Nonstandard Provability for Peano Arithmetic

A Modal Perspective

Paula Henk

Nonstandard Provability for Peano Arithmetic

A Modal Perspective

ILLC Dissertation Series DS-2016-10

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

Cover design by tatlin.net
Printed and bound by GVO Drukkers & Vormgevers B.V.

ISBN: 978–94–6332–103–7

Nonstandard Provability for Peano Arithmetic

A Modal Perspective

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr.ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op vrijdag 16 december 2016, te 12.00 uur

door

Paula Henk

geboren te Tallinn, Estland

Promotiecommisie

Promotors:	Prof. dr. F.J.M.M. Veltman	Universiteit van Amsterdam
	Prof. dr. A. Visser	Universiteit Utrecht
Co-promotors:	Prof. dr. D.H.J. de Jongh	Universiteit van Amsterdam
	Dr. V.Yu. Shavrukov	Independent researcher
Overige leden:	Prof. dr. L.D. Beklemishev	Steklov Mathematical Institute, Moskow
	Prof. dr. J.F.A.K van Benthem	Universiteit van Amsterdam
	Prof. dr. Y. Venema	Universiteit van Amsterdam
	Prof. dr. L.C. Verbrugge	Rijksuniversiteit Groningen
	Dr. R. Iemhoff	Universiteit Utrecht
	Dr. L. Incurvati	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Contents

Acknowledgments

This thesis owes its existence to a constellation of people, ideas, and coincidences, which I feel very lucky to have been a part of.

Thank you, Albert. Your expertise and guidance were essential for this project. I especially want to thank you for our frequent meetings — for sharing your ideas and enthusiasm during lively discussions, as well as for your patience at times when it was needed. Dick, your support and guidance throughout the years have been very important to me. I am especially grateful to you for encouraging me to continue with a PhD project in the first place. Thank you, also, for writing a Dutch translation of the summary. Volodya, I cannot even imagine what this thesis would have been without you playing the role of an oracle — contributing beautiful ideas and surprising insights every now and then. Thank you for your patient and subtle guidance towards the arithmetical completeness proof of GLT, and while clearing up my confusions concerning quantifier complexities in $\mathrm{I}\Delta_0{+}\mathsf{exp}$. Writing an article together was an invaluable experience. Whether it comes to issues visible from mole's or bird's eye view — your thoroughness and high standards are an inspiration to me.

Albert, Dick, and Volodya, I want to say a very special word of thanks to you for the final weeks before submission. Not only did your feedback allow me to considerably improve the thesis, it also gave me the feeling of being part of a supportive team — something which was at least as important during this challenging time.

Thank you, Frank, for listening to my doubts and worries during a difficult time, for not taking them too seriously, and helping me reach the start of less difficult times. It's always a pleasure talking to you.

My stay at the Steklov Institute in Moscow was a turning point of my PhD — the start of its best part. I am very grateful to you Lev, for letting me visit, and making me feel welcome there. I learned so much during our meetings: about specific topics, but even more importantly about how to do logic in the first place. The resulting sense of inspiration and confidence played a crucial role in making the final year of my PhD so enjoyable. Fedja, your beautiful ideas occur in several parts of this thesis, most importantly the connection between Lindstöm logic and slow provability. I want to thank you for these ideas, as well as for many inspiring discussions — I have learned a lot from you.

Mathias, your feedback on the Introduction chapter was extremely helpful for making it read more fluently, as well as actually gentle (as it claims to be). Thank you for this and many other enjoyable conversations about nonstandard models and related issues. Thank you, Julia, for several helpful discussions, in particular for letting me distract you with those just before my deadline. But I am especially grateful for your friendship and continuing support throughout the PhD life in all its aspects - it has been a pleasure sharing this journey with you!

A very special word of thanks goes to my paranymphs. Thank you, Marta, for lovely weekends in Amsterdam and in den Haag, and for many excellent conversations. Your thoroughness and discipline are an inspiration, and I always look forward to spending time with you — no matter what mood I happen to be in! Vigjilenca, thank you for very special times in New York, Florence, and Amsterdam, for infecting me with ambition and kindness, and for your unwavering support. Our conversations always leave me feeling so inspired and optimistic! I feel very lucky to have the two of you by my side.

I am very grateful to all my friends who are still my friends despite of me having seemingly forgotten them during the time of intense writing. Last but not least, I want to thank my parents for their confidence in me and my decisions, as well as for their continuing support, especially during the last week before my submission. Aitäh teile!

Amsterdam — Paula Henk
October, 2016.

Chapter 1

Introduction

This thesis is a modal logical study of nonstandard provability predicates for Peano Arithmetic. We give a gentle introduction to the subject, which the reader familiar with the above notions should feel free to skip.

1.1 Numbers

We all know the natural numbers:

$$0, 1, 2, 3, \ldots$$

We can, of course, do much more with the natural numbers than just list them. They can be added or multiplied. They can also have various properties, for example being even, prime, or the smallest number expressible as a sum of cubes in two different ways.

Certain statements about the natural numbers are *true*, for example $1+1=2$, while certain other statements, for example $0=1$, are false. What we mean by truth, in this thesis, is always truth in the above sense, i.e. truth about the natural numbers.

Some truths are a matter of simple calculation — the stuff of elementary school. Here is an example:

$$97 \cdot 79 = 7663$$

The truth of other statements requires more insight to establish, for example:

"There are infinitely many prime numbers."

But some statements are so complicated that we do not yet know whether they are true or false at all. An example is *Goldbach's conjecture*:

"Every even number greater than 2 can be expressed as the sum of two primes."

If Goldbach's conjecture is true, then — since it is a statement about *infinitely* many numbers — we cannot establish it by simple calculation, similarly to the product of 97 and 79. However, there could still be a *systematic* and *mechanical* way of finding out truths. This brings us to the protagonist of this thesis, the theory of Peano Arithmetic.

1.2 Peano Arithmetic and Gödel's theorems

Peano Arithmetic (PA) is a theory about the natural numbers. We can think of it as a computation device for producing true statements. It consists of:

- some basic *facts*, for example: $n + 1 > 0$ and $n \cdot 0 = 0$ for any n

- some basic *rules* of reasoning, for example *modus ponens*: if φ holds, and φ implies ψ, then also ψ must hold

Applying the basic rules of reasoning to the basic facts, called *axioms*, allows PA to prove more complex statements about the natural numbers. What we mean by a *proof* is a finite sequence of sentences $\psi_0, \ldots, \psi_j$, where each ψ_i is either an axiom, or obtained from the previous sentences by one of the basic rules. We write $\mathsf{PA} \vdash \varphi$ if there exists a proof of the sentence φ.

Many true statements about the natural numbers are provable in PA. For example, $\mathsf{PA} \vdash 97 \cdot 79 = 7663$, but also

$$\mathsf{PA} \vdash \text{ For each prime number } p\text{, there exists a prime number } p' \text{ with } p' > p.$$

The axioms of PA are chosen in such a way that they are obviously true, while the rules of inference are guaranteed to preserve truth. It follows from this that every statement provable in PA is true. In other words, the natural numbers are a *model* of PA; we also say that PA is *sound*. Denote by $\perp$ some contradiction, for example $0 = 1$, and note that since PA is sound, it is clear that $\mathsf{PA} \nvdash \perp$. We say that PA is *consistent*.

It is, of course, good to know that PA only proves true statements. The really interesting question is, however, whether the converse also holds:

Can PA prove *every* true statement?

An answer — a negative one — to this question is given by Gödel's incompleteness theorems. Before explaining the latter, let us note that while we started off describing PA as a tool for reasoning about the natural numbers, the above question makes it an object of mathematical inquiry itself. We have moved from mathematics to *metamathematics*.

The First Incompleteness Theorem According to the First Incompleteness Theorem, there are statements about the natural numbers that are neither provable nor disprovable in PA. To formulate this more precisely, let us write $\neg\varphi$ for the negation of φ. Gödel's First Incompleteness Theorem states that there is a statement φ about the natural numbers for which

$$\mathsf{PA} \nvdash \varphi \text{ and } \mathsf{PA} \nvdash \neg\varphi.$$

We can think of PA as being undecided about φ — considering it possible that φ is true, but also considering it possible that φ is false. Since one of φ and $\neg\varphi$ must be true — in fact, it is clear from Gödel's proof which one it is — it follows that not every true statement is provable in PA.

The Second Incompleteness Theorem The Second Incompleteness Theorem gives a particularly interesting example of a statement that PA is undecided about: the *statement of its own consistency*. Making this precise requires some explanation.

The statement "PA is consistent" is, of course, about PA itself, not about the natural numbers. The method of arithmetisation, developed by Gödel, allows us to overcome this technical obstacle. The idea is to use the natural numbers themselves as *codes* for syntactical objects of PA such as sentences or proofs. This is done in such a way that simple operations on the syntactical objects, for example negating a sentence, become simple functions on the corresponding codes. Similarly, simple properties of the syntactical objects, for example "being a proof of the sentence $2 + 2 = 4$", become simple properties of the respective codes. This idea allows metamathematics to be done inside PA itself.

In particular, there is a formula $\mathrm{Pr}(x)$, the so-called *provability predicate*, that expresses in a natural way basic facts about provability in PA. For example, writing $\ulcorner\varphi\urcorner$ for the code of φ, we have for any φ,

$$\mathsf{PA} \vdash \varphi \text{ if and only if } \mathsf{PA} \vdash \mathrm{Pr}(\ulcorner\varphi\urcorner). \tag{1.1}$$

Roughly speaking, $\mathrm{Pr}(\ulcorner\varphi\urcorner)$ is PA's way of saying that it proves φ. We may thus read (1.1) as: PA proves φ if and only if PA knows that it proves φ.

The sentence $\neg\mathrm{Pr}(\ulcorner\bot\urcorner)$ then expresses that a contradiction is not provable in PA, in other words that PA is consistent. The Second Incompleteness Theorem states:

$$\mathsf{PA} \nvdash \neg\mathrm{Pr}(\ulcorner\bot\urcorner).$$

This may be seen as failure of negative introspection for PA: while PA is consistent — it does not know $\bot$ — it does not know this fact. In other words, PA does not know that it does not know $\bot$.

Hilbert-Bernays-Löb derivability conditions Hilbert and Bernays isolated from Gödel's work certain principles concerning $\mathrm{Pr}(x)$ which are sufficient for proving the Second Incompleteness Theorem. These were later simplified by Löb, and are now collectively referred to as the *Hilbert-Bernays-Löb* (*HBL*) *derivability conditions*:

1. $\mathsf{PA} \vdash \varphi \Rightarrow \mathsf{PA} \vdash \mathrm{Pr}(\ulcorner\varphi\urcorner)$
2. $\mathsf{PA} \vdash \mathrm{Pr}(\ulcorner\varphi \to \psi\urcorner) \to (\mathrm{Pr}(\ulcorner\varphi\urcorner) \to \mathrm{Pr}(\ulcorner\psi\urcorner))$
3. $\mathsf{PA} \vdash \mathrm{Pr}(\ulcorner\varphi\urcorner) \to \mathrm{Pr}(\ulcorner\mathrm{Pr}(\ulcorner\varphi\urcorner)\urcorner)$

Condition (1) is simply one direction of (1.1), while (2) states that PA knows that modus ponens is among its rules of inference. Condition (3) can be seen as positive introspection: if PA knows that it proves φ, then it knows this fact.

Löb's Theorem is a consequence of the HBL-conditions, together with the Fixed Point Lemma. It states that for any φ,

$$\text{if } \mathsf{PA} \vdash \mathrm{Pr}(\ulcorner\varphi\urcorner) \to \varphi, \text{ then } \mathsf{PA} \vdash \varphi.$$

The implication $\mathrm{Pr}(\ulcorner\varphi\urcorner) \to \varphi$ may be seen as PA's way of expressing its own soundness: "if I prove φ, then φ is true". Löb's Theorem states that PA is very modest, making the above claim only in case it can actually prove φ.

Taking $\bot$ for φ, Löb's Theorem tells us that $\mathsf{PA} \vdash \mathrm{Pr}(\ulcorner\bot\urcorner) \to \bot$ implies $\mathsf{PA} \vdash \bot$; in other words that if $\mathsf{PA} \nvdash \bot$, then $\mathsf{PA} \nvdash \neg\mathrm{Pr}(\ulcorner\bot\urcorner)$. The Second Incompleteness Theorem is thus an instance of Löb's Theorem.

1.3 Provability logic

According to the HBL-conditions, PA knows about some features of provability in itself. Löb's Theorem, on the other hand, indicates that there are limits to this knowledge. But what exactly are these limits? How can we describe *all* principles concerning $\mathrm{Pr}(x)$ that are provable in PA?

A beautiful answer to the above question emerges from a *modal* perspective. This brings us to the area of provability logic, where modal logic is used to investigate what formal theories such as PA can prove about provability and other metamathematical notions.

Writing $\Box$ instead of $\mathrm{Pr}(x)$, the modal counterparts of the HBL-conditions are:

$$\begin{array}{rl} (\mathsf{Nec}) & \vdash A \;\Rightarrow\; \vdash \Box A \\ (\mathsf{K}) & \Box(A \to B) \to (\Box A \to \Box B) \\ (4) & \Box A \to \Box\Box A \end{array}$$

A reader familiar with modal logic recognises the modal system K4. The system GL, named after Gödel and Löb, is obtained by adding to K4 the modal counterpart of Löb's rule:

$$\vdash \Box A \to A \;\;\Rightarrow\; \vdash A.$$

GL is alternatively axiomatised by adding to the basic modal logic K the following, known as Löb's axiom: $\Box(\Box A \to A) \to \Box A$.

Axiom (4) follows from Löb's axiom over K. The proof of Löb's Theorem in turn can be seen — modulo the Fixed Point Lemma — as a derivation in K4. There is thus an interplay between Löb's axiom on the one hand, and axiom (4) together with self-reference — in the form of the Fixed Point Lemma — on the other.

Given the HBL-conditions and Löb's Theorem, it is clear that the rules and axioms of GL may be used when reasoning about $\mathrm{Pr}(x)$ in PA. It was shown by Solovay ([Sol76]) that we will not miss anything when doing so: the theorems of GL are exactly the propositional schemata concerning $\mathrm{Pr}(x)$ that are provable in PA. In other words, GL is the provability logic of PA.

In view of Solovay's Theorem we shall usually write $\Box\varphi$ instead of $\mathrm{Pr}(\ulcorner\varphi\urcorner)$. The consistency statement $\neg\mathrm{Pr}(\ulcorner\bot\urcorner)$ thus becomes $\neg\Box\bot$, allowing us to state the Second Incompleteness Theorem as: PA $\nvdash \neg\Box\bot$.

Philosophical significance The importance of provability logic is manifold. In the context of modal logic, provability logic stands out by endowing the modal operator with an unambiguous interpretation. Formal provability, unlike other common interpretations of $\Box$ such as truth, knowledge, or necessity, has a precise mathematical definition. Which modal axioms are the correct ones is thus not a matter of dispute but a matter of proof. For $\mathrm{Pr}(x)$, the relevant proofs of soundness and completeness were provided, respectively, by Löb and Solovay.

From a foundations-of-mathematics point of view, GL embodies — in a direct and simple manner — a substantial body of reasoning leading to the incompleteness theorems. Its principles reflect salient features of formal provability. While $\mathrm{Pr}(x)$ may, at first sight, seem like an unfathomable creature, Solovay's Theorem provides a description that speaks directly to our intuition — $\mathrm{Pr}(x)$ behaves like a modal operator, governed by the principles of GL.

We have been treating formal provability as a kind of epistemic modality for PA. However, a modal perspective makes it clear that the provability modality is very different from our usual notion of knowledge, which can be seen as informal provability. One of the most important modal principles for the latter is $\Box A \to A$ — "If I know something, then it is true". In view of Löb's Theorem, it is clear that this principle is incompatible with GL. Modal analysis thus tells us that the notion of formal provability is essentially different from that of informal provability.

1.4 Feferman provability

Feferman, in his influential paper [Fef60], constructed a curious formula $\triangle_{\mathtt{f}}$, the so-called *Feferman provability predicate*. One could argue that $\triangle_{\mathtt{f}}$, like $\Box$, is a provability predicate for PA: writing $\triangle_{\mathtt{f}}\varphi$ as shorthand for $\triangle_{\mathtt{f}}(\ulcorner\varphi\urcorner)$, we have

$$\mathsf{PA} \vdash \varphi \quad \text{if and only if} \quad \mathsf{PA} \vdash \triangle_{\mathtt{f}}\varphi.$$

At the same time, however, $\mathsf{PA} \vdash \neg\triangle_{\mathtt{f}}\bot$, i.e. PA proves its own Feferman-consistency. Instead of contradicting the Second Incompleteness Theorem, the existence of $\triangle_{\mathtt{f}}$ illustrates the need for a more careful formulation of this result.

In order to have a closer look at the situation, let us point out that the axioms of PA include the induction *schema*, i.e. the induction axiom for every statement φ. The theory $\mathrm{I}\Sigma_n$ is obtained from PA by restricting the use of induction to formulas of certain complexity, depending on n. Thus we have that $\mathsf{PA} = \bigcup_{n\in\omega} \mathrm{I}\Sigma_n$.

The formula $\triangle_{\mathtt{f}}$ defines provability in the theory

$$\mathsf{PA}^{\mathsf{f}} := \bigcup_{n\in\omega}\{\mathrm{I}\Sigma_n \mid \text{for all } m \leqslant n, \mathrm{I}\Sigma_m \text{ is consistent}\}.$$

Since we know that PA is consistent, we also know that each $\mathrm{I}\Sigma_n$ is consistent. But this implies $\mathsf{PA}^{\mathsf{f}} = \bigcup_{n\in\omega} \mathrm{I}\Sigma_n$, i.e. that PA and PA^{f} are, in fact, one and the same theory. It is in this sense that $\triangle_{\mathtt{f}}$ might be claimed to be a provability predicate for PA.

On the other hand, the theory PA^{f} is *consistent by definition*. Indeed, PA^{f} could equivalently be described as the largest consistent subtheory of PA in the sequence $(\mathrm{I}\Sigma_n)_{n\in\omega}$. Given this, it should come as no surprise that its consistency is known to PA, i.e. that $\mathsf{PA} \vdash \neg\triangle_{\mathtt{f}}\bot$.

The system PA^{f} could also be introduced by outlining the following proof-procedure:

1. Enumerate pairs $(\pi_0, \varphi_0), (\pi_1, \varphi_1), (\pi_2, \varphi_2), \ldots$, where π_i is a PA-proof of φ_i.
2. As soon as $\varphi_i = \bot$, determine the amount of induction $\mathrm{I}\Sigma_n$ used in π_i.
3. Backtrack and delete (π, φ), whenever π makes use of $\mathrm{I}\Sigma_{n'}$ for $n' \geqslant n$.
4. Return to step 1, but skip (π, φ) whenever π makes use of $\mathrm{I}\Sigma_{n'}$ for $n' \geqslant n$.

Let us call a proof *stable* if it occurs in our enumeration and is never scratched out. The stable proofs are exactly the proofs of PA^{f}. The catch is that we need to wait infinitely long in order to be sure that a proof is stable.

The above description portrays $\triangle_{\mathtt{f}}$ as being allowed to change its mind about the statements it considers provable. This kind of self-correcting behaviour is arguably closer to the way humans reason than the one embodied by the ordinary provability predicate. As such, Feferman provability is related to the so-called *trial-and-error predicates* and *experimental systems* studied by Putnam ([Put65]) and Jeroslow ([Jer75]). The provability logics of such systems have been studied by Visser ([Vis89]) and Shavrukov ([Sha94]).

1.4.1 Feferman provability as nonstandard provability

Arguably, a notion of provability with built-in consistency is a rather unusual one. But are there precise criteria for separating "strange" provability predicates such as $\triangle_{\mathtt{f}}$ from "normal" ones such as $\Box$?

As a way approaching the above question, note that our argument for the equivalence of PA and PA^{f} relied on the consistency of PA. By the Second Incompleteness Theorem, this argument is therefore not available when reasoning inside PA. Indeed, the equivalence of the two theories is not known to PA: it does not prove $\triangle_{\mathtt{f}}\varphi \leftrightarrow \Box\varphi$ for all φ. In fact, since $\mathsf{PA} \vdash \neg\triangle_{\mathtt{f}}\bot$ and $\mathsf{PA} \nvdash \neg\Box\bot$, it is clear that

$$\mathsf{PA} \nvdash \Box\bot \to \triangle_{\mathtt{f}}\bot.$$

Following Smoryński ([Smo85, p.279]), let us call a provability predicate $\triangle$ for PA *standard* if the schema $\triangle\varphi \leftrightarrow \Box\varphi$ is provable in PA. Thus, while $\triangle_{\mathtt{f}}$ might claim to be a provability predicate for PA, it is not a standard one.

In this thesis, the term *nonstandard provability predicate* is used to refer to the provability predicate $\triangle$ of a theory PA° such that — even though PA° coincides with PA in a strong enough metatheory T — the schema $\triangle\varphi \leftrightarrow \Box\varphi$ is not verifiable in PA. In case of $\triangle_{\mathtt{f}}$, we can take as T any theory that knows about the consistency of PA, for example $\mathsf{PA} + \neg\Box\bot$.

1.4.2 Bimodal provability logic

Our reluctance to accept $\triangle_{\mathtt{f}}$ as a genuine provability predicate could be explained by emphasising that certain natural principles for formal provability are not verifiable for it in PA. In fact, since $\mathsf{PA} \vdash \neg\triangle_{\mathtt{f}}\bot$, at least one of the Hilbert-Bernays-Löb-derivability conditions must fail for $\triangle_{\mathtt{f}}$. While conditions (1) and (2) do hold for $\triangle_{\mathtt{f}}$, condition (3) does not, i.e. there are sentences φ for which

$$\mathsf{PA} \nvdash \triangle_{\mathtt{f}}\varphi \to \triangle_{\mathtt{f}}\triangle_{\mathtt{f}}\varphi.$$

Despite of $\triangle_{\mathtt{f}}$ not obeying the same modal principles as $\Box$, its complete behaviour can nevertheless be described by means of modal logic. The modal logic F, formulated in the language containing a modal operator $\triangle$, is axiomatised by adding to K the following:

(F1) $\neg\triangle\bot$

(F2) $\triangle A \to \triangle((\triangle B \to B) \vee \triangle A)$

Shavrukov ([Sha94, Remark 4.12]) showed that F is the provability logic of $\triangle_{\mathtt{f}}$: the propositional schemata involving $\triangle_{\mathtt{f}}$ that are provable in PA are exactly the theorems of F.

Since $\mathsf{PA} \nvdash \triangle_{\mathtt{f}}\varphi \leftrightarrow \Box\varphi$, it is also interesting to ask which principles concerning the interaction of $\Box$ and $\triangle_{\mathtt{f}}$ are verifiable in PA. Adopting a modal perspective provides a neat answer to this question as well. The bimodal system LF contains the principles of GL for $\Box$, the principles of F for $\vartriangle$, as well as the following:

$$\begin{array}{ll} (\mathsf{T2}) & \Box A \to \vartriangle\Box A \\ (\mathsf{T3}) & \Box A \to \Box\vartriangle A \\ (\mathsf{F3}) & \Box A \leftrightarrow (\vartriangle A \vee \Box\bot) \end{array}$$

It was shown by Shavrukov ([Sha94, Theorem 4.9]) that the PA-provable propositional schemata concerning the interaction of ordinary and Feferman provability are exactly the theorems of LF. We say that LF is the *joint provability logic* of $\Box$ and $\vartriangle_{\mathtt{f}}$.

Philosophical significance While the fact that the provability logic of $\vartriangle_{\mathtt{f}}$ is different from that of $\Box$ might seem to suggest using provability logic as a means for distinguishing natural provability predicates from unnatural ones, such a strategy is not viable in general.

In this thesis, we shall see several nonstandard provability predicates whose provability logic is GL. In this respect, they are indistinguishable from the ordinary provability predicate. Their nonconformity becomes obvious, however, when focusing on their *interaction* with ordinary provability. In some cases, the difference between nonstandard and standard provability is thus only visible from a bimodal point of view.

A bimodal perspective is also useful in the study of standard provability. Solovay's Theorem for ordinary provability is very robust: GL is not just the provability logic of PA, but the provability logic of any reasonable[1] theory. This generality might be seen as a drawback, implying that a simple modal approach does not allow us to distinguish between interesting properties of theories such as finite axiomatisability or essential reflexivity.

The joint behaviour of two provability predicates turns out to be less uniform than that of a single provability predicate alone. There is no system that could justifiably be called *the* bimodal provability logic. For example, if S is a finite extension of T, the joint provability logic of S and T is different than in the case where S proves $\Box_T\varphi \to \varphi$ for all φ. The joint provability logic of two provability predicates thus tells us something about the nature of the relationship between the corresponding theories.

1.5 Overview of the thesis

This thesis is an exploration of certain nonstandard provability predicates and their modal logics, in particular their joint provability logic with the ordinary

[1] What we mean by *reasonable* here is: a Σ_1-sound extension of $\mathrm{I}\Delta_0+\mathsf{exp}$ with a recursively enumerable axiomatisation.

provability predicate $\Box$. The bimodal system GLT, studied in Chapter 3, plays a central role in our thesis. The principles of GLT describe the interaction of $\Box$ with various different nonstandard provability predicates. Chapter 4 is about *fast* and *slow provability*, while Chapter 5 deals with the so-called *supremum adapters*. In Chapter 6 it is shown that the provability logic of a certain supremum adapter is GL.

1.5.1 Fast and slow provability

Chapter 4 is concerned with the theories PA^* and $\mathsf{PA}{\restriction}_{\mathtt{F}}$ that can be seen as a *speeded up* and a *slowed down* version of PA, respectively.

Fast provability The theory PA^*, first studied by Parikh ([Par71]), is obtained by adding to PA the following inference rule, known as *Parikh's rule*:

$$\frac{\Box\varphi}{\varphi}.$$

From the metaperspective, we can see that Parikh's rule is admissible in PA: if $\mathsf{PA} \vdash \Box\varphi$, then by soundness, $\mathsf{PA} \vdash \varphi$. Thus PA^* has exactly the same theorems as PA.

As in the case of PA, we can construct a formula $\triangle_{\mathtt{p}}$ expressing in a natural way the property of being a PA^*-proof. The above argument for the equivalence of PA^* and PA made use of soundness and can therefore, as a consequence of Löb's theorem, not be used when reasoning in PA. Indeed, the schema $\triangle_{\mathtt{p}}\varphi \leftrightarrow \Box\varphi$ turns out not to be verifiable in PA. Thus $\triangle_{\mathtt{p}}$, like $\triangle_{\mathtt{f}}$, is a nonstandard provability predicate for PA.

It follows from the proof of Solovay's Theorem that GL is the provability logic of $\triangle_{\mathtt{p}}$. It was shown by Lindström ([Lin06]) that the joint provability logic of $\Box$ and $\triangle_{\mathtt{p}}$ is the modal system GLT.

One way to understand PA's ignorance about the equivalence of PA and PA^* is to note that some theorems have much shorter proofs in PA^* than in PA — we say that PA^* has *speed-up* over PA. The increase in proof length when converting PA^*-proofs into PA-proofs grows faster than any provably total function of PA. We shall thus refer to the notion of provability specified by $\triangle_{\mathtt{p}}$ as *fast provability*.

Slow provability Exploiting the fact that certain total functions cannot be proven to be total in PA, a notion of *slow provability* can also be defined. Such a notion was introduced by Friedman, Rathjen, and Weiermann in [FRW13]. They consider a certain computable function $\mathtt{F}$ on the natural numbers whose totality is not provable in PA, and the theory defined as

$$\mathsf{PA}{\restriction}_{\mathtt{F}} := \bigcup_{n\in\omega} \{\mathrm{I}\Sigma_n \mid \mathtt{F}(n){\downarrow}\}.$$

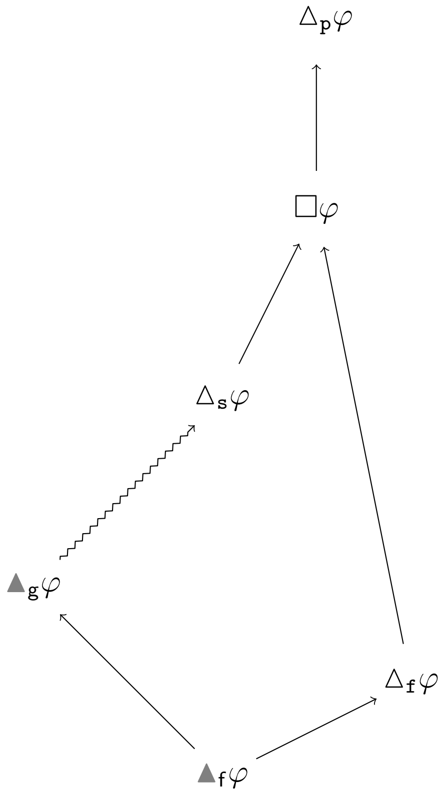

Figure 1.1: The zoo of provability predicates. Arrows indicate provable inclusion in PA. The squiggly arrow indicates provability modulo an index shift.

When proving theorems in $\mathsf{PA}\restriction_{\mathtt{F}}$, we can only use induction for formulas of complexity n after having computed the value of $\mathtt{F}(n)$. The process of proving theorems in $\mathsf{PA}\restriction_{\mathtt{F}}$ is therefore potentially slower than the process of proving theorems in PA — depending on how long the required calculations take.

Nevertheless, since $\mathtt{F}$ is total it is clear that $\mathsf{PA}\restriction_{\mathtt{F}}$ and PA have exactly the same theorems. Arguing in PA, on the other hand, the totality of $\mathtt{F}$ cannot be assumed, and so $\mathsf{PA}\restriction_{\mathtt{F}}$ might seem to be a weaker theory than PA. As these considerations suggest, the provability predicate $\triangle_{\mathtt{s}}$ of $\mathsf{PA}\restriction_{\mathtt{F}}$ is a nonstandard one.

It follows from the proof of Solovay's Theorem that GL is the provability logic of $\triangle_{\mathtt{s}}$. We show that the joint provability logic of $\triangle_{\mathtt{s}}$ and $\Box$ is GLT. Our proof is rather general, and also yields a new proof of the fact that GLT is the joint provability logic of $\Box$ and $\triangle_{\mathtt{p}}$. The joint provability logic of slow and ordinary provability is thus the same as that of ordinary and fast provability.

1.5.2 Supremum adapters

The theories PA^* and $\mathsf{PA}{\restriction_{\mathsf{F}}}$ of Chapter 4 are recursively enumerable (r.e.). In Chapter 5, we study nonstandard provability predicates corresponding theories defined in a "non-r.e." way.

A formula φ is said to be 1-*provable* in a theory T if it is provable in T together with all true Π_1-sentences. We say that T is 1-*inconsistent* if $\bot$ is 1-provable in T. Note that if T is inconsistent, then it is also 1-inconsistent. On the other hand, a theory could be consistent but 1-inconsistent. Let

$$\mathsf{PA}^\mu := \bigcup_{n\in\omega}\{\mathrm{I}\Sigma_n \mid \text{for all } m < n,\ \mathrm{I}\Sigma_m \text{ is 1-consistent}\}.$$

Thus PA^μ is defined similarly to PA^{f}, except for requiring 1-consistency instead of ordinary consistency, and using $<$ instead of $\leqslant$. The theory PA^μ may alternatively be described as $\mathrm{I}\Sigma_\mu$, where μ is the smallest n such that $\mathrm{I}\Sigma_n$ is 1-inconsistent.

Since we know that PA is sound, we also know that it is 1-consistent. Using this, it is clear that PA^μ is in fact the same theory as PA. On the other hand, it follows from the Second Incompleteness Theorem that PA does not prove its own 1-inconsistency, and so the above argument is inaccessible when reasoning in PA. As one might expect, the provability predicate $\blacktriangle_{\mathtt{g}}$ of PA^μ is another nonstandard provability predicate for PA.

We call $\blacktriangle_{\mathtt{g}}$ a *supremum adapter*, because it turns out to be useful for obtaining interpretability-suprema of finite extensions of PA, i.e. theories of the form $\mathsf{PA}+\varphi$. Roughly speaking, the theory $\mathsf{PA}+(\neg\blacktriangle_{\mathtt{g}}\neg\varphi \wedge \neg\blacktriangle_{\mathtt{g}}\neg\psi)$ is the weakest theory that is stronger than each of the theories $\mathsf{PA}+\varphi$ and $\mathsf{PA}+\psi$.

In Chapter 6, it is shown that GL is the provability logic of $\blacktriangle_{\mathtt{g}}$. Since $\blacktriangle_{\mathtt{g}}$ is not the provability predicate of a r.e. theory, this result is not a simple consequence of the proof of Solovay's Theorem, as was the case with $\triangle_{\mathtt{p}}$ and $\triangle_{\mathtt{s}}$. The joint provability logic of $\Box$ and $\blacktriangle_{\mathtt{g}}$ contains GLT, together with an additional modal principle S. Whether the joint provability logic of $\Box$ and $\blacktriangle_{\mathtt{g}}$ is equal to GLT together with S is an open question.

A slight modification of the definition of PA^μ yields another supremum adapter $\blacktriangle_{\mathtt{f}}$. In contrast to $\blacktriangle_{\mathtt{g}}$, the formula $\blacktriangle_{\mathtt{f}}$ behaves according to the principles of the modal system F. Determining the provability logic of $\blacktriangle_{\mathtt{f}}$, as well as its joint provability logic with $\Box$, remain challenges for future work.

1.6 Sources of the material

Much of the material in this thesis has been published elsewhere. Chapter 4 contains a subset of the material contained in [HP16]. Chapter 5 is loosely based on [HV16], but also contains some new results. Chapter 6 is, modulo some minor changes, [HS16].

Chapter 2

Preliminaries

This chapter introduces the central notions and results used in the thesis. We assume the reader to be familiar with first-order logic and its model theory, as well as basic modal logic.

2.1 Arithmetical theories

We work with first-order theories formulated in the language $\mathcal{L}$ of arithmetic containing 0, S, $+$, $\cdot$, and $\leqslant$. We assume a Hilbert-style axiomatisation of first-order logic, with modus ponens as the only rule of inference. Such a system can be found for example in [Fef60, Section 2].

The *standard model* of arithmetic, denoted by $\mathfrak{N}$, are the natural numbers together with the usual arithmetical structure. An $\mathcal{L}$-sentence φ is said to be *true* if $\mathfrak{N} \models \varphi$. We define for each natural number n an $\mathcal{L}$-term $\overline{n}$ by letting $\overline{0} = 0$ and $\overline{n+1} = \mathsf{S}\overline{n}$. Given this, we shall mostly write n instead of $\overline{n}$. Terms of the form $\overline{n}$ are called *numerals*.

An $\mathcal{L}$-formula is *bounded* or Δ_0 (equivalently, Σ_0 or Π_0) if all quantifiers occurring in it are of the form $\exists x \leqslant y$ or $\forall x \leqslant y$. A formula is Σ_{n+1} (Π_{n+1}) if it is of the form $\exists x_1 \dots \exists x_n\, \varphi$ ($\forall x_1 \dots \forall x_n\, \varphi$), with φ a Π_n (Σ_n)-formula. Formulas obtained from Σ_1-formulas by using propositional connectives and bounded quantification are said to be $\Delta_0(\Sigma_1)$.

The basic facts concerning 0, S, $+$, $\cdot$, and $\leqslant$ are given by the axioms of the theory Q of Robinson Arithmetic ([HP93, Definition I.1.1]). The theory Q is Σ_1-complete: it proves every true Σ_1-sentence ([HP93, Theorem I.1.8]).

Given a class Γ of formulas, $\mathrm{I}\Gamma$ is the theory obtained by adding to Q the induction schema for Γ-formulas. For $n > 0$, $\mathrm{I}\Sigma_n$ is finitely axiomatisable ([HP93, Theorem I.2.52]). The theory of Peano Arithmetic (PA) is given as $\bigcup_{n \in \omega} \mathrm{I}\Sigma_n$.

The graph of the exponentiation function x^y is definable in $\mathrm{I}\Delta_0$ by a Δ_0-formula ([HP93, Theorem V.3.15]). To be more precise, there is a Δ_0-formula

$\varphi_{\mathsf{e}}(x, y, z)$ for which:

$$\mathrm{I}\Delta_0 \vdash \varphi_{\mathsf{e}}(x, 0, z) \leftrightarrow z = 1$$
$$\mathrm{I}\Delta_0 \vdash \varphi_{\mathsf{e}}(x, y+1, z) \leftrightarrow \exists w\, (\varphi_{\mathsf{e}}(x, y, w) \wedge z = w \cdot x)$$

Denote by exp the sentence $\forall x\, \forall y\, \exists! z\, \varphi_{\mathsf{e}}(x, y, z)$ stating, intuitively, that exponentiation is a total function. We have that $\mathrm{I}\Sigma_1 \vdash \mathsf{exp}$ ([HP93, I.1.50]). On the other hand, since every Δ_0-defined provably total function of $\mathrm{I}\Delta_0$ is bounded by a polynomial ([HP93, Theorem V.1.4]), it is clear that $\mathrm{I}\Delta_0 \nvdash \mathsf{exp}$. The theory $\mathrm{I}\Delta_0{+}\mathsf{exp}$ is finitely axiomatisable ([HP93, Theorem V.5.6]).

The language of EA (Elementary Arithmetic) is obtained by adding to $\mathcal{L}$ a function symbol exp for binary exponentiation 2^x. The theory EA, given as $\mathrm{I}\Delta_0^{\exp}$ together with the recursive definition of 2^x, is strong enough to formalise almost all of finitary mathematics outside logic.

A formula is *elementary* or $\Delta_0^{\exp}$ if it is Δ_0 in the language of EA. We can also speak of $\Delta_0^{\exp}$-formulas in the context of $\mathrm{I}\Delta_0{+}\mathsf{exp}$: given the Δ_0-formula defining exponentiation as above, we can use the well-known term-elimination algorithm ([Vis92, Section 7.3]) in order to replace terms of the form 2^t with $\mathcal{L}$-formulas. A formula φ is $\Delta_n(T)$ if $T \vdash \varphi \leftrightarrow \sigma$ and $T \vdash \varphi \leftrightarrow \pi$ for some Σ_n-formula σ and Π_n-formula π.

The proof of the following theorem is a minor variation of [GD82, Proposition 2.1].

2.1.1. THEOREM. *Every $\Delta_0^{\exp}$-formula is Δ_1 in $\mathrm{I}\Delta_0{+}\mathsf{exp}$.* ❑

It follows from [GD82, Theorem 3.1] that $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \mathrm{I}\Delta_0^{\exp}$. Since, as is well-known, EA is a conservative extension of $\mathrm{I}\Delta_0{+}\mathsf{exp}$, the two theories can therefore be treated as equivalent for most purposes.

2.1.1 Provably recursive functions

Every primitive recursive relation R is represented in $\mathrm{I}\Delta_0{+}\mathsf{exp}$ by a Σ_1-formula φ_R in the sense that for all $n_0, \ldots, n_k$,

$$(n_0, \ldots, n_k) \in R \text{ iff } \mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \varphi_R(\overline{n_0}, \ldots, \overline{n_k}).$$

According to Kleene's Normal Form Theorem for recursive functions, there exist a primitive recursive relation T (Kleene's T-predicate) and a primitive recursive function U, such that for every recursive function $\mathtt{f}$, there is some e such that for all inputs n,

$$\mathtt{f}(n) \simeq \mathsf{U}\,(\mu y\, \mathsf{T}(e, n, y))\,. \tag{2.1}$$

Above, $\mu y\, \mathsf{T}(e, n, y)$ denotes the smallest number k for which $\mathsf{T}(e, n, k)$ holds, and $\mathtt{g}(m) \simeq \mathtt{h}(m)$ means that $\mathtt{g}(m)$ and $\mathtt{h}(m)$ are either both undefined, or defined and equal.

Making use of (2.1), we can associate to any k-ary recursive function $\mathtt{f}$ a Σ_1-formula $\varphi_{\mathtt{f}}$ defining its graph in $\mathrm{I}\Delta_0+\mathsf{exp}$, i.e. for all $n_1, \ldots, n_k$,

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \varphi_{\mathtt{f}}(\overline{n_1}, \ldots, \overline{n_k}, \overline{\mathtt{f}(n_1, \ldots, n_k)}), \text{ and}$$
$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \exists! z\, \varphi_{\mathtt{f}}(\overline{n_1}, \ldots, \overline{n_k}, z).$$

We assume every recursive function $\mathtt{f}$ to be equipped with such a Σ_1-formula $\varphi_{\mathtt{f}}$. Given a k-ary recursive function $\mathtt{f}$, we denote by $\mathtt{f}(x_1, \ldots, x_k){\downarrow}$ the formula $\exists y\, \varphi_{\mathtt{f}}(x_1, \ldots, x_k, y)$, and say that $\mathtt{f}$ *converges* on input $x_1, \ldots, x_k$. Similarly, we denote by $\mathtt{f}(x_1, \ldots, x_k){\uparrow}$ the formula $\neg \mathtt{f}(x_1, \ldots, x_k){\downarrow}$, and say that $\mathtt{f}$ *diverges* on input $x_1, \ldots, x_n$. We use $\mathtt{f}{\downarrow\downarrow}$ as shorthand for $\forall x_1 \ldots x_k\, \mathtt{f}(x_1, \ldots, x_k){\downarrow}$, and $\mathtt{f}{\uparrow\uparrow}$ as shorthand for $\neg \mathtt{f}{\downarrow\downarrow}$.

Since $\mathrm{I}\Delta_0+\mathsf{exp}$ is Σ_1-sound, it follows from the above that any recursively enumerable (r.e.) set A can be represented in $\mathrm{I}\Delta_0+\mathsf{exp}$ by a Σ_1-formula φ_A in the sense that for all n: $n \in A \Leftrightarrow \mathrm{I}\Delta_0+\mathsf{exp} \vdash \varphi_A(n)$. It was first shown in [EF60] that the above holds for every recursively enumerable (not necessarily sound) extension T of $\mathrm{I}\Delta_0+\mathsf{exp}$.

Suppose that T is an extension of $\mathrm{I}\Delta_0+\mathsf{exp}$. A k-ary recursive function $\mathtt{f}$ is said to be *provably total* in T if for some Σ_1-formula $\varphi_{\mathtt{f}}$ defining its graph in $\mathrm{I}\Delta_0+\mathsf{exp}$, $T \vdash \mathtt{f}{\downarrow\downarrow}$.

The following result was established independently by Parsons ([Par70]), Mints ([Min73]), and Takeuti ([Tak75]):

2.1.2. Theorem. *The provably recursive functions of* $\mathrm{I}\Sigma_1$ *are exactly the primitive recursive functions.* ❑

Whenever $\mathtt{f}$ is provably total in some r.e. theory $T \supseteq \mathrm{I}\Delta_0+\mathsf{exp}$, we assume $\mathtt{f}$ to be equipped with a Σ_1-formula $\varphi_{\mathtt{f}}$ defining its graph in T, and such that $T \vdash \mathtt{f}{\downarrow\downarrow}$. For a characterisation of the provably recursive functions of $\mathrm{I}\Sigma_n$ for $n > 1$, see Theorem 4.2.3.

The class of (Kalmar) elementary functions is the smallest class containing successor, zero, projection, addition, multiplication, subtraction, and closed under composition as well as bounded sums and bounded products ([Ros84]).

According to Kleene's Normal Form Theorem for elementary functions, there exist an elementary relation T' and an elementary function U', such that for every elementary function $\mathtt{h}$, there is some e such that for all inputs n,

$$\mathtt{h}(n) \simeq \mathsf{U}'\left(\mu y\, \mathsf{T}'(e, n, y)\right). \tag{2.2}$$

It can be shown that T' and U' are represented in $\mathrm{I}\Delta_0+\mathsf{exp}$ by elementary formulas. Using (2.2), we can thus associate to every elementary function $\mathtt{h}$ an elementary formula $\varphi_{\mathtt{h}}$ defining its graph in $\mathrm{I}\Delta_0+\mathsf{exp}$. A proof of the following result can be found for example in [SW12, Section 3.1].

2.1.3. Theorem. *The provably recursive functions of* $\mathrm{I}\Delta_0+\mathsf{exp}$ *are exactly the elementary functions.* ❑

In view Theorem 2.1.3 and the remark above it, we assume every elementary function $\mathtt{h}$ to be equipped with an elementary formula $\varphi_{\mathtt{h}}$ defining its graph in $\mathrm{I}\Delta_0+\mathsf{exp}$, and such that $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \mathtt{h}{\downarrow}{\downarrow}$.

2.1.4. THEOREM. ([HP93, Remark I.1.59(3)]) *Let* $\mathtt{f}$, $\mathtt{g}$ *and* $\mathtt{k}$ *be elementary. Suppose that* $\mathtt{h}$ *is defined from* $\mathtt{f}$ *and* $\mathtt{g}$ *by primitive recursion, and majorised by* $\mathtt{k}$. *Then* $\mathtt{h}$ *is elementary, hence provably total in* $\mathrm{I}\Delta_0+\mathsf{exp}$, *and moreover the defining equations of* $\mathtt{h}$ *are provable in* $\mathrm{I}\Delta_0+\mathsf{exp}$. ❑

2.1.2 Arithmetisation of syntax

We assume some standard formalisation of syntactical notions in $\mathrm{I}\Delta_0+\mathsf{exp}$, writing $\ulcorner\varphi\urcorner$ for the code of φ. If the meaning is clear from the context, we shall often identify syntactical objects with their codes, writing φ instead of $\ulcorner\varphi\urcorner$.

Smooth recursively enumerable theories

We introduce the notion of a *smooth* recursively enumerable theory. The reason a recursively enumerable theory is required to be smooth is to ensure that its natural provability predicate is provably equivalent in $\mathrm{I}\Delta_0+\mathsf{exp}$ to a Σ_1-formula.

By a theory T we shall, from now on, mean a pair (Ax_T, τ), where Ax_T is a set containing the non-logical axioms of T, and τ an arithmetical formula representing Ax_T in the standard model $\mathfrak{N}$. We say that τ is an *axiomatisation* of T. If Ax_T is r.e. then, as explained in Section 2.1.1, τ may taken to be Σ_1, and we have for all φ,

$$\varphi \in \mathrm{Ax}_T \Leftrightarrow \mathrm{I}\Delta_0+\mathsf{exp} \vdash \tau(\varphi).$$

We say that T is r.e. if τ is Σ_1. Similarly, T is said to be elementary just in case τ is elementary.

Following [Fef60, Definition 4.1], we define the formula $\mathrm{Pr}_\tau(x)$ expressing in a natural way provability in the theory $T = (\mathrm{Ax}_T, \tau)$:

$$\mathrm{Pr}_\tau(x) := \exists p\, (p = \langle \psi_0, \ldots, \psi_j \rangle \wedge \psi_j = x \wedge \forall i \leqslant j\, (\tag{2.3}$$

$$\lambda(\psi_i) \;\vee\; \exists k, l < j\, \psi_k = \psi_l \to \psi_i \;\vee\, \tau(\psi_i))), \tag{2.4}$$

where λ is an elementary formula representing the axioms of first-order logic in $\mathrm{I}\Delta_0+\mathsf{exp}$. The free variable x of $\mathrm{Pr}_\tau(x)$ is assumed to range over (codes of) $\mathcal{L}$-sentences.

If τ is elementary, it follows from Theorem 2.1.1 that $\mathrm{Pr}_\tau(x)$ is equivalent in $\mathrm{I}\Delta_0+\mathsf{exp}$ to a Σ_1-formula.

The axiom sets of most natural theories are elementary. In this thesis, however, we shall also encounter theories given to us via axiom sets that are r.e. but not elementary. We would like to argue that also in these cases, $\mathrm{Pr}_\tau(x)$ is equivalent to a Σ_1-formula in $\mathrm{I}\Delta_0+\mathsf{exp}$. In other words, $\mathrm{I}\Delta_0+\mathsf{exp}$ should know that the notion of theoremhood in a r.e. theory is r.e. itself.

2.1.5. Remark. This context might remind one of Craig's Theorem ([HP93, III.2.29]): given a r.e. theory $T = (\mathrm{Ax}_T, \tau)$, there is an elementary formula τ' representing Ax_T in $\mathfrak{N}$. The theory $T' = (\mathrm{Ax}_T, \tau')$ is thus elementary and equivalent to T in $\mathfrak{N}$. However, it follows from the results of Visser in [Vis15a] that Craig's Theorem is not verifiable in $\mathrm{I}\Delta_0+\mathsf{exp}$, and so the equivalence of T and T' might not be known to $\mathrm{I}\Delta_0+\mathsf{exp}$.

Let us thus examine the case where the axiomatisation τ of T is Σ_1 but not necessarily elementary. In this case, there is a Δ_0-formula $\tau'(y, x)$ such that $\tau(x) = \exists y\, \tau'(y, x)$. The formula $\mathrm{Pr}_\tau(x)$ is thus equivalent in $\mathrm{I}\Delta_0+\mathsf{exp}$ to a formula of the form

$$\exists p\ (\delta(p) \wedge \forall i \leqslant j\, (\delta'(i) \vee \exists y\, \tau'(y, \psi_i)))\,,$$

with δ and δ' elementary. In the presence of Σ_1-collection, the above formula is provably equivalent to the Σ_1-formula

$$\exists b\, \exists p\ (\delta(p) \wedge \forall i \leqslant j\, (\delta'(i) \vee \exists y < b\, \tau'(y, \psi_i)))\,.$$

Collection for Σ_1-formulas is, however, not provable in $\mathrm{I}\Delta_0+\mathsf{exp}$ ([HP93, Theorem 2.5]). In order to make everything work smoothly in $\mathrm{I}\Delta_0+\mathsf{exp}$, we shall work with theories satisfying a slightly stronger condition than being recursively enumerable.

2.1.6. Definition. An axiomatisation τ is *smooth* if

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall x < u\, \exists y\, \tau'(y, x) \to \exists b\, \forall x < u \exists y < b\, \tau'(y, x).$$

In other words, τ is smooth if $\mathrm{I}\Delta_0+\mathsf{exp}$ proves the collection axiom given by τ. If τ is Σ_1, then its smoothness is exactly what is needed in order to conclude that $\mathrm{Pr}_\tau(x)$ is provably equivalent in $\mathrm{I}\Delta_0+\mathsf{exp}$ to a Σ_1-formula.

We say that $T = (\mathrm{Ax}_T, \tau)$ is smooth just in case τ is smooth. From the above considerations, it is clear that if T is r.e. and smooth, then the provability predicate $\mathrm{Pr}_\tau(x)$ is provably equivalent in $\mathrm{I}\Delta_0+\mathsf{exp}$ to a Σ_1-formula. We note that all elementary theories are smooth.

2.1.7. Convention. Throughout this thesis, we use modal notation for provability predicates. Variants of the symbol $\Box$ are mostly used for natural provability predicates of PA and its fragments. As usual, we use $\Diamond$ to denote the dual of $\Box$, i.e. $\Diamond\varphi$ is written as an abbreviation for $\neg\Box\neg\varphi$. Variants of $\vartriangle$ and its dual $\triangledown$ are mainly reserved for nonstandard provability predicates.

The symbol $\Box_0$ denotes the natural provability predicate of $\mathrm{I}\Delta_0+\mathsf{exp}$, while $\Box$ is written for the natural provability predicate of PA. However some sections specify a local, more general interpretation for the symbol $\Box$.

As usual, $\Box\varphi(\dot{x})$ means that the numeral for the value of x has been substituted for the free variable of the formula φ inside $\Box$. If the intended meaning is clear from the context, we will often write $\Box\varphi(x)$ instead of $\Box\varphi(\dot{x})$.

The HBL-conditions

Let $T = (\mathrm{Ax}_T, \tau)$ be a smooth and recursively enumerable extension of $\mathrm{I}\Delta_0+\mathsf{exp}$, and write $\Box$ for the formula $\mathrm{Pr}_\tau(x)$ defined as in (2.3)-(2.4). In virtue of the smoothness and recursive enumerability of T we may, as explained above, assume that $\Box$ is Σ_1 when reasoning in $\mathrm{I}\Delta_0+\mathsf{exp}$. We shall furthermore assume $\mathrm{I}\Delta_0+\mathsf{exp}$ to know that T is an extension of itself:

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box_0\varphi \to \Box\varphi) \tag{2.5}$$

The theory $\mathrm{I}\Delta_0+\mathsf{exp}$ is provably Σ_1-complete ([HP93, Theorem 4.32]), meaning that for any Σ_1-formula σ, $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall y\,(\sigma(y) \to \Box_0\sigma(\dot{y}))$. Since $\Box$ is Σ_1 it follows from this, together with (2.5), that $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box\varphi \to \Box\Box\varphi)$. From the definition of $\Box$, it is clear that the closure of T under modus ponens is verifiable in $\mathrm{I}\Delta_0+\mathsf{exp}$. Summarising the above observations, we see that $\mathrm{I}\Delta_0+\mathsf{exp}$ verifies the *Hilbert–Bernays–Löb* (*HBL*) *derivability conditions* for $\Box$:

1. $T \vdash \varphi \;\;\Rightarrow\;\; \mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box\varphi$
2. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
3. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box\varphi \to \Box\Box\varphi$

Conditions (2) and (3) also hold when φ and ψ are regarded as internal variables ranging over $\mathcal{L}$-sentences.

We recall the Fixed Point Lemma, first extracted from the proof of Gödel's First Incompleteness Theorem by Carnap ([Car37]).

2.1.8. THEOREM. *Let φ be an $\mathcal{L}$-formula whose free variables are exactly $x_0, \ldots, x_n$. Then there is an $\mathcal{L}$-formula ψ with exactly the same free variables, and such that*

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \psi\,(x_1, \ldots, x_n) \leftrightarrow \varphi\,(\ulcorner\psi\urcorner, x_1, \ldots, x_n)\,.$$ ❑

From the proof of Theorem 2.1.8 it is clear that if φ is $\Sigma_n(\Pi_n)$, then so is ψ. The Fixed Point Lemma, together with the above assumptions on T and $\Box$, are sufficient for proving Löb's Theorem ([Löb55]) for T.

2.1.9. THEOREM. *If $T \vdash \Box\varphi \to \varphi$, then $T \vdash \Box\varphi$.*

Proof: By the Fixed Point Lemma, let ϑ be such that

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \vartheta \leftrightarrow (\Box\vartheta \to \varphi)\,.$$

We reason as follows, using the HBL-conditions for $\Box$:

$$\begin{aligned}
\mathrm{I}\Delta_0+\mathsf{exp} &\vdash \vartheta \to (\Box\vartheta \to \varphi)\\
&\vdash \Box\,(\vartheta \to (\Box\vartheta \to \varphi))\\
&\vdash \Box\vartheta \to \Box\,(\Box\vartheta \to \varphi)\\
&\vdash \Box\vartheta \to (\Box\Box\vartheta \to \Box\varphi)\\
&\vdash \Box\vartheta \to \Box\varphi
\end{aligned}$$

Assuming that $T \vdash \Box\varphi \to \varphi$, we thus have:

$$\begin{aligned} T &\vdash \Box\vartheta \to \varphi \\ &\vdash \vartheta \\ &\vdash \Box\vartheta \\ &\vdash \varphi \end{aligned}$$

by using the properties of ϑ and the HBL-conditions for $\Box$. ❑

An inspection of the proof of Theorem 2.1.9 shows that it can be verified in $\mathrm{I}\Delta_0+\mathsf{exp}$:

2.1.10. COROLLARY. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box(\Box\varphi \to \varphi) \to \Box\varphi$ ❑

The rules and axioms of the modal system GL (Section 2.2) can thus be used when reasoning about $\Box$ in $\mathrm{I}\Delta_0+\mathsf{exp}$.

Provability in $\mathrm{I}\Sigma_x$

The formula $\Box_x$ is the conventional Σ_1-provability predicate for $\mathrm{I}\Sigma_x + \mathsf{exp}$, with x a free variable. We write $\Box$ for the provability predicate of PA, where we assume that for all φ, $\Box\varphi$ is provably equivalent in $\mathrm{I}\Delta_0+\mathsf{exp}$ to $\exists x\, \Box_x\varphi$. Given a natural formalisation of provability in $\mathrm{I}\Sigma_x$, it is clear that:

1. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_x\varphi \wedge x \leqslant y \to \Box_y\varphi$ (monotonicity)
2. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_x(\varphi \to \psi) \to (\Box_x\varphi \to \Box_x\psi)$
3. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_0\varphi \to \Box_x\varphi$

As above, it follows from (3) and provable Σ_1-completeness of $\mathrm{I}\Delta_0+\mathsf{exp}$ that

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_x\varphi \to \Box_x\Box_{\dot{x}}\varphi.$$

Using the Fixed Point Lemma, together with the above, it can be shown that Löb's Theorem for $\mathrm{I}\Sigma_x+\mathsf{exp}$ is verifiable in $\mathrm{I}\Delta_0+\mathsf{exp}$: $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_x(\Box_{\dot{x}}\varphi \to \varphi) \to \Box_x\varphi$.

Partial satisfaction predicates

It is well-known that in $\mathrm{I}\Delta_0+\mathsf{exp}$ there is a partial satisfaction predicate $\mathrm{Sat}_{\Pi_1}(\varphi, y)$ for Π_1-formulas, where y and φ are internal variables ranging, respectively, over assignments and $\mathcal{L}$-formulas. The formula Sat_{Π_1} is Π_1 and satisfies Tarski's conditions ([HP93, Theorem I.2.55]). Defining $\mathrm{Tr}_{\Pi_1}(\varphi)$ to be the formula saying that φ is a sentence and $\forall y\, \mathrm{Sat}_{\Pi_1}(\varphi, y)$, it is clear that Tr_{Π_1} is Π_1, and that for any Π_1-formula $\pi(x)$,

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \pi(x) \leftrightarrow \mathrm{Tr}_{\Pi_1}(\pi(\dot{x})).$$

(By our conventions for the dot notation, $\pi(\dot{x})$ is a sentence from the point of view of Tr_{Π_1}.)

Given a theory T, 1-*provability* refers to provability in T together with all true Π_1-sentences. Using the formula Tr_{Π_1}, we can define the provability predicate $\Box_x^{\Pi_1}$ for 1-provability in $\mathrm{I}\Sigma_x + \mathsf{exp}$:

$$\Box_x^{\Pi_1}\varphi := \exists\pi\,(\mathrm{Tr}_{\Pi_1}(\pi) \wedge \Box_x(\pi \to \varphi)).$$

Similarly, the provability predicate $\Box^{\Pi_1}$ for 1-provability in PA is defined as:

$$\Box^{\Pi_1}\varphi := \exists\pi\,(\mathrm{Tr}_{\Pi_1}(\pi) \wedge \Box(\pi \to \varphi)).$$

It is then clear that for all φ, $\Box^{\Pi_1}\varphi$ is provably equivalent in $\mathrm{I}\Delta_0{+}\mathsf{exp}$ to $\exists x\,\Box_x^{\Pi_1}\varphi$. We note that $\Box_x^{\Pi_1}$ is Σ_2. It is well-known that $\Box_x^{\Pi_1}$ is Σ_2-complete, i.e. that for any Σ_2-formula σ,

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \sigma(y) \to \Box_x^{\Pi_1}\sigma(\dot{y}).$$

It follows from this that $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box_x^{\Pi_1}\varphi \to \Box_x^{\Pi_1}\Box_{\dot{x}}^{\Pi_1}\varphi$. Furthermore, $\mathrm{I}\Delta_0{+}\mathsf{exp}$ verifies that modus ponens is among the rules of inference of $\Box_x^{\Pi_n}$, i.e. we have $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box_x(\varphi \to \psi) \to (\Box_x\varphi \to \Box_x\psi)$. Using the above, together with the Fixed Point Lemma, it can be shown $\mathrm{I}\Delta_0{+}\mathsf{exp}$ verifies Löb's axiom for $\Box_x^{\Pi_n}$: $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box_x^{\Pi_1}(\Box_{\dot{x}}^{\Pi_1}\varphi \to \varphi) \to \Box_x^{\Pi_1}\varphi$.

In [HP93, Theorem I.4.33] it is shown that $\mathrm{I}\Sigma_{k+1}$ proves the consistency of the set of all true Π_{k+2}-sentences. The proof can be formalised $\mathrm{I}\Delta_0{+}\mathsf{exp}$:

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall x\,\forall\varphi\,(\varphi \in \Pi_{x+2} \to \Box_{x+1}(\Box_x\varphi \to \varphi))$$

Since $\mathrm{I}\Sigma_k + \mathsf{exp}$ is axiomatised by a single Π_{k+2}-sentence, it follows from the above that $\mathrm{I}\Sigma_{k+1}$ proves the consistency of $\mathrm{I}\Sigma_k + \mathsf{exp} + \Pi_1$-truth:

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box_{x+1}\neg\Box_{\dot{x}}^{\Pi_1}\bot.$$

We refer to the above properties as *reflection*.

A theory T is said to be *essentially reflexive* if it proves the consistency of each of its finite subtheories, and the same holds for every consistent extension in the same language. It follows from the above that, verifiably in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, the theory PA is essentially reflexive: $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,\forall x\,\Box(\varphi \to \Diamond_x\varphi)$.

2.1.3 Interpretability and arithmetised model theory

The notion of interpretability that we are interested in is that of *relative interpretability*, first introduced and carefully studied by Tarski, Mostowski and Robinson ([TMR53]). Due to the availability of a pairing function in all theories considered in this thesis, it is safe to focus our attention on one-dimensional interpretations.

2.1.11. Definition. Let S and T be first-order theories whose languages are $\mathcal{L}_S$ and $\mathcal{L}_T$. An *interpretation* j of S in T is a tuple $\langle \delta, \tau \rangle$, where δ is an $\mathcal{L}_T$-formula with one free variable, and τ a mapping from relation symbols[1] R of $\mathcal{L}_S$ to formulas R^τ of $\mathcal{L}_T$, where the number of free variables of R^τ is equal to the arity of R. We extend τ to a translation from all formulas of $\mathcal{L}_S$ to formulas of $\mathcal{L}_T$ by requiring:

i. $(R(x_1, \ldots x_n))^\tau = R^\tau(x_1, \ldots x_n)$

ii. $(\varphi \to \psi)^\tau = \varphi^\tau \to \psi^\tau$

iii. $\bot^\tau = \bot$

iv. $(\forall x\, \varphi)^\tau = \forall x\, (\delta(x) \to \varphi^\tau)$

Finally, we require that $T \vdash \exists x\, \delta(x)$, and $T \vdash \varphi^\tau$ for all axioms φ of S.

We write $j : T \rhd S$ if j is an interpretation of S in T, and $T \rhd S$ if $j : T \rhd S$ for some j. We say that T and S are mutually interpretable, and write $T \equiv S$, if $T \rhd S$ and $S \rhd T$.

We are interested in interpretability between finite extensions of PA, i.e. theories of the form $\mathsf{PA} + \varphi$, where φ is an $\mathcal{L}$-sentence. We write $\varphi \rhd \psi$ as an abbreviation for $\mathsf{PA} + \varphi \rhd \mathsf{PA} + \psi$.

Interpretability, like provability, is a syntactical notion, and can therefore be formalised in $\mathrm{I}\Delta_0 + \mathsf{exp}$. We also write $\varphi \rhd \psi$ for the $\mathcal{L}$-sentence expressing that $\mathsf{PA} + \varphi$ interprets $\mathsf{PA} + \psi$, certain that the intended meaning is always clear from the context.

The following result is implicit in [Ore61], and was first explicitly stated in [Háj71] and in [HH72]. Item (iii) was added in [Gua79]. Inspection of the proof shows that it can be verified in $\mathrm{I}\Delta_0 + \mathsf{exp}$.

2.1.12. Theorem (Orey-Hájek Characterisation). $\mathrm{I}\Delta_0 + \mathsf{exp}$ *verifies that for all* φ *and* ψ*, the following are equivalent:*

i. $\varphi \rhd \psi$

ii. $\forall x\, \Box(\varphi \to \Diamond_x \psi)$

iii. $\Box(\psi \to \pi) \to \Box(\varphi \to \pi)$ *for any* Π_1*-sentence* π *(we say that* ψ *is* Π_1*-conservative over* φ*)* ❏

[1] We assume here that S is formulated in a purely relational way. This restriction is not essential – function symbols can be replaced by relation symbols by a well-known algorithm (see [Vis92, Section 7.3]).

Model theory within $I\Sigma_1$

It is well-known that basic model-theoretic notions and proofs can be formalised in $I\Sigma_1$. We recall here the basic definitions and facts concerning the latter, referring the reader to [HP93, Section 4(b)] for a more extensive overview. Reasoning within $I\Sigma_1$, what we mean by a *model* $\mathfrak{M}$ is an interpretation, i.e. formulas δ, φ_0, φ_S, φ_+, $\varphi_\cdot$, and $\varphi_\leqslant$, defining the domain as well as the interpretations of the non-logical symbols in $\mathfrak{M}$. A *full model* is a model with a satisfaction relation for all first-order sentences in the language of the model. Reasoning in $I\Sigma_1$, we always assume models to be full.

We recall some model-theoretic definitions, which can easily seen to be formalisable in $I\Sigma_1$.

If $\mathfrak{M}$ and $\mathfrak{M}'$ are models of $I\Delta_0+\mathsf{exp}$, we say that $\mathfrak{M}$ is an *end-extension* of $\mathfrak{M}'$ (or that $\mathfrak{M}'$ is a *cut* of $\mathfrak{M}$) if $\mathfrak{M}$ is an extension of $\mathfrak{M}'$, and for every $a \in \mathfrak{M}$ and $b \in \mathfrak{M}'$, we have that $\mathfrak{M} \models a < b$ implies $a \in \mathfrak{M}'$.

Given a set Γ of $\mathcal{L}$-formulas, we say that $\mathfrak{M}$ is a Γ-*elementary extension* of $\mathfrak{M}'$ (or that $\mathfrak{M}'$ is a Γ-*elementary substructure* of $\mathfrak{M}$), and write $\mathfrak{M}' \prec_\Gamma \mathfrak{M}$, if for every $\varphi \in \Gamma$ and for all $m_1, \ldots, m_n \in \mathfrak{M}'$,

$$\mathfrak{M} \models \varphi(m_1, \ldots, m_n) \text{ iff } \mathfrak{M}' \models \varphi(m_1, \ldots, m_n).$$

The following general version of the Arithmetised Completeness Theorem follows from Theorems 1.7 and 2.2 of [McA78].

2.1.13. Theorem. *Let $\mathfrak{M}' \models \mathsf{PA}$. If $\mathfrak{M}' \models \Diamond_m^{\Pi_1}\varphi$, where $m \in \mathfrak{M}'$ is nonstandard, then there is an end-extension $\mathfrak{M}$ of $\mathfrak{M}'$ with $\mathfrak{M}' \prec_{\Pi_1} \mathfrak{M}$, $\mathfrak{M} \models \mathsf{PA}$ (from the external point of view), and $\mathfrak{M} \models \varphi$.* ❑

2.2 Modal logic

The language $\mathcal{L}_\Box$ of propositional modal logic is obtained by adding a unary operator $\Box$ to the language of propositional logic. The symbol $\Diamond$ is used as the dual of $\Box$, i.e. as an abbreviation for $\neg\Box\neg$. We shall omit brackets that are superfluous according to the following reading conventions:

$$\Box, \neg \quad \succ \quad \wedge, \vee \quad \succ \quad \rightarrow, \leftrightarrow,$$

where $\succ$ indicates binding strength. Thus $\Box$ and $\neg$ are the strongest, while $\rightarrow$ and $\leftrightarrow$ are the weakest binding operators.

The system K contains all propositional tautologies in the language $\mathcal{L}_\Box$, together with axiom K: $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$. The inference rules of K are modus ponens and necessitation: if $\mathsf{K} \vdash A$, then $\mathsf{K} \vdash \Box A$.

The modal logic GL, named after Gödel and Löb, is obtained by adding to K the following, known as Löb's axiom:

(L) $\Box(\Box A \to A) \to \Box A$

It is well-known (for a proof, see for example [Boo93, Theorem 1.18]) that the "transitivity axiom" (4) is derivable in GL:

2.2.1. LEMMA. $\mathsf{GL} \vdash \Box A \to \Box\Box A$ ❑

Note that it follows from Lemma 2.2.1 that the theorems of GL include the modal counterparts of the HBL-conditions.

Given a modal system L and a set Γ of formulas in the language of L, we write $\Gamma \vdash_{\mathsf{L}} B$ to mean that B is derivable in L from some elements $A_0, \ldots, A_n$ in Γ *without use of necessitation.* If $\Gamma = \{A_0, \ldots, A_n\}$, we also write $A_0, \ldots, A_n \vdash B$ instead of $\Gamma \vdash_{\mathsf{L}} B$. A set Γ of formulas is said to be L-*consistent* if $\Gamma \nvdash \bot$, and *maximal* L-*consistent* if additionally it contains either A or $\neg A$ for every A in the language of L. Lindenbaum's Lemma tells us that every consistent set can be extended to a maximal consistent one. Throughout this section, we write (*maximal*) *consistent* to mean (maximal) GL-consistent. The following basic facts concerning GL will mostly be used without explicit mention:

2.2.2. LEMMA. *i.* $\mathsf{GL} \vdash \Box(A \wedge B) \leftrightarrow (\Box A \wedge \Box B)$

ii. $\mathsf{GL} \vdash \Box\top$

iii. If $A_0, \ldots, A_n \vdash_{\mathsf{GL}} B$, *then* $\Box A_0, \ldots, \Box A_n \vdash_{\mathsf{GL}} \Box B$

iv. If $\mathsf{GL} \vdash A \to B$, *then* $\mathsf{GL} \vdash \Diamond A \to \Diamond B$ ❑

A relation $\prec$ on a set W is *converse well-founded* if for every $S \subseteq W$ with $S \neq \emptyset$, there is some $a \in W$ such that $a \not\prec b$ for all $b \in S$, in other words if there are no infinite ascending $\prec$-chains. A converse well-founded relation is, in particular, irreflexive. We write $a \preceq b$ if either $a \prec b$ or $a = b$.

2.2.3. DEFINITION. A GL-*frame* $\mathcal{F}$ is a tuple $\langle W, \prec \rangle$, where $\prec$ is a transitive converse well-founded relation on W.

2.2.4. DEFINITION. A GL-*model* is a triple $\langle W, \prec, \Vdash \rangle$, where $\langle W, \prec \rangle$ is a GL-frame, and $\Vdash$ a valuation assigning to every propositional letter a subset of W. $\Vdash$ is extended to all formulas of $\mathcal{L}_\Box$ by requiring that it commutes with propositional connectives, and interpreting $\prec$ as the accessibility relation for $\Box$:

$$\mathcal{M}, a \Vdash \Box A \text{ if for all } b \text{ with } a \prec b,\ \mathcal{M}, b \Vdash A.$$

Given a frame $\mathcal{F}$ or a model $\mathcal{M}$ with domain W, we shall often write $a \in \mathcal{F}$ or $a \in \mathcal{M}$ instead of $a \in W$. If $\mathcal{M}$ is clear from the context, we write $w \Vdash A$ instead of $\mathcal{M}, w \Vdash A$. A formula A is *valid in a model* $\mathcal{M}$, we write $\mathcal{M} \Vdash A$, if $a \Vdash A$ for every $a \in \mathcal{M}$. Similarly, A is *valid in a frame* $\mathcal{F}$, we write $\mathcal{F} \Vdash A$, if $\mathcal{M} \Vdash A$ for any model $\mathcal{M} = \langle W, \prec, \Vdash \rangle$ with $\mathcal{F} = \langle W, \prec \rangle$.

2.2.5. Theorem. $\mathsf{GL} \vdash A$ *iff for every finite* GL*-frame* $\mathcal{F}$, $\mathcal{F} \Vdash A$.

Proof: The proof of soundness, i.e. the direction from left to right, is straightforward. An overview of the proof of modal completeness is given below. ❑

2.2.1 Proof of modal completeness

The right to left direction of Theorem 2.2.5 is proven, as usual, by contraposition. Given a sentence A with $\mathsf{GL} \nvdash A$, we shall find a GL-frame $\mathcal{F}$ with $\mathcal{F} \nVdash A$, or in other words a GL-model $\mathcal{M}$ where $w \Vdash \neg A$ for some $w \in \mathcal{M}$. The domain of $\mathcal{F}$ will consist of maximal consistent sets, and $\Vdash$ is the *canonical valuation*, by which we mean the valuation defined as:

$$x \Vdash p :\Leftrightarrow p \in x. \tag{2.6}$$

The assumption $\mathsf{GL} \nvdash A$ implies that there is a maximal consistent set x_0 with $\neg A \in x_0$. Our goal is to extend the equivalence in (2.6) beyond propositional formulas. In particular, we would like to have $x_0 \nVdash A$.

It is well-known that GL is not compact: there is an infinite consistent set that cannot be satisfied at any point on a GL-frame (see for example [Boo93, p.102]). This means that we cannot hope to extend (2.6) to all $\mathcal{L}_\Box$-formulas. Nevertheless, it is possible to extend (2.6) to a set that is big enough in order to ensure $x_0 \nVdash A$.

There are many ways to find a frame $\mathcal{F}$ with the required properties; see for example [JdJ98, Theorem 40]), [BV02, Exercise 4.8.7], or [Boo93, Chapter 5]. In fact, some of them are easier than the proof presented here, which is based on the construction method used in [GJ08]. The aim of the exposition is to prepare the ground for the more involved modal completeness proof in Section 3.4.

2.2.6. Definition. A set $\mathcal{D}$ of formulas is said to be *adequate* if it is finite and closed under subformulas and single negations.

Given A with $\mathsf{GL} \nvdash A$, let $\mathcal{D}$ be an adequate set containing A, and let $\mathcal{F}_0$ be the GL-frame consisting of a single maximal consistent set x_0 with $\neg A \in x_0$. Our goal is to extend $\mathcal{F}_0$ to a GL-frame $\mathcal{F}$ where, letting $\Vdash$ be the canonical valuation, we have for all $x \in \mathcal{F}$ and $B \in \mathcal{D}$,

$$x \Vdash B :\Leftrightarrow B \in x.$$

We call the above equivalence a *truth lemma* (with respect to $\mathcal{D}$). If x_0 contains formulas of the form $\neg\Box C$, we thus need to add new nodes to the frame.

2.2.7. Definition. Let $\mathcal{F}$ be a frame whose domain consists of maximal consistent sets, $x \in \mathcal{F}$, and $\mathcal{D}$ an adequate set. A *$\mathcal{D}$-problem in x* is a formula $\neg\Box B$ where $\neg\Box B \in x \cap \mathcal{D}$, but there is no y with $x \prec y$ and $\neg B \in y$.

If $\mathcal{D}$ is clear from the context, we shall refer to $\mathcal{D}$-problems simply as *problems*. An element of $\mathcal{D}$ is a *problem-formula* if it is has the form $\neg\Box B$. We assume as given some ordering of problem-formulas.

2.2.8. Definition. For maximal consistent sets x and y, let $x \prec_\Box y$ if for every $\mathcal{L}_\Box$-formula B, we have that $\Box B \in x$ implies $B \in y$.

2.2.9. Lemma. *If $x \prec_\Box y$, then any problem-formula in y is contained in x.*

Proof: Suppose that $x \prec_\Box y$ and $\neg\Box B \in y$. Assuming $\Box B \in x$, we would have $\Box\Box B \in x$ by Lemma 2.2.1, and so $\Box B \in y$, a contradiction. Since x is maximal consistent, it must be that $\neg\Box B \in x$. ❑

The proof of the following lemma is similarly straightforward:

2.2.10. Lemma. *If $x \prec_\Box y$ and $y \prec_\Box z$, then $x \prec_\Box z$.* ❑

2.2.11. Definition. A frame $\mathcal{F} = \langle W, \prec\rangle$, where W consists of maximal consistent sets, is *adequate* if for all $x, y \in \mathcal{F}$, we have that $x \prec y$ implies $x \prec_\Box y$.

The proof of the following lemma is completely straightforward.

2.2.12. Lemma. *Let $\mathcal{F}$ be an adequate frame containing no $\mathcal{D}$-problems, and let $\Vdash$ be the canonical valuation. Then $x \Vdash B \Leftrightarrow B \in x$ for all $B \in \mathcal{D}$.* ❑

2.2.13. Lemma. *Let $\mathcal{D}$ be adequate, and x maximal consistent with $\neg\Box B \in x \cap \mathcal{D}$. There is some y with $x \prec_\Box y$ and $\neg B, \Box B \in y$.*

Proof: Using Lindenbaum's Lemma, it suffices to show consistency of the set $\{A \mid \Box A \in x\} \cup \{\Box B, \neg B\}$. Assuming the contrary, there would be $\Box A_0 \ldots, \Box A_n$ in x with $A_0, \ldots, A_n \vdash_{\mathsf{GL}} \Box B \to B$. Using necessitation and Löb's axiom:

$$\Box A_0, \ldots, \Box A_n \vdash_{\mathsf{GL}} \Box(\Box B \to B)$$
$$\Box A_0, \ldots, \Box A_n \vdash_{\mathsf{GL}} \Box B$$

Thus also $\Box B$ should be in x, a contradiction. ❑

We describe an algorithm for eliminating problems in adequate GL-frames. The function f is used to keep track of the order in which problems are eliminated.

The problem elimination algorithm: Let x_0 be maximal consistent and $\mathcal{D}$ adequate. Define $\mathcal{F}_0 = \langle \{x_0, \emptyset\} \rangle$ and $f(0) = x_0$. While some world in the range of f contains a problem in $\mathcal{F}_n = \langle W_n, \prec_n \rangle$, do:

1. Let i be the least such that $f(i)$ is defined and contains a problem, let x be such that $f(i) = x$, and let $\neg\Box B$ be the least problem in x.

2. By Lemma 2.2.13, let y be maximal consistent with $x \prec_{\Box} y$ and $\neg B, \Box B \in y$.

3. Let $f(j) := y$, where j is the least such that $f(j)$ is undefined.

4. Let $W_{n+1} := W_n \cup \{y\}$, and define $\prec_{n+1}$ to be the transitive closure of $\prec_n \cup \{x \prec y\}$.

5. Let $\mathcal{F}_{n+1} := \langle W_{n+1}, \prec_{n+1} \rangle$

Clearly, $\mathcal{F}_0$ is an adequate GL-frame. Using Lemma 2.2.10, it is easy to check that if $\mathcal{F}_n$ is an adequate GL-frame, then so is $\mathcal{F}_{n+1}$.

2.2.14. Lemma. *The problem elimination algorithm terminates.*

Proof: We argue by induction on the number of problem-formulas in x_0. When starting the algorithm, we have $f(0) = x_0$. If x_0 contains no problem-formulas, then it contains no problems, and so the while-loop will never be entered.

So suppose that x_0 contains $n + 1$ problem-formulas. This means that all problems $\neg\Box B$ in x are eliminated during the (at most) first $n + 1$ steps of the algorithm by adding some y with $x \prec y$ and $\neg B, \Box B \in y$. After these steps, we thus have $i > 0$ whenever the while-loop is entered in order to eliminate some problem in $f(i)$. Since $x \prec_{\Box} y$ by construction and $\Box B \in y$ while $\neg\Box B \in x$, it follows by using Lemma 2.2.9 that each y contains at most n problem-formulas. Thus the algorithm, when run on each y, terminates by assumption. ❑

We prove the remaining direction of Theorem 2.2.5.

Proof: Suppose GL $\nvdash A$, let x_0 be maximal consistent with $\neg A \in x_0$, and let $\mathcal{D}$ be an adequate set containing A. Run the problem elimination algorithm on x_0 and $\mathcal{D}$. This yields a finite adequate GL-frame $\mathcal{F}$ free of problems. Letting $\Vdash$ be the canoncial valuation, we thus have $x_0 \nVdash A$ by Lemma 2.2.12. ❑

2.3 Provability logic

We show that GL is the provability logic of any reasonable theory. By a *reasonable theory* we shall, throughout this thesis, mean a Σ_1-sound smooth recursively axiomatised theory $T = (\mathrm{Ax}_T, \tau)$ extending $\mathrm{I}\Delta_0+\mathsf{exp}$, verifiably in $\mathrm{I}\Delta_0+\mathsf{exp}$. We write $\Box$ for the provability predicate of T defined as in Section 2.1.2. As explained there, $\Box$ is Σ_1 in $\mathrm{I}\Delta_0+\mathsf{exp}$, and satisfies the following conditions:

1. $T \vdash \varphi \;\Leftrightarrow\; \mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box\varphi$
2. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
3. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_0\varphi \to \Box\varphi$,

where the right to left direction of (1) is by Σ_1-soundness of $\mathrm{I}\Delta_0+\mathsf{exp}$. We show that if T and $\Box$ are as above, then the propositional schemata involving $\Box$ that are provable in T are exactly the theorems of GL. A precise statement of this result makes use of the following definition:

2.3.1. Definition. Let ϑ be an $\mathcal{L}$-formula with one free variable. A *ϑ-realisation* is a function * from the propositional letters of $\mathcal{L}_\Box$ to $\mathcal{L}$-sentences. The domain of * is extended to all $\mathcal{L}_\Box$-formulas by requiring that it commutes with propositional connectives, and furthermore $(\Box A)^* := \vartheta(\ulcorner A^* \urcorner)$.

A ϑ-realisation is thus a translation from the modal language $\mathcal{L}_\Box$ to the language $\mathcal{L}$ of arithmetic, where the modality $\Box$ is translated by means of the formula ϑ. We note that the values of a ϑ-realisation are determined by its values at the propositional letters of $\mathcal{L}_\Box$. Instead of ϑ-realisations, we shall mostly speak of *arithmetical realisations mapping $\Box$ to ϑ*. It is clear how this notion can be generalised to bimodal languages.

2.3.2. Theorem. *Let $\Box$ be the provability predicate of a reasonable theory T. Then for all $A \in \mathcal{L}_\Box$, $\mathsf{GL} \vdash A$ iff $T \vdash A^*$ for all $\Box$-realisations *.*

The left to right direction of Theorem 2.3.2 is referred to as *arithmetical soundness.* It is an immediate consequence of conditions (1)-(3) above, together with the observation that — in the presence of the Fixed Point Lemma — the latter imply Löb's Theorem for $\Box$ (Theorem 2.1.9). In fact, it is clear that arithmetical soundness of GL with respect to T is already verifiable in $\mathrm{I}\Delta_0+\mathsf{exp}$. The proof of the other direction, i.e. arithmetical completeness, is due to Solovay ([Sol76]). An overview of the proof is given below.

2.3.1 Proof of arithmetical completeness

Given an $\mathcal{L}_\Box$-formula A with $\mathsf{GL} \nvdash A$, we would like to find an arithmetical realisation * mapping the modality $\Box$ to the provability predicate $\Box$, and for which it holds that $T \nvdash A^*$. The proof proceeds by showing that any finite GL-frame can be embedded into $\mathrm{I}\Delta_0+\mathsf{exp}$ in a suitable way.

We write $x : \Box\varphi$ to mean that x is a witness of the Σ_1-sentence $\Box\varphi$, i.e. that x is the code of a T-proof of φ. We assume that every number witnesses the proof of a unique sentence — if any — , noting that this requirement can be satisfied for any reasonable arithmetisation of syntax in $\mathrm{I}\Delta_0+\mathsf{exp}$. For the rest of this section, let us fix a GL-frame $\mathcal{F} = \langle W, \prec\rangle$ with root 0.

2.3.3. DEFINITION. ($\mathrm{I}\Delta_0+\mathsf{exp}$) The function $h : \omega \to W$ is defined by:

$$h(0) = 0$$
$$h(x+1) = \begin{cases} b & \text{if } h(x) \prec b \text{ and } x : \Box L \neq b \\ h(x) & \text{otherwise} \end{cases}$$

The formula $L \neq b$ (see (2.7) below) depends on the formula χ representing h. The self-reference in the definition of h is handled by the Fixed Point Lemma. We note that the definition of h only relies on the gödelnumber of $L \neq b$, and the latter can be obtained from b and $\ulcorner\chi\urcorner$ by a function that is provably total in $\mathrm{I}\Delta_0+\mathsf{exp}$.

It follows from Theorem 2.1.4 — for example, by using that W is finite — that h is elementary and provably total in $\mathrm{I}\Delta_0+\mathsf{exp}$, with its defining equations also provable in $\mathrm{I}\Delta_0+\mathsf{exp}$. We write $L = a$ for the formula

$$\exists x\, h(x) = a \wedge \forall x\, h(x) \preceq a. \tag{2.7}$$

The formula $L = a$ states that a is a $\preceq$-maximal element in the range of h. Given the following lemma, we can think of $L = a$ as saying that a is the limit of h.

2.3.4. LEMMA. *i.* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash x' \leqslant x \to h(x') \preceq h(x)$

ii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \exists! w\, L = w$

Proof: (i) is proven by internal induction on x, using that h is defined by an Δ_0^{exp}-formula. The inductive step follows from the transitivity of $\preceq$, together with the fact that $h(x) \preceq h(x+1)$ by definition.

(ii) Since the relation $\preceq$ is antisymmetric, uniqueness is immediate from the definition of $L = a$. For existence, we show by external induction on the converse of $\prec$ that for all $a \in W$,

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash h(x) = a \to \exists w\, L = w.$$

This is sufficient, since $h(0) = 0$ holds in $\mathrm{I}\Delta_0+\mathsf{exp}$. From (i) we have that

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash h(x) = a \to (\forall x' \geqslant x\, h(x') = a \vee \exists x' \geqslant x\, a \prec h(x'))\,. \tag{2.8}$$

Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, assuming $h(x) = a$. If the first disjunct in (2.8) holds, we have, by using clause (i), $L = a$, while if the second disjunct holds, then $\exists w\, L = w$ by the induction assumption. Thus in either case $\exists w\, L = w$ as required. ❑

2.3.5. LEMMA. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash L = a \wedge a \prec b \to \Diamond L = b$

Proof: Argue in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, assuming $a \prec b$ and $L = a$, noting that the latter implies

$$\forall x\, h(x) \preceq a. \tag{2.9}$$

Suppose that there is some x with $x : \Box L \neq b$. By (2.9) we have $h(x) \preceq a$, whence it follows from the definition of h that $h(x+1) = b$. From (2.9) we get $b \preceq a$, whence $b \prec b$ by transitivity of $\prec$, a contradiction. ❑

2.3.6. LEMMA. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash L = a \neq 0 \to \Box a \prec L$

Proof: Argue in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, assuming $L = a$. Let x be such that $h(x) = a$. Since the latter is a Δ_0^{exp}-formula, we have $\Box_0\, h(\dot{x}) = a$ by Theorem 2.1.1, together with provable Σ_1-completeness of $\mathrm{I}\Delta_0{+}\mathsf{exp}$. Using the definition of L, this implies $\Box_0\, a \preceq L$. By assumption, we have $\Box_0\varphi \to \Box\varphi$ for all φ, whence also

$$\Box a \preceq L.$$

Since $a \neq 0$, it follows from the definition of h that

$$\Box L \neq a.$$

Combining the above yields $\Box a \prec L$ as required. ❑

2.3.7. DEFINITION. Let $\mathcal{M} = \langle \mathcal{F}, \Vdash \rangle$ be a finite GL-model with root w. The model $\mathcal{M}_0$ is obtained by appending a new root 0 to $\mathcal{M}$; the truth values of propositional formulas at 0 are defined to be the same as at w. Apply Definition 2.3.3 to $\mathcal{M}_0$, and define the arithmetical realisation * by letting

$$p^* := \bigvee_{\mathcal{M}_0, a \Vdash p} L = a.$$

2.3.8. LEMMA. *Let $\mathcal{M}$ and * be as in Definition 2.3.7. For all $B \in \mathcal{L}_\Box$, $a \neq 0$,*

$$\text{if } \mathcal{M}, a \Vdash B, \text{ then } \mathrm{I}\Delta_0{+}\mathsf{exp} \vdash L = a \to B^*.$$

Proof: We prove the claim by induction on the complexity of B, simultaneously with

$$\text{if } \mathcal{M}, a \Vdash \neg B, \text{ then } \mathrm{I}\Delta_0{+}\mathsf{exp} \vdash L = a \to \neg B^*.$$

The case of propositional letters is immediate from the the definition of *, together with Lemma 2.3.4(ii). The boolean cases are straightforward.

Suppose now that $\mathcal{M}, a \Vdash \Box C$, i.e. that $\mathcal{M}, b \Vdash C$ for all b with $a \prec b$. By the induction assumption we have $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = b \to C^*$ for all such b, whence $\mathrm{I}\Delta_0+\mathsf{exp} \vdash a \prec L \to C^*$. Note that $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \varphi \Rightarrow T \vdash \varphi \Rightarrow \mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box\varphi$ for all φ. Reasoning as in GL, we thus obtain $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box a \prec L \to \Box C^*$. Using Lemma 2.3.6, the latter implies $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \to \Box C^*$ as required.

Finally, suppose $\mathcal{M}, a \vDash \Diamond C$, and let b be such that $a \prec b$ and $\mathcal{M}, b \vDash C$. By the induction assumption, $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = b \to C^*$. Reasoning as above, we obtain $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Diamond L = b \to \Diamond C^*$, whence $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \to \Diamond C^*$ by Lemma 2.3.5. ❑

2.3.9. Lemma. *i.* $\mathfrak{N} \vDash L = 0$.

ii. For all $a \neq 0$, $T \nvdash L \neq a$, *i.e.* $L = a$ *is consistent with* T

Proof: (i) Let $a \neq 0$. We show $\mathfrak{N} \nvDash \exists x\, h(x) = a$ by induction on the converse of $\prec$. Assume that the claim holds for all $\prec$-successors of a, and suppose $\mathfrak{N} \vDash \exists x\, h(x) = a$. By definition of h, this implies $\mathfrak{N} \vDash \Box L \neq a$, whence $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box L \neq a$ by Σ_1-completeness. We have $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box\varphi \Rightarrow T \vdash \varphi$ for all φ, and thus also

$$T \vdash L \neq a. \tag{2.10}$$

Since $\mathfrak{N} \vDash \exists x\, h(x) = a$, by Σ_1-completeness of T we also have $T \vdash \exists x\, h(x) = a$. This, together with (2.10), implies $T \vdash \exists x\, a \prec h(x)$, whence $\mathfrak{N} \vDash \exists x\, a \prec h(x)$ by Σ_1-soundness of T. The latter, however, contradicts our induction assumption. We conclude that for $a \neq 0$, $\mathfrak{N} \vDash \forall x\, h(x) \neq a$, whence $\mathfrak{N} \vDash \forall x\, h(x) = 0$, and thus $\mathfrak{N} \vDash L = 0$ as required. Closer inspection of the argument reveals that $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = 0 \leftrightarrow \Diamond^{n+1}\top$, where n is the depth of $\mathcal{M}$, see [Vis84].

(ii) By (i) and Lemma 2.3.5, we have $\mathfrak{N} \vDash \Diamond L = a$ for all $a \neq 0$. Now, $T \vdash L \neq a$ would imply $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box L \neq a$, and the latter in turn would imply $\mathfrak{N} \vDash \Box L \neq a$, a contradiction. ❑

We prove the right to left direction of Theorem 2.3.2.

Proof: If $\mathsf{GL} \nvdash A$, then by Theorem 2.2.5 there is a finite rooted GL-model $\mathcal{M}$ with $a \nVdash A$ for some $a \in \mathcal{M}$. Let * be the realisation as in Definition 2.3.7. By Lemma 2.3.8, $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \to \neg A^*$, i.e. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash A^* \to L \neq a$. Since T does not prove $L \neq a$ by Lemma 2.3.9(ii), it therefore cannot prove A^* either. ❑

2.4 Ordinals

In Section 2.4.2, we define transfinite iterations of a given provability predicate. In order to do that, ordinals need to be represented as natural numbers, by means of some *ordinal notation system*. Basic facts concerning the latter are reviewed in Section 2.4.1 below. Transfinite iterations of provability predicates are closely related to Turing-Feferman progressions ([Tur39], [Fef62]). The presentation here is based on [Bek03], which in turn is based on [Sch79].

2.4.1 Ordinal notation systems

Ordinals up to ε_0 can be represented in $I\Delta_0+\mathsf{exp}$ in a natural way by making use of the Cantor Normal Form Theorem:

2.4.1. THEOREM. *For every ordinal $\alpha > 0$, there are unique $\alpha_0 \geqslant \ldots \geqslant \alpha_k$ with*

$$\alpha = \omega^{\alpha_0} + \omega^{\alpha_1} + \ldots + \omega^{\alpha_k}.$$

The above representation of α is called its *Cantor normal form*. For an ordinal α and $n < \omega$, we define ω_n^α inductively by letting

$$\omega_0^\alpha = \alpha,$$
$$\omega_{n+1}^\alpha = \omega^{\omega_n^\alpha}.$$

We write ω_n for ω_n^1. Thus we have that $\omega_0 = 1$, $\omega_1 = \omega$, $\omega_2 = \omega^\omega$, etc. Let $\varepsilon_0 := \sup\{\omega_n \mid n \in \omega\}$. Using that $x \to \omega^x$ is strictly increasing and continuous, it is not difficult to show that $\varepsilon_0 = \omega^{\varepsilon_0}$, and that in fact ε_0 is the smallest fixed point of the function $x \to \omega^x$. We note that ε_0 is countable.

It follows from the above that if $\alpha < \varepsilon_0$, then α has a Cantor normal form with exponents $\alpha_i < \alpha$, and these exponents in turn have Cantor normal forms with yet smaller exponents. Assuming the ordinal 0 to be represented by some fixed term, say 0, an ordinal $0 < \alpha < \varepsilon_0$ can thus be represented by its Cantor normal form

$$\omega^{\alpha_0} + \omega^{\alpha_1} + \ldots + \omega^{\alpha_k},$$

where $\alpha_0 \geqslant \ldots \geqslant \alpha_k$, and each α_i is represented in the same way. More formally, we fix for any ordinal $\alpha < \varepsilon_0$ an $\mathcal{L}$-term built of ω^x, $x + y$, and 0. The relation $<$ and the standard functions of ordinal arithmetic ($x + y$, $x \cdot y$ and ω^x) on Cantor ordinal notations are elementary. It follows from this that basic facts about ordinal arithmetic are provable in $I\Delta_0+\mathsf{exp}$ (in fact, already in $I\Delta_0$ — see [Som95, Section 3.5]).

The Cantor ordinal notation system is the most common way of representing ordinals below ε_0. In order to reason in $I\Delta_0+\mathsf{exp}$ about ordinals beyond ε_0, we introduce the notion of an *elementary linear ordering*.

2.4.2. DEFINITION. ([Bek03]) An *elementary linear ordering* is a pair of elementary formulas $x \in D$ and $x \prec y$, such that $I\Delta_0+\mathsf{exp}$ proves that the relation $\prec$ linearly orders the domain D. An *elementary well-ordering* is an elementary linear ordering which is well-founded in the standard model $\mathfrak{N}$.

Given an elementary linear ordering $(D, \prec)$, we use Greek variables α, β, γ etc. to denote the elements of D. Since D is elementary, these variables can also be used within $\mathrm{I}\Delta_0+\mathsf{exp}$.

Reflexive induction, first used in [Sch79], is very useful for reasoning about elementary linear orderings in theories containing $\mathrm{I}\Delta_0+\mathsf{exp}$.

2.4.3. Lemma ([Bek03, Lemma 2.4]). *Let $(D, \prec)$ be an elementary linear ordering, T a theory containing $\mathrm{I}\Delta_0+\mathsf{exp}$, and $\Box$ a provability predicate for T for which Löb's Theorem holds. Then T is closed under the following* reflexive induction *rule:*

$$\frac{\forall \alpha\ (\Box \forall \beta \prec \dot{\alpha}\, \Phi(\beta) \to \Phi(\alpha))}{\forall \alpha\, \Phi(\alpha)}.$$

Proof: Assuming that $T \vdash \forall \alpha\ (\Box \forall \beta \prec \dot{\alpha}\, \Phi(\beta) \to \Phi(\alpha))$, we find:

$$\begin{aligned} T \vdash \Box \forall \alpha\, \Phi(\alpha) &\to \forall \alpha\, \Box \forall \beta \prec \dot{\alpha}\, \Phi(\beta) \\ &\to \forall \alpha\, \Phi(\alpha) \end{aligned}$$

Thus $T \vdash \forall \alpha\, \Phi(\alpha)$ follows by Löb's Theorem for T. ❑

An ordinal α is *recursive* if there is a recursive well-ordering of a subset of the natural numbers with order type α. Recursive ordinals are exactly the ordinals that have a notation in Kleene's $\mathcal{O}$.

The following result follows from [Som93, Theorem 3], where it is shown that recursive ordinals can even be Δ_0-represented in $\mathrm{I}\Delta_0$:

2.4.4. Theorem. *For any recursive ordinal α, there is an elementary linear ordering with order type α; furthermore the basic functions and relations (for example $\leqslant$, $+$, $\cdot$) on ordinals $\leqslant \alpha$ can also be taken to be elementary.* ❑

2.4.2 Transfinite iterations of provability predicates

We wish to talk about iterations of a given provability predicate $\Box$. Iterations corresponding to natural numbers can be defined in a natural way by letting $\Box^0\varphi := \varphi$, and $\Box^{n+1}\varphi := \Box\Box^n\varphi$. Intuitively, $\Box^n\varphi$ thus means: $\Box$ applied to φ, n-times. We carry out the above definition within $\mathrm{I}\Delta_0+\mathsf{exp}$, and extend it to ordinals beyond ω.

Throughout this section, the symbol $\Box$ is used to denote an arithmetical formula for which we have:

1. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall \varphi\, (\Box_0\varphi \to \Box\varphi)$
2. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall \varphi\, \forall \psi\, (\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi))$
3. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall \varphi\, (\Box\varphi \to \Box\Box\varphi)$

Note that we do *not* assume $\Box$ to be Σ_1, hence condition (3) does not follow from (1) as in Section 2.1.2.

For the rest of this section, we fix an elementary linear ordering $(D, \prec)$. We assume that, verifiably in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, $(D, \prec)$ has a least element, denoted by $\mathsf{0}$. Intuitively speaking, a transfinite iteration of $\Box$ in $\mathrm{I}\Delta_0{+}\mathsf{exp}$ should be a formula $\Box^\alpha$, where $\alpha > \mathsf{0}$ is an ordinal (represented within $\mathrm{I}\Delta_0{+}\mathsf{exp}$ by means of some ordinal notation system), and $\mathrm{I}\Delta_0{+}\mathsf{exp}$ verifies that for all φ,

i. $\Box^{\beta+1}\varphi \leftrightarrow \Box\Box^\beta\varphi$ for $0 < \beta < \alpha$

ii. $\Box^\lambda\varphi \leftrightarrow \exists\beta < \lambda\, \Box^\beta\varphi$ for a limit ordinal $0 < \lambda \leqslant \alpha$

2.4.5. Definition. By the Fixed Point Lemma, let $\Box^\alpha\varphi$ be a formula such that

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box^\alpha\varphi \leftrightarrow \exists\beta \prec \alpha\ \big((\beta = \mathsf{0} \wedge \Box\varphi) \vee (\beta \neq 0 \wedge \Box\Box^\beta\varphi)\big)\,,$$

where α and φ are internal variables ranging over $\{\delta \in D \mid \delta \neq \mathsf{0}\}$ and $\mathcal{L}$-sentences, respectively. Similarly, $\Diamond^\alpha\varphi$ is defined to be a formula such that

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Diamond^\alpha\varphi \leftrightarrow \forall\beta \prec \alpha\ \big((\beta = \mathsf{0} \wedge \Diamond\varphi) \vee (\beta \neq \mathsf{0} \wedge \Diamond\Diamond^\beta\varphi)\big)\,.$$

When reasoning in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, we shall, less formally, assume $\Box^\alpha\varphi$ to satisfy

$$\Box^\alpha\varphi \leftrightarrow \exists\beta \prec \alpha\, \Box\Box^\beta\varphi$$

and similarly for $\Diamond^\alpha\varphi$, having in mind that $\Box^{\mathsf{0}}\varphi$ and $\Diamond^{\mathsf{0}}\varphi$ are defined to be φ.

It is clear from Definition 2.4.5 that if $\Box$ is Σ_{n+1}, then so is $\Box^\alpha\varphi$, and that if $\Diamond$ is Π_{n+1}, then so is $\Diamond^\alpha\varphi$. In [HP16, Lemma 8], it is shown that, verifiably in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, iterations are unique. Thus we can use $\Box^\alpha\varphi$ freely, without specifying the formulas involved.

We define the formulas $\mathsf{s}(\beta, \alpha)$ and $\mathsf{lim}(\alpha)$ stating, respectively, that α is an immediate successor of β, and that α is a limit:

$$\begin{aligned}\mathsf{s}(\beta,\alpha) &:= \beta \prec \alpha \wedge \neg\exists\gamma\ (\beta \prec \gamma \prec \alpha)\\ \mathsf{lim}(\alpha) &:= \alpha \neq \mathsf{0} \wedge \neg\exists\beta\, \mathsf{s}(\beta,\alpha)\end{aligned}$$

The facts given in the following lemma will often be used without explicit mention.

2.4.6. Lemma ($\mathrm{I}\Delta_0{+}\mathsf{exp}$). *For any $\mathcal{L}$-sentences φ, ψ, and $\alpha, \beta \neq \mathsf{0}$:*

i. $\Box^\alpha\varphi \leftrightarrow \neg\Diamond^\alpha\neg\varphi$

ii. $\Box\varphi \to \Box^\alpha\varphi$

iii. $\beta \prec \alpha \to (\Box^\beta\varphi \to \Box^\alpha\varphi)$

iv. $\Box^\alpha(\varphi \to \psi) \to (\Box^\alpha\varphi \to \Box^\alpha\psi)$

v. $\Box^\alpha\varphi \to \Box^\alpha\Box^\alpha\varphi$

vi. $\mathsf{s}(\beta,\alpha) \to \big(\Box^\alpha\varphi \leftrightarrow \Box\Box^\beta\varphi\big)$

vii. $\mathsf{lim}(\lambda) \to \big(\Box^\lambda\varphi \leftrightarrow \exists\alpha \prec \lambda\ (\alpha \neq \mathsf{0} \wedge \Box^\alpha\varphi)\big)$

Proof: By reflexive induction, using the assumptions on $\Box$. In (iii), we take $\Psi(\alpha) := \forall\beta \prec \alpha\,\big(\Box^\beta\varphi \to \Box^\alpha\varphi\big)$. The proofs of (iv) and (v) use (iii); the proof of (vii) uses (vi) and (iii). ❑

Clauses (ii), (iv), and (v) of Lemma 2.4.6 are the analogues of the HBL-conditions for $\Box^\alpha$. Using the Fixed Point Lemma, it can be shown that also

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\Box^\alpha(\Box^\alpha\varphi \to \varphi) \to \Box^\alpha\varphi).$$

2.4.7. Lemma. *Let* $\Box$ *and* $\triangle$ *be provability predicates for which it holds that* $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\Box\varphi \to \triangle\,\varphi)$. *Then* $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\alpha\,\forall\varphi\,(\Box^\alpha\varphi \to \Box^\alpha\varphi)$.

Proof: By reflexive induction. ❑

The results of this section, together with Theorem 2.4.4, imply that we can reason within $\mathrm{I}\Delta_0{+}\mathsf{exp}$ about transfinite iterations corresponding to any recursive ordinal α. It is important to keep in mind, however, that the natural properties of an ordinal α might not be verifiable in $\mathrm{I}\Delta_0{+}\mathsf{exp}$ for an arbitrary representation of α. When using the Cantor ordinal notation system, on the other hand, ordinals below ε_0 can be assumed to have more of the expected properties. The following lemma gives one example.

2.4.8. Lemma. *Suppose that* $(D,\prec)$ *is the Cantor ordinal notation system for ordinals* $\leqslant \varepsilon_0$. *Then* $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,\forall\beta\,\forall\alpha \neq \mathsf{0}\,(\Box^\alpha\Box^\beta\varphi \leftrightarrow \Box^{\beta+\alpha}\varphi)$.

Proof: By reflexive induction. ❑

Chapter 3

A bimodal provability logic

The system GLT of bimodal logic was introduced by Lindström ([Lin06]) in connection with Parikh provability. The study of supremum adapters lead us to the same system independently. This chapter deals with the modal aspects of GLT. We prove completeness with respect to several classes of Kripke frames, one of which yields decidability. We also provide a characterisation of the closed fragment of GLT, and show that the latter is arithmetically complete with respect to a wide class of provability predicates.

3.1 The system GLT

Let $\mathcal{L}_{\Box\vartriangle}$ be the language of propositional modal logic with two modal operators $\Box$ and $\vartriangle$, their duals denoted by $\Diamond$ and $\triangledown$ respectively.

3.1.1. Definition. The axiom schemata of GLT include all propositional tautologies in the language $\mathcal{L}_{\Box\vartriangle}$, the rules and axiom schemata of GL for both $\vartriangle$ and $\Box$, and furthermore:

(T1) $\vartriangle A \to \Box A$

(T2) $\Box A \to \vartriangle\Box A$

(T3) $\Box A \to \Box\vartriangle A$

(T4) $\Box\vartriangle A \to \Box A$

Throughout this chapter, we write $\vdash A$ if $\mathsf{GLT} \vdash A$, and $C_0, \ldots, C_n \vdash A$ if A is derivable in GLT from $C_0, \ldots, C_n$ without use of necessitation. The notion of a maximal GLT-consistent set is defined as in Section 2.2.1. Throughout this section, we refer to (maximal) GLT-consistent sets as simply (maximal) consistent.

3.1.2. REMARK. Necessitation and Löb's axiom for $\Box$ are redundant in the above axiomatisation of GLT. Necessitation for $\Box$ follows from necessitation for $\triangle$ together with (T1). As for Löb's axiom, we note that by (T3),

$$\vdash \Box(\triangle A \to A) \to \Box\triangle(\triangle A \to A). \tag{3.1}$$

We also have $\vdash \Box(\triangle(\triangle A \to A) \to \triangle A)$ by Löb's axiom for $\triangle$ and necessitation for $\Box$, whence $\vdash \Box\triangle(\triangle A \to A) \to \Box\triangle A$ by axiom (K) for $\Box$. Combining this with (3.1) we obtain $\vdash \Box(\triangle A \to A) \to \Box\triangle A$, and so by (T4),

$$\vdash \Box(\triangle A \to A) \to \Box A. \tag{3.2}$$

The desired result follows since $\vdash \Box(\Box A \to A) \to \Box(\triangle A \to A)$ by (T1).

3.1.3. LEMMA. *i.* $\vdash \triangle A \to \Box\triangle A$

ii. $\vdash \triangle A \to \triangle\Box A$

iii. $\vdash \triangle(\Box A \to A) \to \triangle A$

iv. $\vdash \Box(\triangle A \to A) \to \Box A$

Proof: (i) follows from (T1), since $\vdash \triangle A \to \triangle\triangle A$. (ii) is a consequence of (T1) and (T2). (iii) follows by Löb's axiom for $\triangle$, since $\vdash \triangle(\Box A \to A) \to \triangle(\triangle A \to A)$ by (T1). (iv) is (3.2). ❑

3.2 Bimodal provability logics: an overview

Before introducing the semantics for GLT, let us briefly review some related systems. For more information on bimodal provability logics, see for example [JdJ98] or [AB04].

Considering several provability predicates simultaneously is a natural approach for generalising Solovay's Theorem. We shall first consider what can be said about the provability predicates of reasonable theories (as defined in Section 2.3).

The system CS Let $\triangle$ and $\Box$ be the provability predicates of some reasonable theories T_1 and T_2, respectively. The *joint provability logic* of $\triangle$ and $\Box$ is the collection of all propositional schemata concerning $\triangle$ and $\Box$ that are provable in $T_1 \cap T_2$.

It follows from Solovay's Theorem (Theorem 2.3.2) that the joint provability logic of $\triangle$ and $\Box$ contains the principles of GL for both $\triangle$ and $\Box$. Since $\triangle$ and

$\Box$ are Σ_1, it follows from Σ_1-completeness of $I\Delta_0$+exp that their joint provability logic also contains the following:

$$\Box A \to \triangle\Box A$$
$$\triangle A \to \Box\triangle A$$

The system CS contains the rules and axiom schemata of GL for both $\triangle$ and $\Box$, together with the above axioms. CS is modally complete with respect to a Kripke-style semantics, and enjoys the finite model property. Note that $\Box A \to \triangle\Box A$ is axiom (T2). Given Lemma 3.1.3(i), it is thus clear that CS $\subseteq$ GLT.

It was shown by Smoryński ([Smo85, Theorem 4.3.15]) that CS is the provability logic of a certain pair of r.e. extensions of PA. Beklemishev ([Bek92]) showed that there is a pair $(\triangle, \Box)$ of provability predicates for PA itself, such that the joint provability logic of $\triangle$ and $\Box$ is CS. Thus CS can be seen as the *minimal bimodal provability logic.*

The theories with respect to which CS has shown to be arithmetically complete are obtained by using the Fixed Point Lemma. The existence of *natural* theories whose joint provability logic coincides with CS is an open question. Note that if CS is the joint provability logic of T_1 and T_2 as above, then T_1 and T_2 know as little about each other as possible. Naturally occurring theories thus usually turn out to have some non-trivial knowledge about each other.

Bimodal logics for natural pairs of theories The simplest principle of interaction between two theories must be that of *inclusion*, expressed by the following, already familiar to us as axiom (T1) of GLT:

$$\triangle A \to \Box A$$

The system $\mathsf{CSM_0}$ is CS together with the above principle. Clearly, $\mathsf{CSM_0} \subseteq$ GLT. $\mathsf{CSM_0}$, like CS, is complete with respect to a Kripke-style semantics, and has the finite model property. It has been shown by Wolter [Wol98] that all finitely axiomatisable subframe logics containing $\mathsf{CSM_0}$ are decidable. Given that many systems in this class do not have the finite model property, and are not even complete with respect to any class of Kripke frames, this is a rather strong result.

The following principle expresses that not only is T_2 stronger than T_1, it also proves local reflection for T_1:

$$\text{(ER)} \qquad \Box(\triangle A \to A)$$

It is not difficult to see that (T1) follows from (ER) over CS. The system $\mathsf{CSM_1}$ is CS together with (ER). Unlike $\mathsf{CSM_0}$, the system $\mathsf{CSM_1}$ is not complete with respect to any class of standard Kripke frames. Visser ([Vis93]), using topological methods, devised a generalised Kripke semantics for $\mathsf{CSM_1}$ that is well-behaved, even though the models are infinite.

Recall axiom (T4) of GLT: $\Box\triangle A \to \Box A$, and note that it follows from (ER). Thus we have: $\mathsf{CSM}_0 \subseteq \mathsf{GLT} \subseteq \mathsf{CSM}_1$. As we will see, the system GLT is intermediate between CSM_0 and CSM_1 also with respect to its model-theoretic properties: like CSM_0, it has a nice Kripe semantics, however as in the case of CSM_1, the respective frames are infinite. The closed fragment of CSM_1 has been worked out by Visser in [Vis93]. In Section 3.7, we show that the closed fragment of GLT coincides with that of CSM_1.

It was shown by Carlson ([Car86]) that CSM_1 is the joint provability logic of any pair of reasonable theories T_1 and T_2 whenever T_1 and T_2 are sound, and the schema $\triangle\varphi \to \varphi$ is provable in T_2. Examples of such pairs of theories are $(\mathrm{I}\Delta_0+\mathsf{exp}, \mathsf{PA})$, $(\mathrm{I}\Sigma_n, \mathrm{I}\Sigma_{n+1})$, and $(\mathsf{PA}, \mathsf{ZF})$.

The system GLP Recall from Section 2.1.2 that given a theory T, *n-provability* refers to provability in T together with all true Π_n-sentences. We write $\Box_T^{\Pi_n}$ for the provability predicate of this theory. The theory obtained by adding to T all true Π_n-sentences is — in case it is consistent — not recursively enumerable. Accordingly, $\Box_T^{\Pi_n}$ is not Σ_1.

Smoryński ([Smo85, 3.3.9]) showed that the provability logic of each $\Box_T^{\Pi_n}$ is GL. The polymodal logic GLP, introduced by Japaridze ([DJ88]), contains a modality $[n]$ for each n. Its axioms include the rules and axiom schemata of GL for each $[n]$, and the following interaction axioms, for all $n < m$:

$$[n]\varphi \to [m]\varphi$$
$$\langle n\rangle\varphi \to [m]\langle n\rangle\varphi$$

Generalising Japaridze's result, Ignatiev ([Ign93]) showed that GLP is the joint provability logic of $\Box_T^{\Pi_n}$ for all n. A simplified proof of this result was given by Beklemishev ([Bek11]).

It is well-known that GLP is not complete with respect to a standard Kripke-style semantics. The system GLP^- is obtained from GLP by replacing the axiom $\langle n\rangle\varphi \to [m]\langle n\rangle\varphi$ by requiring that for all $n < m$,

$$[n]\varphi \to [m][n]\varphi$$

The above is a generalisation of the principle $\triangle A \to \Box\triangle A$ of CS. The system GLP^- is sound and complete with respect to a nice class of Kripke frames. The latter are also useful for reasoning about GLP, in particular for obtaining arithmetical completeness results.

3.3 Kripke semantics

We prove modal completeness of GLT with respect to Kripke frames containing accessibility relations Q and R for $\triangle$ and $\Box$ respectively. Given binary relations S and S', we write SS' for the relation $\{(a,b) \mid aSc \text{ and } cS'b \text{ for some } c\}$.

3.3.1. Definition. A GLT-*frame* is a triple $\langle W, Q, R\rangle$ with $W \neq \emptyset$, and

i. Q is converse well-founded

ii. Q is transitive

iii. $R\subseteq Q$

iv. $QR\subseteq R$

v. $RQ\subseteq R$

vi. $R\subseteq RQ$

Conditions (i), (iii), and (iv) imply that the R-relation on a GLT-frame is transitive and converse well-founded. Conditions (ii)-(v) are collectively referred to as the GLT *closure-conditions*. We note that (i), (ii), and (vi) imply that GLT-frames satisfy the following:

if $w\ R\ a$, there are $(a_n)_{n<\omega}$ with $a = a_0$ and for all $n, w\ R\ a_n$ and $a_{n+1}\ Q\ a_n$.

If $w\ R\ a$ and $(a_n)_{n<\omega}$ is as above, we say that $(a_n)_{n<\omega}$ *witnesses* the relation $w\ R\ a$. It follows from (ii) and (i) that if $(a_n)_{n<\omega}$ is as above, then $a_n \neq a_m$ for all $n \neq m$. A GLT-frame with at least one R-relation must thus be infinite.

3.3.2. Definition. A GLT-*model* is a quadruple $\langle W, Q, R, \Vdash\rangle$, where $\langle W, Q, R\rangle$ is a GLT-frame, and $\Vdash$ a valuation on W satisfying the usual clauses, with Q and R as the accessibility relations for $\triangle$ and $\Box$ respectively.

3.3.3. Theorem. $\vdash A \quad \Leftrightarrow \quad \mathcal{F} \Vdash A$ *for every* GLT-*frame* $\mathcal{F}$.

The proof of soundness, i.e. the left to right direction of Theorem 3.3.3, is straightforward. Section 3.4 below is dedicated to the proof of the other direction. Soundness of GLT with respect to GLT-frames has the following simple consequence:

3.3.4. Lemma. *i.* $\nvdash \triangledown\top$

ii. $\nvdash \Box\bot \to \triangle\bot$ ❑

3.4 Proof of modal completeness

Our approach here is the same as in Section 2.2.1. Given a formula A with $\nvdash A$, we will construct a GLT-frame consisting of maximal consistent sets, where a truth lemma holds with respect to an adequate set containing A.

As in Section 2.2.1, an *adequate set* of formulas is one that is closed under subformulas and single negations.

3.4.1. DEFINITION. Let $\mathcal{F}$ be a frame whose domain consists of maximal consistent sets, $x \in \mathcal{F}$, and $\mathcal{D}$ adequate. A $\mathcal{D}$-problem in x is one of the two:

i. a formula $\neg\triangle A \in x \cap \mathcal{D}$ such that there is no y with $x\ Q\ y$ and $\neg A \in y$.

ii. a formula $\neg\Box A \in x \cap \mathcal{D}$ such that there is no y with $x\ R\ y$ and $\neg A \in y$.

If $\mathcal{D}$ is clear from the context, we refer to $\mathcal{D}$-problems simply as *problems.* Elements of $\mathcal{D}$ of the form $\neg\triangle B$ or $\neg\Box B$ are called *problem-formulas.*

3.4.2. DEFINITION. For maximal consistent sets x and y, we define $x \prec_\triangle y$ if for every $\mathcal{L}_{\Box\triangle}$-formula A, we have that $\triangle A \in x$ implies $A \in y$, and similarly $x \prec_\Box y$ if for every $\mathcal{L}_{\Box\triangle}$-formula A, we have that $\Box A \in x$ implies $A \in y$.

3.4.3. LEMMA. *If $x \prec_\triangle y$ or $x \prec_\Box y$, then any problem-formula in y is in x.*

Proof: Let $x \prec_\triangle y$, and suppose that $\neg\triangle B \in y$. Supposing for a contradiction that $\triangle B \in x$, we have $\triangle\triangle B \in x$ by axiom (4) for $\triangle$, and so $\triangle B \in y$, a contradiction. If $\neg\Box B \in y$, then assuming $\Box B \in x$ we obtain $\triangle\Box B \in x$ by axiom (T2), whence $\Box B \in y$, again a contradiction. The case of $x \prec_\Box y$ is similar, by using (4) for $\Box$, and Lemma 3.1.3(i) ❑

3.4.4. DEFINITION. A frame whose domain consists of maximal consistent sets is *adequate* if for all x and y, $x\ Q\ y$ implies $x \prec_\triangle y$, and $x\ R\ y$ implies $x \prec_\Box y$.

The proofs of the following lemmas are straightforward:

3.4.5. LEMMA. *Let x, y, and z be maximal consistent.*

i. If $x \prec_\triangle y$ and $y \prec_\triangle z$, then $x \prec_\triangle z$.

ii. if $x \prec_\Box y$, then $x \prec_\triangle y$.

iii. if $x \prec_\triangle y$ and $y \prec_\Box z$, then $x \prec_\Box z$.

iv. if $x \prec_\Box y$ and $y \prec_\triangle z$, then $x \prec_\Box z$. ❑

3.4.6. LEMMA. *Let $\mathcal{F}$ be an adequate frame containing no $\mathcal{D}$-problems, and let $\Vdash$ be the canoncial valuation. Then $x \Vdash B \Leftrightarrow B \in x$ for all $x \in \mathcal{F}$ and $B \in \mathcal{D}$.* ❑

3.4.1 Lemmas for the elimination of problems

The following lemma is proven in exactly the same way as Lemma 2.2.13.

3.4.7. Lemma. *Let $\mathcal{D}$ be adequate, and x maximal consistent with $\neg\triangle B \in x \cap \mathcal{D}$. There is some y with $x \prec_\triangle y$ and $\neg\triangle B, B \in y$.* ❑

Recall that GLT-frames satisfy the following condition:

$$\text{if } w\ R\ a, \text{ there are } (a_n)_{n<\omega} \text{ with } a = a_0 \text{ and for all } n, w\ R\ a_n \text{ and } a_{n+1}\ Q\ a_n.$$

Thus whenever we add to our frame some world a with $w\ R\ a$, we should at some point also add infinitely many worlds as above. We shall add all these worlds *simultaneously* with a. The following lemma guarantees that maximal consistent sets with suitable properties can be found.

3.4.8. Lemma. *Let $\mathcal{D}$ be adequate, and x maximal consistent with $\neg\Box B \in x \cap \mathcal{D}$. There is a sequence $(y_i)_{i<\omega}$ of maximal consistent sets such that $\neg B \in y_0$ and for all i, $x \prec_\Box y_i$, $y_{i+1} \prec_\triangle y_i$, and $\Box B \in y_i$.*

Iterations of modalities are defined inductively: $\triangle^0 A := A$, $\triangle^{n+1} A := \triangle\triangle^n A$. Similarly $\Box^0 A := A$ and $\Box^{n+1} A := \Box\Box^n A$. The symbols $\nabla^n A$ and $\Diamond^n A$ are used as shorthand for $\neg\triangle^n\neg A$ and $\neg\Box^n\neg A$ respectively.

The rest of this subsection is dedicated to the proof of Lemma 3.4.8. Fix an adequate set $\mathcal{D}$ and a maximal consistent set x with $\neg\Box B \in x \cap \mathcal{D}$. We show that there is a sequence $(y_i)_{i<\omega}$ of maximal consistent sets such that for all i, $x \prec_\Box y_i$, $y_{i+1} \prec_\triangle y_i$, and $\neg\triangle^i B, \triangle^{i+1} B \in y_i$. This suffices, for $\vdash \triangle^{i+1} B \to \Box B$ by axioms (T1) and (T4), and $\neg\triangle^0 B = \neg B$.

Let us fix some notation and terminology. Define

$$s_i := \{A \mid \Box A \in x\} \cup \{\neg\triangle^i B, \triangle^{i+1} B\}$$

For $\alpha \in \omega \cup \{\omega\}$, a sequence $(z_i)_{i<\alpha}$ of sets formulas is said to *extend* another such sequence $(z'_i)_{i<\alpha}$ if $z_i \supseteq z'_i$ for all $i < \alpha$.

3.4.9. Definition. A sequence $(z_i)_{i<\omega}$ of sets of formulas is *good* if it extends $(s_i)_{i<\omega}$, and for all k there is a sequence $(y^k_i)_{i\leqslant k}$ of maximal consistent sets extending $(z_i)_{i\leqslant k}$, with $y^k_{i+1} \prec_\triangle y^k_i$ for all $i < k$.

The statement of Lemma 3.4.8 can be rephrased as:

"There exists a good sequence of maximal consistent sets."

We prove the above claim by showing that $(s_i)_{i<\omega}$ is good (Lemma 3.4.10), and that any good sequence can be extended to a good sequence of maximal consistent sets (Lemma 3.4.11).

3.4.10. LEMMA. *The sequence $(s_i)_{i<\omega}$ is good.*

Proof: It suffices to show that for any k, there are maximal consistent sets $(y_i^k)_{i\leqslant k}$ extending $(s_i)_{i\leqslant k}$, and such that $y_{i+1}^k \prec_\vartriangle y_i^k$ for all $i < k$. Let us fix k. We show the existence of y_j^k, for $j \leqslant k$, by induction on $k - j$.

For y_k^k, it suffices to take a maximal consistent extension of s_k. To see that the latter is consistent, note that assuming the contrary implies the existence of some $\Box A_0, \ldots, \Box A_n \in x$ such that $A_0, \ldots, A_n \vdash \vartriangle^{k+1} B \to \vartriangle^k B$, whence

$$\Box A_0, \ldots, \Box A_n \vdash \Box(\vartriangle^{k+1} B \to \vartriangle^k B),$$

and so by Lemma 3.1.3(iv), $\Box A_0, \ldots, \Box A_n \vdash \Box \vartriangle^k B$. By repeated applications of (T4), we obtain $\Box A_0, \ldots, \Box A_n \vdash \Box B$. Since x contains each $\Box A_j$ and $\neg \Box B$ by assumption, this is a contradiction.

For the inductive step, suppose that y_{i+1}^k has the required properties. We would like to show that there is a maximal consistent $y_i^k \supseteq s_i$ with $y_{i+1}^k \prec_\vartriangle y_i^k$. By Lindenbaum's Lemma, it suffices to show that the set $s_i \cup \{C \mid \vartriangle C \in y_{i+1}^k\}$ is consistent. Assuming the contrary, we obtain some $\Box A_0, \ldots, \Box A_n \in x$ and $\vartriangle C_0 \ldots, \vartriangle C_l \in y_{i+1}^k$ with $A_0, \ldots A_n, C_0, \ldots, C_l \vdash \vartriangle^{i+1} B \to \vartriangle^i B$, whence by necessitation and Löb's axiom for $\vartriangle$,

$$\vartriangle A_0, \ldots \vartriangle A_n, \vartriangle C_0, \ldots, \vartriangle C_l \vdash \vartriangle(\vartriangle^{i+1} B \to \vartriangle^i B)$$
$$\vartriangle A_0, \ldots \vartriangle A_n, \vartriangle C_0, \ldots, \vartriangle C_l \vdash \vartriangle^{i+1} B.$$

Each $\Box A_j$ is in x, whence by (T3) also $\Box \vartriangle A_j$ must be in x. Since $s_{i+1} \subseteq y_{i+1}^k$, it follows that $\vartriangle A_j$ is in y_{i+1}^k. Since y_{i+1}^k also contains each $\vartriangle C_j$, it should thus contain $\vartriangle^{i+1} B$. But $\neg \vartriangle^{i+1} B \in s_{i+1} \subseteq y_{i+1}^k$, so this is a contradiction. ❑

3.4.11. LEMMA. *A good sequence can be extended to a good sequence of* maximal consistent *sets.*

Proof: Suppose that $(z_i)_{i<\omega}$ is a good sequence. We carry out a simultaneous Lindenbaum construction for the sets $(z_i)_{i<\omega}$, making sure that the property of being good is preserved when new formulas are added to the sets. Fix some enumeration $(\varphi_j)_{j<\omega}$ of $\mathcal{L}_{\Box\vartriangle}$-formulas, and define $z_i^0 := z_i$. Assuming that $(z_i^n)_{i<\omega}$ is good, define $(z_i^{n+1})_{i<\omega}$ as follows:

Consider n as the pair (j, m). We add either φ_j or $\neg\varphi_j$ to z_m:

Since $(z_i^n)_{i<\omega}$ is good, we have for all k a sequence $(y_i^k)_{i\leqslant k}$ of maximal consistent sets extending $(z_i)_{i\leqslant k}$, and such that $y_{i+1}^k \prec_\vartriangle y_i^k$ for all $i < k$. In particular, we have $y_m^k \supseteq z_m^n$ for all $k \geqslant m$. Consider the sets

$$A = \{k \mid \varphi_j \in y_m^k\} \quad \text{and} \qquad B = \{k \mid \neg\varphi_j \in y_m^k\},$$

and note that at least one of A and B must be infinite.

If A is infinite, $z_m^{n+1} := z_m^n \cup \{\varphi_j\}$, otherwise $z_m^{n+1} := z_m^n \cup \{\neg\varphi_j\}$.

For $i \neq m$, $z_i^{n+1} := z_i^n$.

We claim that $(z_i^{n+1})_{i<\omega}$ is good. The resulting sequence clearly still extends $(s_i)_{i<\omega}$. Our construction ensures that for infinitely many k, the maximal consistent sets witnessing the goodness of $(z_i^n)_{i<\omega}$ are in fact extensions of $(z_i^{n+1})_{i<\omega}$. Noting that if $k' < k$ and $y_0, \ldots y_k$ is a path of length k, then $y_0, \ldots, y_{k'}$ is a path of length k', it is thus clear that $(z_i^{n+1})_{i<\omega}$ is also a good sequence. ❑

3.4.2 Quasi-frames

Given binary relations Q' and R', denote by S the relation

$$(R'^* \cup Q'^*)^* \; R' \; (R'^* \cup Q'^*)^*,$$

where P^* stands for the transitive closure of P. Intuitively, an S-step is any sequence of Q'-steps and R'-steps containing at least one R'-step. We note that if Q' and R' satisfy the GLT closure-conditions, then $S = R'$.

3.4.12. Definition. A frame $\mathcal{G} = \langle W, Q', R' \rangle$, where W consists of maximal consistent sets, is a *quasi-frame* if:

i. $\mathcal{G}$ is adequate

ii. Q' and S are converse well-founded

iii. if $x\ R'\ y$, there is a sequence $(y_i)_{i<\omega}$ with $y = y_0$, and such that for all i, xSy_i, and $y_{i+1}\ Q'\ y_i$.

3.4.13. Lemma. *Any quasi-frame can be extended to an adequate* GLT*-frame.*

Proof: Let $\mathcal{G}' = \langle W, Q', R' \rangle$ be a quasi-frame. If $\mathcal{G}$ is not an adequate GLT-frame, it must be due to a violation of one of the GLT closure-conditions. Thus let us close off to ensure $QQ \subseteq Q$, $R \subseteq Q$, $QR \subseteq R$, and $RQ \subseteq R$. The resulting frame $\mathcal{F} = \langle W, Q, R \rangle$ satisfies, of course, the GLT closure-conditions. It remains to show that it also meets the remaining criteria for being an adequate GLT-frame.

Lemma 3.4.5 tells us that the process of closing off under the GLT closure-conditions preserves adequacy of the underlying frame, whence it is clear that $\mathcal{F}$ is adequate. It remains to show: (1) Q is converse well-founded, (2) any R-relation is witnessed by an appropriate infinite sequence.

For (1), note that if $\mathcal{F} \models x\ Q\ y$, then either $\mathcal{G} \models x\ Q'^*\ y$ or $\mathcal{G} \models xSy$. Since Q' and S are both converse well-founded, so must thus be Q.

For (2), we note that if $\mathcal{F} \models xRy$, then $\mathcal{G} \models xSy$: new R-relations are only created on top of old ones, in order to ensure $QR \subseteq R$ or $RQ \subseteq R$. If $\mathcal{F} \models xRy$, we thus have

$$\mathcal{G} \models x \dots w\ R'\ z \dots y,$$

where the dots denote a — possibly empty — sequence of R' and Q'-steps.

If $x = w$ and $z = y$, then $\mathcal{G} \models x\ R'\ y$, and the existence of a witnessing sequence follows from Definition 3.4.12(iii), together with the fact that in $\mathcal{F}$, $S = R$. So suppose $x \neq w$, and note that then $\mathcal{F} \models x\ Q\ w$ — the GLT closure-conditions ensure that any sequence of Q- or R-steps is itself a Q-step.

If $z = y$, then $\mathcal{G} \models w\ R'\ y$, and hence there is a sequence $(y_i)_{i<\omega}$ with $y = y_0$ and for all i, wSy_i, and $y_{i+1}\ Q'\ y_i$. From $S = R$ it follows that $\mathcal{F} \models w\ R\ y_i$ for all i. Since $\mathcal{F} \models x\ Q\ w\ R\ y_i$ and $\mathcal{F} \models QR \subseteq R$, we thus have $\mathcal{F} \models x\ R\ y_i$ for all i, and so $(y_i)_{i<\omega}$ is an appropriate witnessing sequence for $x\ R\ y$.

Suppose finally that $z \neq y$, whence it must be that $\mathcal{F} \models z\ Q\ y$. Since $\mathcal{G} \models w\ R\ z'$, there is a sequence $(z_i)_{i<\omega}$ with $z = z_0$, and such that for all i, wSz_i, and $z_{i+1}\ Q'\ z_i$. From $S = R$ it follows that $\mathcal{F} \models w\ R\ z_i$ for all i. We define $(y_i)_{i<\omega}$ as: $y_0 = y$, $y_{i+1} = z_i$. Given that $\mathcal{F} \models x\ Q\ w\ R\ z_i$ for all i, $\mathcal{F} \models QR \subseteq R$, and $\mathcal{F} \models x\ R\ y$, we have $x\ R\ y_i$ for all i. Since also $z = y_1\ Q\ y_0$, it is clear that $(y_i)_{i<\omega}$ is an appropriate witnessing sequence for $x\ R\ y$.

3.4.3 The problem elimination algorithm

Let x_0 be maximal consistent and $\mathcal{D}$ adequate. Let p be the number of problem-formulas in $\mathcal{D}$. Define $\mathcal{F}_0 := \langle \{x_0\}, \emptyset, \emptyset \rangle$, and $f(0) := x_0$. While some world in the range of f contains a problem in the frame $\mathcal{F}_n = \langle W_n, Q_n, R_n \rangle$, do the following:

1. Let i be the least such that $f(i)$ is defined and contains a problem, let x be such that $f(i) = x$, and let B be the least problem in x.

2. If B has the form $\neg\triangle C$, then by Lemma 3.4.7 there is a maximal consistent set y with $x \prec_\triangle y$, and $\neg C, \triangle C \in y$.

   Let $W'_{n+1} := W_n \cup \{y\}$; $Q'_{n+1} := Q_n + \{x\ Q\ y\}$, and $R'_{n+1} := R_n$.

   Let $f(k) := y$, where k is the least such that $f(k)$ is undefined.

3. If B has the form $\neg\Box C$, then by Lemma 3.4.8 there is a sequence $(y_i)_{i<\omega}$ with $\neg C \in y_0$ and for all i, $\Box C \in y_i$, $x \prec_\Box y_i$, and $y_{i+1} \prec_\triangle y_i$. Define:

$$\begin{aligned} W'_{n+1} &:= W_n \cup \{y_i \mid i \in \omega\} \\ R'_{n+1} &:= R_n \cup \{(x, y_i) \mid i < \omega\} \\ Q'_{n+1} &:= Q_n \cup \{(y_{i+1}, y_i) \mid i \in \omega\} \end{aligned}$$

   Let k be the least such that $f(k)$ is undefined.

For $p' \leqslant p$, let $f(k+p') := y_i$, where i is the least index such that y_i contains a problem-formula that is not contained in any element in the range of $f(k+j)$ with $j \leqslant p'$.

In other words, we define $f(k), f(k+1), \ldots, f(k+p)$ to be the *first* among $(y_i)_{i<\omega}$ containining *pairwise distinct* problem-formulas.

4. Let $\mathcal{G}$ be the frame $\langle W'_{n+1}, Q'_{n+1}, R'_{n+1}\rangle$. It is easy to check that in both cases, $\mathcal{G}$ is a quasi-frame.

5. By Lemma 3.4.13, let $\mathcal{F}_{n+1}$ be an adequate GLT-frame extending $\mathcal{G}$.

3.4.14. LEMMA. *The problem elimination algorithm terminates, yielding an adequate problem-free* GLT*-frame.*

Proof: To see that the algorithm terminates, argue by induction on the number of problem-formulas in x_0. Since $\mathcal{D}$ is finite, there are finitely many of those. When starting the algorithm, we have $f(0) := x_0$. If x_0 contains no problems-formulas, then it contains no problems, and so the while-loop will never be entered.

Suppose now that x_0 contains $n+1$ problem-formulas. This means that all problems in x are eliminated during the (at most) $n+1$ steps of the algorithm: problems of the form $\neg\triangle B$ are eliminated by adding some z with $x_0\ Q\ z$ and $\neg B, \triangle B \in z$, while problems of the form $\neg\Box B$ are eliminated by adding $(y_i)_{i<\omega}$ with $\neg B \in y_0$ and $\Box B \in y_i$ for all i.

After $n+1$ steps, we thus have $i > 0$ whenever the while-loop is entered in order to eliminate some problem in $f(i)$. It is clear from the construction that only finitely many worlds are added to the domain of f. It follows from Lemma 3.4.3 that all new worlds contain at most n problem-formulas. The induction assumption guarantees that the algorithm terminates when given each of the new worlds as input.

It remains to show that the resulting frame $\mathcal{F}$ is problem-free, i.e. that no $y \in \mathcal{F}$ contains a $\mathcal{D}$-problem. If y is in the range of f, then it is clear that by the end of the construction there are no problems left in y.

So suppose that y is not in the range of f. This means that y is an element of some $(y_i)_{i<\omega}$ that was added to $\mathcal{F}$ during step 3. Let i be such that $y = y_i$, and suppose for a contradiction that B is a problem in y_i. It follows from the definition of f that there is some $j < i$ with $B \in y_j$ and such that y_j is in the range of f — otherwise y_i would itself be in the range of f. Being in the range of f, y_j is problem-free. If B is of the form $\neg\triangle C$, there is thus some z with $y_j\ Q\ z$ and $\neg C \in z$, and if it is of the form $\neg\Box C$, there is some z with $y_j\ R\ z$ and $\neg C \in z$. Since $\mathcal{F}$ is a GLT-frame, we have $y_i\ Q\ y_j$, and therefore $y_i\ Q\ z$ in the first case and $y_i\ R\ z$ in the second, contradicting our assumption that B is a problem in y_i. ❑

We conclude with a proof of the remaining direction of Theorem 3.3.3.

Proof: Suppose $\nvdash A$, let x_0 be maximal consistent with $\neg A \in x_0$, and let $\mathcal{D}$ be an adequate set containing A. Run the problem elimination algorithm on x_0 and $\mathcal{D}$. By Lemma 3.4.14, this yields an adequate GLT-frame $\mathcal{F}$ free of problems. Letting $\Vdash$ be the canoncial valuation, we also have $x_0 \nVdash A$ by Lemma 3.4.6. ❑

3.5 Lindström semantics

The material in this section is based on [Lin06]. We show that GLT is complete with respect to two classes of Kripke frames containing a single accessibility relation $\prec$. Let us first fix some notation. Given a transitive relation $\prec$ on a tree, we write $a \prec_\omega b$ if the order type of the set $\{c \mid a \prec c \prec b\}$ is at least ω^*.

3.5.1. Definition. A *Lindström-frame* is a transitive converse well-founded tree of height $< \omega^2$. A *Lindström-model* is a triple $\langle W, \prec, \Vdash\rangle$ where $\langle W, \prec\rangle$ is a Lindström-frame, and $\Vdash$ a valuation on W satisfying the usual clauses, with $\prec$ and $\prec_\omega$ as the accessibility relations for $\triangle$ and $\Box$ respectively.

Thus the formula $\Diamond A$ is forced at a node a in a Lindström-model if and only if A is forced at a node b that is infinitely many steps away from a.

3.5.2. Theorem. *$\vdash A \Leftrightarrow \mathcal{F} \Vdash A$ for every Lindström frame $\mathcal{F}$.*

Proof: For soundness, note that any Lindström frame is a GL-frame, whence it must satisfy the axioms of GL for $\triangle$. The validity of $\mathsf{K}^\Box$ is straightforward, while that of (T1)-(T4) follows from:

i. $a \prec_\omega b \Rightarrow a \prec b$,

ii. $a \prec b \prec_\omega c \Rightarrow a \prec_\omega c$,

iii. $a \prec_\omega b \prec c \Rightarrow a \prec_\omega c$, and

iv. $a \prec_\omega b \Rightarrow \exists c\, a \prec_\omega c \prec b$.

While (ii) and (iii) hold in virtue of $\prec_\omega$-steps being defined as infinite $\prec$-steps, for (iv) it is crucial that the order type of the set of nodes between a and b is ω^*.

For the proof of completeness, we recall from Section 3.4 the construction of a GLT-frame $\mathcal{F}$ with $\mathcal{F} \nVdash A$, given some A with $\nvdash A$. Forgetting the R-relations and renaming Q to $\prec$, we obtain a transitive converse well-founded tree $\mathcal{F}' = \langle W, \prec\rangle$. By examining the construction, it is clear that the height of $\mathcal{F}'$

is $< \omega^2$. Thus $\mathcal{F}'$ may be viewed as a Lindström-model. To finish our proof, it suffices to show that the new interpretation of $\Box$ coincides with the old one:

$$\mathcal{F} \models x \, R \, y \quad \Leftrightarrow \quad \mathcal{F}' \models x \prec_\omega y.$$

The direction from left to right is clear since $\mathcal{F}$ is a GLT-frame, whence in particular every R-relation on it is witnessed by an appropriate infinite sequence. For the other direction, suppose that $x \prec_\omega y$. By examining the construction of $\mathcal{F}$, we see that there must be some x' with either $x = x'$ or $x \; Q \; x'$, a sequence $(y_n)_{n<\omega}$ that was added to $\mathcal{F}$ in order to eliminate some problem $\neg\Box B$ in x', and finally $y_k \; Q \; y$ for some $k \in \omega$. Thus either $x \; Q \; x' \; R \; y_k \; Q \; y$ or $R \; y_k \; Q \; y$. Since $\mathcal{F}$ is a GLT-frame, we obtain $x \; R \; y$ in either case. ❑

Given a transitive tree $\langle W, \prec \rangle$ and $a, b \in W$, we write $a \prec_n b$ to mean that the order type of the set $\{c \mid a \prec c \prec b\}$ is at least n.

3.5.3. Definition. An *n-Lindström-frame* is a transitive tree with only finite branches. An *n-Lindström-model* is a triple $\langle W, \prec, \Vdash \rangle$ where $\langle W, \prec \rangle$ is a an *n-Lindström-frame*, and $\Vdash$ a valuation on W satisfying the usual clauses with $\prec$ as the accessibility relation for $\triangle$, and:

$$a \Vdash \Box A \quad \Leftrightarrow \quad \text{there is some } n > 0 \text{ such that } \; b \Vdash A \text{ whenever } a \prec_n b.$$

To put it differently, we have $a \Vdash \Box A$ if and only if $a \Vdash \triangle^n A$ for some n, and dually $a \Vdash \Diamond A$ if and only if $a \Vdash \triangledown^n A$ for all n. Thus $\Box$ is, in a sense, an ω-iteration of $\triangle$.

3.5.4. Theorem. *$\vdash A \Leftrightarrow \mathcal{F} \Vdash A$ for every n-Lindström-frame $\mathcal{F}$.*

Proof: The proof of soundness is straightforward; we treat (T4). Suppose that $a \Vdash \Diamond A$, so $a \Vdash \triangledown^n A$ for all n. Then also $a \Vdash \triangledown^{n+1} A$ for all n, i.e. $a \Vdash \triangledown^n \triangledown A$, and thus we have $a \Vdash \Diamond \triangledown A$.

Completeness is proven by a slight modification of the construction in Section 3.4. $\triangledown$-problems are eliminated exactly as before. Given a problem $\neg\Box B \in x$, the proof of Lemma 3.4.8 guarantees the existence of a sequence $(y_n)_{n<\omega}$, with $\neg B \in y_0$, $x \prec_\Box y_n$, $y_{n+1} \prec_\triangle y_n$, and $\Box B \in y_n$ for all n. For each n, we add to our frame the path $x \prec y_i \prec y_{i-1} \ldots \prec y_0$, and close off to ensure the transitivity of $\prec$. Since each added node contains less problems than x, it is clear that the resulting frame has only finite branches. It remains to check that a truth lemma holds at the end of this construction. We check that for all $x \in W$:

$$\neg\Box A \in x \Leftrightarrow \text{for all } n > 0, \text{ there exists } y \text{ with } x \prec_n y_n \text{ and } \; y_n \Vdash \neg A.$$

The direction from left to right holds by construction. For the other direction, assume that for every n, there is a path of length n from x to a node y_n with

$y_n \Vdash \neg A$, whence $\neg A \in y_n$ by the induction assumption. Since the frame has only finite branches, by König's Lemma there must be some x' with either $x = x'$ or $x \prec x'$ such that x' has infinitely many *immediate* successors $(z_n)_{n<\omega}$, and $z_n \prec y_n$ for all n. This can only be the case if the paths starting with the z_n's were added in order to eliminate some problem $\neg C$ in x'. Thus we have $x \prec_\triangle x' \prec_\Box z_n \prec_\triangle y_n$ or $x \prec_\Box z_n \prec_\triangle y_n$, whence $x \prec_\Box y$ by Lemma 3.4.5. But then $\Box B \in x$ would imply $B \in y_n$, contradicting the consistency of y_n. ❑

3.5.5. Lemma. *i. for every* $n > 0$, $\vdash \triangle^n A \to \Box A$

ii. if $\vdash \triangle^n A \to B$ *for every* n, *then* $\vdash \Box A \to B$

Proof: (i) follows from (T1) and (T4). For (ii), we note that if $\nvdash \Box A \to B$, then by Theorem 3.5.4 there is an n-Lindström-model $\mathcal{M}$ where $\Box A \wedge \neg B$ is forced at a node x. By Defintion 3.5.3, we have $x \Vdash \triangle^n A$ for some n. It now follows from the other direction of Theorem 3.5.4 that $\nvdash \triangle^n A \to B$. ❑

3.5.6. Definition. The system GLT^+ contains all propositional tautologies in the language $\mathcal{L}_{\Box\triangle}$, the rules and axiom schemata of GL for $\triangle$, and furthermore $\triangle^n A \to \Box A$ for all $n > 0$. Finally, GLT^+ contains the following rule of inference:

$$\frac{\triangle^n A \to B \text{ for every } n > 0}{\Box A \to B} \tag{3.3}$$

3.5.7. Lemma. GLT^+ *and* GLT *have exactly the same theorems.*

Proof: It follows from Lemma 3.5.5 that all theorems of GLT^+ are also theorems of GLT . For the other direction, note that (T1) is just $\triangle^n A \to \Box A$ with $n = 1$. For (T4), we have $\triangle^n \triangle A \to \Box A$ by the axiom schema of GLT^+, and thus $\Box\triangle A \to \Box A$ by using (3.3). Similarly, for (T3) it suffices to show that for $n > 0$, $\triangle^n A \to \Box\triangle A$. This can be done by induction on n, using (T1) and (T4). The argument for (T2) is similar. ❑

3.6 Decidability

The material in this section is a careful exposition of the results in [Lin06], which in turn are based on the method used in [Bek94].

In order to establish decidability of GLT, we introduce yet another class of Kripke models. We write $\mathrm{S}(A)$ for the set of subformulas of A, and use $y \, \overline{Q} \, z$ as abbreviation for: $y \, Q \, z$ or $y = z$.

3.6.1. Definition. An *A-sound model* is a quadruple $\langle W, Q, R, \Vdash\rangle$, where $\langle W, Q\rangle$ is a transitive irreflexive finite tree, $R \subseteq Q$, $QR \subseteq R$, $RQ \subseteq R$, and the following condition holds for all w, a: if $w\ R\ a$, there is a d with $w\ R\ d\ \overline{Q}\ a$ such that $d \Vdash \triangle B \to B$ for all $B \in \mathrm{S}(A)$. A node d as above said to be *reflexive*. Finally, $\Vdash$ is a valuation on W satisfying the usual clauses, with Q and R as the accessibility relations for $\triangle$ and $\Box$ respectively.

3.6.2. Theorem. *Let* $n = |S(A)|$. *Then* $\vdash A$ *iff* $\mathcal{M} \Vdash A$ *for every A-sound model* $\mathcal{M}$, *where the cardinality of* $\mathcal{M}$ *can taken to be exponential in* n.

Clearly, Theorem 3.6.2 has the following:

3.6.3. Corollary. GLT *is decidable.* ❑

The two directions of Theorem 3.6.2 are proven below as lemmas 3.6.4 and 3.6.5.

3.6.4. Lemma. *If* $\vdash A$, *then* $\mathcal{M} \Vdash A$ *for every A-sound model* $\mathcal{M}$.

Proof: Suppose $\vdash A$, and let $\mathcal{M}$ be an A-sound model. We transform $\mathcal{M}$ into a GLT-model $\mathcal{M}'$ (Definition 3.3.1) in the following way: whenever $w\ R\ a$ holds in $\mathcal{M}$, add to $\mathcal{M}$ an infinite sequence $(a_n)_{n<\omega}$ of nodes and define

i. $a_0\ Q\ d$, where d is a reflexive node witnessing the relation $w\ R\ a$

ii. $w\ R\ a_n$ and $a_{n+1}\ Q\ a_n$ for all n

iii. $a_n \Vdash p$ iff $d \Vdash p$ for all $p \in \mathrm{S}(A)$ and for all n

Close off to guarantee the transitivity of Q, as well as the frame conditions $R \subseteq Q$, $RQ \subseteq R$, and $QR \subseteq R$.

We claim that each of the new nodes $(a_n)_{n<\omega}$ satisfies exactly the same subformulas of A as d, i.e. that item (iii) extends to all $B \in \mathrm{S}(A)$. The boolean cases are straightforward. If $d \Vdash \nabla B$, then there is some b with $d\ Q\ b \Vdash B$. By transitivity of Q, also $a_n\ Q\ b$, and so $a_n \Vdash \nabla B$. If $d \Vdash \triangle\neg B$, then, since d is reflexive, also $d \Vdash \neg B$, whence by induction assumption $a_n \Vdash \neg B$ for all i. It is clear from the construction that each a_n has no Q-successors apart from a_j with $j < n$, and c with $d\ \overline{Q}\ c$, and so it follows that $a_n \Vdash \triangle\neg B$. If $d \Vdash \Diamond B$, let b be such that $d\ R\ b \Vdash B$. Since $QR \subseteq R$ and $a_n\ Q\ d$, also $a_n\ R\ b$, and thus $a_n \Vdash \Diamond B$. It is clear from the construction that the only R-successors of each a_n are those of d, and so $d \Vdash \Box\neg B$ also implies $a_n \Vdash \Box\neg B$ for all n. Using the above, it is clear that for all $B \in \mathrm{S}(A)$,

$$\mathcal{M}' \Vdash B \Leftrightarrow \mathcal{M} \Vdash B. \tag{3.4}$$

Finally, we check that for all $w\ R\ a$, there is a sequence $(a_n)_{n<\omega}$ with $a_{n+1}\ Q\ a_n$ and $w\ R\ a_n$ for all n. By construction, this is clearly the case when $w\ R\ a$ holds already in $\mathcal{M}$. In the remaining cases where $w\ R\ a$ results from closing off under the frame conditions $RQ{\subseteq}R$ and $QR{\subseteq}R$, we can reuse the infinite sequence $(a_n)_{n\in\omega}$ guaranteed to exist for the old R–transition. Thus $\mathcal{M}'$ is a GLT-model. By Theorem 3.3.3 we now have $\mathcal{M}' \Vdash A$, and therefore $\mathcal{M} \Vdash A$ by (3.4). ❑

3.6.5. Lemma. *If $\nvdash A$, then $\mathcal{M} \nVdash A$ for some A-sound model $\mathcal{M}$. The cardinality of $\mathcal{M}$ can taken to be exponential in n, where $n = |\mathrm{S}(A)|$.*

Proof: Let $\nvdash A$, and let x_0 be maximal GLT–consistent with $A \notin x_0$. We construct an A-sound model $\mathcal{M}$ based on an adequate frame (Definition 3.4.2), where a truth lemma holds with respect to all subformulas of A. The construction proceeds exactly as in Section 3.4, except when it comes to the elimination of problems of the form $\neg\Box C$. Whenever $\neg\Box C$ is a problem in x, we add $y_0, \ldots, y_m$ such that $\neg C \in y_0$, $w\ R\ y_n$, $y_{n+1}\ Q\ y_n$ for all $n \leqslant m$, and $y_m \Vdash \triangle B \to B$ for all $B \in \mathrm{S}(A)$. To see that $y_0, \ldots y_m$ with suitable properties exist, we note that

i. By Lemma 3.4.10 there are for every k, $y_0, \ldots y_k$ with $\neg C \in y_0$, and such that for all $n \leqslant k$, $w \prec_{\Box} y_n$ and $y_{n+1} \prec_{\triangle} y_n$. Thus we can define $w\ R\ y_n$ and $y_{n+1}\ Q\ y_n$ without affecting the adequacy of the frame.

ii. Let $n = |\mathrm{S}(A)|$. Whenever we have $y_0, \ldots y_n$, with $y_{i+1} \prec_{\triangle} y_i$ for all $i < n$, one of them must contain $\triangle B \to B$ for all $B \in \mathrm{S}(A)$. This is because for any $B \in \mathrm{S}(A)$, there exists at most one y_i that does not contain $\triangle B \to B$. Then suppose (w.l.o.g.) that $j > i$, $\triangle B \to B \notin y_j$ and $\triangle B \to B \notin y_i$. Thus $y_j \Vdash \triangle B$ and $y_i \nVdash B$. By transitivity of $\prec_{\triangle}$, we have $y_j \prec_{\triangle} y_i$, and so $\triangle B \in y_j$ implies $B \in y_i$, a contradiction. By the pigeonhole principle, there is thus some y_k among $y_0, \ldots y_n$ such that $y_k \Vdash \triangle B \to B$ for all $B \in \mathrm{S}(A)$.

Having added $y_0, \ldots, y_m$ as above, we close off to ensure that Q is transitive, and that we have $R{\subseteq}Q$, $RQ{\subseteq}R$, and $QR{\subseteq}R$. Proceed with the construction until we obtain a model $\mathcal{M}$ containing no problems. Reasoning as in Lemma 3.4.14, we see that the construction must terminate after a finite number of steps. It is then clear that $\mathcal{M}$ is finite and satisfies transitivity of Q, $R{\subseteq}Q$, $RQ{\subseteq}R$, and $QR{\subseteq}R$. It remains to check that every R-transition is witnessed by an appropriate reflexive node. We need to check the cases where the R–relation was created to guarantee that $RQ{\subseteq}R$ or $QR{\subseteq}R$. In both cases, we can reuse the reflexive node guaranteed to exist for the old R-relation.

Thus $\mathcal{M}$ is an adequate A-sound model containing no problems. Using Lemma 3.4.6, we obtain $\mathcal{M}, x_0 \nVdash A$ as required.

As for the cardinality of $\mathcal{M}$, note that elimination of a problem adds at most $n+1$ nodes to the model. It is clear from properties of the added sets that once a problem has been eliminated in a, it will not occur in any successor of a. Thus the resulting model is, at most, an $n \cdot (n+1)$-ary tree of height n, thus its cardinality is at most $(n^2+n)^{n+1}$. ❑

3.7 The closed fragment

In this section we study the closed fragment of GLT. An $\mathcal{L}_{\Box\triangle}$-formula is said to be *closed* if it is built up from $\top$ using boolean connectives, $\triangle$ and $\Box$; in other words if it does not contain propositional letters.

For an ordinal $\alpha < \omega^2$, i.e. $\alpha = \omega \cdot n + m$ for some $n, m \in \omega$, we write $\bot^\alpha$ for the $\mathcal{L}_{\Box\triangle}$–formula $\triangle^m\Box^n\bot$; $\bot^{\omega^2}$ is defined to be $\top$. An $\mathcal{L}_{\Box\triangle}$-formula is in *normal form* if it is a boolean combination of sentences of the form $\bot^\alpha$, where $\alpha \leqslant \omega^2$. We prove the following:

3.7.1. THEOREM. *Every closed formula is provably equivalent in* GLT *to a formula in normal form.*

We give two proofs of Theorem 3.7.1, relying on syntactical and semantical methods respectively. As a result, we obtain completeness of the closed fragment with respect to a rather wide class of arithmetical interpretations.

The closed fragment of GLT coincides with the closed fragment of CMS_1; see [Vis93, Section 9].

3.7.1 Normal form theorem

The sequence $(\bot^\alpha)_{\alpha\leqslant\omega^2}$ is linearly ordered with respect to provability in GLT:

3.7.2. LEMMA. $\alpha \leqslant \beta \leqslant \omega^2 \Leftrightarrow \vdash \bot^\alpha \to \bot^\beta$.

Proof: For the direction from left to right, let $\alpha \leqslant \beta$ with $\alpha = \omega \cdot n + m$ and $\beta = \omega \cdot n' + m'$. In case $n = n'$ and $m \leqslant m'$, it follows from transitivity of $\triangle$ that $\vdash \triangle^m\Box^n\bot \to \triangle^{m'}\Box^n\bot$. So assume $n < n'$. If $m \leqslant m'$, it follows from transitivity of $\triangle$ and $\Box$ that $\vdash \triangle^m\Box^n\bot \to \triangle^{m'}\Box^{n'}\bot$. It remains to check the case that $n < n'$ and $m > m'$:

$$\begin{array}{rlr} \vdash \triangle^m\Box^n\bot & \to \Box\triangle^m\Box^n\bot & \text{Lemma 3.1.3 (i)} \\ & \to \Box\Box^n\bot & \text{axiom (T4)} \\ & \to \Box^{n'}\bot & \text{transitivity of } \Box \\ & \to \triangle^{m'}\Box^{n'}\bot & \text{axiom (T2)} \end{array}$$

We prove the other direction by contraposition. Let $\beta < \alpha$, and suppose that $\vdash \bot^\alpha \to \bot^\beta$. By what was shown above, we have that $\vdash \bot^{\beta+1} \to \bot^\alpha$. Combining the two, we thus have $\vdash \bot^{\beta+1} \to \bot^\beta$, and so $\vdash \bot^\beta$ by Löb's principle for $\triangle$. But this contradicts Theorem 3.5.2 — it is easy to construct a Lindström-model where $\neg\bot^\beta$ is true at a node. ❑

Here is our first proof of Theorem 3.7.1:

Proof: We proceed by induction on the complexity of a closed formula. The base case and the boolean cases are straightforward. In order to consider $\triangle A$ and $\Box A$, assume that $\vdash A \leftrightarrow C$, where C is a formula in normal form. We write C in conjunctive normal form as $C_0 \wedge \ldots \wedge C_k$, where each C_i has the form

$$\bot^{\alpha_0} \vee \ldots \vee \bot^{\alpha_m} \vee \neg\bot^{\beta_0} \vee \ldots \vee \neg\bot^{\beta_n}.$$

By Lemma 3.7.2, $\vdash \bot^\alpha \vee \bot^\beta \leftrightarrow \bot^{\max\{\alpha,\beta\}}$, and $\vdash \bot^\alpha \wedge \bot^\beta \leftrightarrow \bot^{\min\{\alpha,\beta\}}$. Letting $\alpha = \max\{\alpha_0, \ldots, \alpha_m\}$, and $\beta = \min\{\beta_0, \ldots, \beta_m\}$, it follows that C_i is provably equivalent in GLT to $\bot^\beta \to \bot^\alpha$, i.e. to

$$\bot^\alpha \vee \neg\bot^\beta. \tag{3.5}$$

For $\triangle A$, our goal is to find a formula C' in normal form, with $\vdash \triangle C \leftrightarrow C'$, whence also $\vdash \triangle A \leftrightarrow C'$. Note that we have

$$\vdash \triangle C \leftrightarrow \triangle(C_0 \wedge \ldots \wedge C_k) \leftrightarrow \triangle C_0 \wedge \ldots \wedge \triangle C_k.$$

Thus it suffices to find for each $i \leqslant k$ a formula C'_i in normal form with

$$\vdash \triangle C_i \leftrightarrow C'_i,$$

as we can then take $C'_0 \wedge \ldots C'_k$ for C'. Using (3.5), it remains to show that $\triangle(\bot^\beta \to \bot^\alpha)$ is equivalent to a formula in normal form. If $\beta \leqslant \alpha$, we have that $\vdash \bot^\beta \to \bot^\alpha$, and so in this case

$$\vdash \triangle(\bot^\beta \to \bot^\alpha) \leftrightarrow \triangle\top \leftrightarrow \bot^{\omega^2}.$$

Assume now that $\beta > \alpha$. Then $\beta \geqslant \alpha+1$, and so by Lemma 3.7.2, $\vdash \bot^{\alpha+1} \to \bot^\beta$. Using this and Löb's axiom for $\triangle$ (recall that $\triangle^{\alpha+1}\bot = \triangle\bot^\alpha$), we see that

$$\begin{aligned} \vdash \triangle(\bot^\beta \to \bot^\alpha) &\to \triangle(\bot^{\alpha+1} \to \bot^\alpha) \\ &\to \triangle\bot^\alpha \\ &\to \triangle(\bot^\beta \to \bot^\alpha) \end{aligned}$$

Thus in this case $\vdash \triangle(\bot^\beta \to \bot^\alpha) \leftrightarrow \bot^{\alpha+1}$.

It remains to consider $\Box A$. As above, it suffices to show that $\Box(\bot^\beta \to \bot^\alpha)$ is equivalent to a formula in normal form. If $\beta \leqslant \alpha$, then as before,

$$\vdash \Box(\bot^\beta \to \bot^\alpha) \leftrightarrow \Box\top \leftrightarrow \bot^{\omega^2}.$$

So suppose $\beta > \alpha$. Then $\beta \geqslant \alpha + 1$, and so by Lemma 3.7.2, $\vdash \bot^{\alpha+1} \to \bot^\beta$. We reason as follows:

$$\begin{aligned}
\vdash \Box(\bot^{\beta} \to \bot^{\alpha}) &\to \Box(\bot^{\alpha+1} \to \bot^{\alpha}) && \\
&\to \Box\bot^{\alpha} && \text{Lemma 3.1.3(iv)} \\
&\to \Box\Delta^{m}\Box^{n}\bot && \text{rewrite, assuming } \alpha = \omega\cdot n + m \\
&\to \Box^{n+1}\bot && \text{axiom } (\mathsf{T4}) \\
&\to \Box\Delta^{m}\Box^{n}\bot && \text{axiom } (\mathsf{T3}) \\
&\to \Box\bot^{\alpha} && \text{rewrite: } \alpha = \omega\cdot n + m \\
&\to \Box(\bot^{\beta} \to \bot^{\alpha}) &&
\end{aligned}$$

Thus in this case $\vdash \Box(\bot^{\beta} \to \bot^{\alpha}) \leftrightarrow \bot^{\lambda}$, where λ is the smallest limit ordinal bigger than α: if $\alpha = \omega\cdot n + m$, then $\lambda = \omega\cdot(n+1)$. ❑

3.7.2 Universal model

Consider the frame $\mathcal{U} = \langle W, \prec\rangle$, where W is the set of ordinals $\leqslant \omega^2$, and $\prec$ the converse of the usual ordering on ordinals. The relation $\prec$ is transitive and converse well–founded, and so $\mathcal{U}$ is a Lindström-frame (Definition 3.5.1). Since the truth of closed formulas is invariant under valuations, $\mathcal{U}$ can be viewed as a Lindström-model for the set of closed formulas. In this section, we show that $\mathcal{U}$ is in fact a *universal model*: any unprovable closed formula can be refuted on $\mathcal{U}$.

3.7.3. Definition. For a closed formula A, define the set $[\![A]\!] \subseteq \omega^2$ (the extension of A in $\mathcal{U}$) as follows:

$$\begin{aligned}
[\![\top]\!] &= \omega^2 \\
[\![\neg A]\!] &= \omega^2 - [\![A]\!] \\
[\![A \wedge B]\!] &= [\![A]\!] \cap [\![B]\!] \\
[\![\nabla A]\!] &= \{\beta \mid \exists\alpha\ \beta \prec \alpha \text{ and } \alpha \in [\![A]\!]\} \\
[\![\Diamond A]\!] &= \{\beta \mid \exists\alpha\ \beta \prec_{\omega} \alpha \text{ and } \alpha \in [\![A]\!]\}
\end{aligned}$$

Recall from Section 3.5 that $\beta \prec_{\omega} \alpha$ means that the order type of the set $\{\gamma \mid \beta \prec \gamma \prec \alpha\}$ is at least ω^*. The following lemma states that the extension of a closed formula in $\mathcal{U}$ has always a certain simple form, namely it is a finite union of disjoint intervals.

3.7.4. Lemma. *If A is closed, then $[\![A]\!] = \bigcup_{i\in I}[\alpha_i, \beta_i)$, where $\alpha_i, \beta_i \leqslant \omega^2$, and*

i. I is finite

ii. $[\alpha_i, \beta_i) \cap [\alpha_j, \beta_j) = \emptyset$ if $i \neq j$.

Proof: By induction on the complexity of a closed formula. For the base case, it suffices to note that $[\![\top]\!] = [0, \omega^2)$. For the boolean cases, we note that intervals of the form $\bigcup_{i\in I} [\alpha_i, \beta_i)$ are closed under complements and intersections. We argue by induction on the number k of elements in I. It is easy to check that

$$\omega^2 \backslash [\alpha, \alpha') = [0, \alpha) \cup [\alpha', \omega^2)$$
$$[\alpha, \alpha') \cap [\beta, \beta') = [\max\{\alpha, \beta\}, \min\{\alpha', \beta'\})$$

The case where $k > 0$ follows by using the above, and noting that:

$$\bigcup_{i\in I} [\alpha_i, \beta_i) \cap \bigcup_{i\in I} [\gamma_i, \delta_i) = \bigcup_{i\in I} ([\alpha_i, \beta_i) \cap [\gamma_i, \delta_i))$$
$$\omega^2 \backslash \bigcup_{i\in I} [\alpha_i, \beta_i) = \bigcap_{i\in I} (\omega^2 \backslash [\alpha_i, \beta_i))$$

Finally, writing $\lambda_{>\min[\![A]\!]}$ for the smallest limit ordinal greater than $\min[\![A]\!]$:

$$[\![\triangledown A]\!] = [\min[\![A]\!] + 1, \omega^2)$$
$$[\![\Diamond A]\!] = [\lambda_{>\min[\![A]\!]}, \omega^2) .$$

❑

For a world a in a Lindström-model, we define $\mathrm{d}(a)$, *the depth of* a, as the ordinal $\alpha < \omega^2$ such that the longest $\prec$-path starting from a has length α. Note that $\mathrm{d}(\alpha) = \alpha$ for all $\alpha \in \mathcal{U}$. The following lemma states that whether a closed formula is satisfied at a node a in a Lindström-model depends solely on $\mathrm{d}(a)$.

3.7.5. Lemma. *Let $\mathcal{M}$ be a Lindström-model, and B a closed formula. Then we have $\mathcal{M}, a \Vdash B \Leftrightarrow \mathrm{d}(a) \in [\![B]\!]$.*

Proof: By induction on the complexity of B. If B is $\top$, then both sides are true for all a, as $[\![\top]\!] = \omega^2$. The boolean cases follow directly from the definitions. Suppose now that the claim holds for A, and that $\mathcal{M}, a \Vdash \triangledown A$. Then there is some b with $a \prec b$ and $\mathcal{M}, b \Vdash A$. By the induction assumption, $\mathrm{d}(b) \in [\![A]\!]$. Since $a \prec b$, we have that $\mathrm{d}(a) > \mathrm{d}(b)$. Thus in $\mathcal{U}$, we have that $\mathrm{d}(a) \prec \mathrm{d}(b)$ and so $\mathrm{d}(a) \in [\![\triangledown A]\!]$ by definition. If on the other hand $\mathrm{d}(a) \in [\![\triangledown A]\!]$, then there is some β such that $\mathrm{d}(a) \prec \beta$ in $\mathcal{U}$, and such that $\beta \in [\![A]\!]$. Now let b be a node with $\mathrm{d}(b) = \beta$ and $a \prec b$ (such a node must exist because $\mathrm{d}(a) > \beta$). The induction assumption now gives us $\mathcal{M}, b \Vdash A$, and thus $\mathcal{M}, a \Vdash \triangledown A$ as required. The case of $\Diamond A$ is completely analogous. ❑

3.7.6. Lemma. $\vdash B \Leftrightarrow [\![B]\!] = \omega^2$

Proof: If $\vdash B$, then $\mathcal{U} \Vdash B$ follows by soundness of GLT with respect to Lindström-frames (Theorem 3.5.2). For the other direction, suppose that B is a closed formula and $\nVdash B$. By Theorem 3.5.2, there is a Lindström-model $\mathcal{M}$ and a world $a \in \mathcal{M}$ such that $\mathcal{M}, a \nVdash B$. By Lemma 3.7.5, we find that $\mathrm{d}(a) \notin [\![B]\!]$, whence $[\![B]\!] \neq \omega^2$ as required. ❑

3.7.7. LEMMA. *For every $\alpha, \beta < \omega^2$,*

$$[\![\bot^\alpha]\!] = [0, \alpha) \tag{3.6}$$

$$[\![\neg\bot^\alpha]\!] = [\alpha, \omega^2) \tag{3.7}$$

$$[\![\neg\bot^\alpha \wedge \bot^\beta]\!] = [\alpha, \beta) \tag{3.8}$$

Proof: We first show that $[\![\bot^\alpha]\!] = \{\beta \mid \beta < \alpha\}$. Note first that $[\![\bot^0]\!] = [\![\bot]\!] = \emptyset$. If $\alpha = \alpha' + 1$ for some α', we have

$$\begin{aligned} [\![\bot^{\alpha'+1}]\!] = [\![\triangle\bot^{\alpha'}]\!] &= \{\gamma \mid \forall\delta\ \gamma \prec \delta\ \ \delta \in [\![\bot^{\alpha'}]\!]\} \\ &= \{\gamma \mid \forall\delta\ \gamma \prec \delta\ \ \delta \in \{\beta \mid \beta < \alpha'\}\} \qquad \text{i.a.} \\ &= \{\gamma \mid \forall\delta\ \gamma \prec \delta\ \ \alpha' \prec \delta\} \\ &= \{\gamma \mid \gamma < \alpha' + 1\} \end{aligned}$$

Finally, if α is a limit ordinal $\lambda < \omega^2$, then it has the form form $w \cdot (n+1)$ for some n. We reason as follows:

$$\begin{aligned} [\![\bot^\lambda]\!] = [\![\Box^{n+1}\bot]\!] &= \{\gamma \mid \forall\delta\ \gamma \prec_\omega \delta\ \ \delta \in [\![\Box^n\bot]\!]\} \\ &= \{\gamma \mid \forall\delta\ \gamma \prec_\omega \delta\ \ \delta \in \{\beta \mid \beta < \omega \cdot n\}\} \qquad \text{i.a.} \\ &= \{\gamma \mid \forall\delta\ \gamma \prec_\omega \delta\ \ \delta < \omega \cdot n\} \\ &= \{\gamma \mid \gamma < \omega \cdot (n+1)\} \\ &= \{\gamma \mid \gamma < \lambda\} \end{aligned}$$

Thus (3.6) and (3.7) hold. For (3.8), note that

$$[\alpha, \beta) = [0, \beta) \cap [\alpha, \omega^2) = [\![\bot^\beta]\!] \cap [\![\neg\bot^\alpha]\!] = [\![\bot^\beta \wedge \neg\bot^\alpha]\!].$$ ❑

3.7.8. LEMMA. *Let A be closed. If $\nvdash A$, then $\vdash (\bot^{\alpha+1} \wedge \neg\bot^\alpha) \to \neg A$ for some $\alpha < \omega^2$.*

Proof: Let A be a closed with $\nvdash A$. By Lemma 3.7.6, $[\![A]\!] \neq \omega^2$, thus there is some $\alpha < \omega^2$ such that $\alpha \notin [\![A]\!]$. We show that $(\bot^{\alpha+1} \wedge \neg\bot^\alpha) \to \neg A$ is valid on the class of Lindström-frames, whence the desired result follows by Theorem 3.5.2. Let $\mathcal{M}$ be a Lindström-model with $\mathcal{M}, a \Vdash \bot^{\alpha+1} \wedge \neg\bot^\alpha$. Using Lemmas 3.7.5, 3.7.7, we see that $\mathrm{d}(a) \in [\![\bot^{\alpha+1} \wedge \neg\bot^\alpha]\!] = [\alpha, \alpha+1)$, whence $\mathrm{d}(a) = \alpha$. Since $\alpha \notin [\![A]\!]$, we now obtain $a \nVdash A$ by Lemma 3.7.5. ❑

We conclude with another proof of Theorem 3.7.1:

Proof: Let A be a closed formula. By Lemma 3.7.4, we have $[\![A]\!] = \bigcup_{i\in I}[\alpha_n, \beta_n)$ for some finite set I. Let $C := \bigvee_{i\in I}\left(\bot^{\alpha_n} \wedge \neg\bot^{\beta_n}\right)$, and note that C is in normal form. Using Lemma 3.7.7, it is clear that $[\![A]\!] = [\![C]\!]$. Using Definition 3.7.3, it is easy to check that then $[\![A \leftrightarrow C]\!] = \omega^2$, whence $\vdash A \leftrightarrow C$ by Lemma 3.7.12. ❑

3.7.3 Arithmetical completeness

We show that the closed fragment of GLT is arithmetically complete with respect to a wide class of pairs of provability predicates. For the rest of this section, let us fix *arithmetical formulas* $\triangle$ and $\Box$. As the notation suggests, we are interested in arithmetical realisations mapping the modalities $\triangle$ and $\Box$ to the arithmetical formulas $\triangle$ and $\Box$, respectively. We only consider arithmetical realisations for the closed fragment. Since arithmetical realisations (Definition 2.3.1) are distinguished only by the values they assign to propositional constants, it is clear that there is exactly one such realisation $*$.

For the proof of arithmetical completeness, we only need to assume that for every closed formula A:

1. If $\vdash A$, then $\mathsf{PA} \vdash \triangle A^*$
2. If A is an axiom of GLT, then $\mathsf{PA} \vdash A^*$
3. If $\mathfrak{N} \models \Box A^*$, then $\mathsf{PA} \vdash A^*$.

Since $\mathsf{PA} \nvdash \bot$, and $\mathsf{GLT} \vdash \triangle A \to \Box A$, it follows from (2) and (3) that for $\alpha < \omega^2$,

$$\mathfrak{N} \models (\bot^\alpha)^* \leftrightarrow \bot. \tag{3.9}$$

3.7.9. Remark. The arithmetical realisations considered here include the *fast* and *slow realisations* of Chapter 4, as well some of the *supremum adapter realisations* of Chapter 5.

3.7.10. Lemma. *For any closed formula A, if* $\mathsf{GLT} \vdash A$, *then* $\mathsf{PA} \vdash A^*$.

Proof: Immediate from requirements (1) and (2). ❑

We show first that the truth of arithmetical sentences that are translations of closed formulas is decidable.

3.7.11. Theorem. *The set $S := \{\varphi \mid \mathfrak{N} \models \varphi$, and φ is A^* for some closed $A\}$ is decidable.*

Proof: Let φ be an arithmetical sentence, and suppose that φ is A^*, where A is some closed formula. By Theorem 3.7.1 and Lemma 3.7.10, there is a sentence C in normal form, for which we have $\mathsf{PA} \vdash \varphi \leftrightarrow C^*$. In order to find out whether $\varphi \in S$, it therefore suffices to check whether $\mathfrak{N} \models C^*$. Denote by D the result of erasing all occurrences of $\Box$ and $\triangle$ in C, thus D is a boolean combination of $\bot$'s. By (3.9), it is clear that $\mathfrak{N} \models C^*$ if and only if D is a tautology. ❑

3.7.12. Theorem. *For any closed formula A, if $\mathsf{PA} \vdash A^*$, then $\vdash A$.*

Proof: We prove the claim by contraposition. Let A be closed with $\nvdash A$. By Lemma 3.7.8, there is some $\alpha < \omega^2$ with $\vdash (\bot^{\alpha+1} \wedge \neg\bot^{\alpha}) \to \neg A$. By Lemma 3.7.10, we thus have $\mathsf{PA} \vdash (\bot^{\alpha+1} \wedge \neg\bot^{\alpha})^* \to \neg A^*$. Given this, $\mathsf{PA} \vdash A^*$ would imply $\mathsf{PA} \vdash (\bot^{\alpha+1} \to \bot^{\alpha})^*$. Requirements (1) and (2) imply that Löb's rule for $\triangle$ is admissible in PA. Applying the latter to $\mathsf{PA} \vdash (\bot^{\alpha+1} \to \bot^{\alpha})^*$, we obtain $\mathsf{PA} \vdash (\bot^{\alpha})^*$. Since PA is sound, this contradicts (3.9). ❑

Chapter 4

Fast and slow provability

The provability concepts studied in this chapter differ from the ordinary one in terms of speed. Section 4.1 is about Parikh provability, which can be seen as an accelerated version of PA-provability. It was shown by Lindström that the joint provability logic of ordinary and Parikh provability is GLT. Section 4.2 studies the notion of slow provability introduced by Friedman, Rathjen, and Weiermann. In Section 4.3, we show that GLT is also the joint provability logic of slow and ordinary provability. Our proof of arithmetical completeness works for a wide class of pairs of provability predicates, including ordinary and Parikh provability.

4.1 Parikh provability

PA^* is the system obtained by adding to PA the following inference rule, known as *Parikh's rule*:

$$\frac{\Box\varphi}{\varphi},$$

where $\Box$ is the usual provability predicate of PA, and φ an arithmetical sentence. The system PA^* was first studied by Parikh ([Par71]). Parikh's rule is admissible in PA: if $\mathsf{PA} \vdash \Box\varphi$, then by soundness $\mathfrak{N} \models \Box\varphi$. But this means that there exists a PA-proof of φ, i.e. $\mathsf{PA} \vdash \varphi$. Hence it is clear that PA^* has exactly the same theorems as PA.

The above reasoning makes use of soundness and can therefore — as a consequence of Löb's Theorem — not be formalised in PA. Indeed, the equivalence of PA and PA^* cannot be established within PA. In particular, PA considers it possible that it itself is consistent while PA^* is inconsistent (Corollary 4.1.5).

While PA and PA^* have exactly the same theorems, some theorems have much shorter proofs in PA^* than in PA — we say that PA^* has *speed-up* over PA:

4.1.1. Theorem. ([Par71, Theorem 1.3]) Let g be provably recursive in PA. For all $i \geqslant 1$, there exists some n and a formula φ such that there is a PA–proof of $\Box^i\varphi$ of length n, but for no $j < i$ is there a proof of $\Box^j\varphi$ of length $\leqslant \mathsf{g}(n)$. ❑

PA's lack of knowledge concerning the equivalence of PA and PA* can thus be explained by the observation that the function converting PA*-proofs into ordinary proofs grows faster than any provably recursive function of PA.

4.1.1 Modal principles for Parikh provability

We denote by $\triangle_{\mathtt{p}}$ a natural formalisation of PA*-provability; as usual $\triangledown_{\mathtt{p}}\varphi$ is written as shorthand for $\neg\triangle_{\mathtt{p}}\neg\varphi$. Obtained by adding to PA a single sound inference rule, the theory PA* satisfies the criteria of being a reasonable theory, given in Section 2.3. In particular, the formula $\triangle_{\mathtt{p}}$ is Σ_1, and we have

1. $\mathsf{PA}^* \vdash \varphi \quad\Leftrightarrow\quad \mathrm{I}\Delta_0+\mathsf{exp} \vdash \triangle_{\mathtt{p}}\varphi$
2. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\triangle_{\mathtt{p}}(\varphi \to \psi) \to (\triangle_{\mathtt{p}}\varphi \to \triangle_{\mathtt{p}}\psi))$
3. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box_0\varphi \to \triangle_{\mathtt{p}}\varphi)$

Since PA* is reasonable, it follows from Theorem 2.3.2 that GL is the provability logic of $\triangle_{\mathtt{p}}$.

Let us now look at how $\triangle_{\mathtt{p}}$ interacts with the ordinary provability predicate $\Box$. While it is not known to PA that any Parikh proof can be converted into an ordinary proof, it is known — in fact even to $\mathrm{I}\Delta_0+\mathsf{exp}$ — that Parikh proofs can be converted into finite *iterations* of ordinary proofs. The following lemma was originally proven by Lindström ([Lin06, Lemma 1]), with PA in place of $\mathrm{I}\Delta_0+\mathsf{exp}$.

4.1.2. LEMMA. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\triangle_{\mathtt{p}}\varphi \leftrightarrow \Box^\omega\varphi)$

Proof: Reason in $\mathrm{I}\Delta_0+\mathsf{exp}$, fixing some φ. Assume first $\Box^\omega\varphi$ i.e. $\exists x\,\Box\Box^x\varphi$, and suppose that p is a PA-proof of $\Box^x\varphi$. Letting p' be the sequence resulting from p by appending x applications of Parikh's rule, it is clear that p' witnesses $\triangle_{\mathtt{p}}\varphi$.

For the other direction, suppose that $\langle\psi_0, \ldots, \psi_k\rangle$ is a Parikh proof of φ. Thus each ψ_i is either an axiom of PA, or obtained from the previous formulas in the sequence by either an application of modus ponens or of the Parikh rule. We show by using reflexive induction[1] that for all $i \leqslant k$, $\Box^{2i+1}\psi_i$. Fix i, and assume:

$$\forall j < i\,\Box_0\Box^{2j+1}\psi_j \tag{4.1}$$

Consider ψ_i, and use reasoning granted by Lemma 2.4.6. If ψ_i is an axiom, it is clear that $\Box\psi_i$, whence also $\Box^{2i+1}\psi_i$. If ψ_i is obtained by modus ponens, then by (4.1) we have $\Box_0\Box^{2j+1}\psi_j$ and $\Box_0\Box^{2j'+1}(\psi_j \to \psi_i)$ for some $j, j' < i$. Assuming w.l.o.g. $j' < j$, we also have $\Box_0\Box^{2j+1}(\psi_j \to \psi_i)$. Thus $\Box^{2j+2}\psi_j$ and $\Box^{2j+2}(\psi_j \to \psi_i)$, whence also $\Box^{2j+2}\psi_i$ and finally, since $j < i$, $\Box^{2i+i}\psi_i$. If ψ_i is

[1] Arguing in $\mathrm{I}\Sigma_1$, it is easy to show that for all i, $\Box^i\psi_i$. That this can be verified in $\mathrm{I}\Delta_0+\mathsf{exp}$ is less obvious. We thank F. Pakhomov for the neat idea of replacing i with $2i+1$, thus still allowing us to prove the claim by means of a simple reflexive induction.

obtained by Parikh's rule, then by (4.1) there is some $j < i$ with $\Box_0\Box^{2j+1}\Box\psi_i$. It follows that $\Box^{2j+3}\psi_i$. Since $j < i$, we have $2j + 3 \leqslant 2i + 1$, and so $\Box^{2i+1}\psi_i$ as required. ❑

We consider arithmetical realisations mapping the modalities $\triangle$ and $\Box$ to the provability predicates $\Box$ and $\triangle_{\mathtt{p}}$, respectively; let us call such realisations *fast realisations*. Lindström ([Lin06]) showed that GLT is arithmetically sound and complete with respect to fast realisations, i.e. that GLT is the joint provability logic of $\Box$ and $\triangle_{\mathtt{p}}$. Since GL is the provability logic of both $\Box$ and $\triangle_{\mathtt{p}}$, arithmetical soundness of GLT with respect to fast realisations is an immediate consequence of the following:

4.1.3. Lemma. *i.* $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\Box\varphi \to \triangle_{\mathtt{p}}\varphi)$

ii. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\triangle_{\mathtt{p}}\varphi \to \Box\triangle_{\mathtt{p}}\varphi)$

iii. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\triangle_{\mathtt{p}}\varphi \to \triangle_{\mathtt{p}}\Box\varphi)$

iv. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\triangle_{\mathtt{p}}\Box\varphi \to \triangle_{\mathtt{p}}\varphi)$

v. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\exists x\,\triangle_{\mathtt{p}}\Box^x\varphi \to \triangle_{\mathtt{p}}\varphi)$

Proof: (i) and (ii) are immediate from the definition of $\triangle_{\mathtt{p}}$, together with provable Σ_1-completeness of $\Box$. Clauses (iii) and (iv) follow from Lemma 4.1.2, since we have $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box^\omega\varphi \leftrightarrow \Box^\omega\Box\varphi$ by Lemma 2.4.8. Clause (v) follows from (i) together with the right to left direction of Lemma 4.1.2, ❑

Using Lemma 4.1.3, it is clear that fast realisations satisfy the conditions of Theorem 3.7.12; thus we obtain:

4.1.4. Proposition. *The closed fragment of* GLT *is arithmetically complete with respect to fast realisations.*

4.1.5. Corollary. $\triangledown_{\mathtt{p}}\top \trianglerighteq_{\neq} \Diamond\top$.

Proof: Since $\mathsf{GLT} \nvdash \Box\bot \to \triangle\bot$ by Lemma 3.3.4, we have $\mathsf{PA} \nvdash \triangle_{\mathtt{p}}\bot \to \Box\bot$ by Proposition 4.1.4, i.e. $\mathsf{PA} \nvdash \Diamond\top \to \triangledown_{\mathtt{p}}\top$. Since $\triangledown_{\mathtt{p}}\top$ is Π_1, $\Diamond\top \not\trianglerighteq \triangledown_{\mathtt{p}}\top$ follows by Theorem 2.1.12. ❑

4.1.6. Remark. ([Bek03, Appendix B]) Consider the system

$$\mathsf{PA}_\omega := \mathsf{PA} + \{\Diamond^n\top \mid n \in \omega\}.$$

It is clear that, verifiably in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, the consistency of PA_ω is equivalent to $\forall x\,\Diamond\Diamond^x\top$, i.e. to $\Diamond^\omega\top$. By Lemma 4.1.2, the latter is provably equivalent in

$\mathrm{I}\Delta_0+\mathsf{exp}$ to $\triangledown_{\mathsf{p}}\top$, i.e. to the consistency of PA^*. Thus PA^* and PA_ω have the same consistency strength, relative to $\mathrm{I}\Delta_0+\mathsf{exp}$.

It is shown in [Bek05, Corollary 2.34] that PA_ω is, verifiably in $\mathrm{I}\Delta_0+\mathsf{exp}$, mutually interpretable with the following theory

$$\mathsf{PA} + \mathrm{Rfn}(\mathsf{PA}) := \mathsf{PA} + \{\Box\varphi \to \varphi \mid \varphi \text{ is an } \mathcal{L}\text{- sentence}\}.$$

Hence the theories PA^*, PA_ω, and $\mathsf{PA} + \mathrm{Rfn}(\mathsf{PA})$ all have the same consistency strength, relative to $\mathrm{I}\Delta_0+\mathsf{exp}$. From the external point of view, on the other hand, we have that PA^* is equivalent to PA which does not interpret PA_ω or $\mathsf{PA} + \mathrm{Rfn}(\mathsf{PA})$, by the Second Incompleteness Theorem.

Thus we have here a natural example of intensionality: the theory PA^*, while extensionally equivalent to PA, has the same consistency strength as the much stronger theories PA_ω and $\mathsf{PA} + \mathrm{Rfn}(\mathsf{PA})$. Moreover, we have a natural example separating the notion of relative consistency from that of interpretability. While the consistency of PA^* implies (relative to $\mathrm{I}\Delta_0+\mathsf{exp}$) the consistency of PA_ω, PA^* fails to interpret PA_ω.

4.2 Slow provability

The notion of slow provability for PA was introduced by Friedman, Rathjen, and Weiermann ([FRW13]). They showed that the interpretability strength of the corresponding slow consistency statement lies strictly between $\top$ and the ordinary consistency statement $\Diamond\top$ of PA. The slow consistency statement is the first example of a sentence with this property that is natural: its existence does not rely on the Fixed Point Lemma.

The definition of slow provability makes use of a certain recursive function $\mathtt{F}_{\varepsilon_0}$ whose totality is not provable in PA. Consider the following theory:

$$\mathsf{PA}{\restriction}_{\mathtt{F}_{\varepsilon_0}} := \{\mathrm{I}\Sigma_n \mid \mathtt{F}_{\varepsilon_0}(n)\downarrow\}.$$

Since $\mathtt{F}_{\varepsilon_0}$ is recursive, it is clear that $\mathsf{PA}{\restriction}_{\mathtt{F}_{\varepsilon_0}}$ is recursively enumerable. Moreover, since $\mathtt{F}_{\varepsilon_0}$ is total, it is clear that $\mathsf{PA}{\restriction}_{\mathtt{F}_{\varepsilon_0}}$ and PA have exactly the same theorems. Arguing in PA, on the other hand, the totality of $\mathtt{F}_{\varepsilon_0}$ cannot be assumed, and thus $\mathsf{PA}{\restriction}_{\mathtt{F}_{\varepsilon_0}}$ might seem to be a weaker theory than PA. Indeed, as shown in [FRW13], PA considers it possible that it itself is inconsistent while $\mathsf{PA}{\restriction}_{\mathtt{F}_{\varepsilon_0}}$ is consistent.

4.2.1 The fast-growing hierarchy

The function $\mathtt{F}_{\varepsilon_0}$ needed to define slow provability belongs to a certain fast-growing hiearachy, also known as the *extended Grzegorczyk hieararchy.* Following [FRW13], what we mean by this is an ordinal-indexed family of recursive functions $\{\mathtt{F}_\alpha\}_{\alpha\leqslant\varepsilon_0}$. Each of the functions $\{\mathtt{F}_{\omega_n}\}_{n<\omega}$ is provably total in PA. The function $\mathtt{F}_{\varepsilon_0}$ results from diagonalising over the latter, and is not provably total in PA itself.

The idea of using diagonalisation to construct rapidly growing functions goes back to de Bois Reymond ([DBR75, p.365ff]). The functions $\{\mathtt{F}_\alpha\}_{\alpha<\omega}$ are closely related to a family of classes of functions known as the Grzegorczyk hierarchy ([Grz53]). Löb and Wainer ([LW70a], [LW70b]) extended the hierarchy into the transfinite. The exact version of the fast-growing hierarchy used in [FRW13] was introduced by Solovay and Ketonen ([KS81]).

The definition of the fast-growing hierarchy is tailored to ensure that for $\beta > \alpha$, the function $\mathtt{F}_\beta$ grows substantially faster than $\mathtt{F}_\alpha$. The function $\mathtt{F}_0$ is the successor function, $\mathtt{F}_{n+1}$ is defined by using iterations of $\mathtt{F}_n$, and $\mathtt{F}_\omega$ by diagonalising over $\{\mathtt{F}_n\}_{n<\omega}$:

$$\mathtt{F}_\omega(n) = \mathtt{F}_{n+1}(n).$$

The functions $\{\mathtt{F}_n\}_{n<\omega}$ are primitive recursive, and every primitive recursive function is majorised by some function in $\{\mathtt{F}_n\}_{n<\omega}$.

The definition of $\mathtt{F}_{n+1}$ can be straightforwardly adjusted to successor ordinals beyond ω. In order to generalise the limit case to a countable limit ordinal $\lambda > \omega$, we need to introduce the notion of a *fundamental sequence* for λ. By this, we mean a strictly monotone sequence $(\lambda[n])_{n<\omega}$ converging to λ from below, i.e. $\lambda[n] < \lambda[n+1] < \lambda$ for all $n < \omega$, and $\sup\{\lambda[n] \mid n < \omega\} = \lambda$. We consider the standard assignment of fundamental sequences to limit ordinals below ε_0.

4.2.1. Definition. Let $\varepsilon_0[n] := \omega_{n+1}$. For a limit ordinal $\lambda < \varepsilon_0$ with Cantor normal form $\lambda = \omega^{\alpha_0} + \omega^{\alpha_1} + \ldots + \omega^{\alpha_k}$, we define $\lambda[n]$ as follows:

$$\lambda[n] := \begin{cases} \omega^{\alpha_0} + \omega^{\alpha_1} + \ldots + \omega^{(\alpha_k - 1)} \cdot (n+1) & \text{if } \alpha_k \text{ is a successor ordinal} \\ \omega^{\alpha_0} + \omega^{\alpha_1} + \ldots + \omega^{\alpha_k[n]} & \text{if } \alpha_k \text{ is a limit ordinal} \end{cases}$$

Given a function $\mathtt{F} : \mathbb{N} \to \mathbb{N}$, we use exponential notation to denote repeated compositions of $\mathtt{F}$, thus $\mathtt{F}^0(n) = n$, and $\mathtt{F}^{k+1}(n) = \mathtt{F}(\mathtt{F}^k(n))$.

4.2.2. Definition. The fast-growing hierarchy $\{\mathtt{F}_\alpha\}_{\alpha\leqslant\varepsilon_0}$ of recursive functions is given by:

$$\begin{aligned} \mathtt{F}_0(n) &= n + 1 \\ \mathtt{F}_{\alpha+1}(n) &= \mathtt{F}_\alpha^{n+1}(n) \\ \mathtt{F}_\lambda(n) &= \mathtt{F}_{\lambda[n]}(n) \end{aligned}$$

The results in [KS81], together with those in [Par80], imply the following classification of the provably recursive functions of PA:

4.2.3. Theorem. *For $n > 0$, $\mathrm{I}\Sigma_n \vdash \mathtt{F}_\alpha{\downarrow}{\downarrow} \;\Leftrightarrow\; \alpha < \omega_n$.* ❑

In order to reason about the fast-growing hierarchy in $I\Delta_0+\mathsf{exp}$, we assume ordinals $< \varepsilon_0$ to be represented in $I\Delta_0+\mathsf{exp}$ by their Cantor normal forms. Since all functions involved in Definitions 4.2.1, 4.2.2 are elementary, it is clear that these definitions can be carried out in $I\Delta_0+\mathsf{exp}$.

The following essential property of the fast-growing hierarchy was originally stated in [KS81]; verifiability in $I\Delta_0+\mathsf{exp}$ is treated in [FRW13, Lemma 2.3].

4.2.4. LEMMA. $I\Delta_0+\mathsf{exp} \vdash \mathtt{F}_\alpha(x)\!\downarrow\, \wedge\, y < x \;\rightarrow\; \mathtt{F}_\alpha(y)\!\downarrow\, \wedge\, \mathtt{F}_\alpha(x) \geqslant \mathtt{F}_\alpha(y)$, *where* α *ranges over ordinals* $\leqslant \varepsilon_0$. ❑

The natural way of establishing properties of the functions $\{\mathtt{F}_\alpha\}_{\alpha\leqslant\varepsilon_0}$ is by transfinite induction on ordinals $\leqslant \varepsilon_0$. For an ordinal $\alpha \leqslant \varepsilon_0$ and $n < \omega$, we write TI_{Π_n}-α for the following schema:

$$\forall\beta < \alpha\ (\forall\,\gamma < \beta\,\varphi(\gamma) \rightarrow \varphi(\beta)) \rightarrow \forall\gamma < \alpha\,\varphi(\gamma), \tag{4.2}$$

where φ is a Π_n-formula, possibly with additional parameters. Since Π_n-truth is definable in $I\Delta_0+\mathsf{exp}$ (Section 2.1.2), we can consider TI_{Π_n}-α to be a single formula.

It follows from Gentzen's work in [Gen43] that PA proves TI_{Π_n}-α for all n and $\alpha < \varepsilon_0$, and that it does not prove TI_{Π_0}-ε_0. For a treatment of the amount of transfinite induction available in the fragments of PA, see [Som95].

4.2.5. REMARK. In order to show, in the presence of TI_{Π_n}-α, that all ordinals less than α have the property expressed by a Π_n-formula φ, it suffices to show:

$$\forall\beta < \alpha\ (\forall\,\gamma < \beta\,\varphi(\gamma) \rightarrow \varphi(\beta))\,.$$

It is not difficult to see that for every α, it is verifiable in $I\Sigma_1$ that

$$\forall\beta < \alpha\ (\forall\gamma < \beta\,\mathtt{F}_\gamma\!\downarrow\downarrow \rightarrow \mathtt{F}_\beta\!\downarrow\downarrow)\,.$$

It follows that whether $\mathtt{F}_\alpha$ is provably total in some extension T of $I\Sigma_1$ depends on the amount of transfinite induction available in T. Since $\mathtt{F}_\alpha\!\downarrow\downarrow$ is a Π_2-sentence, we have $I\Sigma_1 \vdash \mathrm{TI}_{\Pi_2}$-$\alpha \rightarrow \mathtt{F}_\alpha\!\downarrow\downarrow$. In fact, the above is verifiable in $I\Delta_0+\mathsf{exp}$: $I\Delta_0+\mathsf{exp} \vdash \forall\alpha \leqslant \varepsilon_0\,\Box_1\,(\mathrm{TI}_{\Pi_2}\text{-}\alpha \rightarrow \mathtt{F}_\alpha\!\downarrow\downarrow)$.

4.2.6. THEOREM ([SOM95, THEOREM 4.1]). *For* $0 < m \leqslant n$ *and* $\omega \leqslant \alpha < \varepsilon_0$,

$$I\Delta_0 + \mathrm{TI}_{\Pi_n}\text{-}\alpha \vdash \mathrm{TI}_{\Pi_m}\text{-}\beta \quad\Leftrightarrow\quad \beta < \omega_{n-m}^{\alpha^\omega}.$$ ❑

An inspection of the proof of Theorem 4.2.6 shows that it can be verified in $I\Delta_0+\mathsf{exp}$. In particular, noting that for $x \geqslant 1$, $\omega_{x+1} = \omega_{x-1}^{\omega^{\omega}}$ we have:

4.2.7. PROPOSITION. $I\Delta_0+\mathsf{exp} \vdash \forall x \geqslant 1\,\forall\alpha < \omega_{x+1}\,\Box_0(\mathrm{TI}_{\Pi_{x+1}}\text{-}\omega \rightarrow \mathrm{TI}_{\Pi_2}\text{-}\alpha)$ ❑

4.2.8. REMARK. For proving the following lemma, it is useful to note that

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \omega_{x+1}[z] = \omega_x^{z+1}.$$

To see that, we reason in $\mathrm{I}\Delta_0{+}\mathsf{exp}$. Since ordinals $< \varepsilon_0$ are represented by elementary formulas, we can use induction on x. For $x = 0$, we have by definition:

$$\omega_1[z] = \omega^1[z] = \omega^0 \cdot (z+1) = z+1 = \omega_0^{z+1}.$$

For $x > 0$, we have that $\omega_{x+1} = \omega^{\omega_x}$, where ω_x is a limit ordinal. This means that $\omega_{x+1}[z]$ is defined as $\omega^{\omega_x[z]}$. By the induction assumption, we have that $\omega_x[z] = \omega_{x-1}^{z+1}$. It follows that $\omega_{x+1}[z]$ is equivalent to $\omega^{\omega_{x-1}^{z+1}}$, i.e. to ω_x^{z+1}.

Clause (i) of the following lemma is a formalisation of the right to left direction of Theorem 4.2.3 in $\mathrm{I}\Delta_0{+}\mathsf{exp}$.

4.2.9. LEMMA. *i.* $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall x\, \forall \alpha < \omega_{x+1}\, \Box_{x+1} \mathtt{F}_\alpha{\downarrow}{\downarrow}$

ii. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall x\, \Box_{x+1} \mathtt{F}_{\varepsilon_0}(x){\downarrow}$.

Proof: Argue in $\mathrm{I}\Delta_0{+}\mathsf{exp}$. (i) Fix x, and let $\alpha < \omega_{x+1}$. By Remark 4.2.5, it suffices to show that $\Box_{x+1}\mathrm{TI}_{\Pi_2}\text{-}\alpha$. If $x = 0$, we have that $\alpha = z$ for some $z < \omega$. Note that $\mathrm{TI}_{\Pi_n}\text{-}z$ is the schema $\forall y < z\, (\forall u < y\, \varphi(u) \to \varphi(y)) \to \forall y < z\, \varphi(y)$. By using that $\mathsf{Q} \vdash y < \overline{n} \leftrightarrow y = 0 \vee \ldots \vee y = \overline{n-1}$, the above is easily seen to be verifiable in Q, whence clearly also in $\mathrm{I}\Sigma_1$. Assuming that $x \geqslant 1$, we obtain $\Box_1(\mathrm{TI}_{\Pi_{x+1}}\text{-}\omega \to \mathrm{TI}_{\Pi_2}\text{-}\alpha)$ from Proposition 4.2.7. Note that $\mathrm{TI}_{\Pi_{x+1}}\text{-}\omega$ is induction for Π_{x+1}-formulas, and equivalent to induction for Σ_{x+1}-formulas ([HP93, Theorem 2.4]). Thus it is clear that $\Box_{x+1}\mathrm{TI}_{\Pi_{x+1}}\text{-}\omega$, and therefore also $\Box_{x+1}\mathrm{TI}_{\Pi_2}\text{-}\alpha$.

(ii) Fix an x. By clause (i), $\Box_{x+1}\mathtt{F}_\alpha{\downarrow}{\downarrow}$ whenever $\alpha < \omega_{x+1}$. This implies that for all y, $\Box_{x+1}\mathtt{F}_{\omega_x^{y+1}}(y){\downarrow}$. By Remark 4.2.8 and Definition 4.2.2 we have

$$\mathtt{F}_{\omega_x^{y+1}}(y) = \mathtt{F}_{\omega_{x+1}[y]}(y) = \mathtt{F}_{\omega_{x+1}}(y).$$

Combining this with the above, we see that for all y, $\Box_{x+1}\ \mathtt{F}_{\omega_{x+1}}(y){\downarrow}$. Thus in particular also $\Box_{x+1}\ \mathtt{F}_{\omega_{x+1}}(x){\downarrow}$. Since $\mathtt{F}_{\varepsilon_0}(x) = \mathtt{F}_{\varepsilon_0[x]}(x) = \mathtt{F}_{\omega_{x+1}}(x)$ by Definitions 4.2.1, 4.2.2, this finishes the proof. ❑

Let us recall the system ACA_0 of second-order arithmetic. The axioms of ACA_0 include the axioms of Q, the induction axiom (formulated in the second-order language), as well as the arithmetical comprehension schema:

$$\exists X\, \forall n\, (n \in X \leftrightarrow \varphi(n)),$$

where φ is an $\mathcal{L}$-formula (possibly containing free second-order variables). For more information on ACA_0, the reader is referred to [Sim09, Section III.1].

4.2.10. Theorem ([Sim09, Theorem IX.1.5]). ACA_0 *is a conservative extension of* PA. ❑

The informal version of the following lemma is [FRW13, Corollary 3.8]; which in turn follows from [Som95, Theorem 5.25]. Verifiability in ACA_0 is treated in [Fre16, Theorem 2.15].

4.2.11. Lemma. *The following is verifiable in* ACA_0*: Let* $\mathfrak{M} \models \mathsf{PA}$ *be nonstandard, and let* $a \in \mathfrak{M}$ *be nonstandard. If* $\mathfrak{M} \models \mathtt{F}_{\varepsilon_0}(a+1)\downarrow$*, then for every* n*, there is a cut* $\mathfrak{J}_n$ *of* $\mathfrak{M}$ *with* $\mathfrak{J}_n \models \mathrm{I}\Sigma_{n+1}$ *and* $\mathtt{F}_{\varepsilon_0}(a) < \mathfrak{J}_n < \mathtt{F}_{\varepsilon_0}(a+1)$. ❑

4.2.2 Modal principles for slow provability

The slow provability predicate introduced by Friedman, Rathjen, and Weiermann ([FRW13]) is defined as $\exists x\,(\Box_x\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\downarrow)$. We study variants of slow provability obtained by letting

$$\triangle_{\langle k\rangle}\varphi := \exists x\,(\Box_{x+k}\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\downarrow)\,,$$

where k is a standard integer. For $k < 0$, $x+k$ is defined to be $x \dot{-} |k|$, and $x \dot{-} k$ is defined to be $x+|k|$, where $|k|$ is the absolute value of k, and $x \dot{-} y$ is defined as $\max\{0, x-y\}$. As usual, we write $\triangledown_{\langle k\rangle}$ for the dual of $\triangle_{\langle k\rangle}$, i.e. $\triangledown_{\langle k\rangle}\varphi$ is an abbreviation for $\neg\triangle_{\langle k\rangle}\neg\varphi$. The symbols $\triangle_{\mathtt{s}}$ and its dual $\triangledown_{\mathtt{s}}$ are used as collective names for $\triangle_{\langle k\rangle}$ and $\triangledown_{\langle k\rangle}$, for $k \in \mathbb{Z}$.

Each $\triangle_{\mathtt{s}}$ is a natural provability predicate for some theory

$$\mathsf{PA}\restriction_{\mathtt{F}_{\varepsilon_0}} := \{\mathrm{I}\Sigma_{n+k} \mid \mathtt{F}_{\varepsilon_0}(n)\downarrow\}\,,$$

where $k \in \mathbb{Z}$. We claim that each $\mathsf{PA}\restriction_{\mathtt{F}_{\varepsilon_0}}$ is a reasonable theory, as defined in Section 2.3. Since $\mathtt{F}_{\varepsilon_0}$ is total, it is clear that $\mathsf{PA}\restriction_{\mathtt{F}_{\varepsilon_0}}$ has the same theorems as PA, and is therefore Σ_1-sound. By Σ_1-completeness, we have $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \mathtt{F}_{\varepsilon_0}(n)\downarrow$ for every n, whence in particular $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \mathtt{F}_{\varepsilon_0}(0)\downarrow$. It follows from this, together with monotonicity of $\Box_x$, that $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box_0\varphi \to \triangle_{\mathtt{p}}\varphi)$.

It remains to show that $\mathsf{PA}\restriction_{\mathtt{F}_{\varepsilon_0}}$ is smooth and recursively enumerable. The latter is clear, for $\mathtt{F}_{\varepsilon_0}$ is recursive. For smoothness, we would need to show:

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall x < y\, \mathtt{F}_{\varepsilon_0}(x)\downarrow \to \exists b\,\forall x < y\,\exists c < b\; c\colon \mathtt{F}_{\varepsilon_0}(x)\downarrow, \tag{4.3}$$

where $c\colon \mathtt{F}_{\varepsilon_0}(x)\downarrow$ is written to mean that c is a witness for the Σ_1-formula $\mathtt{F}_{\varepsilon_0}(x)\downarrow$, i.e. that c codes a computation of $\mathtt{F}_{\varepsilon_0}(x)$. Recall that by Lemma 4.2.4,

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \mathtt{F}_{\varepsilon_0}(x)\downarrow \;\wedge\; y < x \;\to\; \mathtt{F}_{\varepsilon_0}(y)\downarrow \;\wedge\; \mathtt{F}_{\varepsilon_0}(x) \geqslant \mathtt{F}_{\varepsilon_0}(y).$$

Given a natural coding of computations, we can assume that if $\mathtt{F}_{\varepsilon_0}(y)\downarrow \geqslant \mathtt{F}_{\varepsilon_0}(x)\downarrow$ and $c : \mathtt{F}_{\varepsilon_0}(y)\downarrow$, then there exists some $c' \leqslant c$ with $c' : \mathtt{F}_{\varepsilon_0}(x)\downarrow$. In other words, that greater outputs require greater computations. In order to show (4.3), argue in $\mathrm{I}\Delta_0+\mathsf{exp}$ and assume $\forall x < y\, \mathtt{F}_{\varepsilon_0}(x)\downarrow$. Given the considerations above, we can always take for b the witness c with $c\colon \mathtt{F}_{\varepsilon_0}(y-1)\downarrow$.

We conclude that $\mathsf{PA}\restriction_{\mathtt{F}_{\varepsilon_0}}$ is a reasonable theory, and we have:

1. $\mathsf{PA}\restriction_{\mathtt{F}_{\varepsilon_0}} \vdash \varphi \;\;\Leftrightarrow\;\; \mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \triangle_{\mathtt{s}}\varphi$
2. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\triangle_{\mathtt{s}}(\varphi \to \psi) \to (\triangle_{\mathtt{s}}\varphi \to \triangle_{\mathtt{s}}\psi))$
3. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall\varphi\,(\Box_0\varphi \to \triangle_{\mathtt{s}}\varphi)$

In particular, it follows from Theorem 2.3.2 that GL is the provability logic of $\triangle_{\mathtt{s}}$.

Let us now look at how slow provability interacts with ordinary provability. The following theorem concerning the joint behaviour of the two is essentially the same as [FRW13, Theorem 4.1].

4.2.12. Theorem. *$\mathsf{PA} \vdash \forall\varphi\,(\Box\triangle_{\mathtt{s}}\varphi \to \Box\varphi)$.*

Proof: We argue by contraposition and show that $\mathsf{PA} \vdash \forall\varphi\,(\Diamond\varphi \to \Diamond\triangledown_{\mathtt{s}}\varphi)$. Using Theorem 4.2.10, it suffices to show $\mathsf{ACA}_0 \vdash \forall\varphi\,(\Diamond\varphi \to \Diamond\triangledown_{\mathtt{s}}\varphi)$. Fix some $k \in \mathbb{Z}$ and argue in ACA_0, assuming $\Diamond\varphi$, where φ is some arithmetical sentence. We shall show $\Diamond\triangledown_{\langle k\rangle}\varphi$.

Since $\Diamond\varphi$, i.e. PA together with φ is consistent, there is a model $\mathfrak{M}$ of PA with $\mathfrak{M} \models \varphi$. If $\mathfrak{M} \models \triangledown_{\langle k\rangle}\varphi$, we are done. So let us assume $\mathfrak{M} \models \triangle_{\langle k\rangle}\neg\varphi$, i.e.

$$\mathfrak{M} \models \Box_{a+k}\neg\varphi \wedge \mathtt{F}_{\varepsilon_0}(a)\!\downarrow$$

for some $a \in \mathfrak{M}$. Since $\mathfrak{M} \models \varphi$, it follows by essential reflexivity of PA that $a + k$ is nonstandard. Given that k is a standard integer, it is clear that a is nonstandard. Since $\mathfrak{M} \models \mathsf{PA}$, we can assume that a is the least x with $\mathfrak{M} \models \Box_{x+k}\neg\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\!\downarrow$. By Lemma 4.2.4, $\mathfrak{M} \models \mathtt{F}_{\varepsilon_0}(a)\!\downarrow \to \forall b < a\, \mathtt{F}_{\varepsilon_0}(b)\!\downarrow$, thus it must be that $\mathfrak{M} \models \Diamond_{(a+k)-1}\varphi$. Applying Lemma 4.2.11, we obtain for every n an initial segment $\mathfrak{J}_n$ of $\mathfrak{M}$ with $\mathfrak{J}_n \models \mathrm{I}\Sigma_{n+1}$ and $\mathtt{F}_{\varepsilon_0}(a-1) < \mathfrak{J}_n < \mathtt{F}_{\varepsilon_0}(a)$. Since $\Diamond_{(a+k)-1}\varphi$ is a Π_1-sentence and $\mathfrak{M} \models \Diamond_{(a+k)-1}\varphi$, we have $\mathfrak{J}_n \models \Diamond_{(a+k)-1}\varphi$ for all n. Since $\mathtt{F}_{\varepsilon_0}(a) \notin \mathfrak{J}_n$, it is clear that $\mathfrak{J}_n \models \mathtt{F}_{\varepsilon_0}(a)\!\uparrow$. Hence $\mathfrak{J}_n \models \triangledown_{\langle k\rangle}\varphi$. We now have $\Diamond_n\triangledown_{\langle k\rangle}\varphi$ for all n, i.e. $\Diamond\triangledown_{\langle k\rangle}\varphi$ as required. ❑

4.2.13. Corollary. *$\mathsf{PA} \vdash \forall\varphi\,(\exists x\,\Box\triangle_{\mathtt{s}}^{x}\varphi \to \Box\varphi)$.*

Proof: Reason in PA. Fix some φ and argue by induction on x. For $x = 0$ the claim is trivial. For the inductive step, note that

$$\Box\triangle_{\mathtt{s}}^{x+1}\varphi \to \Box\triangle_{\mathtt{s}}\triangle_{\mathtt{s}}^{x}\varphi \to \Box\triangle_{\mathtt{s}}^{x} \to \Box\varphi,$$

where the second implication is by Theorem 4.2.12. ❑

We consider arithmetical realisations mapping the modalities $\triangle$ and $\Box$ of GLT to $\triangle_{\mathtt{s}}$ and $\Box$ respectively; let us call such realisations *slow realisations*. Since $\triangle_{\mathtt{s}}$ and $\Box$ both obey the rules and axioms of GL, arithmetical soundness of GLT with respect to slow realisations is an immediate consequence of the following:

4.2.14. LEMMA. *i.* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Delta_{\mathsf{s}}\varphi \to \Box\varphi)$

ii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box\varphi \to \Delta_{\mathsf{s}}\Box\varphi)$

iii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box\varphi \to \Box\Delta_{\mathsf{s}}\varphi)$

iv. $\mathsf{PA} \vdash \forall\varphi\,(\Box\Delta_{\mathsf{s}}\varphi \to \Box\varphi)$

Proof: Clause (i) is immediate from the definition of Δ_{s}, while (ii) follows by provable Σ_1-completeness of $\Box_0$, since $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box_0\varphi \to \Delta_{\mathsf{s}}\varphi)$. (iv) is Theorem 4.2.12. For (iii), fix $k \in \mathbb{Z}$ and argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, supposing $\Box_x\varphi$ for some φ. Suppose first that $k \leqslant x$, and note that then $x \dot{-} k = x - k$. We have:

$$\Box_x\varphi \to \Box_0\Box_x\varphi \to \Box_{(x-k)+1}(\Box_x\varphi \wedge \mathtt{F}_{\varepsilon_0}(x-k)\downarrow) \to \Box_{z+1}(\Box_{z+k}\varphi \wedge \mathtt{F}_{\varepsilon_0}(z)\downarrow),$$

where the first implication is by Σ_1-completeness of $\Box_0$, the second by Lemma 4.2.9(ii), and the third by renaming $x-k$ to z. Thus we have $\Box\Delta_{\langle k\rangle}\varphi$ as required. It remains to consider the case $k > x$. We reason as follows:

$$\Box_x\varphi \to \Box_0\Box_x\varphi \to \Box_{x+1}(\Box_x\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\downarrow) \to \Box_{x+1}(\Box_{x+k}\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\downarrow),$$

where the second implication is by Lemma 4.2.9(ii), and the third by monotonicity. ❑

Given Lemma 4.2.14, it is clear that slow realisations satisfy the conditions of Theorem 3.7.12; thus we have:

4.2.15. THEOREM. *The closed fragment of* GLT *is arithmetically complete with respect to slow realisations.*

4.2.16. COROLLARY. $\Diamond\top \mathrel{\underset{\neq}{\trianglerighteq}} \nabla_{\mathsf{s}}\top \mathrel{\underset{\neq}{\trianglerighteq}} \top$.

Proof: By Lemma 4.2.14(i), $\mathsf{PA} \vdash \Diamond\top \to \nabla_{\mathsf{s}}\top$, whence clearly $\Diamond\top \trianglerighteq \nabla_{\mathsf{s}}\top \trianglerighteq \top$. It remains to show $\top \not\trianglerighteq \nabla_{\mathsf{s}}\top$ and $\nabla_{\mathsf{s}}\top \not\trianglerighteq \Diamond\top$. Since $\Diamond\top$ and $\nabla_{\mathsf{s}}\top$ are Π_1, by Theorem 2.1.12 it suffices to show that $\mathsf{PA} \nvdash \nabla_{\mathsf{s}}\top \to \Diamond\top$ and $\mathsf{PA} \nvdash \nabla_{\mathsf{s}}\top$. This follows by Theorem 4.2.15, for we have $\mathsf{GLT} \nvdash \nabla\top \to \Diamond\top$ and $\mathsf{GLT} \nvdash \nabla\top$ by Lemma 3.3.4. ❑

Transfinite iterations

We conclude with some observations concerning the behaviour of transfinite iterations of slow provability. The following lemma is an immediate consequence of Lemma 4.2.14(i) and Corollary 4.2.13:

4.2.17. LEMMA. $\mathsf{PA} \vdash \forall\varphi\,(\Delta_{\mathsf{s}}^{\omega}\varphi \to \Box\varphi)$ ❑

The following proposition is an analogue of Lemma 4.1.2. While it follows from Proposition 4.2.16 that ordinary proofs cannot always be converted into $\Delta_{\langle k\rangle}$-proofs, for $k \geqslant 2$ they can be converted into *finite iterations* of $\Delta_{\langle k\rangle}$-proofs.

4.2.18. PROPOSITION. *For* $k \geqslant 2$, $\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \Delta^{\omega}_{\langle k\rangle}\varphi)$.

Proof: Let $k \geqslant 2$. The direction from right to left is Lemma 4.2.17. For the other direction, we argue in PA and show by induction on x:

$$\forall\varphi\,(\Box_{x+(k-1)}\varphi \rightarrow \Delta_{\langle k\rangle}\Delta^{x}_{\langle k\rangle}\varphi)$$

By Σ_1-completeness and monotonicity, we have $\forall\varphi\,(\Box_m\varphi \rightarrow \Delta_{\langle k\rangle}\varphi)$ for all standard m. Using this, we have the claim for $x = 0$. For the inductive step, assume $\forall\psi\,(\Box_{x+(k-1)}\psi \rightarrow \Delta_{\langle k\rangle}\Delta^{x}_{\langle k\rangle}\psi)$. Fix some φ with $\Box_{x+k}\varphi$, and reason as follows:

$$\Box_{x+k}\varphi \rightarrow \Box_{(x+k)-1}(\Box_{x+k}\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\!\downarrow) \rightarrow \Box_{(x+k)-1}\Delta_{\langle k\rangle}\varphi \rightarrow \Delta_{\langle k\rangle}\Delta^{x}_{\langle k\rangle}\Delta_{\langle k\rangle}\varphi$$

where the first implication follows by monotonicity and Lemma 4.2.9(ii), noting that $k \geqslant 2$ implies $(x + k) - 1 > x$. The last implication is by the induction assumption. Thus we have $\Delta_{\langle k\rangle}\Delta^{x+1}_{\langle k\rangle}\varphi$ as required. ❑

As was shown independently by Pakhomov and Freund, transfinite iterations of $\Delta_{\langle 1\rangle}$ behave rather differently:

4.2.19. THEOREM ([HP16, THEOREM 10], [FRE16, SECTION 3]).

$$\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \Delta^{\varepsilon_0}_{\langle 1\rangle}\varphi).$$ ❑

It is shown in [HP16, Theorem 11] that the above holds for any $\Delta_{\langle k\rangle}$ with $k \leqslant 1$. Notwithstanding the contrast between Proposition 4.2.18 and Theorem 4.2.19, the provability predicates $\Delta_{\langle k\rangle}$ nevertheless share the same joint provability logic with ordinary provability, as will be shown in the next section.

We conclude this section by mentioning yet another slow provability predicate. Given a recursive function $\mathtt{f}$, write $\Delta_{\mathtt{f}}\varphi$ for the provability predicate defined as $\exists x\,(\Box_x\varphi \wedge \mathtt{f}(x)\!\downarrow)$.

4.2.20. THEOREM ([HP16, THEOREM 12]). *There is a recursive function* $\mathtt{r}$ *with*

$$\mathsf{PA} \vdash \Box\varphi \leftrightarrow \Delta_{\mathtt{r}}\Delta_{\mathtt{r}}\varphi.$$ ❑

The slow provability predicates $\Delta_{\mathtt{r}}$, $\Delta_{\langle k\rangle}$ for $k \geqslant 2$, and $\Delta_{\langle u\rangle}$ for $u \leqslant 1$, may thus be seen as the square root, ω-root, and ε_0-root of ordinary provability, respectively.

4.2.21. QUESTION. Do other roots of ordinary provability exist?

4.3 Arithmetical completeness

We prove completeness of GLT with respect to a wide class of arithmetical realisations, including fast and slow realisations.

4.3.1. Definition. Let T be a Σ_1-sound extension of $\mathrm{I}\Delta_0$+exp, and let $\triangle$ and $\Box$ be Σ_1-formulas. An arithmetical realisation mapping the modalities $\triangle$ and $\Box$ to the formulas $\triangle$ and $\Box$ is a a GLT-*realisation* if the latter satisfy the following conditions:

1. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box\varphi \quad \Rightarrow \quad T \vdash \varphi \quad \Rightarrow \quad \mathrm{I}\Delta_0+\mathsf{exp} \vdash \triangle\varphi$
2. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\Box_0\varphi \to \triangle\varphi)$
3. $\mathrm{I}\Delta_0$+exp verifies the axioms of GL for both $\triangle$ and $\Box$.
4. $\mathrm{I}\Delta_0$+exp verifies axioms (T1), (T2), and (T3) of GLT for $\triangle$ and $\Box$.
5. $T \vdash \exists x\,\Box\triangle^x\varphi \to \Box\varphi$

Condition (5) implies that T verifies axiom (T4) of GLT. It is thus clear that GLT is arithmetically sound with respect to GLT-realisations. We prove the converse:

4.3.2. Theorem. *If* $T \vdash A^*$ *for all* GLT*-realisations* **, then* GLT $\vdash A$.

4.3.3. Remark. From Lemma 4.1.3, it follows that fast realisations are GLT-realisations. Using Lemma 4.2.14 and Corollary 4.2.13, we see that slow realisations are GLT-realisations. In both cases, we take PA for the theory T.

4.3.4. Question. Consider arithmetical realisations mapping the modalities $\triangle$ and $\Box$ to the provability predicates $\triangle_{\mathtt{s}}$ and $\triangle_{\mathtt{p}}$ respectively. Are these realisations GLT-realisations? It is easy to see that such realisations satisfy (1)-(4) of Definition 4.3.1. Do they also satisfy condition (5)? Note that this would follow if the schema $\triangle_{\mathtt{p}}\triangle_{\mathtt{s}}\varphi \to \triangle_{\mathtt{p}}\varphi$ would be verifiable in PA.

Our proof of Theorem 4.3.2 uses the method of Theorem 2.3.2: we show that any Kripke model for GLT can be suitably embedded into T.

4.3.1 A Solovay function

Recall (Definition 3.6.1) the notion of an A-sound GLT-model. For the remainder of this section, fix some A-sound model $\mathcal{M} = \langle W, \prec, \prec_R, \Vdash\rangle$. We assume that $\mathcal{M}$ has a root, i.e. that there is a node $0 \in W$ such that $0 \prec a$ for every $0 \neq a \in W$. As usual, $a \preceq b$ is written as shorthand for $a \prec b \vee a = b$; similarly for $\preceq_R$.

We write $x : \Box\varphi$ to mean that x witnesses the Σ_1-sentence $\Box\varphi$ — intuitively, that x is the code of a $\Box$-proof —, and similarly for $\triangle$. We assume that every number witnesses the proof of a unique sentence — if any — , noting that this requirement can be satisfied for any reasonable arithmetisation of syntax in $\mathrm{I}\Delta_0$+exp.

4.3.5. DEFINITION. ($I\Delta_0 + \mathsf{exp}$) The function $h : \omega \to W$ is defined by:

$$h(0) = 0$$
$$h(x+1) = \begin{cases} b & \text{if } h(x) \prec_R b,\ b \text{ is reflexive, and } x : \Box L \neq b, \text{ else:} \\ c & \text{if } h(x) \prec c \text{ and } x : \Delta L \neq c \\ h(x) & \text{otherwise} \end{cases}$$

The formula $L \neq b$ (see (4.4) below) depends on the formula χ representing h in $I\Delta_0 + \mathsf{exp}$. The self-reference in the definition of h is handled by the Fixed Point Lemma. We note that the definition of h only relies on the gödelnumber of $L \neq b$, and the latter can be obtained from b and $\ulcorner \chi \urcorner$ by a function that is provably total in $I\Delta_0 + \mathsf{exp}$.

It follows from Theorem 2.1.4 — for example, by using that W is finite — that h is elementary and hence provably total in $I\Delta_0 + \mathsf{exp}$, with its defining equations also provable in $I\Delta_0 + \mathsf{exp}$. For $a \in W$, we write $L = a$ for the formula

$$\exists x\, h(x) = a \wedge \forall x\, h(x) \preceq a. \tag{4.4}$$

Using Theorem 2.1.1, we see that $L = a$ is provably equivalent in $I\Delta_0 + \mathsf{exp}$ to a $\Delta_0(\Sigma_1)$-formula. The formula $L = a$ states that a is the $\preceq$-largest element in the range of h. In view of the following lemma, we can think of $L = a$ as saying that a is the limit of h.

4.3.6. LEMMA. *i.* $I\Delta_0 + \mathsf{exp} \vdash y < x \to h(y) \preceq h(x)$

ii. $I\Delta_0 + \mathsf{exp} \vdash \exists! w\, L = w$

Proof: (i) is proven by internal induction on x, using that h is elementary. Since $a \prec_R b$ implies $a \prec b$, it is clear from the definition of h that $h(y) \preceq h(y+1)$. The inductive step thus follows from the transitivity of $\preceq$.

(ii) Since $\preceq$ is antisymmetric, uniqueness is immediate from the definition of L. For existence, we prove by external induction on the converse of $\prec$ that for all $a \in W$, $I\Delta_0 + \mathsf{exp} \vdash h(x) = a \to \exists w\, L = w$. This suffices, because $I\Delta_0 + \mathsf{exp}$ proves that $h(0) = 0$. By clause (i), we have

$$I\Delta_0 + \mathsf{exp} \vdash h(x) = a \to \forall y \geqslant x\, h(y) = a \vee \exists y > x\, a \prec h(y) \tag{4.5}$$

Argue in $I\Delta_0 + \mathsf{exp}$, assuming $h(x) = a$. If the first disjunct in (4.5) holds, we have, by using clause (i), $L = a$. And if the second disjunct holds, then $\exists w\, L = w$ follows by the induction assumption. ❑

4.3.7. LEMMA. $I\Delta_0 + \mathsf{exp} \vdash L = a \wedge a \prec b \to \nabla L = b.$

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, assuming $L = a$ and $a \prec b$. From the definition of L, we have that for all y, $h(y) \preceq a$, whence $h(y) \prec b$ by transitivity of $\preceq$. Thus $x : \triangle L \neq b$ would imply, by definition of h, that $h(x+1) = b$. Since $a \prec b$, this contradicts that $h(y) \preceq a$ for all y. ❑

4.3.8. LEMMA. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \wedge a \prec_R b \rightarrow \Diamond L = b$

Proof: Suppose that $a \prec_R b$. By properties of A-sound models, there exists some reflexive c with $a \prec_R c$ and $c \preceq b$. Arguing as in the proof of Lemma 4.3.7, and using that $a \prec_R c$ implies $a \prec c$, we have $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \rightarrow \Diamond L = c$. If $c = b$, we are done. If $c \prec b$, then $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = c \rightarrow \triangledown L = b$ by Lemma 4.3.7, whence

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Diamond L = c \rightarrow \Diamond \triangledown L = b$$

by using the axioms and rules of GL for $\Box$. Combining the above, we obtain $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \rightarrow \Diamond \triangledown L = b$, and so $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \rightarrow \Diamond L = b$ by $(\mathsf{T3})$. ❑

4.3.9. LEMMA. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \rightarrow \triangle\, a \preceq L$.

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$. Since $L = a$, we have that $h(x) = a$ for some x. Since $h(x) = a$ is elementary, we have have $\Box_0\, h(\dot{x}) = a$ by Theorem 2.1.1 and Σ_1-completeness of $\Box_0$. By definition of L, it is clear that $\Box_0\, (h(\dot{x}) = a \rightarrow a \preceq L)$, whence $\Box_0\, a \preceq L$. Since we have assumed that $\Box_0\varphi$ implies $\triangle\varphi$ for all φ, we obtain $\triangle\, a \preceq L$ as required. ❑

4.3.10. LEMMA. *If $a \neq 0$ is not reflexive, then* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \rightarrow \triangle\, a \prec L$.

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$. If the limit of h is some non-reflexive $a \neq 0$, then h must have moved to a due to some number witnessing $\triangle L \neq a$. By lemma 4.3.9, we also have $\triangle\, a \preceq L$. Combining these, we get $\triangle\, a \prec L$. ❑

4.3.11. LEMMA. *If $a \neq 0$, then* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \rightarrow \exists x\, \Box\, \triangle^x\, a \prec_R L$.

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, assuming $L = a \neq 0$. From Lemma 4.3.9 and $(\mathsf{T1})$ we have $\Box\, a \preceq L$. Since $a \neq 0$, we also have $\Box\, L \neq a$ by the definition of h and $(\mathsf{T1})$. Thus $\Box\, a \prec L$, whence by $(\mathsf{T3})$ and Lemma 2.4.6(iii), for all x,

$$\Box \triangle^x\, a \prec L. \tag{4.6}$$

It therefore suffices to show that if b is such that $a \prec b$ but $a \not\prec_R b$, then $\Box\triangle^x\, L \neq b$ for some x. Using that W is finite, we prove the claim by induction on the number of nodes between a and b. Assume that for all d with $a \prec d \prec b$, there is some x_d with $\Box\triangle^{x_d}\, L \neq d$. We let $x' := \max\{x_d \mid a \prec d \prec b\}$. In case there is no d with $a \prec d \prec b$, we define $x' := 0$. By Lemma 2.4.6(iii), we have $\Box\triangle^{x'} L \neq d$ for all d with $a \prec d \prec b$. We reason in $\Box\triangle^{x'}$:

Let c be such that $L = c$. From (4.6) we have $a \prec c$ and from the induction assumption that $c \not\prec b$. If $c = b$ then, since $a \not\prec_R b$, it must be that $\triangle L \neq b$ — otherwise h would have never moved to b. Suppose now that $c \neq b$. By Lemma 4.3.9 we have $\triangle c \preceq L$. Since $c \not\prec b$, this means that $\triangle L \neq b$ also in this case.

Back in $\mathrm{I}\Delta_0+\mathsf{exp}$, we now have $\Box\triangle^{x'}\triangle L \neq b$, i.e. $\Box\triangle^{x'+1} L \neq b$ as required. ❑

Since we have assumed $T \vdash \exists x\,\Box\triangle^{x}\varphi \rightarrow \Box\varphi$, the following is an immediate consequence of Lemma 4.3.11:

4.3.12. Lemma. *If $a \neq 0$, then $T \vdash L = a \rightarrow \Box\, a \prec_R L$.* ❑

4.3.13. Definition. Let $\mathcal{M} = \langle W, \prec, \prec_R, \Vdash\rangle$ be an A-sound model with root w. The model $\mathcal{M}_0$ is obtained by appending a root 0 to $\mathcal{M}$; the truth values of propositional formulas at 0 are defined to be exactly the same as at w. Apply Definition 4.3.5 to $\mathcal{M}_0$, and define the realisation * by letting:

$$p^* := \bigvee_{\mathcal{M}_0, a \Vdash p} L = a.$$

4.3.14. Lemma. *Let $\mathcal{M}$ and * be as in Definition 4.3.13. Then for every subformula B of A, and for every $a \neq 0$,*

$$\mathcal{M}, a \Vdash B \quad\Rightarrow\quad T \vdash L = a \rightarrow B^*. \tag{4.7}$$

Proof: We use Lemmas 4.3.7-4.3.12 to prove the claim by induction on the structure of B, simultaneously with

$$\mathcal{M}, a \Vdash \neg B \quad\Rightarrow\quad T \vdash L = a \rightarrow \neg B^*.$$

We treat the case of $\triangle B$; the other cases are exactly like in the proof of Lemma 2.3.8. Assume $a \Vdash \triangle B$, i.e. $b \Vdash B$ for all b with $a \prec b$. By the induction assumption, we have $T \vdash L = b \rightarrow B^*$ for all such b, and thus

$$T \vdash a \prec L \rightarrow B^*. \tag{4.8}$$

If a is reflexive, then $a \Vdash \triangle B$ additionally implies $a \Vdash B$, whence (4.8) can be strengthened to $T \vdash a \preceq L \rightarrow B^*$. Using the rules and axioms of GL for $\triangle$, the latter implies $T \vdash \triangle a \preceq L \rightarrow \triangle B^*$, and so $\vdash L = a \rightarrow \triangle B^*$ by Lemma 4.3.9. If a is not reflexive, we obtain from (4.8) by modal reasoning $T \vdash \triangle\, a \prec L \rightarrow \triangle B^*$, and finally $T \vdash \triangle\, a \prec L \rightarrow \triangle B^*$ by Lemma 4.3.10. ❑

4.3.15. Lemma. *i. $\mathfrak{N} \models L = 0$*

ii. for any $a \in W$, the sentence $L = a$ is consistent with T

Proof: (i) Let $a \neq 0$. We show $\mathfrak{N} \not\models \exists x\, h(x) = a$ by induction on the converse of $\prec$. Assume that the claim holds for all $\prec$-successors of a, and suppose $\mathfrak{N} \models \exists x\, h(x) = a$. By definition of h, this implies $\mathfrak{N} \models \Box L \neq a$, whence $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box L \neq a$ by Σ_1-completeness. We have assumed that $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box\varphi$ implies $T \vdash \varphi$ for all φ, and thus also

$$T \vdash L \neq a. \tag{4.9}$$

Since $\mathfrak{N} \models \exists x\, h(x) = a$, by Σ_1-completeness of T we also have $T \vdash \exists x\, h(x) = a$. This, together with (4.9), implies $T \vdash \exists x\, a \prec h(x)$, whence $\mathfrak{N} \models \exists x\, a \prec h(x)$ by Σ_1-soundness of T. The latter, however, contradicts our induction assumption. We conclude that for $a \neq 0$, $\mathfrak{N} \models \forall x\, h(x) \neq a$, whence $\mathfrak{N} \models \forall x\, h(x) = 0$, and thus $\mathfrak{N} \models L = 0$ as required.

(ii) Using Lemma 4.3.7, it follows from (i) that $\mathfrak{N} \models \nabla L = a$ for all $a \neq 0$. Now, $T \vdash L \neq a$ would imply $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \triangle L \neq a$, and the latter in turn would imply $\mathfrak{N} \models \triangle L \neq a$, a contradiction.

❑

We conclude with the proof of Theorem 4.3.2:

Proof: If $\mathsf{GLT} \nvdash A$, then by Lemma 3.6.5, there is some A-sound model $\mathcal{M}$ with $\mathcal{M}, w \nVdash A$ for some $w \in \mathcal{M}$. Let * be as in Definition 4.3.13. By Lemma 4.3.14, $T \vdash L = w \to \neg A^*$. Since T does not prove $L = w$ by Lemma 4.3.15, it therefore cannot prove A^* either. ❑

Chapter 5

Supremum adapters

The first part of this chapter develops the philosophy and technology needed for adding a supremum operator to the interpretability logic ILM of Peano Arithmetic (PA). It is well-known that any two theories extending PA have a supremum in the interpretability ordering. While provable in PA, this fact is not reflected in the theorems of the modal system ILM, due to limited expressive power. We would like to enrich the language of ILM by adding to it a new modality for the interpretability supremum. We explore different options for specifying the exact meaning of the new modality.

Our final proposal involves a unary operator: a certain nonstandard provability predicate that we call *supremum adapter*. We consider several variants of the supremum adapters. Some of them resemble the ordinary provability predicate, in that they behave according to the rules and axioms of GL. Others satisfy the principles of the modal system F, and are in this sense more similar to the Feferman provability predicate. We also study the joint behaviour of a supremum adapter together with ordinary provability. Finally, we make some observations concerning the transfinite behaviours of certain supremum adapters.

5.1 Introduction

Our overall goal in this chapter is to bring closer together two approaches to the study of interpretability. We recall (Definition 2.1.11) that a first-order theory T is said to *interpret* S ($T \rhd S$) if there is some structure-preserving translation from the language of S to the language of T, such that the translations of all theorems of S are provable in T.

Interpretability can be seen as *generalised provability*: T is required to prove everything that S proves modulo some well-behaved translation. As such, it allows us to compare theories talking about different objects, such as PA and Zermelo-Fraenkel set theory (ZF). In fact, the notion of interpretability is a natural candidate for giving a precise meaning to the intuitive idea of one theory being stronger than another one. Seen from a semantic perspective, an interpretation of S in T gives rise to a uniform way of constructing models of S inside models of T. Interpretations therefore give rise to relative consistency proofs.

The study of interpretability may roughly be divided into two traditions, briefly outlined below. A more comprehensive overview can be found in [Vis98].

The lattice of degrees Interpretability as a means of comparison naturally leads one to study the space of all theories ordered by this relation. A *degree* is a collection of all theories that are equally strong as a given theory, i.e. that all mutually interpret each other. We write $[S]$ for the degree of S.

Among the first results concerning interpretability degrees is a strengthening of Gödel's Second Incompleteness Theorem by Feferman. In [Fef60, Theorem 6.5] it is shown that not only is $\Diamond\top$ unprovable, it is also "uninterpretable": $[\mathsf{PA}] \not\rhd [\mathsf{PA} + \Diamond\top]$. Jeroslow ([Jer71, Theorem 3.1, 3.2]) showed that the degrees intermediate between $[\mathsf{PA}]$ and $[\mathsf{PA} + \Diamond\top]$ form a dense partial order. It follows from his work that the interpretability ordering is dense ([Šve78, p.798]). Montague ([Mon62, Theorem 1]) proved the existence of an infinite set of finitely axiomatised subtheories of PA, all of which are mutually incomparable with respect to the interpretability ordering.

A systematic study of interpretability degrees was undertaken independently by Švejdar and Lindström ([HP93, p.402]). Švejdar studied the structure $(\mathcal{V}, \rhd)$ of the degrees theories of the form $\mathsf{PA}+\varphi$, where φ is an arithmetical sentence. He proved, among other things, that this structure is a distributive lattice ([Šve78, Theorem 4.4, 4.7]). Lindström was concerned with the structure of the degrees of *all* r.e. extensions of PA in the same language ([Lin79], [Lin84]), but also showed that the latter is isomorphic to $(\mathcal{V}, \rhd)$ ([Lin79, p.348, Theorem 3]).

Our interest in provability logic makes it natural to focus on $(\mathcal{V}, \rhd)$. Note first that $[\mathsf{PA} + \top]$ is the minimum and $[\mathsf{PA} + \bot]$ the maximum element of this structure. It is easy to see that the infimum of $[\mathsf{PA} + \varphi]$ and $[\mathsf{PA} + \psi]$ in $(\mathcal{V}, \rhd)$ is $[\mathsf{PA} + (\varphi \vee \psi)]$.

5.1.1. Remark. Someone used to Boolean algebras has to think of the lattice $(\mathcal{V}, \rhd)$ as being *upside down*: the weakest theory $\top$ is the bottom element, and the strongest theory $\bot$ is the top element of $(\mathcal{V}, \rhd)$.

In contrast to what the above might lead one to expect, the supremum of $[\mathsf{PA} + \varphi]$ and $[\mathsf{PA} + \psi]$ is in general *not* $[\mathsf{PA} + (\varphi \wedge \psi)]$. Švejdar shows that the supremum of $[\mathsf{PA} + \varphi]$ and $[\mathsf{PA} + \psi]$ may taken to be $[\mathsf{PA} + \vartheta]$, where ϑ is such that $\mathsf{PA} \vdash \vartheta \leftrightarrow \forall x\,(\Diamond_x\vartheta \rightarrow \Diamond_x\varphi \wedge \Diamond_x\psi)$. The existence of ϑ is guaranteed by the Fixed Point Lemma.

Interpretability logic Interpretability, like provability, is a syntactical notion, and can therefore be formalised in the language of arithmetic. We may thus ask:

which statements concerning provability and interpretability (between its finite extensions) are provable in PA?

As in the case of provability alone, this question has an elegant modal logical answer. The system ILM of interpretability logic is obtained by adding to the provability logic GL a binary modality $\rhd$ for interpretability, together with axioms governing the behaviour of $\rhd$ as well as its interaction with $\Box$. It was proven independently by Berarducci ([Ber90]) and Shavrukov ([Sha88]) that the theorems of ILM are exactly the propositional schemata involving formalised provability and interpretability that are provable in PA. We say that ILM is the interpretability logic of PA.

A dissonance Given the two approaches to interpretability, it is natural to ask how they relate to each other. The starting point for the work presented in this chapter is the following question: does the modal logic ILM "know" that $(\mathcal{V}, \rhd)$ is a lattice?

The fact that the infimum in $(\mathcal{V}, \rhd)$ is given by disjunction is indeed reflected in the axioms of ILM. On the other hand, Švejdar's construction of a supremum employs a language much more complex than that available in propositional modal logic. Indeed, it is shown in Section 5.2.2 that interpretability suprema lie beyond the expressive power of ILM.

We note that the issue described above is by far not the only dissonance between the two approaches to interpretability. For example, in the framework of interpretability logic it is natural to consider statements containing nested occurrences of the modalities, such as $\Diamond A \to A \not\rhd \Diamond A$. At the same time, the latter cannot be meaningfully formulated in the context of the lattice of degrees.

We want to boost the expressive power of ILM by adding to it a new modality whose intended interpretation is a supremum operator in $(\mathcal{V}, \rhd)$. As we will see, Švejdar's construction is only one out of many ways of obtaining suprema in this structure. Since each of these can, in principle, be used to specify the meaning of the new modality, we are faced with a design choice.

Overview of this chapter Section 5.2 introduces the modal system ILM, and considers its relation to the interpretability degrees. Švejdar's construction, along with its dual discovered by Visser, are studied in Section 5.3. Section 5.4 introduces our favourite way of interpreting the supremum modality: a combination of conjunction with a certain nonstandard provability predicate, the so-called *supremum adapter*. Section 5.5 is concerned with the provability logic of supremum adapters, as well as their joint provability logic with ordinary provability. Section 5.6 contains some observations about transfinite iterations of supremum adapters.

5.2 Interpretability logic

This section introduces the modal system ILM of interpretability logic. We consider its relation to the lattice of degrees, and show that it fails to express interpretability suprema.

5.2.1 The system ILM

Denote by $\mathcal{L}_{\Box\rhd}$ the language $\mathcal{L}_{\Box}$ together with a binary modality $\rhd$. The operator $\rhd$ binds weaker than $\wedge$, but stronger than $\to$.

5.2.1. Definition. The system ILM contains all propositional tautologies in the language $\mathcal{L}_{\Box\rhd}$, the axiom schemata of of GL, and:

(J1) $\Box(A \to B) \to A \rhd B$

(J2) $(A \rhd B) \wedge (B \rhd C) \to (A \rhd C)$

(J3) $(A \rhd C) \wedge (B \rhd C) \to (A \vee B) \rhd C$

(J4) $A \rhd B \to (\Diamond A \to \Diamond B)$

(J5) $\Diamond A \rhd A$

(M) $A \rhd B \to (A \wedge \Box C) \rhd (B \wedge \Box C)$

The rules of ILM are modus ponens and necessitation for $\Box$.

5.2.2. Lemma. *i.* ILM $\vdash \Box A \leftrightarrow \neg A \rhd \bot$

ii. ILM $\vdash A \rhd A \wedge \Box\neg A$

iii. ILM $\vdash A \rhd B \;\to\; A \;\rhd\; B \wedge \Box\neg A$

Proof: (i) The left to right direction follows by (J1), since $\Box A \to \Box(\neg A \to \bot)$. The other direction follows from (J4), since K $\vdash \Diamond\bot \to \bot$.

(ii) By Löb's axiom, $\Diamond A \to \Diamond(A \wedge \Box\neg A)$. Using neccesitation and (J1), it follows that $\Diamond A \rhd \Diamond(A \wedge \Box\neg A)$. With (J1) we have $A \wedge \Diamond A \rhd \Diamond A$, whence by (J2), $A \wedge \Diamond A \rhd \Diamond(A \wedge \Box\neg A)$. Using (J5), the latter implies $A \wedge \Diamond A \rhd A \wedge \Box\neg A$. Since $A \wedge \Box\neg A \rhd A \wedge \Box\neg A$, it follows by (J3) that $(A \wedge \Box\neg A) \vee (A \wedge \Diamond A) \rhd A \wedge \Box\neg A$. Given (J1), it is clear that $A \rhd (A \wedge \Box\neg A) \vee (A \wedge \Diamond A)$, thus $A \rhd A \wedge \Box\neg A$ follows by one last application of (J3).

(iii) Follows from (M), (J2), and clause (ii). ❑

5.2.3. Definition. An ILM*–frame* is a tuple $\langle W, R, S\rangle$, with $W \neq \emptyset$, R a binary relation on W, and S a set of binary relations indexed by the elements of W, and such that $S_a \subseteq \{(b, c) \mid a\ R\ b, a\ R\ c\}$. Furthermore, *i.* R is transitive and converse well-founded

ii. S_a is transitive

iii. $a\,R\,b \Rightarrow b\,S_a\,b$

iv. $a\,R\,b,\, b\,R\,c \Rightarrow b\,S_a\,c$

v. $S_a R \subseteq R$

An ILM–*model* is a quadruple $\langle W, R, S, \Vdash\rangle$, where $\langle W, R, S\rangle$ is an ILM–frame, and $\Vdash$ a valuation on $\langle W, R, S\rangle$ satisfying the usual clauses, with R as the accessibility relation for $\Box$, and

$$a \Vdash A \rhd B \quad \text{if for all } b \text{ with } a\,R\,b, \text{ there is some } c \text{ with } b\,S_a\,c \Vdash B.$$

5.2.4. Theorem ([dJV90]). *ILM $\vdash A \Leftrightarrow \mathcal{F} \Vdash A$ for every ILM-frame $\mathcal{F}$.* ❑

We consider arithmetical realisations mapping the modality $\rhd$ to an arithmetical formula expressing in a natural way interpretability between theories of the form $\mathsf{PA} + \varphi$, where φ is an arithmetical sentence. The modality $\Box$ is mapped to the ordinary provability predicate $\Box$ of PA.

5.2.5. Theorem ([Sha88], [Ber90]). *For all $A \in \mathcal{L}_{\Box\rhd}$,*

$$\mathsf{ILM} \vdash A \Leftrightarrow \textit{for all arithmetical realisations } {}^*, \mathsf{PA} \vdash A^*.$$ ❑

It follows from Visser's results in ([Vis91]) that the left to right direction, i.e. arithmetical soundness of ILM with respect to PA, is already verifiable in $\mathrm{I}\Delta_0$+exp. We shall thus freely refer to the axioms of ILM when reasoning about formal interpretability in theories containing $\mathrm{I}\Delta_0$+exp.

5.2.2 ILM and the lattice of degrees

Recall the structure $(\mathcal{V}, \rhd)$ of the degrees of finite extensions of PA, ordered by the relation of interpretability. As mentioned in Section 5.1, $(\mathcal{V}, \rhd)$ is a lattice. We now ask whether this is visible from the perspective of ILM.

The fact that $(\mathcal{V}, \rhd)$ is a lower semilattice is indeed reflected in the axioms of ILM. Axioms (J1) and (J2) imply that the ordering given by $\rhd$ is reflexive and transitive, and furthermore that $\varphi \rhd \varphi \vee \psi$ and $\psi \rhd \varphi \vee \psi$, i.e. that $[\varphi \vee \psi]$ is a lower bound of $[\varphi]$ and $[\psi]$ in $(\mathcal{V}, \rhd)$. Axiom (J3) states that it is in fact the *greatest* lower bound, i.e. the infimum of $[\varphi]$ and $[\psi]$. The following theorem implies that $[\varphi \wedge \psi]$ is not, in general, the supremum of $[\varphi]$ and $[\psi]$.

5.2.6. Theorem ([Ore61, Theorem 2.4]). *There is a sentence ρ such that $\top \rhd \rho$ and $\top \rhd \neg\rho$.* ❑

A sentence ρ as in Theorem 5.2.6 is called an *Orey sentence*. We note that an Orey sentence and its negation are both in $[\top]$, and hence also their supremum is an element of $[\top]$, which is clearly not the case for the sentence $\rho \wedge \neg\rho$.

5.2.7. EXAMPLE. The Gödel-sentence of the Feferman provability predicate $\triangle_{\mathtt{f}}$ is, verifiably in PA, an Orey sentence. This can be shown by a simple modal argument found by Visser ([Vis89, p. 177]). A crucial ingredient is the following result by Feferman ([Fef60, Theorem 6.2]), according to which an analogue of (J5) holds with Feferman-consistency instead of ordinary consistency:

$$\mathsf{PA} \vdash \triangledown_{\mathtt{f}}\varphi \rhd \varphi \tag{5.1}$$

Verifiability in PA follows from [Vis91, Section 6]. By the Fixed Point Lemma, let γ be a Gödel-sentence for $\triangle_{\mathtt{f}}$:

$$\mathsf{PA} \vdash \gamma \leftrightarrow \neg\triangle_{\mathtt{f}}\gamma$$

Since $\mathsf{PA} \vdash \gamma \to \neg\triangle_{\mathtt{f}}\gamma$, i.e. $\mathsf{PA} \vdash \gamma \to \triangledown_{\mathtt{f}}\neg\gamma$, by (J1) also $\mathsf{PA} \vdash \gamma \rhd \triangledown_{\mathtt{f}}\neg\gamma$, whence by (5.1) and (J2), $\mathsf{PA} \vdash \gamma \rhd \neg\gamma$. Since $\mathsf{PA} \vdash \neg\gamma \rhd \neg\gamma$, it follows by (J3) that $\mathsf{PA} \vdash \top \rhd \gamma$. It remains to show $\mathsf{PA} \vdash \top \rhd \neg\gamma$. Since $\mathsf{PA} \vdash \triangledown_{\mathtt{f}}\top$ and the modal system K is sound with respect to $\triangle_{\mathtt{f}}$, we have that $\mathsf{PA} \vdash \triangle_{\mathtt{f}}\gamma \to \triangledown_{\mathtt{f}}\gamma$. Combining the above with $\mathsf{PA} \vdash \neg\gamma \to \triangle_{\mathtt{f}}\gamma$ and using (J1), we obtain $\mathsf{PA} \vdash \neg\gamma \rhd \triangledown_{\mathtt{f}}\gamma$. With (5.1), this implies $\mathsf{PA} \vdash \neg\gamma \rhd \gamma$. Thus $\mathsf{PA} \vdash \top \rhd \neg\gamma$ follows by (J3) as above.

The existence of Orey sentences rules out simple conjunctions as interpretability suprema in ILM. Of course, this alone does not prove interpretability suprema to lie beyond the expressive power of ILM. There could be some $\mathcal{L}_{\Box\rhd}$-formula $S(p, q)$ such that for any propositional letter r,

$$\mathsf{ILM} \vdash (r \rhd p) \wedge (r \rhd q) \quad \leftrightarrow \quad r \rhd S(p, q).$$

We show that such a formula S cannot exist already when the above is only required to hold for $r = \top$.

5.2.8. THEOREM. *There is no $\mathcal{L}_{\Box\rhd}$-formula S, for which*

$$\mathsf{ILM} \vdash (\top \rhd p) \wedge (\top \rhd q) \quad \leftrightarrow \quad \top \rhd S(p, q). \tag{5.2}$$

Proof: Suppose for a contradiction that S is as in (5.2). Note first that

$$\mathsf{ILM} \vdash \top \rhd S(p, q) \quad \leftrightarrow \quad \top \rhd S(p, q) \wedge \Box\bot. \tag{5.3}$$

The direction from left to right is Lemma 5.2.2(iii); the other direction follows by using (J1) and (J2). Combining (5.2) and (5.3), we obtain

$$\mathsf{ILM} \vdash (\top \rhd p) \wedge (\top \rhd q) \quad \leftrightarrow \quad \top \rhd S(p, q) \wedge \Box\bot. \tag{5.4}$$

We claim that there is a *purely propositional* formula S' such that

$$\mathsf{ILM} \vdash \Box\bot \to (S(p,q) \leftrightarrow S'(p,q)) \tag{5.5}$$

By induction on the structure of S, it can be shown that such a formula S' exists for *any* $\mathcal{L}_{\Box\rhd}$-formula S. The base case and the boolean cases are straightforward. It remains to consider the cases that $S = \Box C(p,q)$ and $S = C(p,q) \rhd D(p,q)$. Using Theorem 5.2.4, it is easy to see that in both cases we can take $S' = \top$. Combining (5.4) and (5.5), we obtain

$$\mathsf{ILM} \vdash (\top \rhd p) \wedge (\top \rhd q) \quad \leftrightarrow \quad \top \rhd S'(p,q) \wedge \Box\bot. \tag{5.6}$$

Substituting $\top$ for p and q in (5.6), we obtain $\mathsf{ILM} \vdash \top \rhd S'(\top,\top) \wedge \Box\bot$, thus it must be that $S'(\top,\top) = \top$. Substituting $\bot$ for one of p or q yields

$$\mathsf{ILM} \vdash \top \rhd \bot \quad \leftrightarrow \quad \top \rhd S'(p,q) \wedge \Box\bot,$$

whence it must be that $S'(p,q) = \bot$ if either $p = \bot$ or $q = \bot$. Hence $S'(p,q)$ is propositionally equivalent to $p \wedge q$, and we have

$$\mathsf{ILM} \vdash (\top \rhd p) \wedge (\top \rhd q) \quad \leftrightarrow \quad \top \rhd (p \wedge q) \wedge \Box\bot. \tag{5.7}$$

Consider an arithmetical realisation $*$ with $p^* = \rho$ and $q^* = \neg\rho$, where ρ is, verifiably in PA, an Orey sentence (see Example 5.2.7). By Theorem 5.2.5 we have from (5.7) that

$$\mathsf{PA} \vdash (\top \rhd \rho) \wedge (\top \rhd \neg\rho) \quad \leftrightarrow \quad \top \rhd (\rho \wedge \neg\rho) \wedge \Box\bot, \tag{5.8}$$

i.e. $\mathsf{PA} \vdash (\top \rhd \rho) \wedge (\top \rhd \neg\rho) \quad \leftrightarrow \quad \top \rhd \bot$. Since $\mathsf{PA} \vdash (\top \rhd \rho) \wedge (\top \rhd \neg\rho)$ but $\mathsf{PA} \nvdash \top \rhd \bot$, this is a contradiction. ❑

5.3 Binary suprema

Given Theorem 5.2.8, we would like to enrich ILM by adding to it a new modality for interpretability suprema. The methodological points related to this goal are discussed in Subection 5.3.1. Most importantly, we define the notion of a supremum implementation. The rest of this section studies two natural examples of supremum implementations due to Švejdar and Visser, respectively. The purpose of this section is mainly methodological and genealogical. We have therefore decided to omit proofs of most results, referring the interested reader to [HV16].

5.3.1 Methodological considerations

We would like to enrich ILM by adding to it a new binary modality $\oslash$, together with the following:

(Sup) $(C \rhd A) \wedge (C \rhd B) \leftrightarrow (C \rhd A \oslash B)$.

Denote by ILMS the system ILM together with axiom Sup.

5.3.1. Definition. An arithmetical formula σ with two free variables is a *supremum implementation* if for all sentences φ, ψ, χ of the language of arithmetic,

$$\mathsf{PA} \vdash (\chi \rhd \varphi) \wedge (\chi \rhd \psi) \leftrightarrow \chi \rhd \sigma(\varphi, \psi).$$

Given a supremum implementation σ, arithmetical realisations for ILM can be extended to the new language by mapping $\oslash$ to a supremum implementation:

$$(\varphi \oslash \psi)^* := \sigma(\ulcorner \varphi^* \urcorner, \ulcorner \psi^* \urcorner).$$

This guarantees that ILMS is arithmetically sound with respect to a given σ.

For arithmetical completeness, we would like the theorems of ILMS to include *all* propositional schemata involving σ that are known to PA. However, PA's knowledge about a formula σ is far from determined by what is required in Definition 5.3.1. For example, while for any such σ it is clear that $\sigma(\varphi, \psi)$ and $\sigma(\psi, \varphi)$ are, verifiably in PA, mutually interpretable, there is no a priori reason why

$$\sigma(\varphi, \psi) \leftrightarrow \sigma(\psi, \varphi) \tag{5.9}$$

should be provable in PA, or even true. An example of a supremum adapter for which (5.9) fails can be found in [HV16, Appendix A.2]. On the other hand, the supremum implementations we encounter below all satisfy (5.9).

Whether the axiom $A \oslash B \leftrightarrow B \oslash A$ should be added to ILMS therefore depends on which supremum implementation(s) we have in mind. In contrast to formalised provability and interpretability, there is no strong intuition as to what constitutes a *natural* supremum implementation. Thus our choice will depend on practical and esthetical criteria. For example, we prefer implementations that allow for a nice Kripke semantics.

5.3.2 Švejdar's and Visser's suprema

It follows from Theorem 2.1.12 that the supremum of $[\mathsf{PA} + \varphi]$ and $[\mathsf{PA} + \psi]$ in $(\mathcal{D}, \rhd)$, the structure of degrees of all r.e. extensions of PA in the same language, is the following infinite theory:

$$S := \mathsf{PA} + \{\Diamond_n \varphi \wedge \Diamond_n \psi \mid n \in \omega\}.$$

Švejdar's construction can be seen as a way of compressing the information contained in S into a single sentence.

5.3.2. THEOREM ([ŠVE78, THEOREM 4.4]). *By the Fixed Point Lemma, let ϑ be such that*

$$\mathsf{PA} \vdash \vartheta \leftrightarrow \forall x\, (\Diamond_x \vartheta \to \Diamond_x \varphi \wedge \Diamond_x \psi).$$

Then ϑ is (verifiably in $\mathrm{I}\Delta_0$+exp) a supremum of φ and ψ in $(\mathcal{V}, \rhd)$. ❑

The formula ϑ is Π_2. A dual construction, yielding a Σ_2-supremum, was discovered by Visser.

5.3.3. THEOREM. *By the Fixed Point Lemma, let ϑ be such that*

$$\mathsf{PA} \vdash \vartheta \leftrightarrow \exists x\, (\Box_x \neg\vartheta \wedge (\Diamond_x \varphi \wedge \Diamond_x \psi)).$$

Then ϑ is, verifiably in $\mathrm{I}\Delta_0$+exp, a supremum of φ and ψ in $(\mathcal{V}, \rhd)$.

Proof: Argue in $\mathrm{I}\Delta_0$+exp, and let ϑ be as above. We show:

i. $\vartheta \rhd \varphi \;\wedge\; \vartheta \rhd \psi$

ii. $(\chi \rhd \varphi) \wedge (\chi \rhd \psi) \;\to\; \chi \rhd \vartheta$

(i) By Theorem 2.1.12, it suffices to show $\forall x\, \Box(\vartheta \to \Diamond_x \varphi \wedge \Diamond_x \psi)$. We fix some x and argue in $\Box$:

> If ϑ, there is some z with $\Box_z \neg\vartheta$, $\Diamond_z \varphi$, and $\Diamond_z \psi$. By essential reflexivity, we have $\Diamond_x \vartheta$, thus z must be greater than x. By monotonicity, it thus follows that $\Diamond_x \varphi$ and $\Diamond_x \psi$ as required.

(ii) Assume $\chi \rhd \varphi$ and $\chi \rhd \psi$. With Theorem 2.1.12, we see that the latter imply

$$\forall x\, \Box(\chi \to \Diamond_x \varphi \wedge \Diamond_x \psi). \tag{5.10}$$

By properties of ϑ, we have $\Box(\neg\vartheta \leftrightarrow \forall x\; ((\Diamond_x \varphi \wedge \Diamond_x \psi) \to \Diamond_x \vartheta))$, whence it follows from (5.10) that

$$\forall x\, \Box(\chi \wedge \neg\vartheta \to \Diamond_x \vartheta),$$

and thus $\chi \wedge \neg\vartheta \rhd \vartheta$ by Theorem 2.1.12. Since clearly $\chi \wedge \vartheta \rhd \vartheta$, it follows by principle (J3) of ILM that $(\chi \wedge \vartheta) \vee (\chi \wedge \neg\vartheta) \rhd \vartheta$, i.e. $\chi \rhd \vartheta$. ❑

5.3.4. REMARK. Strictly speaking, theorems 5.3.2 and 5.3.3 do not yet provide us with supremum implementations in the sense of Definition 5.3.1. Corresponding to Theorem 5.3.2, we would like to have a formula σ with two free variables, such that for any sentences φ and ψ,

$$\mathsf{PA} \vdash \sigma(\varphi, \psi) \leftrightarrow \forall x\, (\Diamond_x \sigma(\varphi, \psi) \to \Diamond_x \varphi \wedge \Diamond_x \psi),$$

and similarly for Theorem 5.3.3. By using the Fixed Point Lemma with parameters, it is however clear that formulas with the required properties exist.

5.3.3 Some features of Švejdar's and Visser's suprema

Svejdar's and Visser's supremum implementations are obtained as fixed points of the following formulas (the capital Y indicates that we are interested in fixed points with respect to this variable):

$$\forall x\,(\Diamond_x Y \to (\Diamond_x\varphi \wedge \Diamond_x\psi))$$
$$\exists x\,(\Box_x\neg Y \wedge (\Diamond_x\varphi \wedge \Diamond_x\psi))$$

The following theorems state the existence of unique explicit fixed points for these fixed point equations. Theorem 5.3.5 is due to V. Yu. Shavrukov.

5.3.5. THEOREM ([HV16, THEOREM 7]). *Let $\varepsilon(\varphi,\psi)$ be the formula*

$$\forall x\;(\Box_x\neg\varphi \vee \Box_x\neg\psi \to \Box_x\exists y < x\,(\Box_y\neg\varphi \vee \Box_y\neg\psi))\,,$$

where φ and ψ are regarded as internal variables ranging over $\mathcal{L}$-sentences.

i. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \varepsilon(\varphi,\psi) \leftrightarrow \forall x\;(\Diamond_x\,\varepsilon(\varphi,\psi) \to \Diamond_x\varphi \wedge \Diamond_x\psi)$

ii. If $\mathsf{PA} \vdash \vartheta \leftrightarrow \forall x\,(\Diamond_x\vartheta \to \Diamond_x\varphi \wedge \Diamond_x\psi)$*, then* $\mathsf{PA} \vdash \vartheta \leftrightarrow \varepsilon(\varphi,\psi)$. ❑

5.3.6. THEOREM ([HV16, THEOREM 8]). *Let $\varepsilon(\varphi,\psi)$ be the formula*

$$\exists x\;(\Box_x(\Box_x\neg\varphi \vee \Box_x\neg\psi) \wedge (\Diamond_x\varphi \wedge \Diamond_x\psi))$$

where φ and ψ are regarded as internal variables ranging over $\mathcal{L}$-sentences.

i. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \varepsilon(\varphi,\psi) \leftrightarrow \exists x\;(\Box_x\,\neg\varepsilon(\varphi,\psi) \wedge (\Diamond_x\varphi \wedge \Diamond_x\psi))$

ii. If $\mathsf{PA} \vdash \vartheta \leftrightarrow \exists x\,(\Box_x\neg\vartheta \wedge (\Diamond_x\varphi \wedge \Diamond_x\psi))$*, then* $\mathsf{PA} \vdash \vartheta \leftrightarrow \varepsilon(\varphi,\psi)$. ❑

From the point of view of developing a modal logical treatment of the supremum adapters, it is useful to have the following:

5.3.7. THEOREM ([HV16, THEOREM 9].). *Let ε be as in Theorem 5.3.5 or as in 5.3.6. Then*

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box(\varphi \leftrightarrow \varphi') \wedge \Box(\psi \leftrightarrow \psi') \to \Box(\varepsilon(\varphi,\psi) \leftrightarrow \varepsilon(\varphi',\psi')).$$ ❑

Theorem 5.3.7 states that Švejdar's and Visser's supremum are *extensional.* As shown in [HV16, Theorem 10], Visser's supremum is not monotone: there are formulas φ and ψ such that, taking ε as in the statement of Theorem 5.3.6, $\mathsf{PA} \vdash \varphi \to \psi$ but $\mathsf{PA} \nvdash \varepsilon(\varphi,\varphi) \to \varepsilon(\psi,\psi)$. This property makes ε less suitable for being treated as a modal operator. But also Švejdar's supremum has some quirky features.

5.3.8. Remark. Write $\Theta(Y)$ for the formula $\forall x\,(\Diamond_x Y \to \Diamond_x \bot)$, and let σ be such that $\mathsf{PA} \vdash \sigma \leftrightarrow \Theta(\sigma)$. The formula σ is thus an element of $[\bot]$. It is easy to see that $\mathsf{PA} \vdash \sigma \leftrightarrow \bot$. On the other hand,

$$\mathsf{PA} \nvdash \Theta(\sigma) \leftrightarrow \Theta(\bot).$$

To see that, note that $\mathsf{PA} \vdash \Theta(\bot) \leftrightarrow \top$. Assuming $\mathsf{PA} \vdash \Theta(\sigma) \leftrightarrow \Theta(\bot)$, we would thus have:

$$\mathsf{PA} \vdash \bot \leftrightarrow \sigma \leftrightarrow \Theta(\sigma) \leftrightarrow \Theta(\bot) \leftrightarrow \top$$

— a blatant contradiction. We say that the formula $\Theta(Y)$ is not extensional. The above argument shows that every fixed point of $\Theta(Y)$ is equivalent to $\bot$, while $\bot$ itself is not a fixed point of $\Theta(Y)$.

5.4 Supremum adapters

A *supremum adapter* is an arithmetical formula σ containing one free variable, for which $\mathrm{I}\Delta_0{+}\mathsf{exp}$ verifies that for all φ, ψ, and χ,

$$\chi \rhd \varphi \ \wedge\ \chi \rhd \psi \quad\leftrightarrow\quad \chi \rhd \sigma(\varphi) \wedge \sigma(\psi).$$

We think of σ as *adapting* φ and ψ, so that their conjunction can be used to obtain the supremum. The idea and the first examples of supremum adapters are due to V. Yu. Shavrukov. The versions introduced here were obtained by analysing and simplifying the latter.

We define a family of nonstandard provability predicates, each of which can be viewed as the provability predicate of $\mathrm{I}\Sigma_a$, where a is a certain nonstandard number. In Section 5.4.2, it is shown that the consistency statements associated to these provability predicates are supremum adapters. We shall also refer to the provability predicates themselves as supremum adapters even though, strictly speaking, only their duals satisfy the definition given above.

Consider the following theory:

$$\mathsf{PA}^{\mu} := \bigcup_{n\in\omega}\{\mathrm{I}\Sigma_n \mid \text{ for all } m < n,\ \mathrm{I}\Sigma_m \text{ is 1-consistent}\}$$

Provability in PA^{μ} is naturally expressed by the formula $\exists x\,(\Box_x\varphi \wedge \forall y < x\,\Diamond_y^{\Pi_1}\top)$. Since PA is 1-consistent but, as a consequence of the Second Incompleteness Theorem, does not know this, it is clear that $\exists x\,(\Box_x\varphi \wedge \forall y < x\,\Diamond_y^{\Pi_1}\top)$ is a nonstandard provability predicate for PA.

Different from the theories PA^{*} and $\mathsf{PA}{\restriction}_{\mathtt{F}_{\varepsilon_0}}$ of Chapter 4, the theory PA^{μ}, as defined above, is not recursively enumerable. Its natural provability predicate $\exists x\,(\Box_x\varphi \wedge \forall y < x\,\Diamond_y^{\Pi_1}\top)$ is Σ_3; the complexity is due to the Π_1 truth predicate needed to express 1-consistency.

As in the case of slow provability, variants of supremum adapters are obtained by letting:

$$\blacktriangle_{\langle k\rangle}\varphi := \exists x\left(\Box_{x+k}\varphi \wedge \forall y < x\, \Diamond^{\Pi_1}_{y}\top\right),$$

where k is a standard integer. We recall from Chapter 4 that for $k < 0$, $x + k$ is defined to be $x \dot{-} |k|$, and $x \dot{-} k$ is defined to be $x + |k|$, where $|k|$ is the absolute value of k, and $x \dot{-} y := \max\{0, x - y\}$. As usual, we write $\blacktriangledown_{\langle k\rangle}$ for the dual of $\blacktriangle_{\langle k\rangle}$, i.e. as an abbreviation for $\neg\blacktriangle_{\langle k\rangle}\neg\varphi$.

5.4.1 Provability in the least 1-inconsistent subtheory

$\blacktriangle_{\langle k\rangle}$-provability may be seen as provability in $\mathrm{I}\Sigma_{\mu+k}$, where μ is the least x such that $\mathrm{I}\Sigma_x$ is 1-consistent. In order to make this precise, let $\mu = x$ be the formula

$$(\Diamond^{\Pi_1}\top \wedge x = \infty) \vee \left(\Box^{\Pi_1}_{x}\bot \wedge \forall y < x\, \Diamond^{\Pi_1}_{y}\top\right).$$

The uniqueness of μ is provable in $\mathrm{I}\Delta_0+\mathsf{exp}$: $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \mu = x \wedge \mu = y \rightarrow x = y$. As for existence, note that since $\exists x\, \Box^{\Pi_1}_{x}\bot$ is Σ_2 and the least number principle for Σ_2-formulas is equivalent to induction for Σ_2-formulas ([HP93, Theorem I.2.4]), it is clear that $\mathrm{I}\Sigma_2 \vdash \exists x\, \mu = x$. Theorem 5.4.2 below states that $\mathrm{I}\Sigma_1 \nvdash \exists x\, \mu = x$. Given the above, it is easy to see that

$$\mathrm{I}\Sigma_2 \vdash \forall\varphi\left(\blacktriangle_{\langle k\rangle}\varphi \leftrightarrow \Box_{\mu+k}\varphi\right), \tag{5.11}$$

where $\Box_\infty$ is defined to be $\Box$, and for all $k \in \mathbb{Z}$, $\infty + k$ and $\infty \dot{-} k$ are defined to be ∞. In other words, $\blacktriangle_{\langle k\rangle}$-provability coincides with provability in $\mathrm{I}\Sigma_{\mu+k}$, or — in case PA is 1-consistent — with provability in PA.

Each $\blacktriangle_{\langle k\rangle}$ has a "universal" relative: $\vartriangle_{[k]}\varphi := \forall x\left(\Box^{\Pi_1}_{x}\bot \rightarrow \Box_{x+k}\varphi\right)$. It can be shown that each $\vartriangle_{[k]}$, like $\blacktriangle_{\langle k\rangle}$, is a supremum adapter. Our reason for focusing on the "existential" versions $\blacktriangle_{\langle k\rangle}$ is that the $\vartriangle_{[k]}$ are not *sound* provability predicates. In the standard model $\mathfrak{N}$, each $\vartriangle_{[k]}$ is the provability predicate of the inconsistent theory: we have $\mathfrak{N} \models \vartriangle_{[k]}\varphi$ for all φ.

Defining $\Box_{\mu^\bot}$ like $\Box_\mu$, except with the reading that $\Box_\infty$ is the provability predicate of the the inconsistent theory, it is easy to check that

$$\mathrm{I}\Sigma_2 \vdash \forall\varphi\left(\vartriangle_{[k]}\varphi \leftrightarrow \Box_{\mu^\bot+k}\varphi\right). \tag{5.12}$$

The formulas $\Box^{\Pi_1}\bot$, $\blacktriangle_{\langle k\rangle}$, and $\vartriangle_{[k]}$, are definable in terms of each other:

5.4.1. Lemma. *i.* $\mathrm{I}\Sigma_2 \vdash \Box^{\Pi_1}\bot \leftrightarrow \forall\varphi\ \left(\blacktriangle_{\langle k\rangle}\varphi \leftrightarrow \vartriangle_{[k]}\varphi\right)$

ii. $\mathrm{I}\Sigma_2 \vdash \forall\varphi\,(\blacktriangle_{\langle k\rangle}\varphi \leftrightarrow (\Diamond^{\Pi_1}\top \wedge \Box\varphi) \vee (\Box^{\Pi_1}\bot \wedge \vartriangle_{[k]}\varphi))$

iii. $\mathrm{I}\Sigma_2 \vdash \forall\varphi\,(\vartriangle_{[k]}\varphi \leftrightarrow (\blacktriangle_{\langle k\rangle}\varphi \vee \Diamond^{\Pi_1}\top))$

Proof: (i) Argue in $I\Sigma_2$. If $\Box^{\Pi_1}\bot$, then by (5.11) and (5.12) both $\blacktriangle_{\langle k\rangle}$ and $\vartriangle_{[k]}$ are equivalent to $\Box_{\mu+k}$, where μ is the least x such that $\Box^{\Pi_1}_x\bot$. For the other direction, assume that $\Diamond^{\Pi_1}\top$. Then $\vartriangle_{[k]}$ is the provability predicate of the inconsistent theory, and so $\vartriangle_{[k]}\bot$. On the other hand, $\neg\blacktriangle_{\langle k\rangle}\bot$: $\blacktriangle_{\langle k\rangle}\bot$ would imply $\Box\bot$ and thus also $\Box^{\Pi_1}\bot$, contradicting our assumption. The proof of clauses (ii) and (iii) is straightforward, by using (i), (5.11), and (5.12). ❑

We conclude this section with a result due to F. Pakhomov, showing that the existence of μ is not provable in $I\Sigma_1$.

5.4.2. THEOREM. $I\Sigma_1 \nvdash \exists x\ \mu = x$

Proof: We show that there is a model $\mathfrak{M}$ of $I\Sigma_1$ where PA is 1-inconsistent, but there is no smallest a such that $I\Sigma_a$ is inconsistent. $\mathfrak{M}$ is constructed as the union of an ascending chain $(\mathfrak{M}_i)_{i<\omega}$ of models, where $\mathfrak{M}_i \models \Box^{\Pi_1}\bot$ and $\mathfrak{M}_i \models \mathsf{PA}$ for all i. Denote by m_i the least element such that $\mathfrak{M}_i \models \Box^{\Pi_1}_{m_i}\bot$. We shall ensure that for all i,

i. $m_i > m_{i+1}$

ii. $\mathfrak{M}_i \prec_{\Sigma_1} \mathfrak{M}_{i+1}$ ($\mathfrak{M}_{i+1}$ is a Σ_1-elementary extension of $\mathfrak{M}_i$)

We note that $\mathfrak{M} := \bigcup_{i<\omega}\mathfrak{M}_i$ is a model with the desired properties: from (ii) it follows that for all i,

$$\mathfrak{M}_i \prec_{\Sigma_1} \mathfrak{M}.$$

Using this, it is easy to show that $\mathfrak{M} \models I\Sigma_1$ and furthermore for all $a \in \mathfrak{M}$, we have that $\mathfrak{M} \models \Box^{\Pi_1}_a\bot$ if and only if $\mathfrak{M}_i \models \Box^{\Pi_1}_a\bot$ for some i.

It remains to show that a chain $(\mathfrak{M}_i)_{i<\omega}$ with the required properties exists. We proceed by induction on i. Let $\mathfrak{M}_0$ be any model of PA with $\mathfrak{M}_0 \models \Box^{\Pi_1}\bot$. Now suppose that we have constructed a model $\mathfrak{M}_i$ of PA with $\mathfrak{M}_i \models \Box^{\Pi_1}_{m_i}\bot \wedge \Diamond^{\Pi_1}_{m_i-1}\top$. Since $\mathfrak{M}_i \models \Diamond^{\Pi_1}_{m_i-1}\top$, we have $\mathfrak{M}_i \models \Diamond^{\Pi_1}_{m_i-1}\Box^{\Pi_1}_{\dot{m}_i-1}\bot$ by Löb's principle for $\Box^{\Pi_1}_{m_i-1}$. In fact we have $\mathfrak{M}_i \models \Diamond^{\Pi_1}_{m_i-1}(\Box^{\Pi_1}_{\dot{m}_i-1}\bot \wedge \Diamond^{\Pi_1}_{\dot{m}_i-2}\top)$, for $\mathfrak{M}_i \models \Box_{m_i-1}\Diamond^{\Pi_1}_{\dot{m}_i-2}\top$ by reflection. In other words, $\mathfrak{M}_i$ thinks that the theory

$$T := I\Sigma_{m_i-1} + \Pi_1\text{-truth} + \Box^{\Pi_1}_{\dot{m}_i-1}\bot + \Diamond^{\Pi_1}_{\dot{m}_i-2}\top$$

is consistent (where Π_1-truth is to be understood in the sense of $\mathfrak{M}_i$). By Theorem 2.1.13, there is a Σ_1-elementary end-extension $\mathfrak{M}_{i+1}$ of $\mathfrak{M}_i$ with $\mathfrak{M}_{i+1} \models T$ and such that, from the external point of view, $\mathfrak{M}_{i+1} \models \mathsf{PA}$. We now have $\mathfrak{M}_i \prec_{\Sigma_1} \mathfrak{M}_{i+1} \models \mathsf{PA}$ and $\mathfrak{M}_{i+1} \models \Box^{\Pi_1}_{m_i-1}\bot \wedge \Diamond^{\Pi_1}_{m_i-2}\top$. Thus it suffices to put $m_{i+1} = m_i - 1$. ❑

5.4.2 Verification of the target property

We show that each $\blacktriangle_{\langle k\rangle}$ or, strictly speaking its dual $\blacktriangledown_{\langle k\rangle}$, is a supremum adapter:

5.4.3. Theorem. *x where φ, χ, and ψ are regarded as internal variables ranging over $\mathcal{L}$-sentences.*

Proof: It suffices to show:

i. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\ \left(\blacktriangledown_{\langle k\rangle}\varphi \rhd \varphi\right)$

ii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\chi\,\forall\varphi\,\forall\psi\ \left((\chi \rhd \varphi)\ \wedge\ (\chi \rhd \psi)\ \ \to\ \ \chi \rhd \blacktriangledown_{\langle k\rangle}\varphi \wedge \blacktriangledown_{\langle k\rangle}\psi\right)$

The above clauses are proven in Propositions 5.4.6 and 5.4.8 below. ❑

A formula $\vartheta(x)$ is *monotone* (in x) if $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \vartheta(x) \wedge x \leqslant y \to \vartheta(y)$.

5.4.4. Lemma ([Vis15b, Theorem 4.3]). *Suppose that $\vartheta(x)$, possibly with additional free variables, is a monotone Π_2-formula.*

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall y\ \left(\Box_y \exists x \left(\vartheta(x) \wedge \forall u < x\, \Diamond^{\Pi_1}_u \top\right) \leftrightarrow \Box_y \vartheta(y)\right)$$

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, and fix y. Assume first $\Box_y\vartheta(y)$. By reflection, we have $\Box_y \forall u < y\, \Diamond^{\Pi_1}_u \top$, whence clearly $\Box_y \exists x \left(\vartheta(x) \wedge \forall u < x\, \Diamond^{\Pi_1}_u \top\right)$. For the other direction, assume

$$\Box_y \exists x \left(\vartheta(x) \wedge \forall u < x\, \Diamond^{\Pi_1}_u \top\right). \tag{5.13}$$

We show $\Box_y \left(\Box^{\Pi_1}_y \vartheta(y) \to \vartheta(y)\right)$. Argue in $\Box_y$:

> Assume $\Box^{\Pi_1}_y \vartheta(y)$. Suppose for a contradiction that $\neg\vartheta(y)$. Since this is a Σ_2-formula, we have $\Box^{\Pi_1}_y \neg\vartheta(y)$, and thus $\Box^{\Pi_1}_y \bot$. From (5.13), we have some x with $\vartheta(x)$ and $\forall u < x\, \Diamond^{\Pi_1}_u \top$. The latter, together with $\Box^{\Pi_1}_y \neg\vartheta(y)$, implies that $x \leqslant y$. Thus $\vartheta(y)$ follows by monotonicity of ϑ.

Back in $\mathrm{I}\Delta_0+\mathsf{exp}$, we have $\Box_y \left(\Box^{\Pi_1}_y \vartheta(y) \to \vartheta(y)\right)$. From the latter, it follows that $\Box_y \left(\Box_y \vartheta(y) \to \vartheta(y)\right)$, and finally $\Box_y \vartheta(y)$ by Löb's principle for $\Box_y$. ❑

5.4.5. Lemma. *i.* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall x \forall\varphi \left(\Box_x \blacktriangle_{\langle k\rangle}\varphi \leftrightarrow \Box_x \Box_{x+k}\varphi\right)$

ii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall y \forall\varphi \left(\Box_y \varphi \to \Box_{y \dot{-} k} \blacktriangle_{\langle k\rangle}\varphi\right)$

Proof: Reason in $\mathrm{I}\Delta_0+\mathsf{exp}$. (i) Fix x and φ. The statement follows from Lemma 5.4.4 by taking for $\vartheta(x)$ the formula $\Box_{x+k}\varphi$. Note that the latter is Σ_1 — so also Π_2 — and monotone.

(ii) Fix y and φ. Suppose first that $k \leqslant y$. Then $y \dot{-} k = y - k$, and we have:

$$\Box_y \varphi \to \Box_0 \Box_y \varphi \to \Box_{y-k} \Box_y \varphi \to \Box_{y-k} \blacktriangle_{\langle k\rangle}\varphi$$

The first implication is by Σ_1-completeness of $\Box_0$, the second by monotonicity, and the last follows from clause (i) by taking $y-k$ for x. If $k > y$, then $y \dot{-} k = 0$, and we argue as follows:

$$\Box_y\varphi \to \Box_0\Box_y\varphi \to \Box_0\Box_k\varphi \to \Box_0\blacktriangle_{\langle k\rangle}\varphi$$

The second implication is by monotonicity, the third follows from clause (i) by taking 0 for x. ❑

5.4.6. Proposition. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\ \left(\blacktriangledown_{\langle k\rangle}\varphi \rhd \varphi\right)$

Proof: In view of Theorem 2.1.12, it suffices to show

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\forall y\,\Box(\blacktriangledown_{\langle k\rangle}\varphi \to \Diamond_y\varphi)$$

i.e., by contraposition: $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\forall y\,\Box(\Box_y\varphi \to \blacktriangle_{\langle k\rangle}\varphi)$. Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, fixing some y and φ. By Lemma 5.4.5(ii), we have $\Box(\Box_y\varphi \to \Box_{y\dot{-}k}\blacktriangle_{\langle k\rangle}\varphi)$. Since $\blacktriangle_{\langle k\rangle}\varphi$ is Σ_3, we have by reflection $\Box_{(y\dot{-}k)+2}(\Box_{y\dot{-}k}\blacktriangle_{\langle k\rangle}\varphi \to \blacktriangle_{\langle k\rangle}\varphi)$. Combining the above yields $\Box(\Box_y\varphi \to \blacktriangle_{\langle k\rangle}\varphi)$ as required. ❑

5.4.7. Lemma. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall y\,\forall\varphi\,\forall\psi\,\Box_{y+1}\left(\Box_y(\blacktriangle_{\langle k\rangle}\varphi \vee \blacktriangle_{\langle k\rangle}\psi) \to \Box_{y+k}\varphi \vee \Box_{y+k}\psi\right)$

Proof: Reason in $\mathrm{I}\Delta_0+\mathsf{exp}$. Fix φ, ψ, and y, and argue in $\Box_{y+1}$:

Suppose that $\Box_y(\blacktriangle_{\langle k\rangle}\varphi \vee \blacktriangle_{\langle k\rangle}\psi)$, i.e.

$$\Box_y\left(\exists x\left(\Box_{x+k}\varphi \wedge \forall u < x\,\Diamond_u^{\Pi_1}\top\right) \vee \exists x\left(\Box_{x+k}\psi \wedge \forall u < x\,\Diamond_u^{\Pi_1}\top\right)\right)$$

The latter implies

$$\Box_y\exists x\ \left((\Box_{x+k}\varphi \vee \Box_{x+k}\psi) \wedge \forall u < x\,\Diamond_u^{\Pi_1}\top\right)$$

Applying Lemma 5.4.4 with $\vartheta(x) = \Box_{x+k}\varphi \vee \Box_{x+k}\psi$, we get $\Box_y(\Box_{y+k}\varphi \vee \Box_{y+k}\psi)$. Since we are in $\Box_{y+1}$, and $\Box_{y+k}\varphi \vee \Box_{y+k}\psi$ is Σ_1, we obtain $\Box_{y+k}\varphi \vee \Box_{y+k}\psi$ by reflection. ❑

5.4.8. Proposition. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash (\chi \rhd \varphi) \wedge (\chi \rhd \psi)\ \to\ \chi \rhd \blacktriangledown_{\langle k\rangle}\varphi \wedge \blacktriangledown_{\langle k\rangle}\psi$, *where χ, φ, and ψ are regarded as internal variables ranging over sentences.*

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$. Fix χ, φ, and ψ, and assume $\chi \rhd \varphi$ and $\chi \rhd \psi$. By Theorem 2.1.12, we have $\forall y\,\Box(\chi \to \Diamond_y\varphi)$ and $\forall y\,\Box(\chi \to \Diamond_y\psi)$, whence also

$$\forall y\,\Box(\chi \to \Diamond_y\varphi \wedge \Diamond_y\psi).$$

Using contraposition, it follows from Lemma 5.4.7 that

$$\Box_{y+1}\left(\Diamond_{y+k}\varphi \wedge \Diamond_{y+k}\psi \to \Diamond_y(\blacktriangledown_{\langle k\rangle}\varphi \wedge \blacktriangledown_{\langle k\rangle}\psi)\right).$$

Combining the above and using monotonicity of $\Box_y$, we have

$$\forall y\,\Box(\chi \to \Diamond_y(\blacktriangledown_{\langle k\rangle}\varphi \wedge \blacktriangledown_{\langle k\rangle}\psi)),$$

and so finally $\chi \rhd (\blacktriangledown_{\langle k\rangle}\varphi \wedge \blacktriangledown_{\langle k\rangle}\psi)$ by Theorem 2.1.12. ❑

5.5 Modal principles

This section studies the supremum adapters from a modal logical perspective. Section 5.5.1 deals with the provability logic of a supremum adapter, and Section 5.5.2 with the joint provability logic of a supremum adapter together with ordinary provability.

5.5.1 Provability logic

Note first that each $\blacktriangle_{\langle k\rangle}$ is a cross between ordinary and 1-provability. Dropping the Π_1-oracle from the definition of $\blacktriangle_{\langle k\rangle}$ yields the provability predicate:

$$\Box_{[k]}\varphi := \exists x\,(\Box_{x+k}\varphi \wedge \forall y < x\,\Diamond_y\top)\,.$$

For $k \geqslant 0$, $\mathrm{I}\Sigma_1 \vdash \Box_{[k]}\varphi \leftrightarrow \Box\varphi$. The direction from left to right is obvious. For the other direction, argue in $\mathrm{I}\Sigma_1$, and suppose $\Box\varphi$, i.e. $\Box_x\varphi$ for some x. By the least number principle for Σ_1-formulas, we can assume that x is the least with this property. Since $k \geqslant 0$, we have $\Box_{x+k}\varphi$ by monotonicity. If for some $y < x$, $\Box_y\bot$, then also $\Box_y\varphi$, contradicting the minimality of x. Thus $\forall y < x\,\Diamond_y\top$, and so $\Box_{[k]}\varphi$ as required.

In contrast, for $k < 0$, we have $\mathrm{I}\Sigma_1 \vdash \neg\Box_{[k]}\bot$. To see that, note that $\neg\Box_{[k]}\bot$ is the sentence

$$\forall x\;(\Box_{x+k}\bot \rightarrow \exists y < x\,\Box_y\bot)\,.$$

Argue in $\mathrm{I}\Sigma_1$, assuming $\Box_{x+k}\bot$. Since $\neg\Box_0\bot$ by reflection, it must be that $x+k > 0$. Since $k < 0$, it is clear that $x+k < x$, thus we can take $x+k$ for y. It is not difficult to see that the formula $\Box_{[-1]}$ is provably equivalent in $\mathrm{I}\Sigma_1$ to the Feferman provability predicate $\triangle_{\mathtt{f}}$.

Similarly to their pure-blooded relatives, the $\blacktriangle_{\langle k\rangle}$'s are naturally divided into two classes. For $k \geqslant 0$, $\blacktriangle_{\langle k\rangle}$ behaves like the ordinary provability predicate, whereas for $k < 0$, $\blacktriangle_{\langle k\rangle}$ behaves like Feferman provability predicate. We shall, from now on, write $\blacktriangle_{\mathtt{g}}$ for $\blacktriangle_{\langle k\rangle}$ with $k \geqslant 0$, and $\blacktriangle_{\mathtt{f}}$ for $\blacktriangle_{\langle k\rangle}$ with $k < 0$. The symbol $\blacktriangle$ is used as a common name for $\blacktriangle_{\mathtt{g}}$ and $\blacktriangle_{\mathtt{f}}$.

5.5.1. Lemma. *i. If* $\mathsf{PA} \vdash \varphi$, *then* $\mathsf{PA} \vdash \blacktriangle\varphi$

ii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\;(\Box_0\varphi \rightarrow \blacktriangle\varphi)$

iii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\;(\blacktriangle(\varphi \rightarrow \psi) \rightarrow (\blacktriangle\varphi \rightarrow \blacktriangle\psi))$

Proof: (i) follows from (the informal version of) Lemma 5.4.5(ii). For (ii) note that by monotonicity, we have for all k, $\Box_0\varphi \rightarrow (\Box_{0+k}\varphi \wedge \forall y < 0\,\Diamond_y^{\Pi_1}\top)$. Clause (iii) follows by monotonicity together with the HBL-conditions for $\Box_x$. ❑

Recall the provability logic GL, formulated in the language $\mathcal{L}_\Box$ of propositional logic together with a unary modality $\Box$. We consider arithmetical realisations mapping the modality $\Box$ to the formula $\blacktriangle_{\mathtt{g}}$.

5.5.2. LEMMA. *If* $\mathsf{GL} \vdash A$*, then* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash A^*$ *for all arithmetical realisations* *.

Proof: Given Lemma 5.5.1, it suffices to show $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \blacktriangle_{\mathsf{g}}(\blacktriangle_{\mathsf{g}}\varphi \to \varphi) \to \blacktriangle_{\mathsf{g}}\varphi$. The latter follows by the usual argument — see the proof of Theorem 2.1.9 — from $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \blacktriangle_{\mathsf{g}}\varphi \to \blacktriangle_{\mathsf{g}}\blacktriangle_{\mathsf{g}}\varphi$. We reason as follows:

$$\begin{aligned} \blacktriangle_{\langle k\rangle}\varphi &\to \Box_{y+k}\varphi \wedge \forall u < y\, \Diamond_u^{\Pi_1}\top \\ &\to \Box_{(y+k)\dot{-}k}\blacktriangle_{\langle k\rangle}\varphi && \text{Lemma 5.4.5(ii)} \\ &\to \Box_{y}\blacktriangle_{\langle k\rangle}\varphi && \text{since } k \geqslant 0 \\ &\to \Box_{y+k}\blacktriangle_{\langle k\rangle}\varphi && \text{monotonicity} \\ &\to \blacktriangle_{\langle k\rangle}\blacktriangle_{\langle k\rangle}\varphi \end{aligned}$$

The above argument also works when φ is an internal variable ranging over sentences. ❑

Thus GL is arithmetically sound with respect to $\blacktriangle_{\mathsf{g}}$. In Chapter 6, it is shown that GL is the provability logic of $\blacktriangle_{\langle 0\rangle}$.

5.5.3. QUESTION. Is GL arithmetically complete with respect to each $\blacktriangle_{\mathsf{g}}$?

Recall the provability logic F of the Feferman provability predicate, formulated in the language of propositional modal logic containing a unary modality $\vartriangle$, and axiomatised by adding to K the following:

$$\begin{array}{ll} (\mathsf{F1}) & \neg\vartriangle\bot \\ (\mathsf{F2}) & \vartriangle A \to \vartriangle((\vartriangle B \to B) \vee \vartriangle A) \end{array}$$

We consider arithmetical realisations mapping the modality $\vartriangle$ to the formula $\blacktriangle_{\mathsf{f}}$.

5.5.4. LEMMA. *If* $\mathsf{F} \vdash A$*, then* $\mathsf{PA} \vdash A^*$ *for all arithmetical realisations* *.

Proof: In view of Lemma 5.5.1, it suffices to show:

i. $\mathrm{I}\Sigma_1 \vdash \blacktriangledown_{\mathsf{f}}\top$

ii. $\mathsf{PA} \vdash \blacktriangle_{\mathsf{f}}\varphi \to \blacktriangle_{\mathsf{f}}\left((\blacktriangle_{\mathsf{f}}\psi \to \psi) \vee \blacktriangle_{\mathsf{f}}\varphi\right)$

(i) Note that $\blacktriangledown_{\langle k\rangle}\top$ is the sentence $\forall x\,(\Box_{x+k}\bot \to \exists u < x\, \Box_u^{\Pi_1}\bot)$. Argue in $\mathrm{I}\Sigma_1$, assuming $\Box_{x+k}\bot$. The latter implies $\Box_{x+k}^{\Pi_1}\bot$. Since $\Diamond_0^{\Pi_1}\top$ by reflection, it must be that $x + k > 0$. Since $k < 0$, clearly $x + k < x$; thus we may take $x + k$ for u.

(ii) Let $k' < 0$, and fix some ψ and n with $\psi \in \Pi_n$. Recall that μ was defined to be the least x such that $\mathrm{I}\Sigma_x$ is 1-inconsistent. In Section 5.4.1, it was explained that $\blacktriangle_{\langle k'\rangle}$ can be viewed as the provability predicate of $\mathrm{I}\Sigma_{\mu+k'}$:

$$\mathsf{PA} \vdash \forall\varphi\left(\blacktriangle_{\langle k'\rangle}\varphi \leftrightarrow \Box_{\mu+k'}\varphi\right).$$

Since $k' < 0$, $\mu + k'$ is defined to be $\mu \dot{-} k$, where $k = |k'|$. Reasoning in PA, μ is nonstandard, whence $\mu > k$, and so $\mu \dot{-} k = \mu - k$. In order to show $\mathsf{PA} \vdash \blacktriangle_{\langle k' \rangle}\varphi \to \blacktriangle_{\langle k' \rangle}\big((\blacktriangle_{\langle k' \rangle}\psi \to \psi) \vee \blacktriangle_{\langle k' \rangle}\varphi\big)$, it thus suffices to show

$$\mathsf{PA} \vdash \Box_{\mu-k}\varphi \to \Box_{\mu-k}\left((\Box_{\mu-k}\psi \to \psi) \vee \Box_{\mu-k}\varphi\right).$$

We argue in PA, assuming $\Box_{\mu-k}\varphi$ for some φ. Suppose first $\mu = \infty$. In this case, $\Box_\mu\chi \leftrightarrow \Box_{\mu-k}\chi \leftrightarrow \Box\chi$ for all χ. Thus it suffices to show $\Box\Box_{\mu-k}\varphi$. We have:

$$\Box_{\mu-k}\varphi \to \Box\varphi \to \Box_x\varphi \to \Box_{x+k}\blacktriangle_{\langle k' \rangle}\varphi \to \Box\Box_{\mu-k}\varphi,$$

where the third implication follows by Lemma 5.4.5(ii). Let us now assume $\mu < \infty$. We argue in $\Box_{\mu-k}$:

> From the outside world, we have $\Box_{\dot\mu-k}\varphi$ by Σ_1-completeness. By reflection, $\mu \geqslant \dot\mu - k$. We consider two cases:
>
> (*a*) $\mu \geqslant \dot\mu$. In this case $\Box_{\mu-k}\varphi$ follows from $\Box_{\dot\mu-k}\varphi$ by monotonicity.
>
> (*b*) $\mu < \dot\mu$. In this case $\mu - k < \dot\mu - k$. Assuming $\Box_{\mu-k}\psi$, we thus obtain ψ by reflection. Here we also use that ψ is Π_n and so also $\Pi_{\dot\mu+2}$ — since n is standard but $\dot\mu$ nonstandard (from the external point of view). ❑

5.5.5. Question. Is F arithmetically complete with respect to $\blacktriangle_{\mathsf{f}}$?

5.5.2 Supremum adapters and ordinary provability

We establish some principles for the joint provability logic of $\blacktriangle$ and $\Box$.

5.5.6. Lemma. *i.* $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \blacktriangle\varphi \to \Box\varphi$

ii. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box\varphi \to \blacktriangle\Box\varphi$

iii. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box(\Box\psi \to \blacktriangle\varphi) \leftrightarrow \Box(\Box\psi \to \varphi)$

iv. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box\blacktriangle\varphi \leftrightarrow \Box\varphi$

Proof: (ii) By Σ_1-completeness and Lemma 5.5.1(ii):

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box\varphi \to \Box_0\Box\varphi \to \blacktriangle\Box\varphi.$$

(iii) It follows from Theorem 5.4.3 that $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \varphi \equiv \blacktriangledown\varphi$. Since $\Diamond\psi$ is Π_1, Theorem 2.1.12 tells us that

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box(\blacktriangledown\varphi \to \Diamond\psi) \leftrightarrow \Box(\varphi \to \Diamond\psi).$$

(iii) follows from the above by contraposition and renaming φ, ψ. (iv) follows from (iii) by taking $\top$ for ψ. Inspecting the proof, it is clear that clauses (ii)-(iv) also hold when φ and ψ are regarded as internal variables. ❑

Lemmas 5.5.1 and 5.5.6 imply that $\blacktriangle_{\mathsf{g}}$ and $\Box$ satisfy the axioms and rules of GLT, verifiably in $\mathrm{I}\Delta_0{+}\mathsf{exp}$. We can thus apply Theorem 3.7.12 to obtain:

5.5.7. THEOREM. *The closed fragment of* GLT *is arithmetically complete with respect to* $\Box$ *and* $\blacktriangle_{\mathsf{g}}$.

5.5.8. COROLLARY. $\mathsf{PA} \nvdash \Box\bot \to \blacktriangle\bot$

Proof: The statement for $\blacktriangle_{\mathsf{g}}$ follows from Theorem 5.5.7, since $\mathsf{GLT} \nvdash \Box\bot \to \triangle\bot$ by Lemma 3.3.4. As for $\blacktriangle_{\mathsf{f}}$, note that $\mathsf{PA} \vdash \forall\varphi(\blacktriangle_{\langle u\rangle}\varphi \to \blacktriangle_{\langle k\rangle}\varphi)$ whenever $u < k$; therefore $\mathsf{PA} \vdash \forall\,\varphi(\blacktriangle_{\mathsf{f}}\varphi \to \blacktriangle_{\mathsf{g}}\varphi)$. $\mathsf{PA} \vdash \Box\bot \to \blacktriangle_{\mathsf{f}}\bot$ would thus imply $\mathsf{PA} \vdash \Box\bot \to \blacktriangle_{\mathsf{g}}\bot$. ❑

By Lemma 5.5.6(iii), the joint provability logic of $\blacktriangle_{\mathsf{g}}$ and $\Box$ contains the principle

$$\Box(\Box B \to \triangle A) \leftrightarrow \Box(\Box B \to A).$$

It is not difficult to see that $\mathsf{GLT} \vdash \Box(\Box B \to A) \to \Box(\Box B \to \triangle A)$. However, using Theorem 3.3.3, it can be shown that $\mathsf{GLT} \nvdash \Box(\Box B \to \triangle A) \to \Box(\Box B \to A)$. Let S be the principle $\Box(\Box B \to \triangle A) \to \Box(\Box B \to A)$. Recall axiom ($\mathsf{T4}$) of GLT: $\Box\triangle A \to \Box A$. Taking $\top$ for B, we see that ($\mathsf{T4}$) is derivable from (S) over K. Denote by GLS the system obtained from GLT by replacing axiom ($\mathsf{T4}$) with axiom (S).

5.5.9. QUESTION. Does GLS have a Kripke semantics?

5.5.10. QUESTION. Is GLS arithmetically complete with respect to $\Box$ and $\blacktriangle_{\mathsf{g}}$?

5.5.11. QUESTION. What is the joint provability logic of $\Box$ and $\blacktriangle_{\mathsf{f}}$?

5.6 Transfinite iterations

We conclude with some observations about the transfinite iterations of supremum adapters. The results in this section are due to F. Pakhomov.

We note first that every ω-iteration of a supremum adapter proof can be converted into an ordinary proof.

5.6.1. PROPOSITION. $\mathsf{PA} \vdash \forall\varphi\,(\blacktriangle^{\omega}\varphi \to \Box\varphi)$

Proof: Fix $k \in \mathbb{Z}$. We have $\mathsf{PA} \vdash \forall\varphi\,(\blacktriangle^{\omega}_{\langle k\rangle}\varphi \to \exists x\,\Box\blacktriangle^{x}\varphi)$ from the definition of $\blacktriangle^{\omega}_{\langle k\rangle}\varphi$ and Lemma 5.5.6(i). Thus it suffices to show $\mathsf{PA} \vdash \forall x\,\forall\varphi\,(\Box\blacktriangle^{x}_{\langle k\rangle}\varphi \to \Box\varphi)$. We argue in PA, by induction on x. If $x = 0$, then $\blacktriangle^{x}_{\langle k\rangle}\varphi$ is φ, so the claim trivially holds. Assume now that $\forall\psi(\Box\blacktriangle^{x}_{\langle k\rangle}\psi \to \Box\psi)$. To prove the claim for $x+1$, suppose that $\Box\blacktriangle^{x+1}_{\langle k\rangle}\varphi$ for some φ. This implies $\Box\blacktriangle_{\langle k\rangle}\blacktriangle^{x}_{\langle k\rangle}\varphi$, whence by Lemma 5.5.6(iv), $\Box\blacktriangle^{x}\varphi$, and so finally $\Box\varphi$ by the induction assumption. ❑

The remaining results in this section hold for specific supremum adapters only. Section 5.6.1 is concerned with $\blacktriangle_{\langle 0\rangle}$. Recall (Theorem 4.2.19) that for $k \geqslant 1$, the slow provability predicate $\vartriangle_{\langle k\rangle}$ is an ε_0-root of $\Box$, i.e. $\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \vartriangle^{\varepsilon_0}_{\langle k\rangle}\varphi)$. An analogue of the right to left direction is shown for $\blacktriangle_{\langle 0\rangle}$: ε_0-iterations of $\blacktriangle_{\langle 0\rangle}$-proofs can be converted into ordinary proofs (Theorem 5.6.5). However conversion into the opposite direction fails (Theorem 5.6.7).

Section 5.6.2 is concerned with $\blacktriangle_{\langle k\rangle}$, where $k \geqslant 1$. We show that each such $\blacktriangle_{\langle k\rangle}$ is an ω-root of $\Box$: $\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \blacktriangle^{\omega}_{\langle k\rangle}\varphi)$. Recall (Proposition 4.2.18) that for $k \geqslant 2$, the slow provability predicate $\vartriangle_{\langle k\rangle}$ is also an ω-root of $\Box$. The fact that a supremum adapter has this property is, in a sense, more surprising: it implies that $\blacktriangle_{\langle k\rangle}$, while being Σ_3, becomes Σ_1 when iterated up to ω.

5.6.1 $\mathrm{I}\Sigma_\mu$-proofs

Unless indicated otherwise, the Greek variables α, β, γ, etc. range over elements of an elementary linear ordering $(D, \prec)$, as in Section 2.4.1. Ordinals $\alpha \leqslant \varepsilon_0$ are assumed to be represented in $\mathrm{I}\Delta_0{+}\mathsf{exp}$ by their Cantor normal forms. For these ordinals, we write $<$ instead of $\prec$. To improve readability, we write $\blacktriangle$ for $\blacktriangle_{\langle 0\rangle}$ throughout Subsection 5.6.1.

The following lemma can be seen as a generalisation of Lemma 5.4.5(i), according to which $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall x \forall\varphi\ (\Box_x\blacktriangle\varphi \leftrightarrow \Box_x\Box_x\varphi)$.

5.6.2. Lemma. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall x\forall\alpha\forall\varphi\ (\Box_x\blacktriangle^{\alpha}\varphi \leftrightarrow \Box_x\Box^{\alpha}_x\varphi)$

Proof: By reflexive induction, it suffices to show in $\mathrm{I}\Delta_0{+}\mathsf{exp}$ that for every α,

$$\Box_0\forall\beta \prec \alpha\,\forall\varphi\,(\Box_x\blacktriangle^{\beta}\varphi \leftrightarrow \Box_x\Box^{\beta}_x\varphi) \to \forall\varphi\,(\Box_x\blacktriangle^{\alpha}\varphi \leftrightarrow \Box_x\Box^{\alpha}_x\varphi). \tag{5.14}$$

We argue in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, and fix some α. The antecedent of (5.14) implies

$$\Box_0\ \forall\varphi\ \left(\exists\beta \prec \alpha\,\Box_x\blacktriangle^{\beta}\varphi \leftrightarrow \exists\beta \prec \alpha\,\Box_x\Box^{\beta}_x\varphi\right).$$

Using monotonicity, it follows that

$$\forall\varphi\,\Box_x\left(\exists\beta \prec \alpha\,\Box_x\blacktriangle^{\beta}\varphi \leftrightarrow \exists\beta \prec \alpha\,\Box_x\Box^{\beta}_x\varphi\right).$$

With the HBL-conditions for $\Box_x$, we obtain from the latter

$$\forall\varphi\ \left(\Box_x\exists\beta \prec \alpha\,\Box_x\blacktriangle^{\beta}\varphi \leftrightarrow \Box_x\exists\beta \prec \alpha\,\Box_x\Box^{\beta}_x\varphi\right). \tag{5.15}$$

Note that the formula $\exists\beta \prec \alpha\,\Box_y\blacktriangle^{\beta}\varphi$ is Σ_1 and monotone (in y). By Theorem 5.4.4, we thus have

$$\Box_x\exists y\left(\exists\beta \prec \alpha\,\Box_y\blacktriangle^{\beta}\varphi \wedge \forall u < y\,\Diamond^{\Pi_1}_u\top\right) \leftrightarrow \Box_x\exists\beta \prec \alpha\,\Box_x\blacktriangle^{\beta}\varphi. \tag{5.16}$$

The left hand side is logically equivalent to $\Box_x \exists \beta \prec \alpha\, \exists y \left(\Box_y \blacktriangle^\beta \varphi \wedge \forall u < y\, \Diamond_u^{\Pi_1} \top \right)$, i.e. to $\Box_x \exists \beta \prec \alpha\, \blacktriangle \blacktriangle^\beta \varphi$. Thus it follows from (5.16) that

$$\forall \varphi \left(\Box_x \exists \beta \prec \alpha\, \blacktriangle \blacktriangle^\beta \varphi \leftrightarrow \Box_x \exists \beta \prec \alpha\, \Box_x \blacktriangle^\beta \varphi \right).$$

Combining the above with (5.15) we obtain

$$\forall \varphi \left(\Box_x \exists \beta \prec \alpha\, \blacktriangle \blacktriangle^\beta \varphi \leftrightarrow \Box_x \exists \beta \prec \alpha\, \Box_x \Box_x^\beta \varphi) \right)$$

i.e. $\forall \varphi\, (\Box_x \blacktriangle^\alpha \varphi \leftrightarrow \Box_x \Box_x^\alpha \varphi)$ as required. ❑

5.6.3. LEMMA. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall x\, \forall \varphi\, \forall \alpha < \varepsilon_0\ (\Box_x^{\alpha+1} \varphi \to \Box \varphi)$

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, fixing x, φ, and α as above. Argue in $\Box$:

We show first that

$$\forall \beta < \alpha + 1 \left(\forall \gamma < \beta\, (\Box_x^\gamma \varphi \to \varphi) \to (\Box_x^\beta \varphi \to \varphi) \right) \tag{5.17}$$

Fix $\beta < \alpha+1$, assume $\forall \gamma < \beta\, (\Box_x^\gamma \varphi \to \varphi)$, and suppose $\Box_x^\beta \varphi$. If $\beta = 0$, then we have φ immediately from the definition of $\Box_x^\beta \varphi$. If $\beta \neq 0$, then $\exists \gamma < \beta\, \Box_x \Box_x^\gamma \varphi$. Since x is external, we have $\Box_x^\gamma \varphi$ by reflection, and thus φ by assumption.

We are reasoning in PA and $\alpha < \varepsilon_0$. Using TI_{Σ_1}-$(\alpha + 1)$, it thus follows from (5.17) that $\forall \gamma < \alpha + 1\, (\Box_x^\gamma \varphi \to \varphi)$. In particular, $\Box_x^\alpha \varphi \to \varphi$.

Back in $\mathrm{I}\Delta_0+\mathsf{exp}$, we have shown $\Box(\Box_x^\alpha \varphi \to \varphi)$. It follows that $\Box\Box_x^\alpha \varphi \to \Box \varphi$, whence $\Box_x^{\alpha+1} \varphi \to \Box \varphi$ by Lemma 2.4.8. ❑

The following can be seen as a generalisation of Lemma 5.5.6(iv), according to which $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall \varphi\, (\Box \varphi \leftrightarrow \Box \blacktriangle \varphi)$.

5.6.4. LEMMA. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall \alpha < \varepsilon_0 \forall \varphi\ (\Box \varphi \leftrightarrow \Box \blacktriangle^\alpha \varphi)$

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, and let $\alpha < \varepsilon_0$. Suppose first $\Box \varphi$, and let x be such that $\Box_x \varphi$. By Lemma 2.4.6(iii),(vi) we have $\Box_x \Box_x^\alpha \varphi$, whence $\Box_x \blacktriangle^\alpha \varphi$ by Lemma 5.6.2, and so $\Box \blacktriangle^\alpha \varphi$ as required. For the other direction, let x be such that $\Box_x \blacktriangle^\alpha \varphi$. Then $\Box_x \Box_x^\alpha \varphi$ by Lemma 5.6.2, thus $\Box_x^{\alpha+1} \varphi$, and so $\Box \varphi$ by Lemma 5.6.3. ❑

5.6.5. THEOREM. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall \varphi\, (\blacktriangle^{\varepsilon_0} \varphi \to \Box \varphi)$.

Proof: Argue in $I\Delta_0+\mathsf{exp}$. We show that for all $\alpha < \varepsilon_0$, $\blacktriangle\blacktriangle^\alpha\varphi \to \Box\varphi$. Fix $\alpha < \varepsilon_0$. By Lemma 5.5.6(i), we have that $\blacktriangle\blacktriangle^\alpha\varphi$ implies $\Box\blacktriangle^\alpha\varphi$, and so $\Box\varphi$ follows by Lemma 5.6.4. ❑

We now show that ordinary proofs cannot generally be converted into transfinite iterations of $\blacktriangle$-proofs. In the lemmas below, the Greek letters range over elements of an arbitrary elementary linear ordering $(D, \prec)$. We assume that, verifiably in $I\Delta_0+\mathsf{exp}$, $(D, \prec)$ has a least element, denoted by $\mathsf{0}$. We denote by $\mathsf{s}(\alpha)$ the unique immediate successor of α, i.e. an element D for which $\mathsf{s}(\alpha, \beta)$ holds.

5.6.6. Lemma. $I\Delta_0+\mathsf{exp} \ + \ \mathrm{TI}_{\Pi_1}\text{-}\mathsf{s}(\alpha) \ + \ \Diamond^{\Pi_1}\top \vdash \forall x \forall \psi \, (\Diamond_x \psi \to \Diamond_x \Diamond_x^\alpha \psi)$

Proof: Argue in $I\Delta_0+\mathsf{exp} \ + \ \mathrm{TI}_{\Pi_1}\text{-}\alpha \ + \ \Diamond^{\Pi_1}\top$. Fix x and ψ, and assume $\Diamond_x\psi$. We show by transfinite induction on $\beta \prec \mathsf{s}(\alpha)$ that $\Diamond_x\Diamond_x^\beta\psi$. For $\beta = \mathsf{0}$ the claim is immediate from the definition of $\Diamond_x^\beta\psi$. Supposing that β is a successor, let γ be such that $\mathsf{s}(\gamma, \beta)$. By the induction assumption, we have $\Diamond_x\Diamond_x^\gamma\psi$ which by Lemma 2.4.6(vi) is equivalent to $\Diamond_x^\beta\psi$. Since this is a Π_1-sentence, we have $\Box_x^{\Pi_1}\Diamond_x^\beta\psi$. Supposing that $\Box_x\Box_x^\beta\neg\psi$, we would therefore have $\Box_x^{\Pi_1}\bot$, contradicting that $\Diamond^{\Pi_1}\top$. Thus it must be that $\neg\Box_x\Box_x^\beta\neg\psi$, i.e. $\Diamond_x\Diamond_x^\beta\psi$ as required. Finally, suppose that β is a limit ordinal. Then by assumption we have $\forall\delta \prec \beta \, \Diamond_x\Diamond_x^\delta\psi$. Since this is a Π_1-sentence, we obtain $\Diamond_x\forall\delta \prec \beta \, \Diamond_x\Diamond_x^\delta\psi$, i.e. $\Diamond_x\Diamond_x^\beta\psi$, reasoning exactly as above. ❑

5.6.7. Theorem. *If* $\mathrm{TI}_{\Pi_1}\text{-}\mathsf{s}(\alpha)$ *is consistent with* $\mathsf{PA} + \Diamond^{\Pi_1}\top$*, then there is a sentence* φ_α *with* $\mathsf{PA} \nvdash \Box\varphi_\alpha \to \blacktriangle^\alpha\varphi_\alpha$.

Proof: Suppose that $\mathrm{TI}_{\Pi_1}\text{-}\mathsf{s}(\alpha)$ is consistent with $\mathsf{PA} + \Diamond^{\Pi_1}\top$. We show that there is a sentence φ_α and a model $\mathfrak{M}$ of PA with $\mathfrak{M} \models \Box\varphi_\alpha \wedge \neg\blacktriangle^\alpha\varphi_\alpha$. Let

$$\varphi_\alpha := \neg\left(\mathrm{TI}_{\Pi_1}\text{-}\alpha \wedge \Diamond^{\Pi_1}\top\right),$$

and let $\mathfrak{M}$ be a model of PA with $\mathfrak{M} \models \Box\varphi_\alpha \wedge \neg\varphi_\alpha$. To see that $\mathfrak{M}$ exists, note that $\mathsf{PA} \vdash \Box\varphi_\alpha \to \varphi_\alpha$ would imply $\mathsf{PA} \vdash \varphi_\alpha$ by Löb's Theorem, contradicting our assumption that $\mathsf{PA} + \neg\varphi_\alpha$ is consistent.

Since PA is essentially reflexive, we have $\mathfrak{M} \models \Diamond_n\neg\varphi_\alpha$ for every standard n, whence by overspill $\mathfrak{M} \models \Diamond_m\neg\varphi_\alpha$ for some nonstandard m. By Lemma 5.6.6, $\mathfrak{M} \models \Diamond_m\Diamond_m^\alpha\neg\varphi_\alpha$, and therefore $\mathfrak{M} \models \Diamond_m\blacktriangledown^\alpha\neg\varphi_\alpha$ by Lemma 5.6.2. Since $\mathfrak{M} \models \Box\varphi_\alpha$, we have $\mathfrak{M} \models \Diamond_m(\Box\varphi_\alpha \wedge \blacktriangledown^\alpha\neg\varphi_\alpha)$ by Σ_1-completeness of $\Box_m$. In other words $\mathfrak{M}$ thinks that the theory $I\Sigma_m + \Box\varphi_\alpha + \blacktriangledown^\alpha\neg\varphi_\alpha$ is consistent. By Theorem 2.1.13, there is an end-extension $\mathfrak{M}'$ of $\mathfrak{M}$ with $\mathfrak{M}' \models \mathsf{PA}$ (from the external point of view) and $\mathfrak{M}' \models \Box\varphi_\alpha \wedge \blacktriangledown^\alpha\neg\varphi_\alpha$. ❑

Theorem 5.6.7 implies that for every α, there is some sentence φ_α such that $\mathsf{PA} \nvdash \Box\varphi_\alpha \to \blacktriangle^\alpha\varphi_\alpha$. We show next that there is no single φ that works for all such α. This result relies on Turing's Completeness Theorem for ordinal logics. Kleene's $\mathcal{O}$ is a system of ordinal notations; it contains a (not necessarily unique) notation for every recursive ordinal.

5.6.8. THEOREM ([TUR39]). *For any true Π_1-sentence π, there is an ordinal notation $c \in \mathcal{O}$ with $|c| = \omega + 1$, and such that $\mathrm{I}\Sigma_1 \vdash \Diamond_1^c \top \to \pi$.* ❑

An inspection of the proof of Theorem 5.6.8 shows that the ordering corresponding to $|c|$ is elementary. We can thus reason about $\Box^c$ using the results of Section 2.4.2.

5.6.9. THEOREM. *For any φ, there is an ordinal notation $c \in \mathcal{O}$ with $|c| = \omega + 1$, and such that $\mathsf{PA} \vdash \Box\neg\varphi \to \blacktriangle^c\neg\varphi$.*

Proof: If $\mathsf{PA} \vdash \neg\varphi$, then $\mathsf{PA} \vdash \blacktriangle\neg\varphi$, whence by Lemma 2.4.7, $\mathsf{PA} \vdash \blacktriangle^c\neg\varphi$, and so clearly also $\mathsf{PA} \vdash \Box\neg\varphi \to \blacktriangle^c\neg\varphi$. Assume now that $\mathsf{PA} \nvdash \neg\varphi$. Then $\Diamond\varphi$ is a true Π_1-sentence, whence by Theorem 5.6.8 there is some $c \in \mathcal{O}$ with $|c| = \omega + 1$ and $\mathrm{I}\Sigma_1 \vdash \Diamond_1^c\top \to \Diamond\varphi$. By the HBL-conditions for $\Box^c$, we have $\mathrm{I}\Sigma_1 \vdash \Diamond_1^c\varphi \to \Diamond_1^c\top$. Using reflection, we see that $\mathrm{I}\Sigma_1 \vdash \forall\psi\,(\Box_1\psi \to \blacktriangle\psi)$, whence by Lemma 2.4.7,

$$\mathrm{I}\Sigma_1 \vdash \blacktriangledown^c\varphi \to \Diamond_1^c\varphi.$$

Combining the above yields $\mathrm{I}\Sigma_1 \vdash \blacktriangledown^c\varphi \to \Diamond\varphi$ i.e. $\mathrm{I}\Sigma_1 \vdash \Box\neg\varphi \to \blacktriangle^c\neg\varphi$. ❑

5.6.2 ω-roots of ordinary provability

We consider $\blacktriangle_{\langle k\rangle}$ with $k \geqslant 1$. We shall here write $\blacktriangle$ for such $\blacktriangle_{\langle k\rangle}$.

5.6.10. LEMMA. $\mathsf{PA} \vdash \forall\varphi\,(\Box_x\varphi \to \blacktriangle^{x+1}\varphi)$.

Proof: We argue in PA, by induction on x. If $x = 0$, the claim holds by Lemma 5.5.1(ii). So let us assume $\forall\psi\,(\Box_x\psi \to \blacktriangle^{x+1}\psi)$. To show the claim for $x+1$, suppose $\Box_{x+1}\varphi$. By Lemma 5.4.5(ii), $\Box_{x+1}\varphi \to \Box_{(x+1)\dot{-}k}\blacktriangle_{\langle k\rangle}\varphi$. If $k \geqslant 1$, it is clear that $(x+1) \dot{-} k \leqslant x$. Using monotonicity, we thus have $\Box_{x+1}\varphi \to \Box_x\blacktriangle\varphi$, and so $\Box_x\blacktriangle\varphi$. Taking $\blacktriangle\varphi$ for ψ, the induction assumption gives us $\blacktriangle^{x+1}\blacktriangle\varphi$, and so $\blacktriangle^{x+2}\varphi$ by Lemma 2.4.8. ❑

The following is an immediate consequence of Lemmas 5.6.10 and 5.6.1:

5.6.11. PROPOSITION. $\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \blacktriangle^\omega\varphi)$ ❑

5.6.12. QUESTION. What can be said about transfinite iterations of $\blacktriangle_{\langle k\rangle}$ for $k < 0$?

Chapter 6

A Solovay function for the least 1-inconsistent subtheory of PA

The provability predicate $\vartriangle$, introduced in Section 5.4 as $\blacktriangle_{\langle 0 \rangle}$, is defined as:

$$\vartriangle\varphi := \exists x \, \left(\Box_x \varphi \wedge \forall y < x \, \Diamond^{\Pi_1}_y \top \right) .$$

We recall that $\vartriangle$ is useful for obtaining suprema in the lattice of interpretability degrees of finite extensions of PA (Theorem 5.4.3), and that it can be seen as the provability predicate of $\mathrm{I}\Sigma_\mu$, where μ is the least x such that $\mathrm{I}\Sigma_x$ is 1-inconsistent (Section 5.4.1). In this chapter, we show that the provability logic of $\vartriangle$ coincides with that of the ordinary provability predicate of PA, i.e. the Gödel–Löb provability logic GL. While arithmetical completeness of GL with respect to ordinary provability is established by using a single Solovay function, our proof for $\vartriangle$ relies on a uniformly defined infinite sequence of such functions.

6.1 Introduction

We are interested in the PA-provable propositional schemata involving $\vartriangle$, i.e. the provability logic of $\vartriangle$. We consider arithmetical realisations mapping the modality $\Box$ of GL to the provability predicate $\vartriangle$, and show the following:

6.1.1. Theorem. GL $\vdash A$ *iff for all arithmetical realisations* *, PA $\vdash A^*$.

Arithmetical soundness, i.e. the left to right direction of Theorem 6.1.1, is Lemma 5.5.2. The proof of the other direction proceeds, as usual (see the proof of Theorem 2.3.2), by showing that any finite Kripke frame for GL can be suitably embedded into PA. Our proof is closely related to Beklemishev's arithmetical completeness proof for GLP ([Bek11]). The latter uses a sequence $(h_i)_{i<m}$ of Solovay functions, where m is some standard number. In contrast, we need an infinite sequence $(h_y)_{y<\omega}$ of such functions, uniformly defined by a single formula.

We start with an informal description of the construction, based on the view of $\triangle$ as provability in $\mathrm{I}\Sigma_\mu$ (Section 5.4.1). Given a finite GL-frame $\mathcal{F} = \langle W, \prec \rangle$ with root 0, we consider a family of Solovay functions $(h_y)_{y<\omega}$ climbing up the accessibility relation $\prec$ of $\mathcal{F}$. The function h_0 is the usual Solovay function for $\mathrm{I}\Delta_0+\mathsf{exp}$: it starts at 0 and moves upon the emergence of $\mathrm{I}\Delta_0+\mathsf{exp}$-proofs concerning its own limit. Similarly, each h_{y+1} is like the usual Solovay function for $\mathrm{I}\Sigma_{y+1}$, except that it starts where the previous function h_y came to rest.

We write $x : \Box_y\varphi$ to mean that x is (the code of) a $\Box_y$-proof of φ. Given a reasonable coding of proofs, we have that the formula $x : \Box_y\varphi$ is elementary, and furthermore that every x is the proof of at most one sentence.

Letting ℓ_y denote the limit of h_y, we would like the functions $(h_y)_{y<\omega}$, with $h_y : \omega \to W$, to satisfy:

$$h_0(0) = 0, \qquad h_{y+1}(0) = \ell_y, \qquad \text{and}$$
$$h_y(x+1) = \begin{cases} a & \text{if } h_y(x) \prec a \text{ and } x : \Box_y \ell_y \neq a, \\ h_y(x) & \text{otherwise.} \end{cases}$$

We are interested in the value ℓ_μ, i.e. the limit of the function h_μ, where ℓ_μ is defined to be $\lim_{y\to\infty} \ell_y$ in case $\mu = \infty$. We would like to to show that the sentence $\ell_\mu = a$ is a natural arithmetical representative for the node a, in the sense that for some theory $T \subseteq \mathsf{PA}$,

i. if $a \neq 0$, then $T \vdash \ell_\mu = a \to \Box_\mu a \prec \ell_\mu$,

ii. if $a \prec b$, then $T \vdash \ell_\mu = a \to \Diamond_\mu \ell_\mu = b$.

It might seem, at first sight, that T has to be at least as strong as $\mathrm{I}\Sigma_2$: as seen in the previous chapter (Theorem 5.4.2), the existence of μ is not known to $\mathrm{I}\Sigma_1$. Moreover, each function h_{y+1}, as presented above, is genuinely more complex than the usual Solovay function: it is defined by using the limit of h_y, the natural representation of which is at least Σ_2. It is therefore not obvious that the basic properties of h_y and ℓ_y can be verified in $\mathrm{I}\Delta_0+\mathsf{exp}$ or even $\mathrm{I}\Sigma_1$. By tweaking the construction we shall nevertheless succeed in making everything work smoothly in $\mathrm{I}\Delta_0+\mathsf{exp}$.

6.2 A multi-stage Solovay function

Fix a GL-frame $\mathcal{F} = \langle W, \prec \rangle$. We start by defining an auxiliary function $h_{y,a}$ — the Solovay function for $\Box_y$, starting off at node a (where both y and a are parameters represented by free variables).

6.2.1. Definition. ($\mathrm{I}\Delta_0+\mathsf{exp}$) For $y < \omega$, $a \in W$, the function $h_{y,a} : \omega \to W$ is defined by:

$$h_{y,a}(0) = a$$
$$h_{y,a}(x+1) = \begin{cases} b & \text{if } h_{y,a}(x) \prec b \text{ and } x : \Box_y \ell_{\dot{y}} \neq b, \\ h_{y,a}(x) & \text{otherwise.} \end{cases}$$

The formula $\ell_y \neq b$ (Definition 6.2.3 below) depends on the formula χ representing the family of functions $(h_{y,a})_{y<\omega, a\in W}$. The self-reference in the definition of $h_{y,a}$ is handled, as usual, by the Fixed Point Lemma. We note here that the definition of $h_{y,a}$ only relies on the gödelnumber of $\ell_{\dot{y}} \neq b$, and the latter can be obtained from y, b and $\ulcorner\chi\urcorner$ by a function whose totality is known to $\mathrm{I}\Delta_0+\mathsf{exp}$.

It follows from Theorem 2.1.4 — for example, by using that W is finite — that the function $\mathtt{h}(y,a,x) = h_{y,a}(x)$ is elementary ($\Delta_0^{\exp}$) and hence provably total in $\mathrm{I}\Delta_0+\mathsf{exp}$, with its defining equations also provable in $\mathrm{I}\Delta_0+\mathsf{exp}$.

We write $\lim h_{y,a} = b$ for the formula

$$\exists x\ h_{y,a}(x) = b \wedge \forall x\ h_{y,a}(x) \preceq b.$$

Since $h_{y,a}(x)$ is elementary, we have that $\lim h_{y,a} = b$ is provably equivalent in $\mathrm{I}\Delta_0+\mathsf{exp}$ to a $\Delta_0(\Sigma_1)$-formula. The formula $\lim h_{y,a} = b$ states that b is the $\preceq$-largest element in the range of $h_{y,a}$. In view of the following lemma, we can think of $\lim h_{y,a} = b$ as saying that b is the limit of $h_{y,a}$.

6.2.2. LEMMA. *i.* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash x' \leqslant x \to h_{y,a}(x') \preceq h_{y,a}(x)$

ii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \exists! b \lim h_{y,a} = b$

Proof: (i) is proven by internal induction on x. Since $h_{y,a}(x) \preceq h_{y,a}(x+1)$ by definition, the inductive step follows by using the transitivity of $\preceq$.

(ii) Since $\preceq$ is antisymmetric, uniqueness is immediate from the definition of $\lim h_{y,a}$. For existence, we show by external induction on the converse of $\prec$ that for all $c \in W$,

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash h_{y,a}(x) = c \to \exists b \lim h_{y,a} = b.$$

This is sufficient, since $\mathrm{I}\Delta_0+\mathsf{exp}$ proves that $h_{y,a}(0) = a$. From (i) we have that

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash h_{y,a}(x) = c \to (\forall x' \geqslant x\ h_{y,a}(x') = c \vee \exists x' \geqslant x\ c \prec h_{y,a}(x')) . \quad (6.1)$$

Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, assuming $h_{y,a}(x) = c$. If the first disjunct in (6.1) holds then, using (i), we have $\lim h_{y,a} = c$, while if the second disjunct holds, then $\exists b \lim h_{y,a} = b$ by the induction assumption. Thus in either case $\exists b \lim h_{y,a} = b$ as required. ❑

6.2.3. Definition. ($\mathrm{I}\Delta_0$+exp) The formula $\ell_y = a$, with free variable y, is defined as:

$$\exists s\ (s = (l_0, l_1, \ldots, l_y)\ \wedge\ l_0 = \lim h_{0,0}\ \wedge\ \forall z < y\ l_{z+1} = \lim h_{z+1,l_z}\ \wedge\ l_y = a)\,.$$

A sequence s is a *y-witness* if it satisfies the first three conjuncts in the formula above. If s also satisfies the fourth conjunct, then s is a *witness for $\ell_y = a$*. We write $\ell_y \neq a$ for the negation of $\ell_y = a$.

From Definition 6.2.1, it is clear that any y-witness is a sequence of elements of W. Since the latter is finite, the leading existential quantifier in $\ell_y = a$ can be bounded by a term of the form $\exp(k \cdot y)$, where k is a sufficiently large standard number. Recalling that $\lim h_{y,b} = c$ is a $\Delta_0(\Sigma_1)$-formula, we thus see that $\ell_y = a$ is a $\Delta_0^{\exp}(\Sigma_1)$, where the latter is defined like $\Delta_0(\Sigma_1)$, except that the quantifier bounds are allowed to be of the form $\forall x \leqslant \exp z$ and $\exists x \leqslant \exp z$. Reasoning about the formula $\ell_y = a$ by induction on y can therefore be problematic in $\mathrm{I}\Delta_0$+exp. The following lemmas state that several properties of $\ell_y = a$ are nevertheless verifiable in $\mathrm{I}\Delta_0$+exp.

6.2.4. Lemma ($\mathrm{I}\Delta_0$+exp). *For each y, there is at most one y-witness. In particular, $\ell_y = a$ and $\ell_y = b$ imply $a = b$.*

Proof: We reason in $\mathrm{I}\Delta_0$+exp. Suppose that $(l_0, \ldots, l_y)$ and $(l'_0, \ldots, l'_y)$ are both y-witnesses. We prove, by $\Delta_0^{\exp}$-induction that for all $i \leqslant y$, $l_i = l'_i$. By Lemma 6.2.2(ii) it is clear that $l_0 = \lim h_{0,0} = l'_0$. Supposing that $l_i = l'_i$, we have again by Lemma 6.2.2(ii) that $l_{i+1} = \lim h_{i+1,l_i} = \lim h_{i+1,l'_i} = l'_{i+1}$. ❑

If follows from Lemma 6.2.4 that we can treat ℓ_y as a partial function in $\mathrm{I}\Delta_0$+exp.

6.2.5. Lemma. *i.* $\mathrm{I}\Delta_0$+$\mathsf{exp} \vdash x < y \wedge \ell_y = b \to \exists a\, \ell_x = a$

ii. $\mathrm{I}\Delta_0$+$\mathsf{exp} \vdash x < y \wedge \ell_y = b \to \ell_x \preceq b$

iii. $\mathrm{I}\Delta_0$+$\mathsf{exp} \vdash \ell_y = 0 \leftrightarrow \forall z \leqslant y\ \forall x\ h_{z,0}(x) = 0.$

iv. $\mathrm{I}\Delta_0$+$\mathsf{exp} \vdash \ell_y = a \prec b \to \Diamond_y \ell_{\dot{y}} = b$

v. $\mathrm{I}\Delta_0$+$\mathsf{exp} \vdash \ell_y = a \neq 0 \to \Box_y \ell_{\dot{y}} \neq a.$

Proof: We argue in $\mathrm{I}\Delta_0$+exp. (i) Suppose that $(l_0, \ldots, l_y)$ is a witness for $\ell_y = b$. If $x < y$, then clearly $(l_0, \ldots, l_x)$ is a witness for $\ell_x = l_x$, so we can put $a = l_x$.

(ii) Suppose that $(l_0, \ldots, l_y)$ is a witness for $\ell_y = b$. If $x < y$, then an x-witness exists by clause (i), and by Lemma 6.2.4 it is an initial segment of $(l_0, \ldots, l_y)$. We prove by $\Delta_0^{\exp}$-induction on $z \leqslant y$ that $z' < z$ implies $l_{z'} \preceq l_z$. It follows from the relevant definitions, together with Lemma 6.2.2, that $l_z \preceq \lim h_{z+1,l_z} = l_{z+1}$. Assuming $l_{z'} \preceq l_z$, we thus obtain $l_{z'} \preceq l_{z+1}$ by transitivity of $\preceq$.

(iii) We have $\forall z \leqslant y\ \forall x\ h_{z,0}(x) = 0$ iff $\forall z \leqslant y \lim h_{z,0} = 0$ iff $(l_0, \ldots, l_z)$, with all $l_i = 0$, is a witness for $\ell_z = 0$.

(iv) By Definitions 6.2.1 and 6.2.3, together with the transitivity and anti-symmetry of $\preceq$.

(v) Assume that $\ell_y = a \neq 0$, and let $(l_0, \ldots, l_y)$ be the witness for $\ell_y = a$. By Δ_0^{exp}-induction, let $y' \leqslant y$ be minimal such that $l_{y'} = a$. Since $a \neq 0$, by the definition of ℓ_y and $h_{y,c}$, we have that $\Box_{y'}\ell_{y'} \neq a$. We show $\Box_{y'}\ell_y \neq a$, from which $\Box_y \ell_y \neq a$ clearly follows. Argue in $\Box_{y'}$:

> Suppose that $\ell_y = a$. Since $\ell_{y'} \neq a$, we have $y' < y$, and so $\ell_{y'} \prec \ell_y$ by clause (ii). Let b be such that $\ell_{y'} = \lim h_{y',b}$. We thus have $\forall x\ h_{y',b}(x) \prec a$. By Σ_1-completeness, we also have $\Box_{y'}\ell_{y'} \neq a$. By definition of $h_{y',b}$, this implies $\exists x\, h_{y',b}(x) = a$, hence $a \preceq \lim h_{y',b}$ i.e. $\ell_y \preceq \ell_{y'}$, a contradiction. ❑

6.2.6. Lemma. *i.* $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \ell_y = a \neq 0 \to \Box_y^{\Pi_1}\bot$

ii. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Diamond_x^{\Pi_1}\top \to \ell_x = 0$

iii. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \forall x < y\, \Diamond_x^{\Pi_1}\top \to \ell_y = \lim h_{y,0}$

iv. $\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \lim h_{y,0} = a \to \Box_y a \preceq \ell_{\dot{y}}$.

Proof: (i) Argue in $\mathrm{I}\Delta_0{+}\mathsf{exp}$, letting $(l_0, \ldots, l_y)$ be the witness for $\ell_y = a \neq 0$. Using Δ_0^{exp}-induction, we can assume that y is minimal such that $l_y \neq 0$, thus either $y = 0$ or $l_{y-1} = 0$. It follows that $\ell_y = \lim h_{y,0}$, and so

$$\lim h_{y,0} = a. \tag{6.2}$$

Since $a \neq 0$, we have by Lemma 6.2.5(v) that

$$\Box_y \ell_y \neq a. \tag{6.3}$$

Reason in $\Box_y^{\Pi_1}$:

> We claim first that $\ell_y = \lim h_{y,0}$. If $y = 0$, this is clear from the definition. If $y > 0$, we have $\ell_{y-1} = 0$ since, using Lemma 6.2.5(iii), the latter is equivalent to a true Π_1-formula. Since (6.2) is a true conjunction of a Σ_1- and a Π_1-formula, it is also true here, whence it follows that $\ell_y = a$, contradicting (6.3).

(ii) By reflexive induction (Lemma 2.4.3), it suffices to show:

$$\mathrm{I}\Delta_0{+}\mathsf{exp} \vdash \Box_0 \forall z < x\, (\Diamond_z^{\Pi_1}\top \to \ell_z = 0) \to (\Diamond_x^{\Pi_1}\top \to \ell_x = 0).$$

Argue in $\mathrm{I}\Delta_0{+}\mathsf{exp}$. If $x = 0$, then, since $\ell_0 = \lim h_{0,0}$ and the latter exists by Lemma 6.2.2(ii), we have the claim from clause (i) by contraposition. So let

$x > 0$, and suppose that $\Box_0(\Diamond^{\Pi_1}_{x-1}\top \to \ell_{x-1} = 0)$ and $\Diamond^{\Pi_1}_x\top$. Since $\Box_x\Diamond^{\Pi_1}_{x-1}\top$ by reflection, it follows that $\Box_x\ell_{x-1} = 0$. Since $\Diamond^{\Pi_1}_x\top$ is equivalent to Σ_1-reflection for $\Box_x$ and $\ell_{x-1} = 0$ is equivalent to a Π_1-formula by Lemma 6.2.5(iii), we now have $\ell_{x-1} = 0$. But this means that ℓ_x is equal to $\lim h_{x,0}$, and thus it exists by Lemma 6.2.2(ii). Finally, $\ell_x = 0$ follows by contraposition from clause (i).

(iii) Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$, assuming $\forall x < y\, \Diamond^{\Pi_1}_x\top$. For $y = 0$, $\ell_y = \lim h_{y,0}$ holds by definition. For $y > 0$ we have $\ell_{y-1} = 0$ from clause (ii) together with the assumption, and so again by definition $\ell_y = \lim h_{y,\ell_{y-1}} = \lim h_{y,0}$.

(iv) Suppose that $\lim h_{y,0} = a$, whence $\exists x\, h_{y,0}(x) = a$. Reason in $\Box_y$:

> Using reflection, we obtain $\forall x < y\, \Diamond^{\Pi_1}_x\top$, and so $\ell_y = \lim h_{y,0}$ by clause (iii). By Σ_1-completeness we have $\exists x\, h_{y,0}(x) = a$ from outside, and so $a \preceq \ell_y$ by the definition of $\lim h_{y,0}$. ❑

6.2.7. Remark. While, as shown above, ℓ_y is a partial function in $\mathrm{I}\Delta_0+\mathsf{exp}$, its totality is, in general, not provable in $\mathrm{I}\Delta_0+\mathsf{exp}$. Consider the frame with $W = \{0, 1\}$ and $0 \prec 1$. From the definition of $h_{y,a}$ it is clear that

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_y\bot \leftrightarrow \lim h_{y,0} = 1. \tag{6.4}$$

As in the proof of Theorem 5.4.2, we see that there is a model $\mathfrak{M}$ of $\mathrm{I}\Delta_0+\mathsf{exp}$ and a sequence $(m_i)_{i\in\omega}$ of elements of $\mathfrak{M}$, such that

i. $\mathfrak{M} \models \Box_{m_i}\bot$ and $\mathfrak{M} \models m_i > m_{i+1}$ for all i

ii. For all $k \in \mathfrak{M}$, $\mathfrak{M} \models \Box_k\bot$ if and only if for some $i \in \omega$, $m_i < k$

It follows from the above that $\mathfrak{M} \models \Box\bot$, and $\mathfrak{M} \models \Diamond_n\top$ for all standard n.

Let m be any element from $(m_i)_{i\in\omega}$, and suppose for a contradiction that $\ell_m = a$ is witnessed by $s = (l_0, \ldots, l_m)$. If $a = 0$, then by Lemma 6.2.5(ii) also $\ell_{m-1} = 0$; thus $\ell_m = \lim h_{m,0} = 0$. However since $\mathfrak{M} \models \Box_m\bot$, from (6.4) we have that $\lim h_{m,0} = 1$. Thus it must be that $\ell_m = 1$. Let $i \leqslant m$ be the minimal coordinate of s with $l_i = 1 = \ell_i$. Since $\mathfrak{M} \models \Diamond_0\top$, it follows from (6.4) that $i > 0$. Thus $\ell_{i-1} = \lim h_{i-1,0} = 0$, and so $\neg\Box_{i-1}\bot$ by (6.4). On the other hand, $\ell_i = \lim h_{i,0} = 1$, and so $\Box_i\bot$, contradicting the properties of $\mathfrak{M}$.

Write $L = a$ for the formula

$$\exists y\ \big(\ell_y = a \wedge \forall x < y\, \Diamond^{\Pi_1}_x\top\big) \wedge \forall z\ \big(\forall x < z\, \Diamond^{\Pi_1}_x\top \to \ell_z \preceq a\big)$$

stating, intuitively, that $\ell_\mu = a$.

6.2.8. Lemma. *i.* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \exists!a\, L = a$

ii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall x < y\, \Diamond^{\Pi_1}_x\top \to \ell_y \preceq L$

iii. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_z \ell_{\dot z} \preceq L$

iv. $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \Box_z(\Box^{\Pi_1}_{\dot z}\bot \rightarrow L = \ell_{\dot z})$

Proof: (i). Since $\preceq$ is antisymmetric, uniqueness is immediate from the definition. For existence, we show by external induction on the converse of $\prec$ that for all $a \in W$,

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \ell_y = a \wedge \forall x < y\, \Diamond^{\Pi_1}_x \top \rightarrow \exists b\; L = b.$$

We note that this is sufficient, for $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \ell_0 = \lim h_{0,0} \wedge \forall x < 0\, \Diamond^{\Pi_1}_x \top$. Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$. From $\ell_y = a \wedge \forall x < y\, \Diamond^{\Pi_1}_x \top$ we have by Lemmas 6.2.5(ii) and 6.2.6(iii):

$$\forall z > y\; \big(\forall x < z\, \Diamond^{\Pi_1}_x \top \rightarrow \ell_z = a\big) \vee \exists z > y\; \big(\forall x < z\, \Diamond^{\Pi_1}_x \top \wedge a \prec \ell_z\big)\,.$$

If the first disjunct holds, then, using Lemma 6.2.5(ii), we have that $L = a$. And if the second disjunct holds, then $\exists b\; L = b$ follows by the induction assumption.

(ii) Immediate from the definition of L, by using Lemma 6.2.6(iii) to see that ℓ_y exists.

(iii) Within $\Box_z$ we have $\forall x < z\, \Diamond^{\Pi_1}_x \top$ by reflection, and thus $\ell_z \preceq L$ by clause (ii).

(iv) Argue in $\Box_z$, assuming $\Box^{\Pi_1}_z \bot$. By clause (iii) we have that $\ell_z \preceq L$. Suppose for a contradiction that $\ell_z \prec L$. In particular, there is some x with

$$L = \ell_x \wedge \forall y < x\, \Diamond^{\Pi_1}_y \top.$$

Since $\Box^{\Pi_1}_z \bot$, the second conjunct implies that $x \leqslant z$. On the other hand $\ell_z \prec \ell_x$ implies, using Lemma 6.2.5(ii), that $z < x$, a contradiction. ❑

6.2.9. Lemma. *If $a \prec b$, then* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \rightarrow \nabla L = b$.

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$. Assuming $L = a$, we have

$$\forall z\; \big(\forall x < z\, \Diamond^{\Pi_1}_x \top \rightarrow \ell_z \preceq a\big) \tag{6.5}$$

If $\triangle L \neq b$ for some $b \succ a$, then there is some y with $\Box_y L \neq b \wedge \forall x < y\, \Diamond^{\Pi_1}_x \top$. Using (6.5) we have that $\ell_y \preceq a \prec b$. Now $\Box_y$ thinks:

> Suppose that $\ell_y = b$. Since $a \prec b$, we have $b \neq 0$ whence $\Box^{\Pi_1}_y \bot$ by Lemma 6.2.6(i). By Lemma 6.2.8(iv), the latter implies $L = \ell_y$ i.e. $L = b$, a contradiction.

Back in $\mathrm{I}\Delta_0+\mathsf{exp}$, we conclude $\Box_y \ell_y \neq b$, contradicting Lemma 6.2.5(iv). ❑

6.2.10. Lemma. *If $a \neq 0$, then* $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \to \vartriangle a \prec L$.

Proof: Argue in $\mathrm{I}\Delta_0+\mathsf{exp}$. Assume $L = a$, and let y be such that

$$\ell_y = a \wedge \forall x < y \, \Diamond_x^{\Pi_1} \top.$$

It follows from Lemma 6.2.6(iii)–(iv) that $\Box_y a \preceq \ell_y$. Given $a \neq 0$, we have $\Box_y a \neq \ell_y$ by Lemma 6.2.5(v). Combining the above yields $\Box_y a \prec \ell_y$. Since $\Box_y \ell_y \preceq L$ by Lemma 6.2.8(iii), we obtain $\Box_y a \prec L$, whence also $\vartriangle a \prec L$. ❑

6.2.11. Definition. Let $\mathcal{M} = \langle \mathcal{F}, \Vdash \rangle$ be a finite GL-model. The model $\mathcal{M}_0$ is obtained by appending a new root 0 to $\mathcal{M}$; the truth values of propositional formulas at 0 are set arbitrarily. Apply Definition 6.2.1 to $\mathcal{M}_0$, and define the arithmetical * by letting

$$p^* := \bigvee_{\mathcal{M}_0, a \Vdash p} L = a.$$

6.2.12. Lemma. *Let $\mathcal{M}$ and * be as in Definition 6.2.11. For all $B \in \mathcal{L}_\Box$, $a \neq 0$,*

$$\textit{if } \mathcal{M}, a \Vdash B, \textit{ then } \mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \to B^*.$$

Proof: Using Lemmas 6.2.9 and 6.2.10, we prove the claim simultaneously with

$$\text{if } \mathcal{M}, a \Vdash \neg B, \text{ then } \mathrm{I}\Delta_0+\mathsf{exp} \vdash L = a \to \neg B^*$$

by induction on the structure of B. ❑

6.2.13. Lemma. *i. $\mathfrak{N} \models L = 0$, where $\mathfrak{N}$ is the standard model.*

ii. For all $a \neq 0$, $L = a$ is consistent with PA.

Proof: (i) follows from Lemma 6.2.6(i).

(ii) Note that by (i) and Lemma 6.2.9, we have $\mathfrak{N} \models \triangledown L = a$ for all a, whence also $\mathfrak{N} \models \Diamond L = a$ (recall that since $\mathfrak{N} \models \Diamond^{\Pi_1} \top$, we have $\mathfrak{N} \models \forall \varphi\, (\Box \varphi \leftrightarrow \vartriangle \varphi)$). ❑

We prove the remaining direction of Theorem 6.1.1.

Proof: If $\mathsf{GL} \nvdash A$, then by Theorem 2.2.5 there is a finite rooted GL-model $\mathcal{M}$ with $w \nVdash A$ for some w in $\mathcal{M}$. Let * be as in Definition 6.2.11. By Lemma 6.2.12, $\mathrm{I}\Delta_0+\mathsf{exp} \vdash L = w \to \neg A^*$. Since PA does not prove $L \neq w$ by Lemma 6.2.13(ii), it therefore cannot prove A^* either. ❑

Chapter 7

Conclusion

By way of conclusion, let us review the provability predicates studied in this thesis, and point out some connections between them.

7.1 Fast and slow provability

Chapter 4 studied the notions of fast and slow provability. The formula $\triangle_{\mathtt{p}}$ represents provability in the theory PA^* obtained by adding to PA Parikh's rule: "from $\Box\varphi$, infer φ". While Parikh's rule is admissible in PA, it was shown by Parikh ([Par71]) that PA^* has substantial speed-up over PA.

The notion of slow provability predicate for PA was introduced by Friedman, Rathjen, and Weiermann ([FRW13]). They showed that the corresponding slow consistency statement lies strictly between $\top$ and $\Diamond\top$ in terms of interpretability strength (see also Corollary 4.2.16). In fact, the slow consistency statement is the first example of a sentence with this property that is natural, at least in the sense that its existence does not rely on the Fixed Point Lemma. In contrast, the fast consistency statement $\triangledown_{\mathtt{p}}\top$ is strictly stronger in terms of interpretability strength than $\Diamond\top$ (Corollary 4.1.5).

We studied variants of slow provability defined as

$$\triangle_{\langle k\rangle}\varphi := \exists x\ (\Box_{x+k}\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\downarrow),$$

where k is a standard integer. We use $\triangle_{\mathtt{s}}$ as a collective name for these predicates.

It follows from the proof of Solovay's Theorem that GL is the provability logic of both $\triangle_{\mathtt{p}}$ and $\triangle_{\mathtt{s}}$. Lindström ([Lin06]) proved that the joint provability logic of $\triangle_{\mathtt{p}}$ and $\Box$ is GLT. We showed (Theorem 4.3.2) that GLT is also the joint provability logic of $\Box$ and $\triangle_{\mathtt{s}}$. The system GLT thus seems to be suitable for capturing the interaction of two provability predicates, one of which is slower than the other in some sense.

Lindström ([Lin06]) showed that ordinary provability may be seen as an ω-root of Parikh provability: $\mathsf{PA} \vdash \forall\varphi\,(\triangle_{\mathtt{p}}\varphi \leftrightarrow \Box^{\omega}\varphi)$. We have seen (Proposition 4.2.18) that a similar principle characterises the interaction of $\Box$ with $\vartriangle_{\langle k\rangle}$ for $k \geqslant 2$: $\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \vartriangle^{\omega}_{\langle k\rangle}\varphi)$. Pakhomov showed ([HP16]) that for $k \leqslant 1$, $\vartriangle_{\langle k\rangle}$ is an ε_0-root of $\Box$: $\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \vartriangle^{\varepsilon_0}_{\langle k\rangle}\varphi)$. The results about transfinite iterations of slow provability predicates have been independently proven by Freund ([Fre16]).

7.2 Supremum adapters

The supremum adapters studied in Chapter 5 are defined as:

$$\blacktriangle_{\langle k\rangle}\varphi := \exists x\,\left(\Box_{x+k}\varphi \wedge \forall y < x\,\Diamond^{\Pi_1}_{y}\top\right),$$

where k is a standard integer. The reason why $\blacktriangle_{\langle k\rangle}$ is called a supremum adapter is that $\blacktriangledown_{\langle k\rangle}\varphi \wedge \blacktriangledown_{\langle k\rangle}\psi$ is an interpretability supremum of φ and ψ (Theorem 5.4.3):

$$\mathrm{I}\Delta_0+\mathsf{exp} \vdash (\chi \rhd \varphi) \wedge (\chi \rhd \psi) \leftrightarrow \chi \rhd \blacktriangledown_{\langle k\rangle}\varphi \wedge \blacktriangledown_{\langle k\rangle}\psi \tag{7.1}$$

The symbols $\blacktriangle_{\mathsf{f}}$ and $\blacktriangle_{\mathsf{g}}$ are used as collective names for supremum adapters indexed by negative and non-negative integers, respectively.

GL is arithmetically sound with respect to $\blacktriangle_{\mathsf{g}}$ (Lemma 5.5.2), and arithmetically complete with respect to $\blacktriangle_{\langle 0\rangle}$ (Theorem 6.1.1). The joint provability logic of $\Box$ and $\blacktriangle_{\mathsf{g}}$ contains GLT, together with

$$(\mathsf{S}) \qquad \Box(\Box B \to \vartriangle A) \to \Box(\Box B \to A).$$

Whether the joint provability logic of $\Box$ and $\blacktriangle_{\mathsf{g}}$ is GLT together with (S) is, for now, an open question. The modal system F is sound with respect to $\blacktriangle_{\mathsf{f}}$ (Lemma 5.5.4). Determining the provability logic of $\blacktriangle_{\mathsf{f}}$, as well as its joint provability logic with $\Box$, remain challenges for future work.

Consistency statements

It is an immediate consequence of (7.1) that consistency statements corresponding to supremum adapters have the same interpretability degree as $\top$:

$$\blacktriangledown_{\mathsf{g}}\top \equiv \blacktriangledown_{\mathsf{f}}\top \equiv \top.$$

When it comes to logical strength, on the other hand, we need to distinguish between $\blacktriangledown_{\mathsf{g}}\top$ and $\blacktriangledown_{\mathsf{f}}\top$.

Since the modal system F is sound with respect to $\blacktriangle_{\mathsf{f}}$, it is clear that $\mathsf{PA} \vdash \blacktriangledown_{\mathsf{f}}\top$; thus $\blacktriangledown_{\mathsf{f}}\top$ has the same logical strength as $\top$. This implies that $\blacktriangledown_{\mathsf{f}}\top$ is not an Orey sentence: its negation $\blacktriangle_{\mathsf{f}}\bot$ is refutable, whence clearly $\top \not\rhd \blacktriangle_{\mathsf{f}}\bot$.

We have seen (Example 5.2.7) that the Gödel sentence of the Feferman provability predicate $\vartriangle_{\mathsf{f}}$ is an Orey sentence. By a similar argument, it can be shown that the Gödel sentence of $\blacktriangle_{\mathsf{f}}$ is an Orey sentence.

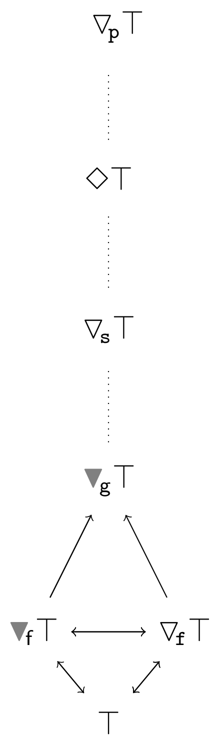

Figure 7.1: The zoo of consistency statements. Arrows stand for the provability ordering; dotted lines separate interpretability degrees.

The provability predicate $\blacktriangle_{\mathtt{g}}$ behaves according to the rules and principles of GL, and so Löb's Theorem holds for it: $\mathsf{PA} \vdash \blacktriangle_{\mathtt{g}}\varphi \to \varphi$ implies $\mathsf{PA} \vdash \varphi$. Given this, it is clear that $\mathsf{PA} \nvdash \blacktriangledown_{\mathtt{g}}\top$.

We claim that $\blacktriangledown_{\mathtt{g}}\top$ is an Orey sentence, i.e. $\mathsf{PA} \vdash \top \rhd \blacktriangle_{\mathtt{g}}\bot$ and $\mathsf{PA} \vdash \top \rhd \blacktriangledown_{\mathtt{g}}\top$. As mentioned above, the latter is an immediate consequence of (7.1). To see that $\mathsf{PA} \vdash \top \rhd \blacktriangle_{\mathtt{g}}\bot$, note that by Löb's axiom, $\mathsf{PA} \vdash \blacktriangledown_{\mathtt{g}}\top \to \blacktriangledown_{\mathtt{g}}\blacktriangle_{\mathtt{g}}\bot$, whence clearly $\mathsf{PA} \vdash \blacktriangledown_{\mathtt{g}}\top \rhd \blacktriangledown_{\mathtt{g}}\blacktriangle_{\mathtt{g}}\bot$. It follows from [Fef60, Theorem 6.2] that $\mathsf{PA} \vdash \blacktriangledown_{\mathtt{g}}\varphi \rhd \varphi$, and so we obtain $\mathsf{PA} \vdash \blacktriangledown_{\mathtt{g}}\top \rhd \blacktriangle_{\mathtt{g}}\bot$. Since $\mathsf{PA} \vdash \top \rhd \blacktriangledown_{\mathtt{g}}\top$, we have $\mathsf{PA} \vdash \top \rhd \blacktriangle_{\mathtt{g}}\bot$ by transitivity of $\rhd$.

Since $\mathsf{PA} \vdash \top \rhd \blacktriangledown_{\mathtt{g}}\top$ but $\mathsf{PA} \nvdash \top \rhd \triangledown_{\mathtt{s}}\top$, it is clear that $\mathsf{PA} \nvdash \blacktriangledown_{\mathtt{g}}\top \to \triangledown_{\mathtt{s}}\top$, i.e. $\mathsf{PA} \nvdash \vartriangle_{\mathtt{s}}\bot \to \blacktriangle_{\mathtt{g}}\bot$. Hence the squiggly arrow in Figure 7.2 is irreversible. Figure 7.1 summarises the relations between our consistency statements.

Transfinite iterations

We have seen (Proposition 5.6.11) that similarly like $\vartriangle_{\langle n\rangle}$ for $n \geqslant 2$, the formula $\blacktriangle_{\langle k\rangle}$, for $k \geqslant 1$, is an ω-root of $\Box$: $\mathsf{PA} \vdash \forall\varphi\,(\Box\varphi \leftrightarrow \blacktriangle^{\omega}_{\langle k\rangle}\varphi)$, and furthermore (Theorem 5.6.5) that $\mathsf{PA} \vdash \forall\varphi\,(\blacktriangle^{\varepsilon_0}_{\langle 0\rangle}\varphi \to \Box\varphi)$. However $\blacktriangle_{\langle 0\rangle}$, unlike $\vartriangle_{\langle 0\rangle}$, is not an ε_0-root of $\Box$: it follows from Theorem 5.6.7 that $\mathsf{PA} \nvdash \forall\varphi\,(\Box\varphi \to \blacktriangle^{\alpha}_{\langle 0\rangle}\varphi)$, whenever α is an element of an arbitrary elementary linear ordering satisfying certain minimal assumptions.

7.2.1. Question. What is the relationship between transfinite iterations of slow provability on the one hand, and those of supremum adapters, on the other?

7.3 Slow provability and supremum adapters

Figure 7.2 gives an overview of our provability predicates. All inclusions except for the one indicated by the squiggly arrow are straightforward. Taking $\bot$ for φ, it follows from the observations made in the previous section (see also Figure 7.1) that all arrows are irreversible. We provide the missing piece of the puzzle.

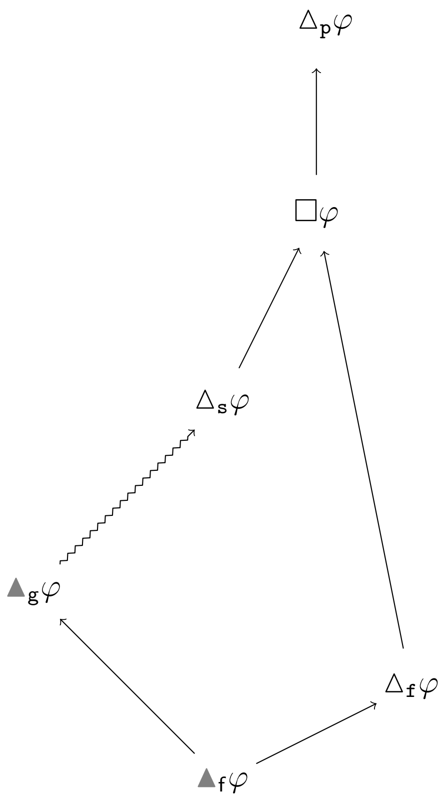

Figure 7.2: The zoo of provability predicates. Arrows indicate provable inclusion in PA. The squiggly arrow indicates provability modulo an index shift.

7.3.1. Remark. Figure 7.2 glosses over the fact that members of the families $\blacktriangle_{\mathtt{f}}$, $\blacktriangle_{\mathtt{g}}$, and $\triangle_{\mathtt{s}}$ are not provably equivalent among themselves. We do have

$\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\,(\vartriangle_{\langle k\rangle}\varphi \to \vartriangle_{\langle k'\rangle}\varphi)$ whenever $k \leqslant k'$, and similarly for $\blacktriangle_{\langle k\rangle}$. However, as witnessed by the behaviour of transfinite iterations of $\blacktriangle_{\langle k\rangle}$ and $\vartriangle_{\langle k\rangle}$ for different values of k, the reverse implication does not hold in general.

The following proposition shows that, modulo an index shift, supremum-adapter proofs can be converted into slow proofs. Recall that the supremum adapter $\blacktriangle_{\langle k\rangle}$ and the slow provability predicate $\vartriangle_{\langle k\rangle}$, with $k \in \mathbb{Z}$, are defined as

$$\blacktriangle_{\langle k\rangle}\varphi := \exists x\ \left(\Box_{x+k}\varphi \wedge \forall y < x\, \Diamond^{\Pi_1}_{y}\top\right)$$
$$\vartriangle_{\langle k\rangle}\varphi := \exists x\ (\Box_{x+k}\varphi \wedge \mathtt{F}_{\varepsilon_0}(x){\downarrow})$$

7.3.2. PROPOSITION. *For all* $k \in \mathbb{Z}$, $\mathrm{I}\Delta_0+\mathsf{exp} \vdash \forall\varphi\left(\blacktriangle_{\langle k\rangle}\varphi \to \vartriangle_{\langle k+2\rangle}\varphi\right)$.

Proof: Let $k \in \mathbb{Z}$. Reason in $\mathrm{I}\Delta_0+\mathsf{exp}$, fixing x and φ with $\Box_{x+k}\varphi \wedge \forall y < x\,\Diamond^{\Pi_1}_{y}\top$. We want to show that $\exists y\ \left(\Box_{y+(k+2)}\varphi \wedge \mathtt{F}_{\varepsilon_0}(y){\downarrow}\right)$. In case $x = 0$ or $x = 1$, this follows by monotonicity, for $\mathtt{F}_{\varepsilon_0}(0){\downarrow}$ by Σ_1-completeness. Assume now that $x \geqslant 2$. Taking $y = x - 2$, we clearly have $\Box_{y+(k+2)}\varphi$. It remains to show $\mathtt{F}_{\varepsilon_0}(x-2){\downarrow}$. Since $\mathtt{F}_{\varepsilon_0}(x-2){\uparrow}$ is a Π_1-formula, its truth would imply $\Box^{\Pi_1}_{x-1}\mathtt{F}_{\varepsilon_0}(x-2){\uparrow}$. By Lemma 4.2.9(ii), we have $\Box_{x-1}\mathtt{F}_{\varepsilon_0}(x-2){\downarrow}$. Combining the above yields $\Box^{\Pi_1}_{x-1}\bot$, contradicting our assumption that $\forall y < x\,\Diamond^{\Pi_1}_{y}\top$. ❑

We recall from Section 5.4.1 that in $\mathrm{I}\Sigma_2$, $\blacktriangle_{\langle k\rangle}$-provability coincides with provability in $\Box_{\mu+k}$, where μ is the least x such that $\Box^{\Pi_1}_{x}\bot$.

Similarly, it is easy to see that $\vartriangle_{\langle k\rangle}$-provability coincides with provability in $\Box_{\nu+k}$, where ν is the greatest x with $\mathtt{F}_{\varepsilon_0}(x){\downarrow}$. If $\mathtt{F}_{\varepsilon_0}{\downarrow\downarrow}$, we let $\nu := \infty$, where $\Box_\infty$ is defined to be $\Box$.

We claim that $\mu \leqslant \nu + 2$. In case $\mathtt{F}_{\varepsilon_0}{\downarrow\downarrow}$, we have $\nu = \infty$, whence it is clear that $\mu \leqslant \nu + 2$. If $\mathtt{F}_{\varepsilon_0}{\uparrow\uparrow}$, then by definition of ν, $\mathtt{F}_{\varepsilon_0}(\nu+1){\uparrow}$. Since the latter is a Π_1-formula and $\Box_{\nu+2}\mathtt{F}_{\varepsilon_0}(\nu+1){\downarrow}$ by Lemma 4.2.9(ii), this implies $\Box^{\Pi_1}_{\nu+2}\bot$.

Since $\mu \leqslant \nu + 2$, it is clear that $\Box_\mu\varphi$ implies $\Box_{\nu+2}\varphi$ and in general that, as stated in Proposition 7.3.2, $\Box_{\mu+k}\varphi$ implies $\Box_{\nu+(k+2)}\varphi$.

Extremely slow provability

Proposition 7.3.2 reveals a certain relationship between supremum adapters and slow provability. The following result, established by Paris ([Par80, Theorem 36]) by means of model-theoretic methods, suggests an even deeper connection between the two:

$$\text{for all } n,\quad \mathrm{I}\Sigma_1 \vdash \Diamond^{\Pi_1}_{n}\top \leftrightarrow \mathtt{F}_{\omega_n}{\downarrow\downarrow}.$$

A proof-theoretic proof of this fact follows from the work of Beklemishev in [Bek03, Theorem 1, Proposition 7.3, Remark 7.4]. It has been shown by Freund that the quantifier ranging over natural numbers can be internalised:

7.3.3. THEOREM ([FRE15]). $\mathrm{I}\Sigma_1 \vdash \forall x\ \left(\Diamond^{\Pi_1}_{x+1}\top \leftrightarrow \mathtt{F}_{\omega_{x+1}}{\downarrow\downarrow}\right)$ ❑

Theorem 7.3.3 yields an alternative characterisation of supremum adapters in terms of the fast-growing hierarchy:

$$I\Sigma_2 \vdash \blacktriangle_{\langle k \rangle}\varphi \leftrightarrow \exists x\, (\Box_{x+k}\varphi \wedge \forall y < x\, \mathtt{F}_{\omega_y}\downarrow\downarrow)$$

Recall that $\triangle_{\langle k \rangle}\varphi = \exists x\, (\Box_{x+k}\varphi \wedge \mathtt{F}_{\varepsilon_0}(x)\downarrow)$. We may thus view $\blacktriangle_{\langle k \rangle}$ as specifying a notion of "extremely slow" provability — it requires certain fast-growing functions to be total, not only to converge on a certain input as does $\triangle_{\langle k \rangle}$.

This thesis may thus be viewed as a study of provability concepts differing from ordinary provability in terms of speed. Parikh provability is a speeded up version of ordinary provability, while both slow provability and supremum adapters are examples of slowed down provability. We solved a number of questions concerning these provability predicates, in particular their behaviour as seen from a modal perspective. However, just as many questions remain open. The latter are listed in Appendix A; our hope is that they give rise to further exciting research.

Appendix A

List of open questions

Modal Logic

The bimodal system GLS is obtained from GLT by replacing axiom ($\mathsf{T4}$) with axiom (S):

$$\Box(\Box B \to \triangle A) \to \Box(\Box B \to A).$$

GLS is arithmetically sound with respect to $\blacktriangle_{\mathtt{g}}$ and $\Box$ (Lemma 5.5.6).

It is easy to see that the axioms of GLS characterise the subclass of GLT-frames satisfying the following property: if wRa, there is a sequence $(a_i)_{i<\omega}$ with $a = a_0$, and for all i: wRa_i, $a_{i+1}Qa_i$, and a_i *has exactly the same* R*-successors as* a. Let us call this class of frames GLS-*frames*.

A.1. Question. Is GLS complete with respect to the class of GLS-frames?

Provability logic

The system GLT is arithmetically sound and complete with respect to slow and ordinary provability, as well as ordinary and fast provability (Theorem 4.3.2). It is thus natural to ask whether it is also the joint provability logic of slow and fast provability:

A.2. Question. Is GLT the joint provability logic of $\triangle_{\mathtt{s}}$ and $\triangle_{\mathtt{p}}$? Consider arithmetical realisations mapping the modalities $\triangle$ and $\Box$ to the provability predicates $\triangle_{\mathtt{s}}$ and $\triangle_{\mathtt{p}}$ respectively. It is clear that such realisations satisfy conditions (1)-(4) of Definition 4.3.1. Do they also satisfy condition (5)? This would follow if the schema $\triangle_{\mathtt{p}}\triangle_{\mathtt{s}}\varphi \to \triangle_{\mathtt{p}}\varphi$ would be provable in PA.

The system GL is arithmetically sound and complete with respect to the supremum adapter $\blacktriangle_{\langle 0\rangle}$ (Theorem 6.1.1).

A.3. Question. Is GL arithmetically complete with respect to each $\blacktriangle_{\mathsf{g}}$?

The modal logic F is arithmetically sound with respect to $\blacktriangle_{\mathsf{f}}$ (Lemma 5.5.4).

A.4. Question. Is F arithmetically complete with respect to $\blacktriangle_{\mathsf{f}}$?

The bimodal system GLS is arithmetically sound with respect to $\Box$ and $\blacktriangle_{\mathsf{g}}$ (Lemma 5.5.6).

A.5. Question. Is GLS arithmetically complete with respect to $\Box$ and $\blacktriangle_{\mathsf{g}}$?

A.6. Question. What is the joint provability logic of $\Box$ and $\blacktriangle_{\mathsf{f}}$?

Transfinite iterations

Provability predicates $\triangle_{\langle k\rangle}$ with $k \leqslant 1$ and $k \geqslant 2$ may be seen, respectively, as ε_0-roots and ω-roots of ordinary provability (Theorem 4.2.19, Proposition 4.2.18). Using similar ideas as in the definition of $\triangle_{\langle k\rangle}$, a square root of ordinary provability can be shown to exist (Theorem 4.2.20).

A.7. Question. Do other roots of ordinary provability exist?

Supremum adapters $\blacktriangle_{\langle k\rangle}$ with $k \geqslant 1$ are also ω-roots of ordinary provability (Proposition 5.6.11). For $\blacktriangle_{\langle 0\rangle}$, it only holds that $\mathsf{PA} \vdash \forall\varphi\,(\blacktriangle^{\varepsilon_0}_{\langle 0\rangle}\varphi \to \Box\varphi)$ (Theorem 5.6.5, Theorem 5.6.7).

A.8. Question. What can be said about transfinite iterations of supremum adapters indexed by negative integers?

We have established (Proposition 7.3.2) the following connection between supremum adapters and slow provability: $\mathsf{PA} \vdash \forall\varphi\,(\blacktriangle_{\langle k\rangle}\varphi \to \triangle_{\langle k+2\rangle}\varphi)$

A.9. Question. What is the relationship between transfinite iterations of slow provability on the one hand, and those of supremum adapters, on the other?

Bibliography

[AB04] S.N. Artemov and L.D. Beklemishev. Provability logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd ed.*, volume 13, pages 229–403. Springer, Dordrecht, 2004.

[Bek92] L. D. Beklemishev. Independent enumerations of theories and recursive progressions. *Siberian Mathematical Journal*, 33(5):760–783, 1992.

[Bek94] L. D. Beklemishev. On bimodal logics of provability. *Annals of Pure and Applied Logic*, 68:115–159, 1994.

[Bek03] L. D. Beklemishev. Proof-theoretic analysis by iterated reflection. *Archive for Mathematical Logic*, 42:515–552, 2003.

[Bek05] L. D. Beklemishev. Reflection principles and provability algebras in formal arithmetic. *Russian Mathematical Surveys*, 60(2):197, 2005.

[Bek11] L. D. Beklemishev. A simplified proof of arithmetical completeness theorem for provability logic GLP. *Proceedings of the Steklov Institute of Mathematics*, 274(1):25–33, 2011.

[Ber90] A. Berarducci. The interpretability logic of Peano arithmetic. *The Journal of Symbolic Logic*, 55:1059–1089, 1990.

[Boo93] G. Boolos. *The Logic of Provability.* Cambridge University Press, Cambridge, 1993.

[BV02] M. Blackburn, P. De Rijke and Y. Venema. *Modal Logic*, volume 53. Cambridge University Press, 2002.

[Car37] R. Carnap. *The Logical Syntax of Language.* London, K. Paul, Trench, Trubner & Co., Ltd., 1937.

[Car86] T. Carlson. Modal logics with several operators and provability interpretations. *Israel Journal of Mathematics*, 54(1):14–24, 1986.

[DBR75] P. Du Bois-Reymond. Über asymptotische Werte, infinitäre Approximationen und infinitäre Auflösungen von Gleichungen. *Mathematische Annalen*, 8:363–414, 1875.

[DJ88] G. Dzhaparidze (Japaridze). The polymodal logic of provability. *Intensional Logics and Logical Structure of Theories*, pages 16–48, 1988. In Russian.

[dJV90] D.H.J. de Jongh and F. Veltman. Provability logics for relative interpretability. In *Mathematical logic, Proceedings of the Heyting 1988 summer school in Varna, Bulgaria*, pages 31–42. Plenum Press, Boston, 1990.

[EF60] A. Ehrenfeucht and S. Feferman. Representability of recursively enumerable sets in formal theories. *Archiv für Mathematische Logik und Grundlagenforschung*, 5(1):37–41, 1960.

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

[Fef62] S. Feferman. Transfinite recursive progressions of axiomatic theories. *The Journal of Symbolic Logic*, 27(3):259–316, 1962.

[Fre15] A. Freund. A uniform characterization of Σ_1-reflection over the fragments of Peano arithmetic. *arXiv preprint arXiv:1512.05122*, 2015.

[Fre16] A. Freund. Slow reflection. *arXiv preprint arXiv: 1601.08214*, 2016.

[FRW13] S.D. Friedman, M. Rathjen, and A. Weiermann. Slow consistency. *Annals of Pure and Applied Logic*, 164(3):382–393, 2013.

[GD82] H. Gaifman and C. Dimitracopoulos. Fragments of Peano's Arithmetic and the MRDP theorem. *Logic and Algorithmic, Monographie No. 30 de L'Enseignement Mathematique*, 3:187–206, 1982.

[Gen43] G. Gentzen. Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie. *Mathematische Annalen*, 119(1):140–161, 1943.

[GJ08] E. Goris and J.J. Joosten. Modal matters in interpretability logic. *Logic Journal of IGPL*, 16(4):371–412, 2008.

[Grz53] A. Grzegorczyk. Some classes of recursive functions. *Rozprawy Matematyczne*, 4:3–45, 1953.

[Gua79] D. Guaspari. Partially conservative extensions of arithmetic. *Transactions of the American Mathematical Society*, 254:47–68, 1979.

[Háj71] P. Hájek. On interpretability in set theories I. *Comm. Math. Univ. Carolinae*, 12:73–79, 1971.

[HH72] M. Hájková and P. Hájek. On interpretability in theories containing arithmetic. *Fundamenta Mathematicae*, 76(2):131–137, 1972.

[HP93] P. Hájek and P. Pudlák. *Metamathematics of First-Order Arithmetic*. Springer-Verlag, 1993.

[HP16] P. Henk and F. Pakhomov. Slow and ordinary provability for Peano Arithmetic. *ArXiv e-prints*, February 2016.

[HS16] P. Henk and V.Yu. Shavrukov. A Solovay function for the least 1-inconsistent subtheory of PA. ILLC Prepublication Series PP-2016-31, ILLC, University of Amsterdam, August 2016.

[HV16] P. Henk and A. Visser. Interpretability Suprema in Peano Arithmetic. ILLC Prepublication Series PP-2016-32, ILLC, University of Amsterdam, August 2016.

[Ign93] K.N. Ignatiev. On strong provability predicates and the associated modal logics. *Journal of Symbolic Logic*, 58(1):249–290, 1993.

[JdJ98] G. Japaridze and D. de Jongh. The logic of provability. In S. Buss, editor, *Handbook of proof theory*, pages 475–546. North-Holland Publishing Co., Amsterdam, 1998.

[Jer71] R. G. Jeroslow. Consistency statements in formal theories. *Fundamenta Mathematicae*, 72:17–40, 1971.

[Jer75] R. G. Jeroslow. Experimental logics and Δ^0_2-theories. *Journal of Philosophical Logic*, 4:253–267, 1975.

[KS81] J. Ketonen and R. Solovay. Rapidly growing Ramsey functions. *Annals of Mathematics*, 113:267–314, 1981.

[Lin79] P. Lindström. Some results on interpretability. In F.V. Jensen, Mayoh B.H., and Moller K.K., editors, *Proceedings of the 5th Scandinavian Logic Symposium 1979*, pages 329–361, Aalborg, 1979. Aalborg University Press.

[Lin84] P. Lindström. On certain lattices of degrees of interpretability. *Notre Dame J. Formal Logic*, 25(2):127–140, 04 1984.

[Lin06] P. Lindström. On Parikh provability - an exercise in modal logic. In H. Lagerlund, S. Lindström, and R. Sliwinski, editors, *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg*, Uppsala Philosophical Studies 53, pages 279–288, 2006.

[Löb55] M.H. Löb. Solution of a problem of Leon Henkin. *Journal of Symbolic Logic*, 20:115–118, 1955.

[LW70a] M.H. Löb and S.S. Wainer. Hierarchies of number-theoretic functions I. *Archiv für mathematische Logik und Grundlagenforschung*, 13(1-2):39–51, 1970.

[LW70b] M.H. Löb and S.S. Wainer. Hierarchies of number-theoretic functions II. *Archiv für mathematische Logik und Grundlagenforschung*, 13(3-4):97–113, 1970.

[McA78] K. McAloon. Completeness theorems, incompleteness theorems and models of arithmetic. *Transactions of the American Mathematical Society*, 239:253–277, 1978.

[Min73] G.E. Mints. Exact estimates of the provability of transfinite induction in the initial segments of arithmetic. *Journal of Soviet Mathematics*, 1(1):85–91, 1973.

[Mon62] R. Montague. Theories incomparable with respect to relative interpretability. *J. Symbolic Logic*, 27(2):195–211, 06 1962.

[Ore61] S. Orey. Relative interpretations. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 7:146–153, 1961.

[Par70] C. Parsons. On a number-theoretic choice schema and its relation to induction. In *Intuitionism and Proof Theory: proceedings of the summer conference at Buffalo N.Y. 1968*, Amsterdam, North–Hollan, pages 459–473, 1970.

[Par71] R. Parikh. Existence and feasibility in arithmetic. *The Journal of Symbolic Logic*, 36(3):494–508, 1971.

[Par80] J.B. Paris. A hierarchy of cuts in models of arithmetic. In Leszek Pacholski, Jedrzej Wierzejewski, and AlecJ. Wilkie, editors, *Model Theory of Algebra and Arithmetic*, volume 834 of *Lecture Notes in Mathematics*, pages 312–337. Springer Berlin Heidelberg, 1980.

[Put65] H. Putnam. Trial and error predicates and a solution to a problem of Mostowski. *The Journal of Symbolic Logic*, 30:146–153, 1965.

[Ros84] H.E. Rose. *Subrecursion: Functions and Hierarchies.* Clarendon Press, Oxford, 1984.

[Sch79] U.R. Schmerl. A fine structure generated by reflection formulas over Primitive Recursive Arithmetic. In M. Boffa, D. van Dalen, and K. McAloon, editors, *Logic Colloquium '78*, Studies in Logic, pages 335–350. North Holland, Amsterdam, 1979.

[Sha88] V.Yu. Shavrukov. The logic of relative interpretability over Peano arithmetic (in Russian). Technical Report No.5, Steklov Mathematical Institute, Moscow, 1988.

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

[Sim09] S.G. Simpson. *Subsystems of Second Order Arithmetic.* Perspectives in Logic. Association for Symbolic Logic, Cambridge University Press, Cambridge, New York, 2009. Second Edition.

[Smo85] C. Smoryński. *Self-Reference and Modal Logic.* Universitext. Springer, New York, 1985.

[Sol76] R.M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 25:287–304, 1976.

[Som93] R. Sommer. Ordinals in bounded arithmetic. In P. Clote and J. Krajiček, editors, *Arithmetic, Proof Theory and Computational Complexity*, pages 185–196. Clarendon Press, Oxford, 1993.

[Som95] R. Sommer. Transfinite induction within Peano Arithmetic. *Annals of Pure and Applied Logic*, 76(3):231 – 289, 1995.

[Šve78] V. Švejdar. Degrees of interpretability. *Commentationes Mathematicae Universitatis Carolinae*, 19:789–813, 1978.

[SW12] H. Schwichtenberg and S. S. Wainer. *Proofs and Computations.* Cambridge University Press, New York, NY, USA, 1st edition, 2012.

[Tak75] G. Takeuti. *Proof Theory.* Studies in Logic and the Foundations of Mathematics, vol. 81. North-Holland, Amsterdam, 1975.

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

[Tur39] A.M. Turing. Systems of logic based on ordinals. *Proceedings of the London Mathematical Society*, 2(1):161–228, 1939.

[Vis84] A. Visser. The provability logics of recursively enumerable theories extending Peano arithmetic at arbitrary theories extending Peano arithmetic. *Journal of Philosophical Logic*, 13:97–113, 1984.

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

[Vis91] A. Visser. The formalization of interpretability. *Studia Logica*, 51:81–105, 1991.

[Vis92] A. Visser. An inside view of EXP. *The Journal of Symbolic Logic*, 57:131–165, 1992.

[Vis93] A. Visser. A course on bimodal provability logic. *Annals of Pure and Applied Logic*, pages 109–142, 1993.

[Vis98] A. Visser. An Overview of Interpretability Logic. In M. Kracht, M. de Rijke, H. Wansing, and M. Zakharyaschev, editors, *Advances in Modal Logic, vol 1*, CSLI Lecture Notes, no. 87, pages 307–359. Center for the Study of Language and Information, Stanford, 1998.

[Vis15a] A. Visser. Oracle bites theory. Logic Group Preprint Series 324, Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, `http://www.phil.uu.nl/preprints/lgps/`, 2015.

[Vis15b] A. Visser. The semantic completion of Peano Arithmetic is Π_1-complete. Unpublished note, 2015.

[Wol98] F. Wolter. All finitely axiomatizable subframe logics containing the provability logic CSM_0 are decidable. *Archive for Mathematical Logic*, 37(3):167–182, 1998.

Samenvatting

Dit proefschrift behelst een studie van niet-standaard bewijsbaarheidspredikaten voor de Peano-Rekenkunde (PA). Onder een niet-standaard bewijsbaarheidspredikaat verstaan we een bewijsbaarheidspredikaat van een theorie die samenvalt met PA vanuit extern standpunt, maar niet op een wijze die verifieerbaar is in PA. Ons perspectief wordt vormgegeven door de modale logica: het doel is te bepalen welke modale principes het gedrag van onze niet-standaard bewijsbaarheidspredikaten in PA beschrijven.

Hoofdstuk 3 behandelt de bimodale bewijsbaarheidslogica GLT. We laten zien dat GLT zonder dat het de eindige model eigenschap heeft beslisbaar is, en volledig met betrekking tot diverse natuurlijke klassen Kripke-frames. We geven een karakterisering van het gesloten fragment van GLT, en tonen de aritmetische volledigheid aan met betrekking tot een ruime klasse van bewijsbaarheidspredikaten.

Hoofdstuk 4 gaat over theorieën die verkregen zijn uit PA door het versnellen of vertragen van gewone bewijsbaarheid. We laten zien dat GLT aritmetisch volledig is met betrekking tot een ruime klasse van bewijsbaarheidspredikaten, waaronder zowel gewone en snelle als langzame en gewone bewijsbaarheid.

Hoofdstuk 5 bestudeert zogenaamde supremum-adapters. Deze bewijsbaarheidspredikaten worden gebruikt bij het verkrijgen van interpreteerbaarheidssuprema van enkele uitbreidingen van PA. We bespreken eerst de methodologische problemen die het verrijken van de interpreteerbaarheidslogica ILM van PA met supremum-operatoren met zich meebrengt. De supremum-adapters verschaffen een passende oplossing voor die vraagstukken. We stellen een aantal modale principes vast van deze operatoren, en bestuderen het gedrag van transfiniete iteraties ervan.

In Hoofdstuk 6 wordt aangetoond dat de bewijsbaarheidslogica van een bepaalde supremum-adapter de Gödel-Löb bewijsbaarheidslogica GL is.

Tenslotte, in Hoofdstuk 7, leggen we een aantal verbanden tussen de supremum-adapters en de langzame bewijsbaarheidspredikaten.

Abstract

This thesis is a study of nonstandard provability predicates for Peano Arithmetic (PA). By a *nonstandard provability predicate*, we mean the provability predicate of a theory that coincides with PA from the external point of view, however not verifiably in PA. Our perspective is shaped by modal logic: the goal is to determine which modal principles govern the behaviour of our nonstandard provability predicates in PA.

Chapter 3 deals with the bimodal provability logic GLT. We show that GLT — while lacking the finite model property — is decidable, and complete with respect to several natural classes of Kripke frames. We provide a characterisation of the closed fragment of GLT, and establish its arithmetical completeness with respect to a wide class of provability predicates.

Chapter 4 is concerned with theories obtained from PA by speeding up or slowing down ordinary provability. We show that GLT is arithmetically complete with respect to a wide class of provability predicates, including ordinary and fast, as well as slow and ordinary provability.

Chapter 5 studies the so-called *supremum adapters*. These provability predicates are useful for obtaining interpretability suprema of finite extensions of PA. We first discuss methodological issues arising from the enterprise of adding supremum operators to the interpretability logic ILM of PA. The supremum adapters provide a convenient solution. We establish some modal principles for these operators, and study the behaviour of their transfinite iterations.

In Chapter 6 it is shown that the provability logic of a certain supremum adapter is the Gödel-Löb provability logic GL.

Finally, in Chapter 7 we establish some connections between the supremum adapters and the slow provability predicates.

Titles in the ILLC Dissertation Series:

ILLC DS-2009-01: **Jakub Szymanik**
Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language

ILLC DS-2009-02: **Hartmut Fitz**
Neural Syntax

ILLC DS-2009-03: **Brian Thomas Semmes**
A Game for the Borel Functions

ILLC DS-2009-04: **Sara L. Uckelman**
Modalities in Medieval Logic

ILLC DS-2009-05: **Andreas Witzel**
Knowledge and Games: Theory and Implementation

ILLC DS-2009-06: **Chantal Bax**
Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.

ILLC DS-2009-07: **Kata Balogh**
Theme with Variations. A Context-based Analysis of Focus

ILLC DS-2009-08: **Tomohiro Hoshi**
Epistemic Dynamics and Protocol Information

ILLC DS-2009-09: **Olivia Ladinig**
Temporal expectations and their violations

ILLC DS-2009-10: **Tikitu de Jager**
"Now that you mention it, I wonder...": Awareness, Attention, Assumption

ILLC DS-2009-11: **Michael Franke**
Signal to Act: Game Theory in Pragmatics

ILLC DS-2009-12: **Joel Uckelman**
More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains

ILLC DS-2009-13: **Stefan Bold**
Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.

ILLC DS-2010-01: **Reut Tsarfaty**
Relational-Realizational Parsing

ILLC DS-2010-02: **Jonathan Zvesper**
Playing with Information

ILLC DS-2010-03: **Cédric Dégremont**
The Temporal Mind. Observations on the logic of belief change in interactive systems

ILLC DS-2010-04: **Daisuke Ikegami**
Games in Set Theory and Logic

ILLC DS-2010-05: **Jarmo Kontinen**
Coherence and Complexity in Fragments of Dependence Logic

ILLC DS-2010-06: **Yanjing Wang**
Epistemic Modelling and Protocol Dynamics

ILLC DS-2010-07: **Marc Staudacher**
Use theories of meaning between conventions and social norms

ILLC DS-2010-08: **Amélie Gheerbrant**
Fixed-Point Logics on Trees

ILLC DS-2010-09: **Gaëlle Fontaine**
Modal Fixpoint Logic: Some Model Theoretic Questions

ILLC DS-2010-10: **Jacob Vosmaer**
Logic, Algebra and Topology. Investigations into canonical extensions, duality theory and point-free topology.

ILLC DS-2010-11: **Nina Gierasimczuk**
Knowing One's Limits. Logical Analysis of Inductive Inference

ILLC DS-2010-12: **Martin Mose Bentzen**
Stit, Iit, and Deontic Logic for Action Types

ILLC DS-2011-01: **Wouter M. Koolen**
Combining Strategies Efficiently: High-Quality Decisions from Conflicting Advice

ILLC DS-2011-02: **Fernando Raymundo Velazquez-Quesada**
Small steps in dynamics of information

ILLC DS-2011-03: **Marijn Koolen**
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

ILLC DS-2011-04: **Junte Zhang**
System Evaluation of Archival Description and Access

ILLC DS-2011-05: **Lauri Keskinen**
Characterizing All Models in Infinite Cardinalities

ILLC DS-2011-06: **Rianne Kaptein**
Effective Focused Retrieval by Exploiting Query Context and Document Structure

ILLC DS-2011-07: **Jop Briët**
Grothendieck Inequalities, Nonlocal Games and Optimization

ILLC DS-2011-08: **Stefan Minica**
Dynamic Logic of Questions

ILLC DS-2011-09: **Raul Andres Leal**
Modalities Through the Looking Glass: A study on coalgebraic modal logic and their applications

ILLC DS-2011-10: **Lena Kurzen**
Complexity in Interaction

ILLC DS-2011-11: **Gideon Borensztajn**
The neural basis of structure in language

ILLC DS-2012-01: **Federico Sangati**
Decomposing and Regenerating Syntactic Trees

ILLC DS-2012-02: **Markos Mylonakis**
Learning the Latent Structure of Translation

ILLC DS-2012-03: **Edgar José Andrade Lotero**
Models of Language: Towards a practice-based account of information in natural language

ILLC DS-2012-04: **Yurii Khomskii**
Regularity Properties and Definability in the Real Number Continuum: idealized forcing, polarized partitions, Hausdorff gaps and mad families in the projective hierarchy.

ILLC DS-2012-05: **David García Soriano**
Query-Efficient Computation in Property Testing and Learning Theory

ILLC DS-2012-06: **Dimitris Gakis**
Contextual Metaphilosophy - The Case of Wittgenstein

ILLC DS-2012-07: **Pietro Galliani**
The Dynamics of Imperfect Information

ILLC DS-2012-08: **Umberto Grandi**
Binary Aggregation with Integrity Constraints

ILLC DS-2012-09: **Wesley Halcrow Holliday**
Knowing What Follows: Epistemic Closure and Epistemic Logic

ILLC DS-2012-10: **Jeremy Meyers**
Locations, Bodies, and Sets: A model theoretic investigation into nominalistic mereologies

ILLC DS-2012-11: **Floor Sietsma**
Logics of Communication and Knowledge

ILLC DS-2012-12: **Joris Dormans**
Engineering emergence: applied theory for game design

ILLC DS-2013-01: **Simon Pauw**
Size Matters: Grounding Quantifiers in Spatial Perception

ILLC DS-2013-02: **Virginie Fiutek**
Playing with Knowledge and Belief

ILLC DS-2013-03: **Giannicola Scarpa**
Quantum entanglement in non-local games, graph parameters and zero-error information theory

ILLC DS-2014-01: **Machiel Keestra**
Sculpting the Space of Actions. Explaining Human Action by Integrating Intentions and Mechanisms

ILLC DS-2014-02: **Thomas Icard**
The Algorithmic Mind: A Study of Inference in Action

ILLC DS-2014-03: **Harald A. Bastiaanse**
Very, Many, Small, Penguins

ILLC DS-2014-04: **Ben Rodenhäuser**
A Matter of Trust: Dynamic Attitudes in Epistemic Logic

ILLC DS-2015-01: **María Inés Crespo**
Affecting Meaning. Subjectivity and evaluativity in gradable adjectives.

ILLC DS-2015-02: **Mathias Winther Madsen**
The Kid, the Clerk, and the Gambler - Critical Studies in Statistics and Cognitive Science

ILLC DS-2015-03: **Shengyang Zhong**
Orthogonality and Quantum Geometry: Towards a Relational Reconstruction of Quantum Theory

ILLC DS-2015-04: **Sumit Sourabh**
Correspondence and Canonicity in Non-Classical Logic

ILLC DS-2015-05: **Facundo Carreiro**
Fragments of Fixpoint Logics: Automata and Expressiveness

ILLC DS-2016-01: **Ivano A. Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

ILLC DS-2016-04: **Johannes Marti**
Interpreting Linguistic Behavior with Possible World Models

ILLC DS-2016-05: **Phong Lê**
Learning Vector Representations for Sentences - The Recursive Deep Learning Approach

ILLC DS-2016-06: **Gideon Maillette de Buy Wenniger**
Aligning the Foundations of Hierarchical Statistical Machine Translation

ILLC DS-2016-07: **Andreas van Cranenburgh**
Rich Statistical Parsing and Literary Language

ILLC DS-2016-08: **Florian Speelman**
Position-based Quantum Cryptography and Catalytic Computation

ILLC DS-2016-09: **Teresa Piovesan**
Quantum entanglement: insights via graph parameters and conic optimization