ANOTHER LOOK AT THE SECOND INCOMPLETENESS THEOREM

ALBERT VISSER

Abstract. In this paper we study proofs of some general forms of the Second Incompleteness Theorem. These forms conform to the Feferman format, where the proof predicate is fixed and the representation of the axiom set varies. We extend the Feferman framework in one important point: we allow the interpretation of number theory to vary.

1. A Dialogue

Alcibiades: Hi, Socrates. You don't know how happy am to see you. I am thoroughly confused and you're just the man to liberate me of this annoying puzzlement.

Socrates: I am flattered that such a popular young person still needs an old man who is not even on twitter. As you know, I think confusion is a good thing. It is an important step on the road to insight. What is your puzzlement about?

Alcibiades: Well, you remember that we were taught about the Second Incompleteness Theorem in the Gymnasium? A theory cannot prove its own consistency? Arithmetization? Great stuff. I worked hard and I dare say that I obtained a decent understanding of the proof.

Socrates: I do remember you did very well in the exam.

Alcibiades: However, now I have been reading Feferman's paper *Arithmetization of Metamathematics*. He gives an example of an axiomatization of Peano Arithmetic such that Peano Arithmetic can prove its own consistency with respect to that axiomatization.

Socrates: I commend you on your diligence. Reading the *Arithmetization* is an important step on the road to wisdom. Let no one say ever again that Alcibiades is only an irresponsible rascal and party animal. But, to be honest, I still do not see the source of your puzzlement. The Second Incompleteness Theorem is applicable when certain conditions are fulfilled and Feferman's clever axiomatization does not fulfill these conditions. That's how it is able to escape Gödel's Second.

Alcibiades: But, you see, Socrates, I seem to be able to prove that any axiomatization, under minimal conditions, must obey Gödel's Second. Moreover, the proof is very simple, just an application of the Compactness Theorem.

Date: July 19, 2017.

2000 *Mathematics Subject Classification.* 03A05, 03B25, 03F25, 03F30, 03F45.

Key words and phrases. Second Incompleteness Theorem, Formal Theories.

I am grateful to Richard Heck whose questions and remarks gave me the idea for this paper. I thank Jetze Zoethout for helpful insights.

Socrates: Bring it on.

Alcibiades: Here it is. Suppose a consistent theory, for the given axiomatization, proves its own consistency. Let us call this theory *Theory*. Then, by compactness, there must be a finitely axiomatized sub-theory that already proves the consistency statement for our original theory. I will call this sub-theory simply *Sub-theory*. Since Sub-theory proves the consistency of Theory, it must also prove its own consistency. So, we have a finitely axiomatized theory that proves its own consistency. But, with the finite axiomatization, we can have no Fefermanian funny business, so, the Second Incompleteness Theorem applies in its full glory and we have a contradiction. It follows that Theory cannot prove its own consistency after all.

Socrates: I can see why you are puzzled. But, here, can you explain to me why we can infer that Sub-theory proves its own consistency, from the fact that it proves the consistency of Theory?

Alcibiades: Isn't that obvious? The consistency of the whole implies the consistency of the part.

Socrates: You are certainly right about that. It is not only true that the consistency of the whole implies the consistency of the part, what is more: theories with a modicum of arithmetic verify this important principle.

Alcibiades: But if this is right, then what can be wrong with my argument?

Socrates: To understand these matters, we have to carefully distinguish the internal perspectives of the theories we are considering and our own external perspective. There are three perspectives here: ours, the perspective of Theory and the perspective of Sub-theory.

What we have seen is that the principle that the consistency of the part is implied by the consistency of the whole is validated from all perspectives.

We also know that Sub-theory is part of Theory in our perspective. What we need is that Sub-theory knows —this is the relevant perspective for your inference— that it is a part of Theory. How does it know that?

Alcibiades: It knows that by proving, for each of its axioms, that the axiom is in the internally represented axiom set of Theory. To be able to speak about provability-in-Theory inside Theory at all, we have to agree that our internal representation of the axiom set is such that, for any axiom of Theory, it proves that the axiom is in the axiom set as internally represented.

Socrates: Admirably said. So, Theory knows of each of its axioms that it is an axiom according to the internal representation. Thus, Theory knows that Sub-theory is part of Theory. But how does it follow that Sub-theory knows of each of its own axioms that that axiom is also an axiom of Theory? One would expect that Sub-theory, being a finite part, cannot automatically do everything that Theory can.

Alcibiades: I start to see some light at the end of the tunnel, but let me still try to push the argument a bit further. Clearly, Sub-theory, when it is very weak, need not be able to do this. But, for my argument, we can take the finite part as large as we want. So simply take a finite sub-theory that can verify that its axioms are axioms of the original theory according to the internal representation.

Socrates: But if you extend the finite sub-theory, you also make the task of the finite sub-theory heavier: it has to prove of more axioms that they are in the axiom set as internally presented. Could it not happen that no finite sub-theory can do this?

Alcibiades: By the dog and by Zeus, Socrates, it seems that you unraveled the mystery. We can escape my argument in case, for every given finite set of axioms of Theory that proves the consistency of Theory, we need more axioms than there are in the given set to verify that the axioms of the given set are indeed axioms according to the given presentation.

Socrates: If you analyze Feferman's clever example, you will see that this is precisely what is going on there.

Alcibiades: I am relieved. Now I can go to Aristophanes' party this evening without having to think about the darned problem all the time.

Socrates: I am glad that I was able to contribute to that.

2. Introduction

The present paper is, in the sense, a footnote to Feferman's great arithmetization paper [Fef60]. More precisely, it is a footnote to the part of Feferman's paper that is concerned with the Second Incompleteness Theorem. In fact the present paper started with the puzzlement voiced by Alcibiades in the dialogue.

We present some general versions of the the Second Incompleteness Theorem. Clearly, the Second Incompleteness Theorem can be generalized in many ways and in many directions. For example, the Löb Conditions provide one such generalization: roughly, suppose N interprets the Tarski-Mostowski-Robinson theory R in U, in symbols $N : \mathsf{R} \lhd U$, and suppose the U-predicate $\Box$ satisfies the Löb Conditions w.r.t. sentences coded in the N-numerals, then $\Box$ also satisfies Löb's Principle. Another important generalization is Feferman's Theorem of the interpretability of inconsistency. Incompleteness, in this generalization, is not failure to prove consistency, but rather the ability to build an internal model of the theory itself plus its inconsistency statement.

Our generalization takes a different direction. We keep —in the Feferman style— our proof-predicate and its arithmetization fixed and vary the formula α representing an axiom-set. There is one extra feature that we allow to vary that is constant in Feferman's paper: the interpretation of arithmetic. Thus, we do not just have to specify the predicate α that defines the axiom set but also an interpretation that tells us where the numbers live. This leads us to an device $\mathcal{A}$, *the presentation*, that translates the language of arithmetic plus an extra predicate A into the language of the given theory. The notion of presentation is worked out in Section 4.

We provide, in Section 5, a sufficient condition on the presentation $\mathcal{A}$ for the validity of the Second Incompleteness Theorem for $\mathcal{A}$, to wit: *being a uniform presentation*, roughly, there are arbitrarily large finite approximations U_0 of the theory U, such that $\mathcal{A}$ semi-numerates the axioms of U_0 in U_0.

A major special case is formed by the Σ^0_1-presentations. Here the axiom set of U is numerated in U by a Σ^0_1-formula σ relativized to a suitable interpretation N of arithmetic. We discuss Σ^0_1-presentations in Section 6. We zoom in on the case where the arithmetic involved is EA. We provide some examples to liberate the reader of the impression that the case at hand is already completely clear. Let us say that a theory is Σ^0_1-numerable if there is a Σ^0_1-presentation that numerates an axiom set of the theory in the theory.

- We have an example of a finitely axiomatized theory B and a Σ^0_1-formula γ that numerates the axiom set in $N_0 : \mathsf{EA} \lhd A$ but not in $N_1 : \mathsf{EA} \lhd A$. For N_0 the Second Incompleteness Theorem applies, but for N_1, we have that B proves the consistency of the theory axiomatized by γ. See Example 6.3.
- We have example of a Σ^0_1-presentation that enumerates the axioms of a given theory U in U, but not uniformly. See Subsection 6.1.
- We have an example that shows that Σ^0_1-numerable theories need not be recursively enumerable but can be arbitrarily complex. See Subsection 6.1.
- We have an example of a Σ^0_1-presentation of Elementary Arithmetic EA that defines a finite axiom set in reality and is believed by EA to define a finite axiom set, for which the Löb conditions fail. Specifically, EA cannot prove the formalized Second Incompleteness Theorem for this presentation.[1] See Subsection 6.2.

We use a variant of our main theorem to prove the full second incompleteness theorem for Σ^0_1-numerations of the axiom set for the case that the interpreted number theory is (at least) EA.

In Section 7, we have a brief look at some salient examples of non-Σ^0_1-presentations over PA. We discuss the well-known Feferman predicate. We give an example where we have a uniform presentation for which the Löb Conditions cannot be verified.

In Appendix A, we provide the basics of translations and interpretations.

3. Preliminaries

In this section, we present some some basic definitions.

3.1. **Theories.** In the present paper, a theory is given by a signature and a set of sentences of that signature closed under deduction. We only consider finite signatures. I guess we can allow countable signatures but if we allow these we need some constraints on the effectiveness of the presentation. Our signatures are officially relational but, since we have a p-time term-elimination algorithm, for most purposes, we can pretend that we have terms.

We allow the theory to be of any complexity. It might be Π^{77}_{88}, or it might be outside any known hierarchy.

An axiom set X for a theory U is simply a set of sentences of the signature of the theory such that the deductive closure $\overline{X}$ of X is equal to U. The axioms for identity will be treated as part of the theory and not of the logic. Thus, for example, an axiomatization of Elementary Arithmetic is suppose to include axioms that imply the usual axioms for identity.

We employ some special theories in this paper. These are the Tarski-Mostowski-Robinson theory R (see [TMR53]), Robinson's arithmetic Q (see [TMR53]), Buss' theory S^1_2 (see [Bus86], [HP93]), Elementary Arithmetic EA (also known as EFA or

[1] This example is a good caveat for Alcibiades' naive use of 'finitely axiomatized theory'. What he has in mind is something like a representation of the axiom set that involves a finite disjunction of formulas of the form $(x = \ulcorner A \urcorner)^N$. For this very special presentation the relevant part of his argument works, but, as the example illustrates, there are readings of 'finitely axiomatizable' that share all the pitfalls of intensionality.

$\mathrm{I}\Delta_0 + \mathsf{exp}$, see [HP93]) and Peano Arithmetic (see [HP93], [Kay91]). For all these theories we work with variants in the signature of arithmetic Ar.

3.2. **Translations and Interpretations.** We treat the definitions of this subsection in more detail in Appendix A.

A *translation* τ from signature Θ to signature Ξ is basically a mapping from the predicates of Θ to Ξ-formulas. The translation can be lifted to the whole Θ-language such that the result commutes with the connectives of predicate logic.

There are two extra features. The first is domain relativization. The translation provides a domain. When we lift the translation to the full language the translated quantifiers are relativized to this domain.

The second feature is the treatment of variables. The simplest possibility is that our translation is 1-dimensional and parameter-free. In this case, we simply demand that the translation of an n-ary P is of the form $A(v_0, \ldots, v_{n-1})$ with all free variables among those shown, where the v_i are from a fixed infinite list $v_0, v_1, \ldots$ We then translate $P(x_0, \ldots, x_{n-1})$ by $A(x_0, \ldots, x_{n-1})$ (under an appropriate convention to handle variable-clashes). In case we allow parameters, the translation of P has the form $A(w_0, \ldots, w_{k-1}, v_0, \ldots, v_{n-1})$, for a fixed k. The w_i are supposed to be distinct from the v_j. When we lift the translation to the full Θ-language, we have to take care that none of the parameters gets bound. Finally, a translation could have dimension $m > 1$. In this case we send P to $A(\vec{v}_0, \ldots, \vec{v}_{n-1})$, where the $\vec{v}_i$ are pairwise disjoint sequences of variables of length m. In this case $P(x_0, \ldots, x_{n-1})$ is translated to $A(\vec{x}_0, \ldots, \vec{x}_{n-1})$ and we need some bookkeeping to assign sequences of variables of the translated language to variables of the translating language. We can combine more-dimensionality with parameters in the obvious way.

We write B^τ for the τ-translation of B of the Θ-language in the Ξ-language.

An *interpretation* K of a theory U in a theory V is a triple $\langle U, \tau, V\rangle$, where τ translates the signature of U into the signature of V. We demand that, for all U-sentences A, if $U \vdash A$, then $V \vdash A^\tau$.[2] We write $K : U \rhd V$ or $K : V \lhd U$ in case $\langle U, \tau, V\rangle$ is an interpretation. Par abus de langage, we write B^K for B^{τ_K}.

3.3. **Arithmetization.** We work in the Feferman-style, where the arithmetization of provability is fixed. It is somewhat ironic that the arithmetization given by Feferman as the one that should be fixed once and for all is one that we cannot adopt. The size of a Feferman code is superexponential (!) in the length of the formula. We want a code to be of order $2^{P(n)}$, where P is a polynomial and n is the length of the formula. In this way we can work naturally with our coding in Buss' theory S^1_2. Thus, we fix an efficient coding. The codings of [Bus86] or of [Zam96] would do, or, more precisely, a reworking of those codings for the proper arithmetical language.[3]

The default in this paper is to use efficient numerals: these simulate dyadic notation. We use $\ulcorner A \urcorner$ ambiguously for the Gödel number of A and for the numeral of the Gödel number of A. We will employ Smoryński's dot notation. E.g. $\ulcorner A(\dot{x})\urcorner$ stands for the arithmetization of the function that sends a formula $A(x)$ and a

[2] In case we have parameters the definition should be slightly expanded.

[3] Of course, the choice of an arithmetical basis rather than a set theory or a theory of strings or a theory of binary trees is for a large part a legacy thing. However, some methods like Craig's Trick work most naturally with our choice. The same holds for Rosser-style arguments.

number n to the Gödel number of the result of substituting the numeral $\underline{n}$ of n for x in $A(x)$.

We use $\Box_\alpha A$ for the arithmetization of the provability of A from axioms whose Gödel numbers are in α. In other words, $\Box_\alpha A$ means $\mathsf{prov}_\alpha(\ulcorner A \urcorner)$. Similarly, $\Box A(\dot{x})$ means $\mathsf{prov}_\alpha(\ulcorner A(\dot{x}) \urcorner)$. In the notation $\Box^\tau_\alpha A$ we relativize the proof and the interpretation of provability to τ, but we leave α non-interpreted. In contrast, when we write $(\Box_\alpha A)^\tau$ the definition of the axiom set α is also translated. So $(\Box_\alpha A)^\tau$ means the same as $\Box^\tau_{\alpha^\tau} A$.

4. A Framework

In this section we develop a basic framework for working with presentations.

Let the signature of arithmetic be Ar and let the signature of arithmetic extended with a unary predicate A be Ar^+. *A presentation* $\mathcal{A}$ is a translation from Ar^+ to a signature Θ. The symbol A stands for the axiom set. To keep our exposition simple, we work with a parameter-free $\mathcal{A}$, but it is easy to adapt the development to the case with parameters.

4.1. Presentations and Axiom Sets.

Let X be a set of Θ-sentences. We write:

- $\mathsf{G}_\mathcal{A}(X) := \{A \in \mathsf{sent}_\Theta \mid X \vdash (\mathsf{A}(\ulcorner A \urcorner))^\mathcal{A}\}$.
- $\mathsf{H}_\mathcal{A}(X) := \mathsf{G}_\mathcal{A}(X) \cap X$.

We note that, since we assumed nothing about the translation $\mathcal{A}$ and alow any input X, the function $\mathsf{G}_\mathcal{A}$ might, for certain values, be completely silly. Numerals might not even exist ...

We write $\mathsf{cl}(X)$ or $\overline{X}$ for $\{A \in \mathsf{sent}_\Theta \mid X \vdash A\}$. We write $X_0 \subseteq_{\mathsf{fin}} X$ for $X_0 \subseteq X$ and X_0 is finite.

Clearly, $\mathsf{G}_\mathcal{A}(X) = \mathsf{G}_\mathcal{A}(\overline{X})$. We also have, by the compactness property of predicate logic, that $\mathsf{G}_\mathcal{A}(X) = \bigcup\{\mathsf{G}_\mathcal{A}(X_0) \mid X_0 \subseteq_{\mathsf{fin}} X\}$. In other words, $\mathsf{G}_\mathcal{A}$ is Scott-continuous. It follows that $\mathsf{G}_\mathcal{A}$ is monotonic and that $\mathsf{G}_\mathcal{A}$ commutes with unions of directed sets of sets of sentences. Similarly, for $\mathsf{H}_\mathcal{A}$.

A set of Θ-sentences X is $\mathcal{A}$-*complete* iff $X \subseteq \mathsf{G}_\mathcal{A}(X)$. Shifting the perspective, we will also say that $\mathcal{A}$ *semi-numerates* X in X. We observe that $\mathcal{A}$-complete sets are closed under arbitrary unions.

The set X is $\mathcal{A}$-*sound* iff $\mathsf{G}_\mathcal{A}(X) \subseteq X$. In case X is both $\mathcal{A}$-sound and $\mathcal{A}$-complete, we say that $\mathcal{A}$ *numerates* X in X.

Here are some basic insights.

Theorem 4.1. *Suppose U is axiomatized by X and X is $\mathcal{A}$-complete. Then $X \subseteq \mathsf{H}_\mathcal{A}(U)$ and $\mathsf{H}_\mathcal{A}(U)$ is $\mathcal{A}$-complete. In other words, if U has an $\mathcal{A}$-complete axiomatization, then $\mathsf{H}_\mathcal{A}(U)$ is the maximal $\mathcal{A}$-complete axiomatization of U.*

Proof. Suppose $\overline{X} = U$ and $X \subseteq \mathsf{G}_\mathcal{A}(X)$. We have $X = \mathsf{H}_\mathcal{A}(X) \subseteq \mathsf{H}_\mathcal{A}(U)$. Moreover, $\mathsf{G}_\mathcal{A}(\mathsf{H}_\mathcal{A}(U)) \supseteq \mathsf{G}_\mathcal{A}(\mathsf{H}_\mathcal{A}(X)) = \mathsf{G}_\mathcal{A}(X) = \mathsf{G}_\mathcal{A}(U) \supseteq \mathsf{H}_\mathcal{A}(U)$. ❑

Theorem 4.2. *Suppose X is $\mathcal{A}$-complete and X axiomatizes U. Suppose further that U finitely axiomatizable. Then, there is a finite $\mathcal{A}$-complete $X_0 \subseteq X$ that axiomatizes U.*

Proof. Suppose Y_0 is a finite set of axioms for U. Since, X axiomatizes U, there is, by compactness, a finite set $X_0 \subseteq X$ that implies Y_0. Thus, X_0 axiomatizes U. It follows that

$$\mathsf{G}_{\mathcal{A}}(X_0) = \mathsf{G}_{\mathcal{A}}(\overline{X_0}) = \mathsf{G}_{\mathcal{A}}(\overline{X}) = \mathsf{G}_{\mathcal{A}}(X) \supseteq X \supseteq X_0.$$

So, X_0 is $\mathcal{A}$-complete. ❑

We say that X is *uniformly $\mathcal{A}$-complete* when X is the union of all its $\mathcal{A}$-complete finite subsets, in other words, if $X = \bigcup\{X_0 \subseteq_{\mathsf{fin}} X \mid X_0 \subseteq \mathsf{G}_{\mathcal{A}}(X_0)\}$. We also say that $\mathcal{A}$ *uniformly semi-numerates* X in X.

We note that uniform $\mathcal{A}$-completeness implies $\mathcal{A}$-completeness. Moreover, uniform $\mathcal{A}$-completeness is closed under arbitrary unions.

We say that $\mathcal{A}$ is *a uniform semi-presentation for U* iff U has a uniformly $\mathcal{A}$-complete axiomatization.

Finally, $\mathcal{A}$ is *a uniform presentation for U* if $\mathcal{A}$ is uniform semi-presentation for U and U is $\mathcal{A}$-sound.

We first prove the analogue of Theorem 4.1

Theorem 4.3. *Suppose U is axiomatized by X and X is uniformly $\mathcal{A}$-complete. Then $X \subseteq \mathsf{H}_{\mathcal{A}}(U)$ and $\mathsf{H}_{\mathcal{A}}(U)$ is uniformly $\mathcal{A}$-complete. In other words, if $\mathcal{A}$ is a uniform semi-presentation for U, then $\mathsf{H}_{\mathcal{A}}(U)$ is the maximal uniformly $\mathcal{A}$-complete axiomatization.*

Proof. Clearly, $X \subseteq \mathsf{H}_{\mathcal{A}}(U)$. Suppose $A \in \mathsf{H}_{\mathcal{A}}(U)$. Since X axiomatizes U, it follows that there are $B_0, \ldots, B_{n-1}$ in X such that $B_0, \ldots, B_{n-1} \vdash A \wedge (\mathsf{A}(\ulcorner A \urcorner))^{\mathcal{A}}$. By uniform $\mathcal{A}$-completeness, we can find $\mathcal{A}$-complete finite $Z_i \subseteq X$, such that $B_i \in Z_i$. Let X_0 be the union of the Z_i, for $i < n$. Then X_0 is a finite, $\mathcal{A}$-complete subset of X. It follows that $X_0 \cup \{A\}$ is a finite, $\mathcal{A}$-complete subset of $\mathsf{H}_{\mathcal{A}}(U)$. ❑

Theorem 4.4. *$\mathcal{A}$ is a uniform semi-presentation for U iff, for every $A \in U$, there is a finite $\mathcal{A}$-complete $X_0 \subseteq U$ such that $X_0 \vdash A$.*

Proof. "⇒" Suppose $\mathcal{A}$ is a uniform semi-presentation for U. Let X be a uniformly $\mathcal{A}$-complete axiomatization of U. Suppose $A \in U$. It follows that $X \vdash A$. Hence, reasoning as in the proof of Theorem 4.3, we find an $\mathcal{A}$-complete finite $X_0 \subseteq X$ such that $X_0 \vdash A$.

"⇐" Suppose for every A in U there is a finite $\mathcal{A}$-complete $X_0 \subseteq U$ such that $X_0 \vdash A$. Let X be the union of all finite $\mathcal{A}$-complete $X_0 \subseteq U$. Clearly X axiomatizes U. Moreover, by definition, X is uniformly $\mathcal{A}$-complete. ❑

4.2. **Ordering on Presentations.** We use, working in the language of arithmetic enriched with A, modal notations like $\Box_\alpha A$ for $\mathsf{prov}_\alpha(\ulcorner A \urcorner)$. Since we did not explicitly stipulate that α consists of exclusively of codes of Θ-sentences, in our definition of prov we simply ignore the numbers that are not codes of Θ-sentences. Alternatively we could stipulate that all non-codes are codes of $\top$. Thus, 'the consistency of α' will be expressed as $\Diamond_\alpha \top$. We note that $\Box$ and $\Diamond$ are dependent of Θ. We write $\Box^\Theta$ and $\Diamond^\Theta$, if we want to make that dependence visible.

We write $\mathcal{A}{\restriction}\mathsf{Ar}$ for the restriction of $\mathcal{A}$ to the arithmetical language. For $\tau : \mathsf{Ar} \to \Theta$, let Δ_τ be the set of all presentations $\mathcal{A} : \mathsf{Ar}^+ \to \Theta$ such that $\mathcal{A}{\restriction}\mathsf{Ar} = \tau$.

Suppose we have a Θ-theory U and a translation $\tau : \mathsf{Ar} \to \Theta$ such that $U \vdash (\mathsf{S}^1_2)^\tau$. We define a preordering of presentations in Δ_τ as follows:

- $\mathcal{A} \leq_{U,\tau} \mathcal{B}$ iff $U \vdash \forall \vec{x} \in \delta_\tau\, ((\mathsf{prov}_\mathsf{A})^{\mathcal{A}}(\vec{x}) \to (\mathsf{prov}_\mathsf{A})^{\mathcal{B}}(\vec{x}))$.
- $\mathcal{A} =_{U,\tau} \mathcal{B}$ iff $\mathcal{A} \leq_{U,\tau} \mathcal{B}$ and $\mathcal{A} \geq_{U,\tau} \mathcal{B}$.

Here the notation $(\Box_D)^\rho$ means that both provability and the axiom set D are translated via ρ. This in contrast to the notation $\Box_D^\rho$, where provability is relativized to δ_ρ but D is left untranslated.

I guess we could drop the restriction to τ by defining a category in stead of a preordering, where an arrow is a U-definable embedding between presentations. However, for the present paper such a category does not seem to be relevant.

4.3. **Operations on Presentations.** In the present subsection, we study ways of transforming presentations into other presentations.

Let $\tau : \mathsf{Ar} \to \Theta$ be a translation. We write $\tau[\mathsf{A} := \alpha]$ for the translation from Ar^+ to Θ, which is equal to τ on Ar and where $\mathsf{A}(x)$ is translated to $\delta_\tau(\vec{x}) \wedge \alpha(\vec{x})$.

We extend the $[\mathsf{A} := \alpha]$ notation to presentations, by writing: $\mathcal{A}[\mathsf{A} := \alpha]$ for $(\mathcal{A}{\restriction}\mathsf{Ar})[\mathsf{A} := \alpha]$. So here $[\mathsf{A} := \alpha]$ becomes a *reset to* rather than a *set to.*

4.3.1. *Arithmetical Axiom Sets.* Suppose $\alpha(x)$ is an arithmetical formula with only x free. Let $\tau : \mathsf{Ar} \to \Theta$ be a translation. We write $\alpha \uparrow \tau$ for $\tau[\mathsf{A} := \alpha^\tau(\vec{x})]$.

Let $X_0 := \{A_0, \ldots, A_{k-1}\}$ be a finite set of Θ-sentences. We write β_{X_0} for $\bigvee_{i<k} x = \ulcorner A_i \urcorner$. Here, of course we should fix some order for the elements of X_0 and have some convention on how to put the brackets. However, since all such choices lead to formulas that are equivalent over predicate logic we will not worry about them. We define $X_0 \Uparrow \tau := \beta_{X_0} \uparrow \tau$. We note that $X_0 \Uparrow \tau$ only functions as intended in a context where τ is used to interpret at least a suitable subtheory of Presburger Arithmetic, say Add, or, in case we use unary or 'tally' numerals, a suitable theory of successor, say Succ. We have the obvious result:

Theorem 4.5. *Let X_0 be a finite set of Θ-sentences. Suppose $X_0 \vdash \mathsf{Add}^\tau$. Then, X_0 is $X_0 \Uparrow \tau$-complete.* (*If we use 'tally' numerals', we already have the result with* Succ *in stead of* Add.)

4.3.2. *Union of Presentations.* Two presentations $\mathcal{A}$ and $\mathcal{B}$ are *compatible* if their restrictions to Ar are the same, in other words, if $\mathcal{A}{\restriction}\mathsf{Ar} = \mathcal{B}{\restriction}\mathsf{Ar}$. The operation $+$ is defined on compatible pairs $\mathcal{A}$, $\mathcal{B}$. It is simply the intensional counterpart of union of axiom sets. Of course, to do this meaningfully, we cannot switch numbers.

We define:

$$\begin{aligned}\mathcal{A} + \mathcal{B} := \mathcal{A}[\,\mathsf{A} \;\; := \;\; & (\exists y \leqslant x\,(x = \mathsf{disj}(\ulcorner \bot \urcorner, y) \wedge \mathsf{A}(y)))^{\mathcal{A}} \;\vee \\ & (\exists y \leqslant x\,(x = \mathsf{conj}(\ulcorner \top \urcorner, y) \wedge \mathsf{A}(y)))^{\mathcal{B}}\,]\end{aligned}$$

The disjunction with $\bot$ and the conjunction with $\top$ are added to make addition a bi-functor with respect to $\preceq_{U,\tau}$. To have this we need to be able to distinguish the sources of the axioms effectively in the proof. (It would be nice to have an example to show that the naive definition does not work.)

We have:

Theorem 4.6. *Suppose $\tau : \mathsf{Ar} \to \Theta$ and $U \vdash (\mathsf{S}^1_2)^\tau$. We have, for $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \Delta_\tau$,*

a. The operation $+$ restricted to Δ_τ is monotonic w.r.t. $\preceq_{U,\tau}$.
b. $\mathcal{A} \preceq_{U,\tau} \mathcal{A} + \mathcal{B}$ and $\mathcal{B} \preceq_{U,\tau} \mathcal{A} + \mathcal{B}$.
c. $\mathcal{A} + \mathcal{A} =_{U,\tau} \mathcal{A}$.
d. $(\mathcal{A} + \mathcal{B}) + \mathcal{C} =_{U,\tau} \mathcal{A} + (\mathcal{B} + \mathcal{C})$.
e. $\mathcal{A} + \mathcal{B} =_{U,\tau} \mathcal{B} + \mathcal{A}$.

Proof. The only item that deserves some attention here is (a). Suppose e.g. $\mathcal{A} \preceq_{U,\tau} \mathcal{A}'$ and $\mathcal{B} \preceq_{U,\tau} \mathcal{B}'$. We extend the signature of arithmetic with A, B, A', B'. Let $\mathcal{C}$ be the translation from the new signature to Θ that is τ when restricted to Ar, where $\mathsf{A}^{\mathcal{C}} := \mathsf{A}^{\mathcal{A}}$, $\mathsf{B}^{\mathcal{C}} := \mathsf{A}^{\mathcal{B}}$, ${\mathsf{A}'}^{\mathcal{C}} := \mathsf{A}^{\mathcal{A}'}$, ${\mathsf{B}'}^{\mathcal{C}} := \mathsf{A}^{\mathcal{B}'}$. We write inside $\mathcal{C}$,

$$(\alpha + \beta)(x) := \exists y \leqslant x\, ((x = \mathsf{disj}(\ulcorner\bot\urcorner, y) \wedge \alpha(y)) \vee (x = \mathsf{conj}(\ulcorner\top\urcorner, y) \wedge \beta(y))).$$

Note that $U \vdash \forall x \in \delta_\tau\, (\mathsf{A}^{\mathcal{A}+\mathcal{B}}(\vec{x}) \leftrightarrow (\mathsf{A} + \mathsf{B})^{\mathcal{C}}(\vec{x}))$, and similarly for $\mathsf{A}' + \mathsf{B}'$.

We reason in U inside $\mathcal{C}$. Suppose $\mathsf{prov}_{\mathsf{A}+\mathsf{B}}(x)$. Let p be a witnessing proof. Let z be a conjunction of the C such that an axiom of the form $(\bot \vee C)$ is used in p and let w be a conjunction of the D such that an axiom of the form $(\top \wedge D)$ is used in p. We note that by definition the C are A-axioms and the D are B-axioms. It is easy to see that these conjunctions exist in S^1_2. We now transform p into a proof q in predicate logic of $\mathsf{imp}(\mathsf{conj}(z, w), x)$. It is easy to check that the transformation of p to q is available in S^1_2. Clearly, we have $\mathsf{prov}_{\mathsf{A}}(z)$ and $\mathsf{prov}_{\mathsf{B}}(w)$. It follows that $\mathsf{prov}_{\mathsf{A}'}(z)$ and $\mathsf{prov}_{\mathsf{B}'}(w)$. Ergo, we can construct a proof witnessing $\mathsf{prov}_{\mathsf{A}'+\mathsf{B}'}(x)$. ❑

Theorem 4.7. *Suppose $\mathcal{A}$ and $\mathcal{B}$ are compatible. Suppose further $\mathcal{A}$ and $\mathcal{B}$ are semi-presentations for U, respectively V. Then, $\mathcal{A} + \mathcal{B}$ is a semi-presentation for $\overline{(U \cup V)}$. Similarly, $\mathcal{A}$ and $\mathcal{B}$ are uniform semi-presentations for U, respectively V. Then, $\mathcal{A} + \mathcal{B}$ is a uniform semi-presentation for $\overline{(U \cup V)}$.*

We leave the trivial proofs to the reader.

We write $\mathcal{A} \oslash A$ for $\mathcal{A} + (\{A\} \Uparrow (\mathcal{A}{\restriction}\mathsf{Ar}))$. We note that for any translation $\tau : \mathsf{Ar} \to \Theta$ and any finite set $X_0 = \{A_0, \ldots, A_{k-1}\}$ of Θ-sentences, we have that $X_0 \Uparrow \tau$ is equivalent over predicate logic to $\bot \uparrow \tau \oslash A_0 \oslash \ldots \oslash A_{k-1}$.

We have:

Theorem 4.8. *Suppose $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. Then, $U \vdash (\Box_{\mathsf{A}} B)^{\mathcal{A} \oslash A} \leftrightarrow (\Box_{\mathsf{A}}(A \to B))^{\mathcal{A}}$.*

The proof is trivial.

4.3.3. *Deductive Closure.* We define the deductive closure of $\mathcal{A}$ as

$$\overline{\mathcal{A}} := \mathsf{Th}(\mathcal{A}) := \mathcal{A}[\mathsf{A} := (\mathsf{prov}_{\mathsf{A}})^{\mathcal{A}}(\vec{x})].$$

We have the obvious:

Theorem 4.9. *Suppose $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. Let $\tau := \mathcal{A}{\restriction}\mathsf{Ar}$. Then,*

a. *The operation Th restricted to Δ_τ is monotonic w.r.t. $\preceq_{U,\tau}$. In other words, Th is an endo-functor of the preorder category given by $\preceq_{U,\tau}$.*
b. *$U \vdash \forall \vec{x} \in \delta_{\mathcal{A}}\, (\mathsf{A}^{\mathcal{A}}(\vec{x}) \to \mathsf{A}^{\mathsf{Th}(\mathcal{A})}(\vec{x}))$, and, hence $\mathcal{A} \preceq_{U,\tau} \mathsf{Th}(\mathcal{A})$.*
c. *If $\mathcal{A}$ is a semi-presentation of U, then so is $\mathsf{Th}(\mathcal{A})$.*
d. *If $\mathcal{A}$ is a uniform semi-presentation of U, then so is $\mathsf{Th}(\mathcal{A})$.*

Suppose $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. Let $\tau := \mathcal{A}{\restriction}\mathsf{Ar}$. It is important to note that we do not generally have: $\mathsf{Th}(\mathcal{A}) \preceq_{U,\tau} \mathcal{A}$. A counterexample will be give in Example 4.13.

4.3.4. *Craigification.* We consider a presentation $\mathcal{A} : \mathsf{Ar}^+ \to \Theta$. We will say that $\mathcal{A}$ is an E-presentation if $\mathsf{A}^{\mathcal{A}}(\vec{x})$ has the form $\alpha(x) := \exists \vec{y} \in \delta_{\mathcal{A}}\, B(\vec{y}, \vec{x})$. Throughout this subsubsection we will assume that $\mathcal{A}$ is an E-presentation.

We extend Ar^+ with a new binary predicate symbol B which we translate as B. Let the resulting translation be $\mathcal{A}^*$. We can now write $\alpha(\vec{x})$ as $(\exists y\, \mathsf{B}(y, x))^{\mathcal{A}^*}$, where we assume that the single variable x is translated to the sequence of variables $\vec{x}$. We define:

$$\alpha^\star(\vec{x}) :\leftrightarrow (\exists u \leqslant x\, \exists y \leqslant x\, (x = \mathsf{conj}(u, \ulcorner \dot{y} = \dot{y} \urcorner) \wedge \mathsf{B}(y, u)))^{\mathcal{A}^*}.$$

We define the Craigification of $\mathcal{A}$ by: $\mathsf{Cr}(\mathcal{A}) := \mathcal{A}[\mathsf{A} := \alpha^\star]$.

Unfortunately, Craigification is irredeemably syntactic. It is generally not a functor w.r.t. $\preceq_{U,\tau}$.

It will be convenient for readability to have the following presentation $\mathcal{A}^\star$ at hand. We expand Ar with A, B and $\mathsf{A}^\star$ and extend $\mathcal{A}^*$ to $\mathcal{A}^\star$ on the new signature by setting the translation of $\mathsf{A}^\star$ to $\alpha^\star$.

Theorem 4.10. *Suppose $\mathcal{A}$ is an E-presentation and $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. Let τ be $\mathcal{A}{\restriction}\mathsf{Ar}$. Then, $\mathsf{Cr}(\mathcal{A}) \preceq_{U,\tau} \mathcal{A}$. In terms of $\mathcal{A}^\star$, this says:*

$$U \vdash (\forall x\, (\mathsf{prov}_{\mathsf{A}^\star}(x) \to \mathsf{prov}_{\mathsf{A}}(x)))^{\mathcal{A}^\star}.$$

Proof. We reason in U inside $\mathcal{A}^\star$. Suppose $\mathsf{prov}_{\mathsf{A}^\star}(x)$. Let p be a witnessing proof. We zoom in on an occurrence axiom C from $\mathsf{A}^\star$ in p. This axiom is of the form $\ulcorner D \wedge \dot{y} = \dot{y} \urcorner$, where D is an axiom from A. We replace the sub-proof consisting of C by the obvious sub-proof of C from D, and similarly for all other axiom-occurrences in p, thus obtaining an A-proof p' of x. To show that this is possible, we have to verify that the transformation $p \mapsto p'$ is p-time. To see this, we note that the length of the subproof replacing C is just twice the length of C plus some standard overhead m. Thus the length of p' will be estimated by two times the length of p plus the length of p times m. In other words, the length of of p' is bounded by $m + 1$ times the length of p. This yields a polynomial bound on p'. ❑

The converse of Theorem 4.10 does not hold, as will be illustrated in the Example 4.13 at the end of this subsection. We collect some further properties of Craigification.

Theorem 4.11. *Suppose $\mathcal{A}$ is an E-presentation $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. We have:*

a. *If $\mathcal{A}$ semi-represents U, then U is $\mathsf{Th}(\mathsf{Cr}(\mathcal{A}))$-complete, in other words, U is closed under the necessitation rule for $\Box^{\mathsf{Cr}(\mathcal{A})}_{\mathsf{A}}$.*
b. *If $\mathcal{A}$ uniformly semi-represents U, then U is uniformly $\mathsf{Th}(\mathsf{Cr}(\mathcal{A}))$-complete.*

Proof. We expand $\mathcal{A}$ to $\mathcal{A}^\star$ as in the proof of Theorem 4.10.

Ad (a): Let X be some axiomatization of $\mathcal{A}$ that is $\mathcal{A}$-complete. So we have: if $A \in X$, then $U \vdash (\mathsf{A}(\ulcorner A \urcorner))^{\mathcal{A}^\star}$. In other words, $U \vdash (\exists y\, \mathsf{B}(y, \ulcorner A \urcorner))^{\mathcal{A}^\star}$.

Suppose $U \vdash B$. Then, for some $A_0, \ldots, A_{n-1} \in X$, we have $A_0, \ldots, A_{n-1} \vdash B$. Say the Gödelnumber of the proof is p. We find:

$$U \vdash (\bigwedge_{i<n} \exists y_i\, \mathsf{B}(y_i, \ulcorner A_i \urcorner))^{\mathcal{A}^\star} \text{ and } U \vdash (\mathsf{proof}_{\beta_{\{A_0,\ldots,A_{n-1}\}}}(\underline{p}, \ulcorner B \urcorner))^{\mathcal{A}^\star}.$$

We want to show $U \vdash (\Box_{\mathsf{A}^\star} B)^{\mathcal{A}^\star}$.

We reason in U inside $\mathcal{A}^\star$. Let y_i be such that $\mathsf{B}(y_i, \ulcorner A_i \urcorner)$ for $i < n$. It follows that $\ulcorner A_i \wedge \dot{y}_i = \dot{y}_i \urcorner$ is in $\mathsf{A}^\star$. We note that the map $y \mapsto \ulcorner A_i \wedge \dot{y}_i = \dot{y}_i \urcorner$ is p-time, thanks to our use of efficient numerals. We now transform $\underline{p}$ to an $\mathsf{A}^\star$-proof q by replacing each sub-proof of $\underline{p}$ that consists of an A-axiom A_i by a proof of A_i from $\ulcorner A_i \wedge \dot{y}_i = \dot{y}_i \urcorner$. We note that the length of q is bounded by the length of $y^\star$ the maximum of the y_i times a standard constant. We know that $y^\star$ exists since n is standard. Thus, the transformation $\underline{p} \mapsto q$ is available within the resources of S^1_2.

Ad (b). Let X be an uniformly $\mathcal{A}$-complete axiomatization of U. Let A be any sentence of U. Let $X_0 \subseteq_{\mathsf{fin}} X$ be such that $X_0 \vdash A$, $X_0 \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$, and X_0 is $\mathcal{A}$-complete. It follows from (a) that $\overline{X}_0$ is $\mathsf{Th}(\mathsf{Cr}(\mathcal{A}))$-complete, and hence that $X_0 \cup \{A\}$ is $\mathsf{Th}(\mathsf{Cr}(\mathcal{A}))$-complete. From this it is immediate that U is uniformly $\mathsf{Th}(\mathsf{Cr}(\mathcal{A}))$-complete. ❑

The next property is shows that Cr is a kind of left-inverse of Th. Regrettably, this does not have functorial meaning because of the irredeemably syntactic character of Cr.

Theorem 4.12. *Suppose* $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. *Let* $\tau = \mathcal{A}{\restriction}\mathsf{Ar}$. *Then,* $\mathsf{Cr}(\mathsf{Th}(\mathcal{A})) =_{U,\tau} \mathcal{A}$.

Proof. We extend the signature Ar^+ with an extra predicate $\widetilde{\mathsf{A}}$ that is interpreted as $\mathsf{A}^{\mathsf{Cr}(\mathsf{Th}(\mathcal{A}))}$. Thus we obtain a translation $\widetilde{\mathcal{A}}$. We reason in U, inside $\widetilde{\mathcal{A}}$. We note that

$$\widetilde{\mathsf{A}}(B) \leftrightarrow \exists C \leqslant B\, \exists p \leqslant B\, (B = \ulcorner C \wedge \dot{p} = \dot{p} \urcorner \wedge \mathsf{proof}_{\mathsf{A}}(p, C)).$$

We have to show that, for any A, $\Box_{\widetilde{\mathsf{A}}} A$ iff $\Box_{\mathsf{A}} A$.

From left to right: Suppose q is an $\widetilde{\mathsf{A}}$-proof of A. Consider any axiom-occurrence B_i in q. B_i is of the form $\ulcorner C_i \wedge \dot{p}_i = \dot{p}_i \urcorner$, where p_i is an A-proof of C_i. We replace the occurrence of B_i by first the A-proof p_i of C_i and then the inference from C_i to $\ulcorner C_i \wedge \dot{p}_i = \dot{p}_i \urcorner$. We may assume that the length of p_i is of the same order as the length of $\ulcorner \dot{p}_i \urcorner$. It follows that the length of the replacement is bounded by four times the length of p_i with some standard overhead.

Now when we replace all $\widetilde{\mathsf{A}}$-axiom occurrences in the manner prescribed, we see that the length of the new A-proof, say $p^\star$, will be bound by k times the length of q, for some standard k. Thus, the transformation $q \mapsto p^\star$ is available with the resources of S^1_2.

From right to left: Let p be an A-proof of A. It follows that $\ulcorner A \wedge \dot{p} = \dot{p} \urcorner$ is in $\widetilde{\mathsf{A}}$. So we can take as $\widetilde{\mathsf{A}}$-proof the inference from $\ulcorner A \wedge \dot{p} = \dot{p} \urcorner$ to A. ❑

Here is an example that separates a Craigification from the original axiomatization and the original axiomatization from its theoretization.

Example 4.13. In [Vis15], we proved the following. Suppose that C is a single axiom that axiomatizes EA. We take $\beta(x) := (x = \ulcorner C \urcorner)$ and $\mathcal{B} := \beta \uparrow \mathsf{Id}_{\mathsf{Ar}}$. Let B be a single statement that is equivalent to Σ^0_1-collection over EA. Then $\mathsf{EA} + \neg B \vdash \Box^{\mathsf{Th}(\mathcal{B})}_{\mathsf{A}} \bot$.

On the other hand we have $\mathsf{Cr}(\mathsf{Th}(\mathcal{B})) =_{\mathsf{EA},\mathsf{Id}_{\mathsf{Ar}}} \mathcal{B}$. But, $\mathsf{EA} + \neg B \nvdash \Box^{\mathcal{B}}_{\mathsf{A}} \bot$, since $\mathsf{EA} + \neg B$ is Π_3-conservative over EA, by the results of Paris & Kirby in [PK78].

It follows that $\mathsf{EA} \nvdash \Box^{\mathsf{Th}(\mathcal{B})}_{\mathsf{A}} \bot \to \Box^{\mathsf{Cr}(\mathsf{Th}(\mathcal{B}))}_{\mathsf{A}} \bot$, and, thus,

$$\mathsf{Cr}(\mathsf{Th}(\mathcal{B})) =_{\mathsf{EA},\mathsf{Id}_{\mathsf{Ar}}} \mathcal{B} \precneqq_{\mathsf{EA},\mathsf{Id}_{\mathsf{Ar}}} \mathsf{Th}(\mathcal{B}).$$

Thus, taking $\mathsf{Th}(\mathcal{B})$ as our original axiomatization, we have that the Craigification is strictly below the original axiomatization. In Example 6.16, we provide a second example of this phenomenon. If we take $\mathcal{B}$ as our original axiomatization, we find that the original axiomatization is strictly below its theoretization. ❑

5. The Second Incompleteness Theorem à la Alcibiades

We are ready and set to give the corrected Alcibiades argument. This will be done in Subsection 5.1. In the succeeding subsections, we will provide some variations and strengthenings.

5.1. **The Basic Version.** We have:

Theorem 5.1. *Suppose $\mathcal{A}$ uniformly semi-presents the Θ-theory U. Then, whenever $U \vdash (\mathsf{S}^1_2 + \Diamond^{\Theta}_{\mathsf{A}}\top)^{\mathcal{A}}$, then U is inconsistent.*

In a different formulation: *suppose $K : U \rhd (\mathsf{S}^1_2 + \Diamond^{\Theta}_{\mathsf{A}}\top)$. Suppose further that τ_K is a uniform semi-presentation of U. Then, U is inconsistent.*

Proof. The theory S^1_2 is finitely axiomatizable, say B is a single axiom for it. So, if $U \vdash (\mathsf{S}^1_2 + \Diamond^{\Theta}_{\mathsf{A}}\top)^{\mathcal{A}}$, it follows that $U \vdash (B \wedge \Diamond_{\mathsf{A}}\top)^{\mathcal{A}}$. Then, for some $\mathcal{A}$-complete finite $X_0 \subseteq U$, we have $X_0 \vdash (B \wedge \Diamond_{\mathsf{A}}\top)^{\mathcal{A}}$. It follows that $X_0 \vdash (B \wedge \Diamond_{\beta_{X_0}}\top)^{\mathcal{A}}$. We now apply the Second Incompleteness Theorem for finitely axiomatized theories and find that $\overline{X_0}$ is inconsistent. Hence, *a fortiori*, U is inconsistent. ❑

We note that the above proof is fully constructive.

5.2. **A Slightly Stronger Version.** In this subsection, we prove a strengthening of Theorem 5.1. In Subsection 6.3, we will see a case where Theorem 5.2 rather than Theorem 5.1 is needed.

Theorem 5.2. *Suppose $\mathsf{Th}(\mathcal{A})$ uniformly semi-presents the Θ-theory U. Then, whenever $U \vdash (\mathsf{S}^1_2 + \Diamond^{\Theta}_{\mathsf{A}}\top)^{\mathcal{A}}$, then U is inconsistent.*

In a different formulation: *suppose $K : U \rhd (\mathsf{S}^1_2 + \Diamond^{\Theta}_{\mathsf{A}}\top)$. Suppose further that $\mathsf{Th}(\tau_K)$ is a uniform semi-presentation of U. Then, U is inconsistent.*

Proof. Let B be a single axiom for S^1_2. If $U \vdash (\mathsf{S}^1_2 + \Diamond_{\mathsf{A}}\top)^{\mathcal{A}}$, it follows that $U \vdash (B \wedge \Diamond_{\mathsf{A}}\top)^{\mathcal{A}}$. Then, for some $\mathsf{Th}(\mathcal{A})$-complete finite $X_0 \subseteq U$, we have $X_0 \vdash (B \wedge \Diamond_{\mathsf{A}}\top)^{\mathcal{A}}$. We find that, for any $C \in X_0$, we have $X_0 \vdash \Box^{\mathcal{A}}_{\mathsf{A}}C$. It follows, by a feasible transformation of proofs, using that X_0 is standardly finite, that:

$$X_0 \vdash (\forall x \in \mathsf{sent}_{\Theta}\, (\mathsf{prov}_{\beta_{X_0}}(x) \to \mathsf{prov}_{\mathsf{A}}(x)))^{\mathcal{A}}.$$

A fortiori, $X_0 \vdash (\Diamond_{\mathsf{A}}\top \to \Diamond_{\beta_{X_0}}\top)^{\mathcal{A}}$, and, thus, $X_0 \vdash (\mathsf{S}^1_2 + \Diamond_{\beta_{X_0}}\top)^{\mathcal{A}}$. We now apply the Second Incompleteness Theorem for finitely axiomatized theories and find that $\overline{X_0}$ is inconsistent. Hence, *a fortiori*, U is inconsistent. ❑

5.3. **Löb's Rule.** We prove closure under Löb's Rule under the appropriate conditions.

Theorem 5.3. *Suppose $\mathcal{A}$ uniformly semi-presents the Θ-theory U and $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. Then, whenever $U \vdash (\Box_{\mathsf{A}}C)^{\mathcal{A}} \to C$, we have $U \vdash C$.*

We give two proofs. The proof of closure under Löb's Rule from the Second Incompleteness Theorem, works in the context of our version of the Second Incompleteness Theorem. This proof is ascribed to Saul Kripke.

First Proof. Suppose $\mathcal{A}$ uniformly semi-presents the Θ-theory U and $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. Suppose further that $U \vdash (\Box_{\mathsf{A}} C)^{\mathcal{A}} \to C$. It follows that $U + \neg C \vdash (\Diamond_{\mathsf{A}} \neg C)^{\mathcal{A}}$. By Theorems 4.5 and 4.7, the theory $U + \neg C$ is uniformly $\mathcal{A} \oslash \neg C$-complete. By Theorem 4.8, We have $U \vdash (\Diamond_{\mathsf{A}} \neg C)^{\mathcal{A}} \leftrightarrow (\Diamond_{\mathsf{A}} \top)^{\mathcal{A} \oslash \neg C}$. Thus, we may conclude that $U + \neg C \vdash (\Diamond_{\mathsf{A}} \top)^{\mathcal{A} \oslash \neg C}$ and $\mathcal{A} \oslash \neg C$ is a uniform semi-presentation of $U + \neg C$. It follows, by Theorem 5.1, that $U + \neg C \vdash \bot$, and, hence, $U \vdash C$. ❑

We can also prove the desired result directly.

Second Proof. Suppose $\mathcal{A}$ uniformly semi-presents the Θ-theory U and $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. Let B be a single axiom for S^1_2. Suppose further that $U \vdash (\Box_{\mathsf{A}} C)^{\mathcal{A}} \to C$. So, $U \vdash B^{\mathcal{A}} \wedge ((\Box_{\mathsf{A}} C)^{\mathcal{A}} \to C)$. Then, for some $\mathcal{A}$-complete finite $X_0 \subseteq U$, we have $X_0 \vdash B^{\mathcal{A}} \wedge ((\Box_{\mathsf{A}} C)^{\mathcal{A}} \to C)$. It follows that $X_0 \vdash B^{\mathcal{A}} \wedge ((\Box_{\beta_{X_0}} C)^{\mathcal{A}} \to C)$. We now apply Löb's Rule for finitely axiomatized theories and find that $X_0 \vdash C$. Hence, *a fortiori*, $U \vdash C$. ❑

We note that the second proof has the advantage that it is fully constructive. We can strengthen the previous theorem a bit.

Theorem 5.4. *Suppose* $\mathsf{Th}(\mathcal{A})$ *uniformly semi-presents the* Θ*-theory* U *and* $U \vdash (\mathsf{S}^1_2)^{\mathcal{A}}$. *Then, whenever* $U \vdash (\Box_{\mathsf{A}} C)^{\mathcal{A}} \to C$, *we have* $U \vdash C$.

The proof is a slight modification of the proofs above using either Theorem 5.2 or a variation on the proof of Theorem 5.2.

6. Σ^0_1-PRESENTATIONS

In this section, we will consider the case of Σ^0_1-numerations of the axiom set. This is, of course, in part, the case discussed in Feferman's [Fef60]. Before proceeding, let us state and prove our version of the traditional version of the Second Incompleteness Theorem for Σ^0_1-numerations.

Theorem 6.1. *Suppose* $\sigma(x)$ *is a* Σ^0_1*-formula that numerates the axioms of* U *in the standard model. Then, if* $U \rhd (\mathsf{S}^1_2 + \Diamond_\sigma \top)$, *then* U *is inconsistent.*

Proof. Suppose $N : U \rhd (\mathsf{S}^1_2 + \Diamond_\sigma \top)$. It follows that, for some finite sub-theory U_0 of U, we have $U_0 \vdash (\mathsf{S}^1_2)^{\tau_N}$. It immediately follows, by Σ^0_1-completeness, that $\sigma \upharpoonright \tau_N$ semi-numerates the axioms of U in U_0. So, *a fortiori*, $\sigma \upharpoonright \tau_N$ uniformly semi-numerates the axioms of U in U. ❑

In Appendix B, we give four alternative proofs of Theorem 6.1.

Since Feferman's set-up was less general than ours, the result does not summarize everything to be said. There are three issues to contend with.

The first issue is that Σ^0_1-numerations of the axiom set in the given theory need not be uniform. We provide an example of this phenomenon in Subsection 6.1. In this subsection we will also show that a theory whose axioms are numerated by a Σ^0_1-formula can be as complex as we like: for every set of numbers Z there is a such a theory U such that Z is reducible to U.

The second issue is, that we do not know whether we have Σ^0_1-completeness in weak theories like S^1_2. This problem is connected to questions concerning the collapse of the polynomial hierarchy. We want to sidestep this issue. There are two ways to do this. The first is to replace Σ^0_1 by $\exists\Sigma^{\mathsf{b}}_1$. In the context of the stronger theory EA,

Σ^0_1 and $\exists\Sigma^{\mathsf{b}}_1$ coincide modulo provable equivalence. So this approach does not differ from the classical one as soon as we have EA. The second way is more simpleminded: just work with EA as our basic theory. We will choose this last option.

The third issue is that, in the absence of Σ^0_1-collection, Σ^0_1-axiomatized theories need not satisfy the Löb conditions. We provide an example of this fact in Subsection 6.2. In case we do have Σ^0_1-collection, we do have the following theorem.

Theorem 6.2. *Suppose $\sigma(x)$ is Σ^0_1 and and $N : U \rhd (\mathsf{EA} + \mathrm{B}\Sigma^0_1 + \Diamond_\sigma\top)$. Suppose further that $\mathcal{S} := \sigma \uparrow \tau_N$ semi-numerates the axioms of U in U, then U is inconsistent.*

Proof. The predicate $(\Box_{\mathsf{A}} A)^{\mathcal{S}}$ satisfies the Löb conditions. ❑

Finally, in Subsection 6.3, we improve Theorem 6.2 by presenting two proofs of the Second Incompleteness Theorem for Σ^0_1-numerations in the EA-case.

To wet the reader's appetite, here is a first example of the behaviour of Σ_1-numerations.

Example 6.3. Let C be a single axiom for EA and let $\beta(x) :\leftrightarrow x = \ulcorner C \urcorner$. We consider the theory $A := \mathsf{EA} + \Box_\beta\bot$. Since, A is a finitely axiomatized sequential theory, there is a faithful interpretation $K : A \rhd_{\mathsf{faith}} \mathsf{EA}$. Let $\tau := \tau_K$. (This result is due to Harvey Friedman. See [Vis05] for an exposition.) Let $B := A + \Diamond^\tau_\beta\top$. So, B is a consistent theory.

Let
$$\gamma(x) :\leftrightarrow \beta(x) \vee (\Box_\beta\bot \wedge (x = \ulcorner\Box_\beta\bot\urcorner \vee x = \ulcorner\Diamond^\tau_\beta\top\urcorner)).$$
We set $\mathcal{S} := \gamma(x) \uparrow \mathsf{Id}_{\mathsf{Ar}}$ and $\mathcal{T} := \gamma \uparrow \tau$.

Clearly, $\mathcal{S}$ numerates the axiom set $\{C, \Box_\beta\bot, \Diamond^\tau_\beta\top\}$ of B in B. Hence, $B \nvdash \Diamond^{\mathcal{S}}_{\mathsf{A}}\top$.

On the other hand, $B \vdash (\forall x\,(\gamma(x) \leftrightarrow \beta(x)))^\tau$, and, hence, $B \vdash (\Diamond_\gamma\top)^\tau$, or, in other words, $B \vdash (\Diamond_{\mathsf{A}}\top)^{\mathcal{T}}$. ❑

6.1. **A Non-Uniform Σ_1-Numeration.** Let C be a single axiom for EA and let $\beta := (x = \ulcorner C \urcorner)$. By the Gödel Fixed Point Theorem, we find a formula $R(x)$ such that:
$$\mathsf{EA} \vdash R(x) \leftrightarrow \Box_{\beta + \bigwedge_{y<x} R(y)} \neg R(x) \leqslant \Box_{\beta + \bigwedge_{y<x} R(y)} R(x).$$
We consider the theory U axiomatized by $X := \{C\} \cup \{R(\underline{n}) \mid n \in \omega\}$.

The theory U is consistent, since $R(\underline{n})$ is a Rosser sentence for $\mathsf{EA} + \{R(\underline{k}) \mid k < n\}$.

Let α be the following predicate: $\alpha(x) :\leftrightarrow \beta(x) \vee \exists y < x\,(x = \ulcorner R(\dot{y})\urcorner \wedge R(y+1))$. Let $\mathcal{A} := \alpha \uparrow \mathsf{Id}_{\mathsf{Ar}}$. We clearly have: $\mathsf{EA} \vdash \alpha(\ulcorner R(\underline{n})\urcorner) \leftrightarrow R(\underline{n+1})$. Since U is consistent, it follows that $\mathcal{A}$ numerates X in U.

Theorem 6.4. *$\mathcal{A}$ is not a uniform numeration of X in U.*

Proof. Let X_0 be a finite subset of X. Without loss of generality, we may assume that X_0 contains an axiom larger than C. Let $R(\underline{n})$ be the largest axiom in X_0. Suppose $X_0 \vdash \alpha(\ulcorner R(\underline{n})\urcorner)$. It follows that $\mathsf{EA} + \{R(\underline{k}) \mid k \leqslant n\} \vdash \alpha(\ulcorner R(\underline{n})\urcorner)$. Hence, $\mathsf{EA} + \{R(\underline{k}) \mid k \leqslant n\} \vdash R(\underline{n+1})$. Quod non. ❑

Remark 6.5. We can adapt the ideas around the construction of the non-uniform Σ^0_1-axiomatization to produce a very complex U which still numerates its own axioms with a Σ^0_1-formula. Let C and β be as in the proof of Theorem 6.4.

Let Z be any set of natural numbers. By a result due, independently to Mostowski, Feferman, Scott and Kripke (see e.g. [Kri62]), there is a Σ^0_1-formula $S^*(x)$ such that

$$\mathsf{EA} + \{S^*(\underline{n}) \mid n \in Z\} + \{\neg\, S^*(\underline{m}) \mid m \notin Z\}$$

is consistent. We consider the theory $X := C + \{S^*(\underline{n}) \mid n \in Z\}$. It follows that $n \in Z$ iff $X \vdash S^*(n)$. It is easy to see that

$$\zeta(x) := (\beta(x) \vee \exists y < x\,(x = \ulcorner S^*(\dot{y})\urcorner \wedge S^*(y)))$$

numerates X in $U := \overline{X}$. ❑

6.2. **Failure of the Löb Conditions for a Σ^0_1-axiomatization of** EA**.** In this subsection we provide a curious example. We provide a Σ^0_1-formula $\sigma(x)$ with the following properties.

- The formula σ defines the axioms of EA, and, hence, numerates the axioms of EA in EA.
- EA knows that σ defines a finite set of axioms.
- EA knows that the theory defined by σ is between EA and $\mathsf{EA} + \Diamond_\beta \top$, where β is a standard axiomatization of EA.
- EA does not prove the Löb condtions for $\Box_\sigma$, what is more EA does not prove the formalized Second Incompleteness Theorem for σ.

We work, for the moment in EA. Our first order of business is to define a Kripke model $\mathcal{K}$. Let p be any (possibly non-standard) number. Our model has nodes $0, \ldots, p+1$. We set $x \prec y$ iff $x = 0$ and $1 \leqslant y \leqslant p+1$. Let C be a single axiom for EA and let $\beta(x) := x = \ulcorner C \urcorner$. We define the usual Solovay function h_p on $\mathcal{K}$ for β.

- $\ell_p = 0$ iff $\forall x\, h(x) = 0$.
- $\ell_p = y$, if $0 \prec y$ and $\exists x\, h(x) = y$.
- $h_p(0) = 0$,
- $h_p(y+1) := \begin{cases} x & \text{if } h(y) \prec x \text{ and } \mathsf{proof}_\beta(y, \ulcorner \ell_{\dot{p}} \neq \dot{x} \urcorner) \\ h_p(y) & \text{otherwise} \end{cases}$

We find the following:

Lemma 6.6. *We have:*

a. $\mathsf{EA} \vdash \Box_\beta(\Box_\beta \bot \leftrightarrow \bigvee_{x \leqslant p} \ell_{\dot{p}} = (x \dot{+} 1))$.
b. $\mathsf{EA} \vdash (x \leqslant p \wedge \Box_\beta \ell \neq x \dot{+} 1) \to \Box_\beta \bot$.

Proof. The proof follows the usual lines of a proof of Solovay's Theorem. Since our model is so simple, there are some short cuts possible. ❑

We use that over EA we have a Σ_1-predicate $\mathsf{def}(y, z)$ such that an element a is Σ_1-definable iff, for some number k, $\mathsf{def}(\underline{k}, z)$ defines a. We follow Paris & Kirby ([PK78]) in defining def as follows.[4] Let $\mathsf{T}(e, w, x)$ be Kleene's T-predicate where T is Δ_0. Here 'e' is the place for the index of a partial recursive function, 'w' is the place for the sequence of arguments and 'x' is the place for the computation. We use

[4]Paris & Kirby defined a somewhat more general version with parameters.

ε for the (code of) the empty sequence. We assume that there is a result-extracting rudimentary U such that $\mathsf{U}(x)$ is the result of the computation. We take:

$$\mathsf{def}(s,y) :\leftrightarrow \exists v\, (\mathsf{T}(s,\varepsilon,v) \wedge \mathsf{U}(v)=y).$$

We proceed to specify σ. Let π be the usual axiomatization of PA.

$$\begin{array}{lll}\sigma(x) & :\leftrightarrow & \beta(x) \ \vee \\ & & \exists p\, (\, \mathsf{proof}_\pi(p,\bot) \wedge \forall q < p\, \neg\, \mathsf{proof}_\pi(q,\bot) \ \wedge \\ & & \exists y \leqslant p\, (x = \ulcorner \ell_{\dot p} \neq (y \dot{+} 1)\urcorner \wedge \exists s < p\, \mathsf{def}(s,y)))\end{array}$$

It is convenient to have a partial term $\mathfrak{p}$ (as defined symbol) that stands for the smallest inconsistency proof of PA (as given by π) if it exists. We note that $\mathfrak{p}$ is 'rigid' over EA:

Lemma 6.7. $\mathsf{EA} \vdash \mathfrak{p} = x \to \Box_\beta\, \mathfrak{p} = \dot{x}$.

With our new notation we can rewrite σ as:

$$\sigma(x) :\leftrightarrow \beta(x) \vee (\mathfrak{p}\downarrow \wedge \exists y \leqslant \mathfrak{p}\, (x = \ulcorner \ell_{\mathfrak{p}} \neq (y \dot{+} 1)\urcorner \wedge \exists s < \mathfrak{p}\ \mathsf{def}(s,y))).$$

Without any worries about the meaning of the second conjunct of the definition of σ, we can already prove some important claims about σ. Let $\mathcal{S} : \sigma \uparrow \mathsf{Id}_{\mathsf{Ar}}$. We write $\preceq := \preceq_{\mathsf{EA},\mathsf{Id}_{\mathsf{Ar}}}$.

Lemma 6.8. *We have:*

a. $\mathcal{S}$ *is a uniform presentation for* EA*. Consequently, we have* $\mathsf{EA} \nvdash \Diamond_\sigma \top$.
b. $\mathsf{EA} \vdash \forall A\, ((\Box_\beta A \to \Box_\sigma A) \wedge (\Box_\sigma A \to \Box_\beta(\Diamond_\sigma \top \to A)))$.
In other words, $\beta \preceq \sigma \preceq (\beta + \Diamond_\sigma \top)$.

Proof. Ad (a): Clearly, we have $\mathsf{EA} \vdash \sigma(\ulcorner C \urcorner)$. Conversely suppose $\mathsf{EA} \vdash \sigma(\ulcorner D \urcorner)$ and $C \neq D$. It follows that $\mathsf{EA} \vdash \mathfrak{p}\downarrow$. Quod non.

Ad (b). We reason in EA. The first conjunct is immediate. Suppose $\Box_\sigma A$. In case $\mathfrak{p}\uparrow$, it follows that β and σ coincide, and hence $\Box_\beta A$ and, *a fortiori*, $\Box_\beta(\Diamond_\sigma \top \to A)$). Suppose $\mathfrak{p}\downarrow$. It follows, by Lemma 6.6(a), that $\Box_\sigma(\Diamond_\sigma \top \to \bigwedge_{x \leqslant \mathfrak{p}} \ell_{\mathfrak{p}} \neq (x+1))$. Hence, $\Box_{\sigma + \Diamond_\sigma \top}$ extends $\Box_\sigma$. ❑

We now proceed to 'compute' $\Box_\sigma \bot$ and $\Box_\sigma \Diamond_\sigma \top$. The result of the computation will be expressed as closed terms constructed from $\Box_\beta$ and a special propositional constant $\mathsf{S}^\star$. We define $\mathsf{S}^\star$ as follows.

$$\begin{array}{lll}\mathsf{S}^\star & :\leftrightarrow & \exists p\, (\, \mathsf{proof}_\pi(p,\bot) \wedge \forall q < p\, \neg\, \mathsf{proof}_\pi(q,\bot) \ \wedge \\ & & \quad \forall y \leqslant p\, \exists s < p\, \mathsf{def}(s,y)\,).\end{array}$$

In our $\mathfrak{p}$-notation: $\mathsf{S}^\star :\leftrightarrow (\mathfrak{p}\downarrow \wedge \forall y \leqslant \mathfrak{p}\, \exists s < \mathfrak{p}\ \mathsf{def}(s,y))$. We note that $\mathsf{S}^\star$ is $\Sigma^0_{1,1}$, i.e. it can be rewritten, modulo provable equivalence, to a formula of the form $\exists x\, \forall y \leqslant t(x) \exists z\, S^\star_0(x,y.z)$, where $S^\star_0(x,y.z)$ is elementary. Here is a basic insight about $\mathsf{S}^\star$.

Lemma 6.9. $\mathsf{EA} + \mathfrak{p}\downarrow \vdash \Box_\beta \neg\, \mathsf{S}^\star$*, and, hence,* $\mathsf{EA} + \mathsf{S}^\star \vdash \Box_\beta \neg\, \mathsf{S}^\star$.

Proof. Let us write $\mathsf{def}^z(s,y)$ for $\exists v \leqslant z\, (\mathsf{T}(s,\varepsilon,v) \wedge \mathsf{U}(v)=y)$. Let

$$D(x) := (\forall y \leqslant x\, \exists s < x\ \mathsf{def}(s,y) \to \exists z\, \forall y \leqslant x\, \exists s < x\ \mathsf{def}^z(s,y)).$$

We work in EA. It is easily seen that D is closed under zero and successor. Let I be a cut shortening D. Here we assume that I is downward closed w.r.t. $\leqslant$ and is closed under successor, addition, multiplication and ω_1.

We note that $D(x)$ is equivalent to $\neg\,(\forall y \leqslant x\, \exists s < x\ \mathsf{def}(s,y))$, since EA verifies the Δ_0-pigeon hole principle. See e.g. [HP93], p. 42.

We have by a result of Pudlák (the *outside large, inside small* principle) that $\forall x\, \Box_\beta\, x \in I$. See [Pud85]. Suppose $\mathfrak{p}\downarrow$. It follows that $\Box_\beta\, \mathfrak{p} \in I$. Hence, $\Box_\beta \neg\, \mathsf{S}^\star$. ❑

The considerations from the proof of Theorem 6.13 show that $\mathsf{EA} \nvdash \mathsf{S}^\star \to \Box_\beta\bot$. Thus, it follows that $\mathsf{EA} \nvdash \mathsf{S}^\star \to \Box_\beta \mathsf{S}^\star$. We may conclude that EA does not prove $\Sigma_{1,1}$-completeness. The following result characterizes $\Box_\sigma$ under the assumption $\mathsf{S}^\star$.

Lemma 6.10. $\mathsf{EA} + \mathsf{S}^\star \vdash \Box_\sigma A \leftrightarrow \Box_\beta(\Diamond_\beta\top \to A)$.

Proof. We reason in $\mathsf{EA} + \mathsf{S}^\star$. We remind the reader that:

$$\sigma(x) :\leftrightarrow \beta(x) \vee (\mathfrak{p}\downarrow \wedge\, \exists y \leqslant \mathfrak{p}\ x = \ulcorner \ell_{\mathfrak{p}} \neq (y \dot{+} 1)\urcorner \wedge \exists s < \mathfrak{p}\ \mathsf{def}(s,y)).$$

In combination with $\mathsf{S}^\star$ this yields:

$$\sigma(x) \leftrightarrow \beta(x) \vee \exists y \leqslant \mathfrak{p}\ x = \ulcorner \ell_{\mathfrak{p}} \neq (y \dot{+} 1)\urcorner.$$

It follows that $\Box_\sigma A$ iff $\Box_{\beta + \bigwedge_{y\leqslant\mathfrak{p}}(\ell_{\mathfrak{p}}\neq y)} A$, hence, we find by Lemma 6.6(a), $\Box_\sigma A$ iff $\Box_\beta(\Diamond_\beta\top \to A)$. ❑

The following lemma gives our calculation of $\Box_\sigma\bot$.

Lemma 6.11. *We have:*

a. $\mathsf{EA} + \mathsf{S}^\star \vdash \Box_\sigma\bot \leftrightarrow \Box_\beta\Box_\beta\bot$.
b. $\mathsf{EA} + \neg\,\mathsf{S}^\star \vdash \Box_\sigma\bot \leftrightarrow \Box_\beta\bot$.
c. $\mathsf{EA} \vdash \Box_\sigma\bot \leftrightarrow ((\mathsf{S}^\star \wedge \Box_\beta\Box_\beta\bot) \vee \Box_\beta\bot)$.

Proof. (a): This is immediate from Lemma 6.10.

(b): The right-to-left direction is immediate. We prove the left-to-right direction. We work in $\mathsf{EA} + \neg\,\mathsf{S}^\star + \Box_\sigma\bot$. By Lemma 6.8(b), we get $\Box_\beta(\Diamond_\beta\top \to \bot)$. Hence, $\Box_\beta\Box_\beta\bot$, and so $\Box_\pi\bot$. Thus, $\mathfrak{p}\downarrow$.

Now $\neg\,\mathsf{S}^\star$, in combination with $\mathfrak{p}\downarrow$, tells us that there is an $i \leqslant \mathfrak{p}$, such that, for all $s < \mathfrak{p}$, we have $\neg\,\mathsf{def}(s,i)$. It follows that every σ axiom is either C or of the form $\ell_{\mathfrak{p}} \neq (j+1)$, where $j \leqslant p$ and $j \neq i$. Thus, $\Box_\sigma\bot$ yields: $\Box_\beta\neg \bigwedge_{j\leqslant\mathfrak{p},\, j\neq i} \ell_{\mathfrak{p}} \neq (j+1)$. It follows that $\Box_\beta\, \ell_{\mathfrak{p}} \neq (i+1)$. By Lemma 6.6(b), we find, as promised, $\Box_\beta\bot$.

(c) is immediate from (a) and (b). ❑

Here is our calculation of $\Box_\sigma\Diamond_\sigma\top$.

Lemma 6.12. *We have:*

a. $\mathsf{EA} + \mathsf{S}^\star \vdash \Box_\sigma\Diamond_\sigma\top$.
b. $\mathsf{EA} + \neg\,\mathsf{S}^\star \vdash \Box_\sigma\Diamond_\sigma\top \leftrightarrow \Box_\beta\bot$.
c. $\mathsf{EA} \vdash \Box_\sigma\Diamond_\sigma\top \leftrightarrow (\mathsf{S}^\star \vee \Box_\beta\bot)$.

Proof. Ad (a): We work in $\mathsf{EA} + \mathsf{S}^\star$. By Lemma 6.10, we find that $\Box_\sigma\Diamond_\sigma\top$ is equivalent with $\Box_\beta(\Diamond_\beta\top \to \Diamond_\sigma\top)$. By Lemma 6.9, we have $\Box_\beta\neg\,\mathsf{S}^\star$. Hence, by Lemma 6.11(b) and necessitation for $\Box_\beta$, we have $\Box_\beta(\Diamond_\sigma\top \leftrightarrow \Diamond_\beta\top)$. We may conclude that $\Box_\sigma\Diamond_\sigma\top$.

Ad (b): We reason in EA. The right-to-left direction is easy. We treat left-to-right. Suppose $\Box_\sigma \Diamond_\sigma \top$.

Suppose $\mathfrak{p}\uparrow$. In this case we find $\Box_\beta \Diamond_\sigma \top$ and hence $\Box_\beta \Diamond_\beta \top$. We may conclude $\Box_\beta \bot$.

Suppose $\mathfrak{p}\downarrow$. Since $\neg\,\mathsf{S}^\star$, It follows that, for some $i \leqslant \mathfrak{p}$, we have

$$\Box_\beta (\bigwedge_{j \leqslant \mathfrak{p}, j \neq i} \ell_{\mathfrak{p}} \neq (j+1) \to \Diamond_\sigma \top).$$

Ergo, $\Box_\beta(\ell_{\mathfrak{p}} = (i+1) \to \Diamond_\sigma \top)$. On the other hand, $\Box_\beta(\ell_{\mathfrak{p}} = (i+1) \to \Box_\beta \bot)$. Hence, $\Box_\beta(\ell_{\mathfrak{p}} = (i+1) \to \Box_\sigma \bot)$. We may conclude $\Box_\beta(\ell_{\mathfrak{p}} \neq (i+1))$ and, thus, $\Box_\beta \bot$.

Ad (c): (c) is immediate from (a) and (b). ❑

Theorem 6.13. $\mathsf{EA} \nvdash \Box_\sigma \Diamond_\sigma \top \to \Box_\sigma \bot$.

Our proof is a simple adaptation of a proof of Paris & Kirby ([PK78]). See also [Kay91].

Proof. Let $\mathcal{M}$ be a model of $\mathsf{PA} + \Box_\pi \bot$. Let $\mathcal{N}$ be the model given by all Σ_1-definable elements of $\mathcal{M}$. We can easily see that $\mathcal{N}$ is a Σ^0_1-elementary submodel of $\mathcal{M}$. It follows that Π^0_2-sentences are downwards preserved from $\mathcal{M}$ to $\mathcal{N}$. Hence, $\mathcal{N} \models \mathsf{EA}$. Reflecting on the construction of $\mathcal{N}$, we have $\mathcal{N} \models \mathsf{S}^\star$.

It follows, by Lemma 6.12(a), that $\mathcal{N} \models \Box_\sigma \Diamond_\sigma \top$. If we would have $\mathcal{N} \models \Box_\sigma \bot$, it would follow, by Lemma 6.11(a), that $\mathcal{N} \models \Box_\beta \Box_\beta \bot$. Since $\Box_\beta \Box_\beta \bot$ is Σ^0_1, we would find that $\mathcal{M} \models \Box_\beta \Box_\beta \bot$.[5] But this is impossible, since $\mathcal{M} \models \mathsf{PA}$ and PA proves reflection for $\Box_\beta$. ❑

Let G be the Gödel sentence for $\Box_\sigma$.

Corollary 6.14. $\mathsf{EA} \nvdash \Box_\sigma G \to \Box_\sigma \Box_\sigma G$.

Proof. If we would have $\mathsf{EA} \vdash \Box_\sigma G \to \Box_\sigma \Box_\sigma G$, then the usual reasoning for the formalized proof of the Second Incompleteness Theorem for $\Box_\sigma$ would go through, contradicting Theorem 6.13. ❑

Open Question 6.15. Here are four questions.

A. What are the possible provability logics of Σ^0_1-numerations of the axioms of EA over EA?
B. What are the possible closed fragments of provability logics of Σ^0_1-numerations of the axioms of EA over EA?
C. What is the provability logic of $\Box_\sigma$ over EA?
D. What is the closed fragment of the provability logic of $\Box_\sigma$ over EA?

❑

Example 6.16. Let $\mathcal{S}$ be the presentation constructed in this subsection. We note that $\mathsf{Cr}(\mathcal{S})$ is an elementary axiomatization. Hence, $\mathsf{Cr}(\mathcal{S})$ does satisfy the formalized Second Incompleteness Theorem over EA. It follows that $\mathcal{S} \neq_{\mathsf{EA},\mathsf{Id}_{\mathsf{Ar}}} \mathsf{Cr}(\mathcal{S})$. ❑

[5] We note that $\Box_\sigma \bot$ is not *prima facie* Σ^0_1 since it has $\Sigma^0_{1,1}$-form.

6.3. **The Second Incompleteness Theorem for Σ^0_1-presentations.** We prove the Second Incompleteness Theorem for Σ^0_1-presentations of U in U. As we have seen, in Subsection 6.1, Σ^0_1-presentations need not be uniform. Thus, we cannot use Theorem 5.1. Fortunately, if $\mathcal{S}$ is a Σ^0_1-semi-presentation, then $\mathsf{Th}(\mathcal{S})$ is a uniform Σ^0_1-semi-presentation. So, we may apply Theorem 5.2.

Theorem 6.17. *Suppose U is consistent and $N : U \rhd \mathsf{EA}$. Let $\sigma(x)$ be a Σ^0_1-formula and suppose that $\mathcal{S} := \sigma \uparrow \tau_N$ is a semi-presentation of U in U. Then, $U \nvdash (\Diamond_{\mathsf{A}}\top)^{\mathcal{S}}$.*

Proof. By Theorem 5.2, it is sufficient to show that $\mathsf{Th}(\mathcal{S})$ is a uniform semi-presentation of U in U.

Suppose X axiomatizes U and $\mathcal{S}$ semi-numerates X in U. Let C be a single axiom for EA. Let Y_0 be a finite subset of U. Let X_0 be a finite subset of X-axioms sufficient for deriving the sentences in $Y_0 \cup \{C^N\}$. We take

$$Y_1 := Y_0 \cup \{C^N\} \cup \{\sigma^N(\ulcorner B \urcorner) \mid B \in X_0\}.$$

We note that the $\sigma^N(\ulcorner C \urcorner)$ for C in X_0 are in U since $\mathcal{S}$ semi-numerates X. So $Y_1 \subseteq^{\mathsf{fin}} U$. Consider any $A \in Y_1$. In case $A \in Y_0 \cup \{C^N\}$, we have a proof of A from X_0, so certainly, $Y_1 \vdash (\Box_\sigma A)^N$, i.o.w. $Y_1 \vdash (\Box_{\mathsf{A}} A)^{\mathcal{S}}$. If A in $\{\sigma^N(\ulcorner C \urcorner) \mid C \in X_0\}$, then $Y_1 \vdash \sigma^N(\ulcorner C \urcorner)$. It follows by the fact that $C^N \in Y_1$ in combination with Σ^0_1-completeness in EA, that $U \vdash (\Box_S S^N(\ulcorner C \urcorner))^N$, so $U \vdash (\Box_{\mathsf{A}} A)^{\mathcal{S}}$. ❑

We remind the reader that if we want to push down our result from EA to S^1_2 without solving any complexity theoretic problems, then we can replace Σ^0_1 by $\exists\Sigma^{\mathsf{b}}_1$.

There is an alternative proof Theorem 6.17 that runs as follows.

Second proof of Theorem 6.17. Suppose $U \vdash (\Diamond_{\mathsf{A}}\top)^{\mathcal{S}}$. It follows, by Theorem 4.10, that $U \vdash (\Diamond_{\mathsf{A}}\top)^{\mathsf{Cr}(\mathcal{S})}$. By, Theorem 4.11(a), we see that U is closed under the necessitation rule for $(\Box_{\mathsf{A}})^{\mathsf{Cr}(\mathcal{S})}$. Inspection of the definition of $\mathsf{Cr}(\cdot)$ shows that $\mathsf{A}^{\mathsf{Cr}(\mathcal{S})}$ is elementary. Hence, $(\Box_{\mathsf{A}})^{\mathsf{Cr}(\mathcal{S})}$ satisfies the Löb conditions. It follows that $U \vdash \bot$. Quod non. ❑

Remark 6.18. The attentive reader may wonder whether the alternative proof is not more successful than I make it here. Does not the argument establish the theorem with S^1_2 substituted for EA? I do not think so. The point is that $\Box^{\mathsf{Cr}(\mathcal{S})}_{\mathsf{A}}$ is still Σ^0_1 rather than $\exists\Sigma^{\mathsf{b}}_1$. There are tricks to improve Craigification to yield a Δ^{b}_0-axiom set, but I suspect that these tricks manage to obstruct the proof of Theorem 4.11(a). Whatever is the case, there is more to explore here. ❑

7. Extensions of Peano Arithmetic: Examples

In this section we provide various examples of applications of our main theorem for theories extending Peano Arithmetic.

7.1. **Feferman Provability.** The contraposition of Theorem 5.1 tells us that if U is consistent and $M : U \rhd (\mathsf{S}^1_2 + \Diamond_{\mathsf{A}}\top)$, then $\mathcal{A} := \tau_M$ is not a uniform semi-presentation of U. The interesting case is of course, when $\mathcal{A}$ is a semi-presentation.

We consider the theory PA. Let π be a usual arithmetization of the axiom set of PA. We define π_y by $\pi_y(x) \leftrightarrow \pi(x) \wedge x \leqslant y$. Feferman defines a predicate $\widetilde{\pi}$ by $\widetilde{\pi}(x) :\leftrightarrow \pi(x) \wedge \Diamond_{\pi_x}\top$. The Feferman Presentation is $\mathcal{F} := \widetilde{\pi} \uparrow \mathsf{Id}_{\mathsf{Ar}}$. Since PA is reflexive, it follows that $\mathcal{F}$ is, indeed, a presentation for PA. On the other hand, trivially, we have $\mathsf{PA} \vdash (\Diamond_{\mathsf{A}}\top)^{\mathcal{F}}$. Thus, $\mathcal{F}$ is not uniform.

Feferman provability $\Box_{\widetilde{\pi}}$ has been studied intensively by provability logicians. See [Mon78], [Vis89] and [Sha94]. Shavrukov in [Sha94] characterizes the bimodal provability logic of $\Box_{\pi}$ and $\Box_{\widetilde{\pi}}$.

Applications of Feferman provability are based on the use of Feferman consistency as basis for a Henkin-style interpretations existency result. See [Vis11] and [VisXX]. The full generality of Feferman's method is only realized if we admit also restricted provability where we restrict our proofs to proofs only involving formulas below a given complexity. In a sense, we leave the Feferman framework there, since we employ other notions of proof than the standard one.

7.2. **Oracle Provability.** We consider the theory U axiomatized by the set X consisting of the usual PA-axioms plus all true Π^0_n-sentences. We define $\pi^{[n]} :\leftrightarrow \pi(x) \vee \mathsf{true}_{\Pi^0_n}(x)$.

It is easy to see that $\mathcal{P}^{[n]} := \pi^{[n]}(x) \uparrow \mathsf{Id}_{\mathsf{Ar}}$ uniformly enumerates X in U. Hence, we have the Second Incompleteness Theorem for $\mathcal{P}^{[n]}$. We can also verify this fact by showing that $\Box_{\mathsf{A}}^{\mathcal{P}^{[n]}}$ satisfies the Löb Conditions.

The provability logic of the predicates $\mathcal{P}^{[n]}$ is Japaridze's Logic. It has been studied in great detail by provability logicians. See e.g. [Jap85] and [Bek06].

7.3. **Failure of the Formalized Second Incompleteness Theorem over Peano Arithmetic.** Recently Taishi Kurahashi characterized the provability logics of some Σ^0_2-provability predicates. See [Kur17] and [KurXX]. One of his results is that there is a Σ_2-numeration of the axioms of PA such that the provability logic for that numeration is precisely K. Thus, Kurahashi provides an example of a Σ^0_2-axiomatization for which we do have the Second Incompleteness Theorem, but which does not satisfy the Löb-conditions. We provide a quick example of the same phenomenon here. Our aim is far more modest than Kurahashi's, we just provide failure of the Second Incompleteness Theorem and not a characterization of a provability logic.

Suppose the usual representation of the axioms of PA is π. As before, we define π_y by $\pi_y(x) :\leftrightarrow (\pi(x) \wedge x \leqslant y)$ and the Feferman axiomatization $\widetilde{\pi}(x) :\leftrightarrow \pi(x) \wedge \Diamond_{\pi_x}\top$. We define: $\pi^*(x) :\leftrightarrow \pi(x) \wedge \forall y \leqslant x\, \neg\, \mathsf{proof}_{\pi}(y, \Box_{\pi}\bot)$. Finally, let $(\widetilde{\pi} \vee \pi^*)(x) :\leftrightarrow \widetilde{\pi}(x) \vee \pi^*(x)$.

We (locally) write: $\vdash$ for provability in PA, $\Box := \Box_{\pi}$, $\triangle_0 := \Box_{\widetilde{\pi}}$, $\triangle_1 := \Box_{\pi^*}$, $\triangle := \Box_{\widetilde{\pi} \vee \pi^*}$.

It is easy to see that $\widetilde{\pi} \vee \pi^*$ uniformly semi-numerates $\mathsf{ax}_{\mathsf{PA}}$ in PA, since π^* does so. We have the stronger: if $A \in \mathsf{ax}_{\mathsf{PA}}$, then $\mathsf{PA}_k \vdash (\widetilde{\pi} \vee \pi^*)(A)$, where k is sufficiently large such that $\mathsf{PA}_k \vdash \mathsf{EA}$. Thus, $\mathsf{PA} \nvdash \triangledown\top$. We note that, since $\widetilde{\pi}$ and π^* are PA-provably subsets of π, we even have that $\widetilde{\pi} \vee \pi^*$ uniformly numerates $\mathsf{ax}_{\mathsf{PA}}$ in PA.

We show that PA does not prove the arithmetized Second Incompleteness Theorem for $\triangle$. Thus, *a fortiori*, we do not have the Löb conditions for $\triangle$.

We enumerate some salient facts.

a. $\vdash \triangle A \leftrightarrow (\triangle_0 A \vee \triangle_1 A)$.
b. $\vdash \triangle A \to \Box A$

c. $\vdash \Diamond\top \to (\triangle A \leftrightarrow \Box A)$.
d. $\vdash \triangle\bot \leftrightarrow \triangle_1\bot$.
e. $\vdash \Box\Box\bot \to \Box(\triangle A \leftrightarrow \triangle_0 A)$.
f. $\vdash \Box\Box\bot \to \Box\nabla\top$.
g. $\vdash \neg\Box\Box\bot \to \neg\Box\nabla\top$.
h. $\vdash \Box\nabla\top \leftrightarrow \Box\Box\bot$.

The items (a,b,c,d) are trivial.

Ad (e): We reason in PA. Suppose $\Box\Box\bot$. Let x be the smallest witness of $\Box\Box\bot$. We reason inside $\Box$. We find that $\triangle A$ is equivalent to $\triangle_0 A$ or $\Box_{x-1}A$. Since, by the reflexiveness of PA, $\Diamond_{x-1}\top$, we see that $\Box_{x-1}A$ implies $\triangle_0 A$, and we are done.

(f) is an immediate consequence of (e).

(g) is by formalizing the reasoning for $\nvdash \nabla\top$.

(h) follows by combining (f) and (g).

Using the above facts, we can show that $\nvdash \triangle\nabla\top \to \triangle\bot$. Suppose (†) $\vdash \triangle\nabla\top \to \triangle\bot$. We have:

$$\begin{aligned} \Box\Box\bot \wedge \Diamond\top &\vdash_{(\mathrm{f})} \Box\nabla\top \\ &\vdash_{(\mathrm{c})} \triangle\nabla\top \\ &\vdash_{(\dagger)} \triangle\bot \\ &\vdash_{(\mathrm{b})} \Box\bot \end{aligned}$$

It follows that $\Box\Box\bot \vdash \Box\bot$, and hence $\vdash \Box\bot$. Quod non. Thus, PA does not verify the Second Incompleteness Theorem for $\triangle$. It follows that $\triangle \to \triangle\triangle$ also fails over PA.

8. Beyond Reduction to Finite

We have provided a reasonably general version of the Second Incompleteness Theorem. Is this general version, the last word? *No.* As a matter of principle, there is no last word, there is nothing like a most general version of a theorem.

In the first place there are always new insights. Secondly, salient versions of a theorem are often dependent on further interests. In our case, we already know some limitations of what we do in the paper.

We have already seen, in the alternative proof of Theorem 6.17, that the pattern of proving the Second Incompleteness Theorem for a presentation by proving the same theorem for a 'weaker' presentation extends beyond the reduction to finite axiomatizability. Thus, our scheme can be extended.

In fact, the most important use of this kind of methodology does not fully fall under our scheme: Pudlák's proof of the Second Incompleteness Theorem for recursively enumerable extensions of Q. Here is an outline of the proof.

Theorem 8.1. *Let U be axiomatized by $X := \{C \in \mathsf{sent}_\Theta \mid \mathbb{N} \models \sigma(\ulcorner C \urcorner)\}$, where $\sigma(x)$ is a Σ^0_1-formula. Suppose $U \rhd (\mathsf{Q} + \Diamond_\sigma\top)$. Then, U is inconsistent.*

Proof. Suppose $N : U \rhd (\mathsf{Q} + \Diamond_\sigma\top)$. One can construct a definable cut N_0 of N such that $N_0 : U \rhd \mathsf{S}^1_2$. (See e.g. [HP93].) Since, N_0 is a cut and $\Diamond_\sigma\top$ is $\Pi_{1,1}$, i.e. of the form $\forall x\,\exists y < t(x)\,\forall z\,S_0(x,y,z)$, where S_0 is Δ_0, we find $U \vdash (\Diamond_\sigma\top)^{N_0}$. We may now apply Theorem 6.1. ❑

We note that Pudlák's argument transcends our framework since it involves the interplay between two interpretations of a weak arithmetic. In his paper [Pud85], Pudlák discusses a bimodal logic where two modalities work together to prove the Second Incompleteness Theorem for one of them. In fact, Pudlák addresses a problem that we did not touch upon in this paper: how to deal with inefficient Gödel numberings. As far as I know this proposal was never seriously explored further.

One thing suggested by Pudlák argument is to develop an ordering of presentations that also covers cases where the restriction to the arithmetical repertoire is not constant. This would involve definable initial mappings between such representations.

The Fefermanian restriction to a fixed proof-predicate excludes a lot of interesting contexts where consistency statements occur. There is a lot to say about cut-free, Herbrand and versions of restricted provability. For example, finitely axiomatized sequential theories prove their own cut-free consistency on a definable cut. There is the question whether, for example, S^1_2 proves its own cut-free consistency. The proper generalization of Feferman provability involves restricted consistency statements. Pavel Pudlák studied finitistic versions of consistency proofs. Etcetera.

We hope our paper gave the reader an impression of the generality of the Second Incompleteness Phenomenon.

References

[Bek06] L.D. Beklemishev. The worm principle. In Z. Chatzidakis, P. Koepke, and W. Pohlers, editors, *Logic Colloquium '02*, volume 27 of *Lecture Notes in Logic*, pages 75–95. A.K. Peters and CRC press, Natick, Massachusetts, 2006.

[Bus86] S.R. Buss. *Bounded Arithmetic*. Bibliopolis, Napoli, 1986.

[Fef60] S. Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, 49:35–92, 1960.

[HP93] P. Hájek and P. Pudlák. *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer, Berlin, 1993.

[Jap85] G. Japaridze. The polymodal logic of provability. In *Intensional Logics and Logical Structure of Theories: Material from the fourth Soviet–Finnish Symposium on Logic, Telavi*, pages 16–48, 1985.

[Kay91] R. Kaye. *Models of Peano Arithmetic*. Oxford Logic Guides. Oxford University Press, 1991.

[Kri62] S.A. Kripke. "flexible" predicates of formal number theory. *Proceedings of the American Mathematical Society*, 13(4):647–650, 1962.

[Kur17] T. Kurahashi. Arithmetical completeness theorem for modal logic K. *Studia Logica*, 2017.

[KurXX] T. Kurahashi. Arithmetical soundness and completeness by Σ_2 numerations. 20XX.

[Mon78] F. Montagna. On the algebraization of a Feferman's predicate (the algebraization of theories which express Theor; X). *Studia Logica*, 37:221–236, 1978.

[PK78] J.B. Paris and L.A.S. Kirby. Σ_n-collection schemas in arithmetic. In A. Macintyre, L. Pacholski, and J.B. Paris, editors, *Logic Colloquium '77*, pages 199–209. North–Holland, 1978.

[Pud85] P. Pudlák. Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, 50(2):423–441, 1985.

[Sha94] V.Yu. Shavrukov. A smart child of Peano's. *Notre Dame Journal of Formal Logic*, 35:161–185, 1994.

[TMR53] A. Tarski, A. Mostowski, and R.M. Robinson. *Undecidable theories*. North–Holland, Amsterdam, 1953.

[Vis89] A. Visser. Peano's smart children: A provability logical study of systems with built-in consistency. *Notre Dame Journal of Formal Logic*, 30(2):161–196, 1989.

[Vis05] A. Visser. Faith & Falsity: a study of faithful interpretations and false Σ^0_1-sentences. *Annals of Pure and Applied Logic*, 131(1–3):103–131, 2005.
[Vis11] A. Visser. Can we make the Second Incompleteness Theorem coordinate free. *Journal of Logic and Computation*, 21(4):543–560, 2011. First published online August 12, 2009, doi: 10.1093/logcom/exp048.
[Vis14] A. Visser. Peano Corto and Peano Basso: A study of local induction in the context of weak theories. *Mathematical Logic Quarterly*, 60(1–2):92–117, 2014.
[Vis15] A. Visser. Oracle bites theory. In S. Gosh and J. Szymanik, editors, *The Facts Matter. Essays on Logic and Cognition in Honour of Rineke Verbrugge*, pages 133–147. College Publications, 2015.
[VisXX] A. Visser. The interpretation existence lemma. In *Feferman on Foundations*, Outstanding Contributions to Logic. Springer, 20XX. To appear.
[Zam96] D. Zambella. Notes on Polynomially Bounded Arithmetic. *The Journal of Symbolic Logic*, 61(3):942–966, 1996.

Appendix A. Translations and Interpretations

We present the notion of *m-dimensional interpretation without parameters.* There are two extensions of this notion: we can consider piecewise interpretations and we can add parameters. We just treat the ordinary m-dimensional case without parameters here.

Consider two signatures Σ and Θ. An m-dimensional translation $\tau : \Sigma \to \Theta$ is a quadruple $\langle \Sigma, \delta, \mathcal{F}, \Theta\rangle$, where $\delta(v_0, \ldots, v_{m-1})$ is a Θ-formula and where for any n-ary predicate P of Σ, $\mathcal{F}(P)$ is a formula $A(\vec{v}_0, \ldots, \vec{v}_{n-1})$ in the language of signature Θ, where $\vec{v}_i = v_{i0}, \ldots, v_{i(m-1)}$. Both in the case of δ and A all free variables are among the variables shown. Moreover, if $i \neq j$ or $k \neq \ell$, then v_{ik} is syntactically different from $v_{j\ell}$.

We demand that we have $\vdash \mathcal{F}(P)(\vec{v}_0, \ldots, \vec{v}_{n-1}) \to \bigwedge_{i<n} \delta(\vec{v}_i)$. Here $\vdash$ is provability in predicate logic. This demand is inessential, but it is convenient to have.

We allow identity to be translated to a formula that is not identity.

We define B^τ as follows:

- $(P(x_0, \ldots, x_{n-1}))^\tau := \mathcal{F}(P)(\vec{x}_0, \ldots, \vec{x}_{n-1})$.
- $(\cdot)^\tau$ commutes with the propositional connectives.
- $(\forall x\, A)^\tau := \forall \vec{x}\, (\delta(\vec{x}\,) \to A^\tau)$.
- $(\exists x\, A)^\tau := \exists \vec{x}\, (\delta(\vec{x}\,) \wedge A^\tau)$.

There are two worries about this definition. First, what variables $\vec{x}_i$ on the side of the translation A^τ correspond with x_i in the original formula A? The second worry is that substitution of variables in δ and $\mathcal{F}(P)$ may cause variable clashes. These worries are never important in practice: we choose 'suitable' sequences $\vec{x}$ to correspond to variables x, and we avoid clashes by α-conversions. However, if we want to give precise definitions of translations and, for example, of composition of translations these problems come into play. These problems are clearly solvable, but they are beyond the scope of this paper.

Instead of introducing τ explicitly as being $\langle \Sigma, \delta, \mathcal{F}, \Theta\rangle$, we will write e.g. δ_τ for the δ of τ, and $P_\tau := \mathcal{F}_\tau(P)$.

We specify the identity translation and composition of translations.

- id_Σ is the identity translation. We take $\delta_{\mathsf{id}_\Sigma}(v) := v = v$ and $\mathcal{F}(P) := P(\vec{v}\,)$.
- We can compose translations. Suppose $\tau : \Sigma \to \Theta$ and $\nu : \Theta \to \Lambda$. Then $\nu \circ \tau$ or $\tau\nu$ is a translation from Σ to Λ. We define:
  - $\delta_{\tau\nu}(\vec{v}_0, \ldots, \vec{v}_{m_\tau - 1}) := \bigwedge_{i<m_\tau} \delta_\nu(\vec{v}_i) \wedge (\delta_\tau(v_0, \ldots, v_{m_\tau-1}))^\nu$.

– $P_{\tau\nu}(\vec{v}_{0,0},\ldots,\vec{v}_{0,m_\tau-1},\ldots,\vec{v}_{n-1,0},\ldots,\vec{v}_{n-1,m_\tau-1}) :=$
$\bigwedge_{i<n,j<m_\tau} \delta_\nu(\vec{v}_{i,j}) \wedge (P(v_0,\ldots,v_{n-1})^\tau)^\nu$.

A translation relates signatures; an interpretation relates theories. An interpretation $K : U \to V$ is a triple $\langle U,\tau,V\rangle$, where U and V are theories and $\tau : \Sigma_U \to \Sigma_V$. We demand: for U-sentences A, if $U \vdash A$, then $V \vdash A^\tau$.

We can define the identity interpretation and composition of interpretations as follows.

- $\mathsf{ID}_U : U \to U$ is the interpretation $\langle U, \mathsf{id}_{\Sigma_U}, U\rangle$.
- Suppose $K : U \to V$ and $M : V \to W$. Then, $KM := M \circ K : U \to W$ is $\langle U, \tau_M \circ \tau_K, W\rangle$.

Appendix B. Alternative Proofs for Theorem 6.1

Here is Theorem 6.1 again.

Theorem 6.1. *Suppose $\sigma(x)$ is a Σ^0_1-formula that numerates the axioms of U in the standard model. Then, if $U \rhd (\mathsf{S}^1_2 + \Diamond_\sigma\top)$, then U is inconsistent.*

Alternative Proof 1. Suppose $N : U \rhd (\mathsf{S}^1_2 + \Diamond_\sigma\top)$. Then, $\mathsf{S}^1_2 \vdash \Box_\sigma(\bigwedge \mathsf{S}^1_2 \wedge \Diamond_\sigma\top)^N$. Let $\gamma(x) :\leftrightarrow x = \ulcorner\bigwedge \mathsf{S}^1_2 \wedge \Diamond_\sigma\top\urcorner$. We find: $\mathsf{S}^1_2 \vdash \Box_\gamma\bot \to \Box_\sigma\bot$, In other words, $\mathsf{S}^2_1 + \Diamond_\sigma\top \vdash \Diamond_\gamma\top$. Since we have the Löb conditions for $\Box_\gamma$ over $\mathsf{S}^2_1 + \Diamond_\sigma\top$ we obtain $\mathsf{S}^2_1 + \Diamond_\sigma\top \vdash \bot$, by the Second Incompleteness Theorem. Hence, $\mathsf{S}^1_2 \vdash \Box_\sigma\bot$. But, then $U \vdash \bot$. ❑

Alternative Proof 2. Suppose $N : U \rhd (\mathsf{S}^1_2 + \Diamond_\sigma\top)$. Let N_0 be a cut in N on which we have $\mathsf{S}^1_2 + \mathrm{B}\Sigma_1$ with the additional property that $U \vdash \forall x \in \delta_{N_0}\, 2^x \in \delta_N$. We claim that:

$$U \vdash \forall A \in \delta_{N_0}\,((\Box_\sigma A)^{N_0} \to (\Box_\sigma(\Box_\sigma A)^{N_0})^N).$$

We briefly sketch the idea for the verification of this last equation. First, inside N_0 we can transform $\Box_\sigma A$ from a $\Sigma^0_{1,1}$-formula to a Σ^0_1-formula, say $(\Box_\sigma A)^\star$. We can construct $(\Box_\sigma A)^\star$ in such a way that $\mathsf{S}^1_2 \vdash (\Box_\sigma A)^\star \to \Box_\sigma A$. Secondly we can estimate the transformation of a witness x of a Σ^0_1-formula S to a witness of $\Box_\sigma S^{N_0}$ as of order $2^{k\cdot x\cdot|S|}$, for standard k. Note that this uses that we have a standardly finite verification of $(\mathsf{S}^1_2)^{\tau_{N_0}}$ in U. So, if we start with a witness p in N_0 of $(\Box_\sigma A)^\star$, we have a witness $p^\star$ of $\Box_\sigma(\Box_\sigma A)^{\star N_0}$ in N. We can transform $p^\star$ easily to a witness $\widetilde{p}$ of $\Box_\sigma(\Box_\sigma A)^{N_0}$ in N.

By the Gödel Fixed Point Lemma, we find G such that $\mathsf{S}^1_2 \vdash G \leftrightarrow \neg\Box_\sigma G^{N_0}$.

We now reason in U as follows. Suppose $(\Box_\sigma G^{N_0})^{N_0}$. Then, we have both $(\Box_\sigma G^{N_0})^N$ and $(\Box_\sigma(\Box_\sigma G^{N_0})^{N_0})^N$. By the fixed point equation, it follows that $(\Box_\sigma\neg G^{N_0})^N$, and, hence, $(\Box_\sigma\bot)^N$.

On the other hand, we have $(\Diamond_\sigma\top)^N$. So, we may conclude $\bot$. Eliminating our assumption, we obtain $\neg(\Box_\sigma G^{N_0})^{N_0}$, or, in other words, $(\neg\Box_\sigma G^{N_0})^{N_0}$. The fixed point equation gives us: G^{N_0}.

We leave U. We have shown $U \vdash G^{N_0}$. Since $\mathcal{T} := \sigma \upharpoonright \tau_{N_0}$ semi-numerates X in U, we find $U \vdash (\Box_\sigma G^{N_0})^{N_0}$ and, so, $U \vdash (\neg G)^{N_0}$. So $U \vdash \bot$. ❑

Alternative Proof 3. Suppose $N : U \rhd (\mathsf{S}^1_2 + \Diamond_\sigma \top)$. Suppose $\sigma(x)$ is written in the form $\exists u\, \sigma_0(u,x)$, where σ_0 is Δ_0. Without loss of generality we may assume that $x < u$ is implied by $\sigma_0(u,x)$. We define:

$$\alpha(x) :\leftrightarrow \exists w < x\, \exists u < |x| \exists y < |x|\, (w = \mathsf{tally}(y) \wedge \sigma_0(u,y) \wedge x = \mathsf{conj}(\mathsf{id}(w,w),y)).$$

Here tally computes the standard *non*-efficient numeral. We assume that our coding is such that $2^y \leqslant \mathsf{tally}(y)$. We note that all terms in $\sigma_0(u,y)$ can be bounded by polynomials in $|x|$. So, α is Σ^{b}_1. By the Σ^{b}_1-collection principle, we find that $\mathsf{proof}_\alpha(y)$ is Σ^{b}_1 (modulo S^1_2-provable equivalence). See [Bus86], p53. It follows that $\Box_\alpha$ is $\exists\Sigma^{\mathsf{b}}_1$.

We set $\mathcal{S} := \sigma(x) \uparrow \tau_N$ and $\mathcal{T} := \alpha(x) \uparrow \tau_N$. We find that $\mathcal{T} \preceq_{U,\tau_K} \mathcal{S}$. Hence, $U \vdash (\Diamond_{\mathsf{A}} \top)^{\mathcal{T}}$.

Since σ represents the axioms of U in the standard model, it follows that $\mathsf{Th}(\mathcal{T})$ semi-numerates U. The reason is that S^1_2 verifies the existence of $\mathsf{tally}(\underline{n})$, for standard n. Thus, $(\Box_{\mathsf{A}})^{\mathcal{T}}$ satisfies the Löb conditions. It follows that $U \vdash \bot$. ❑

Alternative Proof 4. Suppose $N : U \rhd (\mathsf{S}^1_2 + \Diamond_\sigma \top)$.

In [Vis14], we show that the theory Peano Basso is locally cut-interpretable in PA^-. We use the following consequence of this fact: the theory

$$W := \mathsf{S}^1_2 + \mathrm{B}\Sigma^0_1 + \{S \to S^I \mid I \text{ is an } \mathsf{S}^1_2\text{-definable cut and } S \text{ is a } \Sigma^0_1\text{-sentence}\}.$$

is locally cut-interpretable in S^1_2. It follows, by the downwards preservation of $\Pi_{1,1}$-sentences, that $T := W + \Diamond_\sigma \top$ is locally cut-interpretable in $\mathsf{S}^1_2 + \Diamond_\sigma \top$.

Let θ be a Δ^{b}_0 standard representation of the axioms of T. We define $\theta_n(x) :\leftrightarrow \theta(x) \wedge x \leqslant n$. We have that U as axiomatized by σ interprets W_n axiomatized by θ_n. Let us, *ad hoc*, write this as $\mathsf{int}[\sigma,\theta_n]$. Now, by $\Sigma^0_{1,1}$-completeness (in the metalanguage), we have $U \vdash (\mathsf{int}[\sigma,\theta_n])^N$. It follows that $U \vdash (\Diamond_{\theta_n} \top)^N$.

By the Interpretation Existence Lemma (see [VisXX]), we obtain an interpretation $M : U \rhd T$. It is easily seen that, in U, we have the Löb Conditions for $\Box^M_\sigma$, which yields a contradiction. ❑

Department of Philosophy, Utrecht University, Janskerkhof 13, 3512BL Utrecht, The Netherlands

E-mail address: a.visser@uu.nl