \pi'_U(\sigma(\vec{x}))]\big)\Big).$$
This sentence is provable in $\mathbf{PA}$.
- If φ is $\forall \vec{x}(\Box\psi(\vec{x}) \to \psi(\vec{x}))$, then $\pi'_U(\varphi)$ is
$$\forall \vec{x}\big(\mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}})\urcorner) \wedge \pi'_U(\psi(\vec{x})) \to \pi'_U(\psi(\vec{x}))\big),$$
and this is obviously provable in $\mathbf{PA}$.
- If φ is $\forall \vec{x}(\Box\psi(\vec{x}) \to \Box\Box\psi(\vec{x}))$, then $\pi_U(\varphi)$ is
$$\forall \vec{x}\Big([\mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}})\urcorner) \wedge \pi'_U(\psi(\vec{x}))] \\ \to [\mathrm{Pr}_U(\ulcorner \Box\psi(\dot{\vec{x}})\urcorner) \wedge \mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}})\urcorner) \wedge \pi'_U(\psi(\vec{x}))]\Big),$$
and this is provable in $\mathbf{PA}$.
- If $\varphi(\vec{x})$ is derived from $\psi(\vec{x})$ by Nec, then $\varphi(\vec{x}) \equiv \Box\psi(\vec{x})$. Since $T \vdash \psi(\vec{x})$, by the induction hypothesis, $\mathbf{PA} \vdash \pi'_U(\psi(\vec{x}))$. Also, $\mathbf{PA} \vdash \mathrm{Pr}_U(\ulcorner \psi(\dot{\vec{x}})\urcorner)$ because U is an extension of T. Thus, $\mathbf{PA} \vdash \pi'_U(\varphi(\vec{x}))$. □

As in the proof of Proposition 6.47, we can prove the following proposition from Propositions 6.45 and 6.50.

Proposition 6.51 (Cf. [21, TB]). *Let T be an $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* $\mathbf{PA}(\mathbf{S4})$ *into* $\mathbf{PA}_\Box$*, and let U be any r.e. extension of T. If $\mathbb{N} \models \pi'_U(\varphi)$ for all $\varphi \in U \setminus T$, then U is $\mathcal{L}_A(\Box)$-sound.*

Corollary 6.52. *The theories* $\mathbf{PA}(\mathbf{KT})$ *and* $\mathbf{PA}(\mathbf{S4})$ *are $\mathcal{L}_A(\Box)$-sound.*

The alternative $\mathcal{L}_A(\Box)$-soundness of $\mathbf{PA}(\mathbf{S4})$ is already proved by Shapiro [21, TB'].

Corollary 6.53. *Let T be an r.e. $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* $\mathbf{PA}(\mathbf{S4}) + \{\Box\Diamond\varphi \to \varphi \mid \varphi$ *is an $\mathcal{L}_A$-sentence$\}$ into* $\mathbf{PA}_\Box$*. Then, T is $\mathcal{L}_A(\Box)$-sound.*

Proof. By Proposition 6.51, it suffices to show that for any $\mathcal{L}_A$-sentence φ, if $\Box\Diamond\varphi \to \varphi \in T$, then $\mathbb{N} \models \pi'_T(\Box\Diamond\varphi \to \varphi)$. Suppose that $\Box\Diamond\varphi \to \varphi \in T$ and $\mathbb{N} \models \pi'_T(\Box\Diamond\varphi)$. Then, $T \vdash \Diamond\varphi$, and $T \vdash \Box\Diamond\varphi$. Since $T \vdash \Box\Diamond\varphi \to \varphi$, we have $T \vdash \varphi$. Since T is a subtheory of $\mathbf{PA}(\mathbf{Triv})$, T is an conservative extension of $\mathbf{PA}$ by Corollary 2.4. Then, $\mathbf{PA} \vdash \varphi$ because φ is an $\mathcal{L}_A$-sentence. By the $\mathcal{L}_A$-soundness of $\mathbf{PA}$, we have $\mathbb{N} \models \varphi$ and hence $\mathbb{N} \models \pi'_T(\varphi)$. We have proved $\mathbb{N} \models \pi'_U(\Box\Diamond\varphi \to \varphi)$. □

In contrast to Corollary 6.53, we have the following proposition which is a refinement of Proposition 4.28.1.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

PROPOSITION 6.54. *Let T be a consistent r.e. $\mathcal{L}_A(\Box)$-theory extending the theory* $\mathbf{PA}(\mathbf{KT}) + \{\Box\Diamond\varphi \to \varphi \mid \varphi$ *is a* $\Sigma(\mathrm{B})$*-sentence*$\}$. *Then, T does not have* (B, Σ_1)-DP.

Proof. Let φ be a Σ_1 sentence such that $T \nvdash \varphi$ and $T \nvdash \neg\varphi$. Then, $T \nvdash \Box\neg\varphi$ because $T \vdash \Box\neg\varphi \to \neg\varphi$. Since $T \vdash \Box\Diamond\varphi \to \Diamond\varphi$, we have $T \vdash \Diamond\Box\neg\varphi \vee \Diamond\varphi$. Then, $T \vdash \Diamond(\Box\neg\varphi \vee \varphi)$ and hence $T \vdash \Box\Diamond(\Box\neg\varphi \vee \varphi)$. Since $\Box\neg\varphi \vee \varphi$ is a $\Sigma(\mathrm{B})$ sentence, we obtain $T \vdash \Box\neg\varphi \vee \varphi$ because $\Box\Diamond(\Box\neg\varphi \vee \varphi) \to (\Box\neg\varphi \vee \varphi)$ is an axiom of T. We have shown that $T \nvdash \Box\neg\varphi$, $T \nvdash \varphi$, and $T \vdash \Box\neg\varphi \vee \varphi$. This means that T does not have (B, Σ_1)-DP. $\Box$

6.2. $\Sigma(\mathrm{B})$***-soundness and weak*** $\Sigma(\mathrm{B})$***-soundness.*** We then export the notion of the Σ_1-soundness to modal arithmetic. This is easy to do since we have already introduced the class $\Sigma(\mathrm{B})$ corresponding to Σ_1 in modal arithmetic. Here we further introduce another type of translation ρ_T, which is different from π_T.

DEFINITION 6.55 (ρ-translations). *Let T be any r.e. $\mathcal{L}_A(\Box)$-theory. We define a translation ρ_T of $\mathcal{L}_A(\Box)$-formulas into $\mathcal{L}_A$-formulas inductively as follows*:

1. *If φ is an $\mathcal{L}_A$-formula, then* $\rho_T(\varphi) :\equiv \varphi$.
2. *ρ_T preserves logical connectives and quantifiers.*
3. $\rho_T(\Box\varphi(\vec{x})) :\equiv \mathrm{Pr}_T(\ulcorner \Box\varphi(\dot{\vec{x}})\urcorner)$.

With respect to $\Sigma(\mathrm{B})$ sentences, there is the following relationship between the translations π_T and ρ_T.

PROPOSITION 6.56. *Let T be any r.e. $\mathcal{L}_A(\Box)$-theory.*

1. *For any* $\Sigma(\mathrm{B})$*-sentence* φ, $\mathbb{N} \models \pi_T(\varphi) \to \rho_T(\varphi)$.
2. *If T is closed under the box elimination rule, then for any* $\Sigma(\mathrm{B})$*-sentence* φ, $\mathbb{N} \models \rho_T(\varphi) \to \pi_T(\varphi)$.

Proof. These statements are proved by induction on the construction of φ. We only prove the case of $\varphi \equiv \Box\psi$.

1. If $\mathbb{N} \models \pi_T(\Box\psi)$, then $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\psi\urcorner)$. Then, $T \vdash \psi$. By the rule NEC, $T \vdash \Box\psi$. Then, $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\Box\psi\urcorner)$, and hence $\mathbb{N} \models \rho_T(\Box\psi)$.

2. If $\mathbb{N} \models \rho_T(\Box\psi)$, then $T \vdash \Box\psi$. By the box elimination rule, $T \vdash \psi$. Hence, $\mathbb{N} \models \pi_T(\Box\psi)$. $\Box$

We strengthen the usual Σ_1-completeness theorem of **PA** as follows.

THEOREM 6.57 ($\Sigma(\mathrm{B})$-completeness theorem). *Let T be any r.e. extension of* $\mathbf{PA}_\Box$. *Then, for any* $\Sigma(\mathrm{B})$ *sentence* φ, *if* $\mathbb{N} \models \rho_T(\varphi)$, *then* $T \vdash \varphi$.

Proof. We prove the theorem by induction on the construction of φ.

- If φ is a Σ_1 sentence, then the statement immediately follows from the usual Σ_1-completeness of **PA** because $\rho_T(\varphi)$ is exactly φ.
- If φ is of the form $\Box\psi$, then $\mathbb{N} \models \rho_T(\Box\psi)$ means $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\Box\psi\urcorner)$, and hence $T \vdash \Box\psi$.
- If φ is one of the forms $\psi \wedge \sigma$, $\psi \vee \sigma$, $\exists x\psi$, and $\forall x < t\ \psi$, then the proof is straightforward by the induction hypothesis. $\Box$

In the light of Proposition 6.56 and Theorem 6.57, we introduce the following two different types of the notion of $\Sigma(\mathrm{B})$-soundness.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

Definition 6.58. *Let T be any r.e. $\mathcal{L}_A(\Box)$-theory.*

- *T is said to be* $\Sigma(\mathrm{B})$-sound *if for any* $\Sigma(\mathrm{B})$ *sentence* φ, *if* $T \vdash \varphi$, *then* $\mathbb{N} \models \pi_T(\varphi)$.
- *T is said to be* weakly $\Sigma(\mathrm{B})$-sound *if for any* $\Sigma(\mathrm{B})$ *sentence* φ, *if* $T \vdash \varphi$, *then* $\mathbb{N} \models \rho_T(\varphi)$.

Lemma 6.59. *For any r.e. $\mathcal{L}_A(\Box)$-theory T, the following are equivalent*:

1. *T is* $\Sigma(\mathrm{B})$-*sound.*
2. *T is weakly* $\Sigma(\mathrm{B})$-*sound and T is closed under the box elimination rule.*

Proof. By Proposition 6.56, it suffices to show that $\Sigma(\mathrm{B})$-soundness implies the box elimination rule. Suppose that T is $\Sigma(\mathrm{B})$-sound. Let φ be any $\mathcal{L}_A(\Box)$-sentence such that $T \vdash \Box\varphi$. By the $\Sigma(\mathrm{B})$-soundness of T, $\mathbb{N} \models \pi_T(\Box\varphi)$ and hence $\mathbb{N} \models \mathrm{Pr}_T(\ulcorner\varphi\urcorner)$. We obtain $T \vdash \varphi$. □

We are ready to prove an analogue of Guaspari's theorem.

Theorem 6.60. *Let T be an r.e. extension of* $\mathbf{PA}_\Box$.

1. *If T contains* $\mathbf{PA}(\mathbf{K})$, $T \nvdash \Box\bot$, *and T has* $\Sigma(\mathrm{B})$-DP, *then T is weakly* $\Sigma(\mathrm{B})$-*sound.*
2. *If T is weakly* $\Sigma(\mathrm{B})$-*sound, then T has* $\Sigma(\mathrm{B})$-EP.

Proof. 1. We prove by induction on the construction of φ that for any $\Sigma(\mathrm{B})$ sentence φ, if $T \vdash \varphi$, then $\mathbb{N} \models \rho_T(\varphi)$.

- If φ is a Σ_1 sentence, then $\rho_T(\varphi) \equiv \varphi$. Suppose $T \vdash \varphi$. Since $\Sigma(\mathrm{B})$-DP implies (B, Σ_1)-DP, T is Σ_1-sound by Proposition 4.21. Therefore, $\mathbb{N} \models \rho_T(\varphi)$.
- If φ is of the form $\Box\psi$, then $\rho_T(\varphi) \equiv \mathrm{Pr}_T(\ulcorner\Box\psi\urcorner)$. Suppose $T \vdash \Box\psi$. Then, obviously $\mathbb{N} \models \rho_T(\varphi)$.
- If φ is of the form $\psi \wedge \sigma$ or $\forall x < t\, \psi$, then the proof is straightforward from the induction hypothesis.
- If φ is $\psi \vee \sigma$, then $\rho_T(\varphi) \equiv \rho_T(\psi) \vee \rho_T(\sigma)$. Suppose $T \vdash \psi \vee \sigma$. Then, by $\Sigma(\mathrm{B})$-DP, $T \vdash \psi$ or $T \vdash \sigma$. By the induction hypothesis, $\mathbb{N} \models \rho_T(\psi)$ or $\mathbb{N} \models \rho_T(\sigma)$. Hence, $\mathbb{N} \models \rho_T(\varphi)$.
- If φ is $\exists x \psi(x)$, then $\rho_T(\varphi) \equiv \exists x \rho_T(\psi(x))$. Suppose $T \vdash \exists x \psi(x)$. Since $T \nvdash \Box\bot$, by Theorem 5.38, T has $\Sigma(\mathrm{B})$-EP. Then, there exists a natural number n such that $T \vdash \psi(\overline{n})$. By the induction hypothesis, $\mathbb{N} \models \rho_T(\psi(\overline{n}))$. Therefore, $\mathbb{N} \models \rho_T(\varphi)$.

2. Let $\varphi(x)$ be any $\Sigma(\mathrm{B})$ formula having no free variables except x such that $T \vdash \exists x \varphi(x)$. By the weak $\Sigma(\mathrm{B})$-soundness of T, $\mathbb{N} \models \rho_T(\exists x \varphi(x))$. Then, for some natural number n, $\mathbb{N} \models \rho_T(\varphi(\overline{n}))$. By Theorem 6.57, $T \vdash \varphi(\overline{n})$. □

Corollary 6.61. *Let T be an r.e. extension of* $\mathbf{PA}_\Box$.

1. *If T contains* $\mathbf{PA}(\mathbf{K4})$, *T is consistent, and T has* MDP, *then T is* $\Sigma(\mathrm{B})$-*sound.*
2. *If T is* $\Sigma(\mathrm{B})$-*sound, then T has* MEP.

Proof. 1. Since T contains $\mathbf{PA}(\mathbf{K4})$ and T has MDP, by Proposition 5.39, T has $\Sigma(\mathrm{B})$-DP. Also, T is closed under the box elimination rule by Proposition 4.17. Then, $T \nvdash \Box\bot$ by the consistency of T. By Theorem 6.60.1, T is weakly $\Sigma(\mathrm{B})$-sound. By Lemma 6.59, T is $\Sigma(\mathrm{B})$-sound.

2. By Lemma 6.59, T is weakly $\Sigma(\mathrm{B})$-sound and is closed under the box elimination rule. By Theorem 6.60.2, T has $\Sigma(\mathrm{B})$-EP. Therefore, by Proposition 4.17, T has MEP. □

Since the $\mathcal{L}_A(\Box)$-soundness implies the $\Sigma(\mathrm{B})$-soundness, we obtain the following corollary from Propositions 6.47 and 6.51.

COROLLARY 6.62. 1. *Let T be an $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* $\mathbf{PA}(\mathbf{GL})$ *into* $\mathbf{PA}_\Box$, *and let U be any r.e. extension of T. If $\mathbb{N} \models \pi_U(\varphi)$ for all $\varphi \in U \setminus T$, then U has* MEP.

2. *Let T be an $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of* $\mathbf{PA}(\mathbf{S4})$ *into* $\mathbf{PA}_\Box$, *and let U be any r.e. extension of T. If $\mathbb{N} \models \pi'_U(\varphi)$ for all $\varphi \in U \setminus T$, then U has* MEP.

In particular, $\mathbf{PA}_\Box$, $\mathbf{PA}(\mathbf{K})$, $\mathbf{PA}(\mathbf{KT})$, $\mathbf{PA}(\mathbf{K4})$, $\mathbf{PA}(\mathbf{S4})$, *and* $\mathbf{PA}(\mathbf{GL})$ *have* MEP.

By Lemma 6.59, each $\Sigma(\mathrm{B})$-sound theory is also weakly $\Sigma(\mathrm{B})$-sound. Therefore, $\mathbf{PA}_\Box$, $\mathbf{PA}(\mathbf{K})$, $\mathbf{PA}(\mathbf{KT})$, $\mathbf{PA}(\mathbf{K4})$, $\mathbf{PA}(\mathbf{S4})$, and $\mathbf{PA}(\mathbf{GL})$ also have $\Sigma(\mathrm{B})$-EP. Recall that $\mathbf{PA}(\mathbf{Verum})$ also has $\Sigma(\mathrm{B})$-EP (Corollary 5.30).

Here we give another sufficient condition for a theory to have $\Sigma(\mathrm{B})$-EP. First, we prove an analogue of Proposition 6.46 with respect to ρ-translations.

PROPOSITION 6.63. *Let T be any $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of the form $\forall\vec{x}(\Box\psi_0 \wedge \cdots \wedge \Box\psi_{k-1} \to \Box\psi_k)$ into* $\mathbf{PA}_\Box$. *Then, for any $\mathcal{L}_A(\Box)$-formula φ, if $T \vdash \varphi$, then for any r.e. extension U of T,* $\mathbf{PA} \vdash \rho_U(\varphi)$.

Proof. Let U be any r.e. extension of T. As in the proof of Proposition 2.3, we prove by induction on the length of proofs of φ in T that for any $\mathcal{L}_A(\Box)$-formula φ, if $T \vdash \varphi$, then $\mathbf{PA} \vdash \rho_U(\varphi)$. We only give proofs of the following two cases.

- The case $\varphi \equiv \forall\vec{x}\big(\Box\psi_0(\vec{x}) \wedge \cdots \wedge \Box\psi_{k-1}(\vec{x}) \to \Box\psi_k(\vec{x})\big)$: Since U is an extension of T, we have $\mathbf{PA} \vdash \forall\vec{x}\, \mathrm{Pr}_U(\ulcorner \Box\psi_0(\dot{\vec{x}}) \wedge \cdots \wedge \Box\psi_{k-1}(\dot{\vec{x}}) \to \Box\psi_k(\dot{\vec{x}})\urcorner)$, and hence $\mathbf{PA}$ proves
$$\forall\vec{x}\big(\mathrm{Pr}_U(\ulcorner\Box\psi_0(\dot{\vec{x}})\urcorner) \wedge \cdots \wedge \mathrm{Pr}_U(\ulcorner\Box\psi_{k-1}(\dot{\vec{x}})\urcorner) \to \mathrm{Pr}_U(\ulcorner\Box\psi_k(\dot{\vec{x}})\urcorner)\big).$$
This sentence is exactly $\rho_U(\varphi)$.
- If $\varphi(\vec{x})$ is derived from $\psi(\vec{x})$ by NEC, then $\varphi(\vec{x}) \equiv \Box\psi(\vec{x})$. Since $T \vdash \Box\psi(\vec{x})$, $U \vdash \Box\psi(\vec{x})$, and hence $\mathbf{PA} \vdash \mathrm{Pr}_U(\ulcorner\Box\psi(\dot{\vec{x}})\urcorner)$. Thus, $\mathbf{PA} \vdash \rho_U(\varphi(\vec{x}))$. □

COROLLARY 6.64. *Let T be any $\mathcal{L}_A(\Box)$-theory obtained by adding some axioms of the form $\forall\vec{x}(\Box\psi_0 \wedge \cdots \wedge \Box\psi_{k-1} \to \Box\psi_k)$ into* $\mathbf{PA}_\Box$, *and let U be any r.e. extension of T. If $\mathbb{N} \models \rho_U(\varphi)$ for all $\varphi \in U \setminus T$, then U has* $\Sigma(\mathrm{B})$-EP.

Proof. Suppose that $\mathbb{N} \models \rho_U(\varphi)$ for all $\varphi \in U \setminus T$. As in the proof of Proposition 6.47, it follows from Proposition 6.63 that U is weakly $\Sigma(\mathrm{B})$-sound. Then, by Theorem 6.60, U has $\Sigma(\mathrm{B})$-EP. □

6.3. Applications. In this subsection, as applications of our results we have obtained so far, we show two non-implications between the properties. Corollary 5.30 shows that in general, $\Sigma(\mathrm{B})$-DP does not imply MDP. The first application shows that this is also true for theories that do not contain $\mathbf{PA}(\mathbf{Verum})$.

PROPOSITION 6.65. 1. *There exists an r.e. theory T such that* $\mathbf{PA}(\mathbf{S4}) \vdash T \vdash \mathbf{PA}(\mathbf{K4})$, *$T \nvdash \Box\bot$, T has* $\Sigma(\mathrm{B})$-EP, *and T does not have* MDP;

2. *There exists an r.e. theory T such that* $\mathbf{PA}(\mathbf{Verum}) \vdash T \vdash \mathbf{PA}(\mathbf{GL})$, *$T \nvdash \Box\bot$, T has* $\Sigma(\mathrm{B})$-EP, *and T does not have* MDP.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

Proof. 1. Let T be the theory $\mathbf{PA}(\mathbf{K4}) + \{\Box\neg\Box\bot\}$. Since T is a subtheory of $\mathbf{PA}(\mathbf{Triv})$, we have $T \nvdash \Box\bot$. By Corollary 6.64, T has $\Sigma(\mathrm{B})$-EP. Suppose, towards a contradiction, that $T \vdash \neg\Box\bot$. Then, by the $\Sigma(\mathrm{B})$-deduction theorem (Corollary 3.14), $\mathbf{PA}(\mathbf{K4}) \vdash \Box\neg\Box\bot \to \neg\Box\bot$. Then, $\mathbf{PA}(\mathbf{Verum}) \vdash \Box\neg\Box\bot \to \neg\Box\bot$. Since $\mathbf{PA}(\mathbf{Verum}) \vdash \Box\neg\Box\bot \wedge \Box\bot$, this contradicts the consistency of $\mathbf{PA}(\mathbf{Verum})$. Therefore, $T \nvdash \neg\Box\bot$. Since $T \vdash \Box\neg\Box\bot$, T does not have MDP.

2. Let $T := \mathbf{PA}(\mathbf{GL}) + \{\Box\Box\bot\}$. By Corollary 6.64, T has $\Sigma(\mathrm{B})$-EP. Suppose, towards a contradiction, that T proves $\Box\bot$. By the $\Sigma(\mathrm{B})$-deduction theorem, $\mathbf{PA}(\mathbf{GL}) \vdash \Box\Box\bot \to \Box\bot$. Then, $\mathbf{PA}(\mathbf{GL}) \vdash \Box(\Box\Box\bot \to \Box\bot)$ and hence $\mathbf{PA}(\mathbf{GL}) \vdash \Box\Box\bot$. By MDP of $\mathbf{PA}(\mathbf{GL})$ (Corollary 6.62 and Proposition 4.18.1), $\mathbf{PA}(\mathbf{GL}) \vdash \bot$. This is a contradiction. Therefore, $T \nvdash \Box\bot$. Since $T \vdash \Box\Box\bot$, T does not have MDP. □

Unlike the notion of the soundness of $\mathcal{L}_A$-theories, Proposition 6.65.1 shows that the $\mathcal{L}_A(\Box)$-soundness is not preserved by taking a subtheory because $\mathbf{PA}(\mathbf{S4})$ is $\mathcal{L}_A(\Box)$-sound but T is not $\Sigma(\mathrm{B})$-sound.

The second application shows that $\Sigma(\mathrm{B})$-DC does not imply B-DP in general.

PROPOSITION 6.66. *There exists a consistent r.e. extension T of* $\mathbf{PA}(\mathbf{K4})$ *satisfying the following two conditions*:

1. *T is* $\Sigma(\mathrm{B})$-DC.
2. *T does not have* B-DP.

Proof. Let φ be a Gödel sentence of $\mathbf{PA}$. Let $T := \mathbf{PA}(\mathbf{K4}) + \{\Box\varphi \vee \Box\neg\varphi\}$, $T_0 := \mathbf{PA}(\mathbf{K4}) + \{\Box\varphi\}$, and $T_1 := \mathbf{PA}(\mathbf{K4}) + \{\Box\neg\varphi\}$. By the $\Sigma(\mathrm{B})$-deduction theorem, it is shown that for any $\mathcal{L}_A(\Box)$-formula ψ,

$$T \vdash \psi \text{ if and only if both } T_0 \vdash \psi \text{ and } T_1 \vdash \psi. \tag{6}$$

Suppose, towards a contradiction, $T_0 \vdash \Box\bot$. By the $\Sigma(\mathrm{B})$-deduction theorem, $\mathbf{PA}(\mathbf{K4})$ proves $\Box\varphi \to \Box\bot$. Since this is also provable in $\mathbf{PA}(\mathbf{Triv})$, by Proposition 2.3, we have $\mathbf{PA} \vdash \alpha(\Box\varphi) \to \alpha(\Box\bot)$. Then, $\mathbf{PA} \vdash \neg\varphi$, a contradiction. Similarly, we can prove $T_1 \nvdash \Box\bot$.

1. Let ψ be any $\Sigma(\mathrm{B})$ sentence such that $T \vdash \psi \vee \mathrm{Pr}_T(\ulcorner\psi\urcorner)$. Then, for $i \in \{0,1\}$, $T_i \vdash \psi \vee \mathrm{Pr}_T(\ulcorner\psi\urcorner)$ by (6). By Corollary 6.64, T_i has $\Sigma(\mathrm{B})$-EP, and hence has $(\Sigma(\mathrm{B}), \Sigma_1)$-DP. By Corollary 5.32, T_i is $\Sigma(\mathrm{B})$-DC. Therefore, $T_i \vdash \psi$. By (6), we obtain $T \vdash \psi$. Thus, T is also $\Sigma(\mathrm{B})$-DC.

2. If $T \vdash \Box\neg\varphi$ or $T \vdash \Box\varphi$, then $T_0 \vdash \Box\bot$ or $T_1 \vdash \Box\bot$ by (6). This is a contradiction. Therefore, $T \nvdash \Box\neg\varphi$ and $T \nvdash \Box\varphi$. On the other hand, $T \vdash \Box\varphi \vee \Box\neg\varphi$. Thus, T does not have B-DP. □

§7. Problems.

In the present paper, several properties related to the modal disjunction property in modal arithmetic are introduced, and the relationships between them are studied. However, some of the properties have not yet been separated in some particular situation. In this section, we list several unsolved problems for further study.

In Section 4, we introduced B-DP and $\Delta(\mathrm{B})$-DP. For theories which are closed under the box elimination rule, these properties are equivalent. However, we have not yet been successful in clarifying whether they are equivalent or not in general. We propose the following problem.

PROBLEM 7.67. 1. *Does there exist an* $\mathcal{L}_A(\Box)$*-theory which has* $\Delta(\mathrm{B})$-DP *but does not have* B-DP?

2. *For each* $n \geq 2$, *does there exist an* $\mathcal{L}_A(\Box)$*-theory which has* B^n-DP *but does not have* B^{n+1}-DP?

For any Σ_1-unsound r.e. extension T of $\mathbf{PA}(\mathbf{Verum})$, T has B-DP but does not have $\Sigma(\mathrm{B})$-DC (see Propositions 4.25 and 5.29). On the other hand, for consistent r.e. extensions of $\mathbf{PA}(\mathbf{S4})$, B-DP implies $\Sigma(\mathrm{B})$-DC by Propositions 4.17 and 5.39 and Corollary 5.32. We have not yet been sure whether B-DP yields $\Sigma(\mathrm{B})$-DC in general when $T \nvdash \Box\bot$.

PROBLEM 7.68. *Does there exist an* $\mathcal{L}_A(\Box)$*-theory* T *such that* $T \nvdash \Box\bot$, T *has* B-DP, *and* T *is not* $\Sigma(\mathrm{B})$-DC?

In the statement of Proposition 5.39, it is assumed that T is an extension of $\mathbf{PA}(\mathbf{K4})$. It is not settled yet whether $\mathbf{PA}(\mathbf{K4})$ can be replaced by $\mathbf{PA}(\mathbf{K})$ in the statement.

PROBLEM 7.69. *In the statements of Proposition* 5.39 *and Corollaries* 5.40 *and* 6.61.1, *can* $\mathbf{PA}(\mathbf{K4})$ *be replaced by* $\mathbf{PA}(\mathbf{K})$?

Proposition 6.65.1 shows that there exists an $\mathcal{L}_A(\Box)$-unsound subtheory T of $\mathbf{PA}(\mathbf{S4})$. Related to this fact, we propose the following problem.

PROBLEM 7.70. *Does there exist an* $\mathcal{L}_A(\Box)$*-unsound r.e. subtheory of* $\mathbf{PA}(\mathbf{GL})$?

Proposition 6.66 shows that $\Sigma(\mathrm{B})$-DC does not imply B-DP. We are not successful in determining whether the theory T in the proof of Proposition 6.66 is closed in the box elimination rule. We then propose the following problem.

PROBLEM 7.71. *Does there exist a consistent r.e.* $\mathcal{L}_A(\Box)$*-theory* T *such that* T *is closed under the box elimination rule,* T *is* $\Sigma(\mathrm{B})$-DC, *and* T *does not have* MDP?

REMARK 7.72. *Notice that if we define* T *to be the theory* $\mathbf{PA}(\mathbf{K4}) + \{\Box\varphi \vee \Box\neg\varphi\}$ *for a* Π_1 *Gödel sentence* φ *of* $\mathbf{PA}(\mathbf{K4})$, *then* T *is not closed under the box elimination rule. This is because* $T \vdash \Box(\Box\varphi \vee \neg\varphi)$ *and* $T \nvdash \Box\varphi \vee \neg\varphi$. *For, if* $T \vdash \Box\varphi \vee \neg\varphi$, *then* $\mathbf{PA}(\mathbf{K4}) \vdash \Box\neg\varphi \to (\Box\varphi \vee \neg\varphi)$. *By Proposition* 6.46, $\mathbf{PA} \vdash \mathrm{Pr}_{\mathbf{PA}(\mathbf{K4})}(\ulcorner\neg\varphi\urcorner) \to (\mathrm{Pr}_{\mathbf{PA}(\mathbf{K4})}(\ulcorner\varphi\urcorner) \vee \neg\varphi)$. *Since* $\mathrm{Pr}_{\mathbf{PA}(\mathbf{K4})}(\ulcorner\varphi\urcorner)$ *is* $\mathbf{PA}(\mathbf{K4})$*-equivalent to* $\neg\varphi$, *we have* $\mathbf{PA}(\mathbf{K4}) \vdash \mathrm{Pr}_{\mathbf{PA}(\mathbf{K4})}(\ulcorner\neg\varphi\urcorner) \to \neg\varphi$. *By Löb's theorem,* $\mathbf{PA}(\mathbf{K4}) \vdash \neg\varphi$. *This contradicts the* Σ_1*-soundness of* $\mathbf{PA}(\mathbf{K4})$.

Acknowledgments. The authors would like to thank the anonymous referees for their valuable comments and suggestions on earlier versions of this paper. This work was partly supported by JSPS KAKENHI Grant Number JP19K14586.

BIBLIOGRAPHY

[1] Boolos, G. (1993). *The Logic of Provability*. Cambridge: Cambridge University Press.

[2] Buchholz, W. (1993). Mathematische Logik II. Available from: http://www.mathematik.uni-muenchen.de/buchholz/articles/LogikII.ps.

[3] Došen, K. (1990). Modal translations of Heyting and Peano arithmetic. *Publications de l'Institut Mathématique. Nouvelle Série*, **47**, 13–23.

[4] Flagg, R. C. & Friedman, H. (1986). Epistemic and intuitionistic formal systems. *Annals of Pure and Applied Logic*, **32**, 53–60.

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press

[5] Friedman, H. (1975). The disjunction property implies the numerical existence property. *Proceedings of the National Academy of Sciences of the United States of America*, **72**, 2877–2878.

[6] Friedman, H. & Sheard, M. (1989). The equivalence of the disjunction and existence properties for modal arithmetic. *The Journal of Symbolic Logic*, **54**(4), 1456–1459.

[7] Gentzen, G. (1934–1935). Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, **39**(2–3), 176–210, 405–431.

[8] Gödel, K. (1933). An interpretation of the intuitionistic propositional calculus. In Feferman, S., editor. *Collected Works*, Vol. 1. Oxford: Oxford University Press, pp. 301–303.

[9] Goodman, N. D. (1984). Epistemic arithmetic is a conservative extension of intuitionistic arithmetic. *The Journal of Symbolic Logic*, **49**, 192–203.

[10] Guaspari, D. (1979). Partially conservative extensions of arithmetic. *Transactions of the American Mathematical Society*, **254**, 47–68.

[11] Halbach, V. & Horsten, L. (2000). Two proof-theoretic remarks on EA + ECT. *Mathematical Logic Quarterly*, **46**(4), 461–466.

[12] Hughes, G. E. & Cresswell, M. J. (1996). *A New Introduction to Modal Logic*. London: Routledge.

[13] Jensen, D. C. & Ehrenfeucht, A. (1976). Some problem in elementary arithmetics. *Fundamenta Mathematicae*, **92**, 223–245.

[14] Kleene, S. C. (1945). On the interpretation of intuitionistic number theory. *The Journal of Symbolic Logic*, **10**, 109–124.

[15] Kurahashi, T. (2018). On partial disjunction properties of theories containing Peano arithmetic. *Archive for Mathematical Logic* **57**(7–8), 953–980.

[16] ———. (2020). A note on derivability conditions. *The Journal of Symbolic Logic*, **85**(3), 1224–1253.

[17] Makinson, D. (1971). Some embedding theorems for modal logic. *Notre Dame Journal of Formal Logic*, **12**, 252–254.

[18] Rautenberg, W. (2010). *A Concise Introduction to Mathematical Logic* (third revised and enlarged edition). London: Springer.

[19] Reinhardt, W. N. (1985). Absolute versions of incompleteness theorems. *Nôus*, **19**(3), 317–346.

[20] ———. (1986). Epistemic theories and the interpretation of Gödel's incompleteness theorems. *Journal of Philosophical Logic*, **15**, 427–474.

[21] Shapiro, S. (1985). Epistemic and intuitionistic arithmetic. In *Intensional Mathematics* (Stewart Shapiro (ed.)). North-Holland: Amsterdam. Studies in Logic and the Foundations of Mathematics, Vol. 113, pp. 11–46.

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

GRADUATE SCHOOL OF SYSTEM INFORMATICS
KOBE UNIVERSITY
1-1 ROKKODAI, NADA, KOBE 657-8501, JAPAN
E-mail: kurahashi@people.kobe-u.ac.jp

INDEPENDENT SCHOLAR

https://doi.org/10.1017/S1755020322000363 Published online by Cambridge University Press