A simplified proof of arithmetical completeness theorem for provability logic **GLP**

L. Beklemishev*

Steklov Mathematical Institute
Gubkina str. 8, 119991 Moscow, Russia
e-mail: bekl@mi.ras.ru

March 11, 2011

Abstract

We present a simplified proof of Japaridze's arithmetical completeness theorem for the well-known polymodal provability logic **GLP**. The simplification is achieved by employing a fragment **J** of **GLP** that enjoys a more convenient Kripke-style semantics than the logic considered in the papers by Ignatiev and Boolos. In particular, this allows us to simplify the arithmetical fixed point construction and to bring it closer to the standard construction due to Solovay.

UDK: 510.652+510.643
Key words: Provability logic, formal arithmetic

1 Basic notions

Logics GLP and J. The polymodal provability logic **GLP** is formulated in the language of the propositional calculus enriched by new unary connectives [0], [1], [2], ..., called *modalities*. The dual connectives $\langle n\rangle$, for all $n \in \mathbb{N}$, are treated as abbreviations: $\langle n\rangle\varphi$ means $\neg[n]\neg\varphi$.

The logic **GLP** is given by the following axiom schemata and inference rules.

Axiom schemata: (i) Tautologies of classical propositional calculus;

(ii) $[n](\varphi \to \psi) \to ([n]\varphi \to [n]\psi)$;

(iii) $[n]([n]\varphi \to \varphi) \to [n]\varphi$ (Löb's axiom);

(iv) $\langle m\rangle\varphi \to [n]\langle m\rangle\varphi$, for $m < n$;

(v) $[m]\varphi \to [n]\varphi$, for $m \leqslant n$ (monotonicity axiom);

Inference rules: modus ponens, $\vdash \varphi \ \Rightarrow \vdash [n]\varphi$.

*Supported by the Russian Foundation for Basic Research, Russian Presidential Council for Support of Leading Scientific Schools, and the STCP-CH-RU project "Computational proof theory".

Thus, for each modality $[n]$ the logic **GLP** contains the axioms (i)–(iii) and the inference rules of Gödel–Löb provability logic **GL**, whereas the schemata (iv) and (v) tie together different modalities. A well-known consequence of the axioms (i)–(iii) is the *transitivity schema*

$$[n]\varphi \to [n][n]\varphi.$$

System **J** is obtained from **GLP** by replacing the monotonicity axiom (v) by the following two axiom schemata easily derivable in **GLP** using the transitivity schema:

(vi) $[m]\varphi \to [n][m]\varphi$, for $m < n$;

(vii) $[m]\varphi \to [m][n]\varphi$, for $m < n$.

Unlike **GLP**, the logic **J** is complete w.r.t. a natural class of finite Kripke frames described below.

Kripke frames. A *Kripke frame* for the language of polymodal logic is a structure $(W; R_0, R_1, \ldots)$, where W is a nonempty set, and R_k are binary relations on W. The elements of W are usually called the *worlds.* A frame is called finite if W is finite and $R_k = \varnothing$, for all but finitely many $k \geqslant 0$.

A *valuation* v on a frame W maps every propositional variable p to a subset $v(p) \subseteq W$ called the *truth set of* p. A *Kripke model* is a Kripke frame together with a valuation on it.

Let $\mathcal{W} = (W, v)$ be a Kripke model. By induction on the build-up of φ we define a relation φ *is valid in a world* x *of* $\mathcal{W}$ (denoted $\mathcal{W}, x \vDash \varphi$).

1. $\mathcal{W}, x \vDash p \iff x \in v(p)$, if p is a variable;
2. $\mathcal{W}, x \vDash (\varphi \wedge \psi) \iff (\mathcal{W}, x \vDash \varphi$ and $\mathcal{W}, x \vDash \psi)$,

   $\mathcal{W}, x \vDash \neg\varphi \iff \mathcal{W}, x \nvDash \varphi$,

   and similarly for the other boolean connectives;
3. $\mathcal{W}, x \vDash [n]\varphi \iff \forall y \in W\ (xR_n y \Rightarrow \mathcal{W}, y \vDash \varphi)$.

Write $\mathcal{W} \vDash \varphi$ if $\forall x \in W\ \mathcal{W}, x \vDash \varphi$.

For the axioms of **J** to be valid in a given Kripke frame for any valuation of variables, we impose some restrictions on the relations R_k. A transitive binary relation R on a set W is called *noetherian*, if there is no infinite chain of elements of W of the form $a_0 R a_1 R a_2 \ldots$ Note that if W if finite the condition of R being noetherian on W is equivalent to its irreflexivity.

A frame $(W; R_0, R_1, \ldots)$ is called a **J**-*frame* if

1. R_k is transitive and noetherian, for all $k \geqslant 0$;
2. $\forall x, y\ (xR_n y \Rightarrow \forall z\ (xR_m z \Leftrightarrow yR_m z))$, for $m < n$;
3. $\forall x, y, z\ (xR_m y \ \ yR_n z \Rightarrow xR_m z)$, for $m < n$.

A **J**-*model* is any Kripke model based on a **J**-frame. By induction on the derivation length one can easily prove the following lemma.

Lemma 1.1 *For any formula φ, if $\mathbf{J} \vdash \varphi$ then $\mathcal{W} \vDash \varphi$, for any $\mathbf{J}$-model $\mathcal{W}$.*

The converse statement, that is, the completeness theorem for **J** with respect to the class of all (finite) **J**-models, is proved in ref. [2].

Proposition 1.2 *For any formula φ, if $\mathbf{J} \nvdash \varphi$ then there is a finite $\mathbf{J}$-model $\mathcal{W}$ such that $\mathcal{W} \nvDash \varphi$.*

Let us call a *root* of **J**-frame $(W; R_0, R_1, \dots)$ an element $r \in W$ for which

$$\forall x \in W \; \exists k \geqslant 0 \; r R_k x.$$

The standard reasoning shows that Proposition 1.2 can be strengthened to the following statement.

Proposition 1.3 *For any formula φ, if $\mathbf{J} \nvdash \varphi$ then there is a finite $\mathbf{J}$-model $\mathcal{W}$ with a root r such that $\mathcal{W}, r \nvDash \varphi$.*

Proof. Assume $\mathbf{J} \nvdash \varphi$ and consider a **J**-model $\mathcal{W}$ and a point $x_0 \in W$ such that $\mathcal{W}, x_0 \nvDash \varphi$. Let $\mathcal{W}_0$ denote a submodel generated by x_0, that is, the restriction of $\mathcal{W}$ to the subset W_0 consisting of all points $y \in W$ reachable from x_0 by moving along the relations R_k (including x_0 itself). More formally, we set $y \in W_0$ if there is a finite sequence of elements of W of the form

$$x_0 R_{n_0} x_1 R_{n_1} x_2 R_{n_3} \dots R_{n_k} x_{k+1} = y.$$

It is not difficult to check the following properties.

1. $\mathcal{W}_0$ is a **J**-model;
2. $\forall x \in W_0 \; (\mathcal{W}, x \vDash \psi \iff \mathcal{W}_0, x \vDash \psi)$, for any formula ψ;
3. $\forall y \in W_0 \setminus \{x_0\} \; \exists k \; x_0 R_k y$.

The last claim is easy to verify by induction on the number of steps from x_0 to y using the following property of **J**-frames:

$$u R_m v R_n z \;\Rightarrow\; u R_{\min(m,n)} z.$$

Thus, x_0 is a root of $\mathcal{W}_0$ and $\mathcal{W}_0, x_0 \nvDash \varphi$. ⊠

Formal arithmetic and provability predicates. First order theories in the language $(0, ', +, \cdot, =)$ containing the axioms of Peano arithmetic PA will be called *arithmetical*. It is known that one can introduce terms for all primitive recursive functions into the language of PA. Δ_0 will denote the class of arithmetical formulas having only bounded occurrences of quantifiers, that is, occurrences of the form

$$\forall x{\leqslant}t\,\varphi \overset{\text{def}}{\Longleftrightarrow} \forall x\,(x \leqslant t \to \varphi),$$
$$\exists x{\leqslant}t\,\varphi \overset{\text{def}}{\Longleftrightarrow} \exists x\,(x \leqslant t \wedge \varphi),$$

where the term t does not contain the variable x.

The classes of Σ_n and Π_n-formulas are defined by induction: Δ_0-formulas are considered as both Σ_0 and Π_0-formulas. Σ_{n+1}*-formulas* are those of the form $\exists\vec{x}\ \varphi(\vec{x},\vec{y})$, where φ is a Π_n-formula; Π_{n+1}*-formulas* are those of the form $\forall\vec{x}\,\varphi(\vec{x},\vec{y})$, where φ is a Σ_n-formula.

We assume some fixed standard primitive recursive gödel numbering of the language of arithmetic. The gödel number of an object τ (term, formula, etc.) is denoted $\ulcorner\tau\urcorner$. We will also be reasoning about gödel numbers of terms and formulas in the formal context of PA. We will use the following abbreviations.

A natural number n will be denoted by the numeral $\overline{n} = 0'\ldots'$ (n primes) in the language of PA; for a given formula φ, the expression $\ulcorner\varphi\urcorner$ will be understood as the corresponding numeral $\overline{\ulcorner\varphi\urcorner}$.

We will also consider within PA primitive recursive families of formulas φ_n depending on a parameter $n \in \mathbb{N}$. In this case, the expression $\ulcorner\varphi_x\urcorner$ will be understood as a primitive recursive definable term (with a free variable x) whose value for a given n is the gödel number of φ_n. In particular, $\ulcorner\varphi(\overline{x})\urcorner$ is a term for the function that, given n, computes the gödel number of the result of substitution of $\overline{n}$ into φ. Following Boolos [3], we shall use $\varphi[\psi]$ as an abbreviation for $\varphi(\ulcorner\psi\urcorner)$, and $\varphi[\psi_x]$ as that for $\varphi(\ulcorner\psi_x\urcorner)$. The intuitive meaning of these expressions is that the formula ψ (respectively, ψ_x) satisfies the property expressed by φ.

It will also be convenient for us to assume that a new sort of variables $\alpha, \beta, \ldots$ ranging over the set of gödel numbers of arithmetical formulas is introduced into the language of arithmetic. Formulas containing variables α, β, ... are treated as the abbreviations for their natural translations into the standard language of PA. We will use the abbreviations $\ulcorner\varphi_\alpha\urcorner$ and $\varphi[\psi_\alpha]$ for families of formulas parametrized by the new sort of variables in the same sense as for the numerical parameters.

Let T be an arithmetical theory, and let $\mathsf{Prov}(\alpha)$ be an arithmetical formula with a single free variable α. Following [4], we call Prov a *provability predicate of level n over T* if, for any arithmetical sentences φ, ψ,

1. $\mathsf{Prov} \in \Sigma_{n+1}$;
2. $T \vdash \varphi$ implies $T \vdash \mathsf{Prov}[\varphi]$;
3. $T \vdash \mathsf{Prov}[\varphi \to \psi] \to (\mathsf{Prov}[\varphi] \to \mathsf{Prov}[\psi])$;

4. $\varphi \in \Sigma_{n+1}$ implies $T \vdash \varphi \to \mathsf{Prov}[\varphi]$.

Notice that Conditions 1 and 4 imply the so-called *third condition of Löb*:

$$T \vdash \mathsf{Prov}[\varphi] \to \mathsf{Prov}[\mathsf{Prov}[\varphi]].$$

Recall that an arithmetical theory T is called *sound*, if $T \vdash \varphi$ implies $\mathbb{N} \vDash \varphi$, for each arithmetical sentence φ. Similarly, a provability predicate Prov is called *sound*, if $\mathbb{N} \vDash \mathsf{Prov}[\varphi]$ implies $\mathbb{N} \vDash \varphi$.

If a theory T is sound, then by Conditions 2 and 3 the set

$$U = \{\varphi : \mathbb{N} \vDash \mathsf{Prov}[\varphi]\}$$

is a deductively closed set of sentences, containing T. One can consider Prov as a formula expressing the provability in a (generally, non-r.e.) theory U.

A sequence π of formulas Prov_0, Prov_1, ... is called a *strong sequence of provability predicates over* T, if there is a sequence of natural numbers $r_0 < r_1 < r_2 < \cdots$ such that, for any $n \geqslant 0$,

1. Prov_n is a provability predicate of level r_n over T;

2. $T \vdash \mathsf{Prov}_n[\varphi] \to \mathsf{Prov}_{n+1}[\varphi]$, for any sentence φ.

We remind two standard examples of strong sequences of provability predicates over PA.

It is known that, for each $n \geqslant 1$, there is a Π_n-truthdefinition for the class of Π_n-formulas in PA, that is, an arithmetical Π_n-formula $\mathsf{True}_{\Pi_n}(\alpha)$ expressing in a natural way the predicate "*α is the gödel number of a true arithmetical Π_n-sentence.*"

Let $\mathsf{Prov}_{\mathsf{PA}}$ denote the usual gödelian provability predicate for PA (of level 0). Define $\mathsf{Prov}_0 = \mathsf{Prov}_{\mathsf{PA}}$ and $\mathsf{Prov}_n(\alpha) = \exists\beta\, (\mathsf{True}_{\Pi_n}[\beta] \land \mathsf{Prov}_0[\beta \to \alpha])$, for $n > 0$. Notice that $\mathsf{Prov}_n(\alpha)$ expresses that α is provable in the theory axiomatized over PA by the set of all true Π_n-sentences; this predicate has level n.

Another strong sequence of provability predicates is defined by the closure of PA under the n-fold application of the ω-rule. Define $\mathsf{Prov}_0 = \mathsf{Prov}_{\mathsf{PA}}$ and

$$\mathsf{Prov}_{n+1}(\alpha) = \exists\beta\, (\forall x\, \mathsf{Prov}_n[\beta(\overline{x})] \land \mathsf{Prov}_n[\forall x\, \beta(x) \to \alpha]).$$

Notice that Prov_n has level $2n$.

One can associate with each provability predicate of level n an analog of the predicate "*y is a proof of α*," that is, a Π_n-formula $\mathsf{Prf}(\alpha, y)$ such that

$$\mathsf{PA} \vdash \mathsf{Prov}(\alpha) \leftrightarrow \exists y\, \mathsf{Prf}(\alpha, y).$$

Without loss of generality we can assume that Prf is chosen in such a way that each number y is a proof of at most one formula α, and that any provable formula has arbitrarily long proofs. These properties are assumed to be verifiable in PA.

Arithmetical interpretation of GLP. Fix a strong sequence of provability predicates π over a theory T. An *arithmetical realization* is any function f mapping each propositional variable p of the language of **GLP** to some arithmetical sentence $f(p)$. For a given π, any arithmetical realization is uniquely extended to a map f_π defined on the set of all modal formulas as follows:

1. $f_\pi(p) = f(p)$, for any variable p;
2. $f_\pi(\varphi \wedge \psi) = (f_\pi(\varphi) \wedge f_\pi(\psi))$, $f_\pi(\neg\varphi) = \neg f_\pi(\varphi)$ and similarly for the other boolean connectives;
3. $f_\pi([n]\varphi) = \mathsf{Prov}_n(\ulcorner f_\pi(\varphi)\urcorner)$, for each $n \geqslant 0$.

The arithmetical formula $f_\pi(\varphi)$ is called a *translation* of the modal formula φ under the realization f. By induction on the length of proof of a formula φ it is easy to establish the soundness of **GLP** under the arithmetical interpretation with respect to any strong sequence of provability predicates over T.

Lemma 1.4 *If* **GLP** $\vdash \varphi$ *then* $T \vdash f_\pi(\varphi)$*, for any arithmetical realization* f*.*

Proof. The only nontrivial element in the proof of this lemma is to show that the translation of Löb's axiom is provable under every arithmetical realization. Such a proof is obtained with the aid of the fixed point lemma similarly to the proof of Löb's theorem for the standard provability predicate in PA. ⊠

The converse statement is valid under a sufficiently broad assumption of the soundness of the considered provability predicates and constitutes the essence of Japaridze's arithmetical completeness theorem for **GLP** [5, 4].

2 Arithmetical completeness theorem

Theorem 1 *Suppose a theory* T *and all the provability predicates* Prov_n *of a strong sequence* π *are sound. Then, for any modal formula* φ *such that* **GLP** $\nvdash \varphi$*, there is an arithmetical realization* f *for which* $T \nvdash f_\pi(\varphi)$*.*

Thus, **GLP** is the provability logic for any strong sequence of sound provability predicates. The remaining part of the paper is devoted to a proof of this theorem.

Our proof of Theorem 1 follows a general approach suggested in the fundamental paper of Solovay [7] and uses some additional ideas contained in ref. [5, 4].

Given a modal formula φ, let $M(\varphi)$ denote

$$\bigwedge_{i<s} \bigwedge_{m_i<j\leqslant n} ([m_i]\varphi_i \to [j]\varphi_i),$$

where $[m_i]\varphi_i$ for $i < s$ ranges over all subformulas of φ of the form $[k]\psi$ and $n = \max_{i<s} m_i$. Further, let

$$M^+(\varphi) = M(\varphi) \wedge \bigwedge_{i\leqslant n} [i]M(\varphi).$$

Obviously, $\mathbf{GLP} \vdash M^+(\varphi)$.

Theorem 1 is a consequence of the following theorem which also characterizes **GLP** in terms of Kripke models.

Theorem 2 *For any modal formula φ, the following statements are equivalent:*

(i) $\mathbf{GLP} \vdash \varphi$;

(ii) $\mathbf{J} \vdash M^+(\varphi) \to \varphi$;

(iii) *for each finite* $\mathbf{J}$*-model* $\mathcal{W}$, $\mathcal{W} \vDash M^+(\varphi) \to \varphi$;

(iv) *for any arithmetical realization* f, $T \vdash f_\pi(\varphi)$.

Note that the implication (iii)⇒(ii) follows from Proposition 1.2, the implication (i)⇒(iv) follows from Lemma 1.4, and (ii)⇒(i) is obvious. Thus, for a proof of Theorem 2 it is sufficient to establish the implication (iv)⇒(iii).

We argue by contradiction and fix a formula φ and a finite $\mathbf{J}$-model $\mathcal{W}_0$ with a root r such that $\mathcal{W}_0, r \nvDash M^+(\varphi) \to \varphi$. Since $\mathcal{W}_0, r \vDash M^+(\varphi)$ and r is the root, we can conclude that

$$\mathcal{W}_0 \vDash [k]\psi \to [j]\psi,$$

for each subformula $[k]\psi$ of the formula φ and every $j > k$.

As in the standard Solovay construction, for an embedding of the model $\mathcal{W}_0 = (W_0; R_0^\circ, R_1^\circ, \dots, v_0)$ into arithmetic we assume that $W_0 = \{1, 2, \dots, N\}$, for some N, and we adjoin a new root 0 to $\mathcal{W}_0$, that is, we define a new frame $(W; R_0, R_1, \dots)$ by putting $W = W_0 \cup \{0\}$, $R_0 = R_0^\circ \cup \{\langle 0, x\rangle : x \in W_0\}$ and $R_k = R_k^\circ$, for $k > 0$. We also let $v(p) = v_0(p)$ for all variables p. Clearly, in the model $\mathcal{W} = (W; R_0, R_1, \dots, v)$, we have $\mathcal{W}, r \nvDash \varphi$ and $\mathcal{W}, x \vDash [k]\psi \to [j]\psi$, for all subformulas $[k]\psi$ of φ, all $j > k$ and all $x \neq 0$.

Define:

$$\begin{array}{rcl} R_k(x) & = & \{y \in W : xR_k y\}, \\ R_k^*(x) & = & \{y \in W : xR_i y \text{ for some } i \geqslant k\}, \\ \widetilde{R}_k(x) & = & R_k^*(x) \cup \bigcup \{R_k^*(z) : x \in R_{k+1}^*(z)\}. \end{array}$$

We are going to specify an arithmetical realization f by assigning to each $x \in W$ an arithmetical sentence S_x and by letting

$$f(p) = \bigvee_{x \in v(p)} S_x. \qquad (*)$$

The sentences S_x will have to satisfy the following requirements.

(S1) $T \vdash \bigvee_{x \in W} S_x$; $T \vdash \neg(S_x \wedge S_y)$, for all $x \neq y$;

(S2) $T \vdash S_x \to \neg \mathsf{Prov}_k[\neg S_y]$, for all $xR_k y$;

(S3) $T \vdash S_x \to \mathsf{Prov}_k[\bigvee_{y \in \widetilde{R}_k(x)} S_y]$, for all $x \neq 0$;

(S4) $\mathbb{N} \vDash S_0$.

Suppose for now that the requirements (S1)–(S3) are met and a realization f is defined by $(*)$.

Lemma 2.1 *For each subformula θ of formula φ and each $x \in W_0$*

(a) $\mathcal{W}, x \vDash \theta$ implies $T \vdash S_x \to f_\pi(\theta)$;

(b) $\mathcal{W}, x \nvDash \theta$ implies $T \vdash S_x \to \neg f_\pi(\theta)$.

Proof. We jointly prove (a) and (b) by induction on the build-up of the formula θ. If θ is a variable or has the form $\theta_1 \wedge \theta_2$ or $\neg\theta_1$, the statements follow from the definition of f together with (S1).

Let $\theta = [k]\psi$. For a proof of (b) assume that $\mathcal{W}, x \nvDash [k]\psi$. Then there is a $y \in W_0$ such that xR_ky and $\mathcal{W}, y \nvDash \psi$. By the induction hypothesis $T \vdash S_y \to \neg f_\pi(\psi)$. It follows that $T \vdash f_\pi(\psi) \to \neg S_y$ and $T \vdash \mathsf{Prov}_k[f_\pi(\psi)] \to \mathsf{Prov}_k[\neg S_y]$. By (S2) we have $T \vdash S_x \to \neg\mathsf{Prov}_k[\neg S_y]$, whence $T \vdash S_x \to \neg\mathsf{Prov}_k[f_\pi(\psi)]$.

For a proof of (a) assume $\mathcal{W}, x \vDash [k]\psi$. First, we show that

$$\forall w \in \widetilde{R}_k(x)\, \mathcal{W}, w \vDash \psi.$$

Let $w \in \widetilde{R}_k(x)$, then for some z we have $x \in R^*_{k+1}(z) \cup \{z\}$ and $w \in R^*_k(z)$. Obviously, $\forall y \in R_k(x)\, \mathcal{W}, y \vDash \psi$, and since $R_k(z) = R_k(x)$ we have $\mathcal{W}, z \vDash [k]\psi$. By the construction of $\mathcal{W}_0$ we also have $\mathcal{W}, z \vDash [k]\psi \to [m]\psi$, for any $m > k$, whence $\mathcal{W}, z \vDash [m]\psi$, for all $m \geqslant k$. Since $w \in R^*_k(z)$, it follows that $\mathcal{W}, w \vDash \psi$.

Since $\forall w \in \widetilde{R}_k(x)\, \mathcal{W}, w \vDash \psi$, by the induction hypothesis we obtain $T \vdash S_w \to f_\pi(\psi)$, for all $w \in \widetilde{R}_k(x)$, and therefore $T \vdash (\bigvee_{w \in \widetilde{R}_k(x)} S_w) \to f_\pi(\psi)$. Using the derivability conditions it follows that

$$T \vdash \mathsf{Prov}_k[\bigvee_{w \in \widetilde{R}_k(x)} S_w] \to \mathsf{Prov}_k[f_\pi(\psi)],$$

whence $T \vdash S_x \to \mathsf{Prov}_k[f_\pi(\psi)]$ by condition (S3). ⊠

As a corollary of this lemma we obtain $T \vdash S_r \to \neg f_\pi(\varphi)$. If $T \vdash f_\pi(\varphi)$, then $T \vdash \neg S_r$. It follows that $T \vdash \mathsf{Prov}_0[\neg S_r]$, and hence $T \vdash \neg S_0$ by (S2). Since T is sound, we have $\mathbb{N} \nvDash S_0$ which contradicts (S4). Thus, our assumption $T \vdash f_\pi(\varphi)$ is wrong, that is, Lemma 2.1 implies Theorem 2.

To complete the proof of Theorem 2 it remains for us to construct arithmetical sentences S_x satisfying (S1)–(S4). A number m is called the *rank* of a model $\mathcal{W}$ if $R_m \neq \varnothing$ and, for all $k > m$, $R_k = \varnothing$. For each $k \leqslant m$, we are going to construct arithmetical functions $h_k : \mathbb{N} \to W$ with the aid of the arithmetical fixed point theorem. Informally speaking, the function h_0 is the “usual” Solovay function for the frame (W, R_0) and the provability predicate Prov_0, whereas the functions h_{k+1} are its analogues for the frames (W, R_{k+1}) and the predicates Prov_{k+1} linked by the condition that the initial value $h_{k+1}(0)$ is set to the limit value of the function h_k, for all $k < m$.

More formally, let ℓ_k denote the limit of the function h_k. A sentence S_x is defined as a formalization of the statement $\ell_m = \overline{x}$, that is,

$$S_x \stackrel{\text{def}}{\Longleftrightarrow} \exists n_0 \, \forall n \geqslant n_0 \, H_m(n, \overline{x}),$$

where $H_m(n, x)$ is a formula expressing $h_m(n) = x$. We notice that $\ulcorner S_x \urcorner$ is a primitive recursive function of x and $\ulcorner H_m \urcorner$.

We would like to construct the functions h_k (provably in T) satisfying the following conditions, for all $k \leqslant m$:

$$h_0(0) = 0 \text{ and } h_k(0) = \ell_{k-1}, \text{ if } k > 0; \tag{1}$$

$$h_k(n+1) = \begin{cases} z, & \text{if } h_k(n) R_k z \text{ and } \mathsf{Prf}_k(\ulcorner \neg S_z \urcorner, n), \\ h_k(n), & \text{otherwise.} \end{cases} \tag{2}$$

A direct formalization of conditions (1) and (2) yields a sequence of arithmetical formulas A_k, for $k \leqslant m$, defining a system of $m+1$ equations in the unknown formulas $H_0, \ldots, H_m$:

$$\begin{cases} T \vdash H_0(n, x) \leftrightarrow A_0(\ulcorner H_m \urcorner, n, x), \\ T \vdash H_{k+1}(n, x) \leftrightarrow A_{k+1}(\ulcorner H_m \urcorner, \ell_k, n, x), \text{for all } k < m. \end{cases} \tag{3}$$

Notice that ℓ_k is formally expressible via H_k, and thus via $A_k(\ulcorner H_m \urcorner, \ell_{k-1}, n, x)$. Hence, we can successively substitute each equation in the system (3), starting from the first, into the next one, and thereby eliminate all the definitions of ℓ_k. Eventually, we obtain one equation defining H_m as a fixed point amenable to an application of the standard fixed point lemma:

$$T \vdash H_m(n, x) \leftrightarrow A'_m(\ulcorner H_m \urcorner, n, x).$$

Having constructed H_m, converse substitutions yield all the other formulas H_k, for $k < m$, satisfying (3).

Notice that A_k belongs to the class $\Delta_0(\Sigma_{r_k})$, since $\mathsf{Prf}_k \in \Pi_{r_k}$. Unwinding the definitions of ℓ_{k-1} shows that each formula H_k belongs to the class Σ_{r_k+1} (modulo equivalence in PA). We also immediately obtain the following lemma.

Lemma 2.2 *For any $k \geqslant 0$,*

(i) $T \vdash \forall n \, \exists! x \in W \, H_k(n, x)$;

(ii) $T \vdash \forall i, j, z \, (i < j \wedge h_k(i) = z \to h_k(j) \in R_k(z) \cup \{z\})$;

(iii) $T \vdash \exists! x \in W \, \ell_k = x$;

(iv) $T \vdash \forall z \, (\exists n \, h_k(n) = z \to \ell_m \in R_k^*(z) \cup \{z\})$.

The following lemma completes the proof of the theorem.

Lemma 2.3 *Conditions* (S1)–(S4) *are satisfied for the set of sentences* $\{S_x : x \in W\}$ *thus constructed.*

Proof. Condition (S1) follows from Lemma 2.2 (iii).

(S2) is obtained by formalizing the following argument in T.

Suppose S_x, then either $\ell_k = x$, or $x \in R^*_{k+1}(\ell_k)$. Notice that in both cases $R_k(x) = R_k(\ell_k)$. Pick a number n_0 such that $\forall n \geqslant n_0\ h_k(n) = \ell_k$. Assume $\mathsf{Prov}_k[\neg S_y]$, for some $y \in R_k(x)$. Then there is an $n_1 \geqslant n_0$ such that $\mathsf{Prf}_k(\ulcorner \neg S_y \urcorner, n_1)$. Since $\ell_k R_k y$ and $h_k(n_1) = \ell_k$, by the definition of h_k we obtain $h_k(n_1+1) = y$, a contradiction. Thus, $\neg\mathsf{Prov}_k[\neg S_y]$.

For any $A \subseteq W$ let $\ell_m \in A$ denote the sentence $\bigvee_{y\in A} S_y$. For a proof of (S3) we have to show, for each $x \neq 0$, that

$$T \vdash S_x \to \mathsf{Prov}_k[\ell_m \in \widetilde{R}_k(x)].$$

We formalize in T the following argument.

Assume S_x, where $x \neq 0$, and let $z = \ell_k$. Then $x \in R^*_{k+1}(z) \cup \{z\}$. By the definition of $\widetilde{R}_k$ this implies $R^*_k(z) \subseteq \widetilde{R}_k(x)$ and, since this property is definable by a Δ_0-formula, $\mathsf{Prov}_k[R^*_k(z) \subseteq \widetilde{R}_k(x)]$. Hence,

$$\mathsf{Prov}_k[\ell_m \in R^*_k(z)] \to \mathsf{Prov}_k[\ell_m \in \widetilde{R}_k(x)]. \tag{4}$$

On the other hand, since $\ell_k = z$ we have $\exists n\, h_k(n) = z$. The latter statement is definable by an arithmetical Σ_{r_k+1}-formula, hence $\mathsf{Prov}_k[\exists n\ h_k(n) = \overline{z}]$. By Lemma 2.2 (iv), for any $u \in W$,

$$T \vdash \exists n\ h_k(n) = \overline{u} \to \ell_m \in R^*_k(u) \cup \{u\}.$$

It follows that

$$\mathsf{Prov}_k[\exists n\ h_k(n) = \overline{u}] \to \mathsf{Prov}_k[\ell_m \in R^*_k(u) \cup \{u\}].$$

In particular, for $u = z$ we obtain $\mathsf{Prov}_k[\ell_m \in R^*_k(z) \cup \{z\}]$.

Now we notice that $z \neq 0$ (since $x \neq 0$). Since $\exists n\ h_k(n) = z$ we have $\mathsf{Prov}_k[\neg S_z]$; in fact, h_k cannot reach z unless $\mathsf{Prov}_k[\neg S_z]$. Thus, $\mathsf{Prov}_k[\ell_m \in R^*_k(z)]$, and using (4) we obtain $\mathsf{Prov}_k[\ell_m \in \widetilde{R}_k(x)]$, as required.

(S4) By the induction on k we show that $\mathbb{N} \vDash \ell_k = 0$, for all $k \leqslant m$. If $\ell_k = z \neq 0$ holds in the standard model of arithmetic, then $\mathsf{Prov}_k[\neg S_z]$, since by the induction hypothesis (for $k > 0$) there holds $h_k(0) = \ell_{k-1} = 0$. Since Prov_k is sound, it follows that $\ell_k \neq z$. Thus, $\ell_k = 0$. ⊠

3 Some generalizations

The requirement that T together with all the provability predicates π be sound is natural but stronger than is necessary for the validity of the arithmetical completeness theorem for **GLP**.

A sequence π of provability predicates is called *strongly consistent over* T if the theory

$$T + \{\mathsf{Con}_n : n < \omega\}$$

is consistent, where Con_n denotes $\neg\mathsf{Prov}_n[\bot]$. Obviously, every sequence of sound provability predicates over T is strongly consistent. The converse is, in general, not true. In fact, if Prov_n is the provability predicate for the theory axiomatized over T by the set of all true Π_n-sentences (as above), then Con_n is equivalent to the so-called uniform Σ_n-reflection principle over T. Thus, $T + \{\mathsf{Con}_n : n < \omega\}$ is equivalent to the extension of T by the full uniform reflection schema, that is, to $T+\mathsf{RFN}(T)$ (see [6, 1]). It is easy to give examples of unsound theories T for which, nevertheless, the theory $T + \mathsf{RFN}(T)$ is consistent.

For example, let T be a sound r.e. theory, and let T' denote $T + \mathsf{RFN}(T)$. Consider the theory $U = T + \neg\mathsf{Con}(T')$. Clearly, U is unsound. On the other hand, since $\neg\mathsf{Con}(T') \in \Sigma_1$ we have

$$T \vdash \mathsf{Con}_n(U) \leftrightarrow (\mathsf{Con}_n(T) \wedge \neg\mathsf{Con}(T')),$$

for each $n < \omega$. Thus, if $U+\mathsf{RFN}(U) \vdash \bot$ then, for some $n < \omega$, $U+\mathsf{Con}_n(U) \vdash \bot$ and therefore $T + \neg\mathsf{Con}(T') + \mathsf{Con}_n(T) \vdash \bot$. From this we can conclude that $T + \mathsf{Con}_n(T) \vdash \mathsf{Con}(T')$ contradicting Gödel's second incompleteness theorem.

Theorem 3 *The conclusion of Theorem 2 holds for any strong sequence of provability predicates π strongly consistent over T.*

Proof. We employ the same construction as in the proof of Theorem 2. Since T is no longer assumed to be sound, the property (S4) does not necessarily hold. Nevertheless, Lemmas 2.1, 2.2 and 2.3 remain valid as before.

Let $d_k(x)$ denote the depth of a point $x \in \mathcal{W}$ in the sense of the binary relation R_k, that is, $d_k(x) = \sup\{d_k(y)+1 : y \in R_k(x)\}$. We prove the following properties.

Lemma 3.1 *Let m be the rank of $\mathcal{W}$. Then, for all $x \in W_0$,*

(i) $T \vdash S_x \to \mathsf{Prov}_m[\ell_m \in R_m(x)]$;

(ii) $T \vdash S_x \to \mathsf{Prov}_m^k[\bot]$, $k = d_m(x) + 1$.

Proof. Statement (i) essentially amounts to a particular case of (S3) for $k = m$. In fact, since R_{m+1} is empty,

$$\widetilde{R}_m(x) = R_m^*(x) \cup \{R_m^*(z) : x \in R_{m+1}^*(z)\} = R_m^*(x) = R_m(x).$$

Statement (ii) is obtained from (i) by an obvious induction on $d_m(x)$. ⊠

For the proof of Theorem 3 we reason as follows. As before, we infer from $T \vdash f_\pi(\varphi)$ and Lemma 2.1 that $T \vdash \neg S_0$, and hence $T \vdash \bigvee_{z\in W_0} S_z$ by (S1). On the other hand, by the previous lemma, for each $z \in W_0$, $T \vdash S_z \to \mathsf{Prov}_m^k[\bot]$, where $k = \sup\{d_m(x) + 1 : x \in W_0\}$. Therefore, we obtain $T \vdash \mathsf{Prov}_m^k[\bot] \vdash \mathsf{Prov}_{m+1}[\bot]$, that is, T cannot be strongly consistent. ⊠

Now we give a simplified proof of Japaridze's arithmetical completeness theorem for the so-called truth polymodal provability logic. Let **GLPS** denote the extension of the set of theorems of **GLP** by the schema $[n]\varphi \to \varphi$, for all formulas φ and all $n < \omega$, and with modus ponens as a sole inference rule.

Let $H(\varphi)$ denote the formula $\bigwedge_{i<s}([n_i]\varphi_i \to \varphi_i)$, where the formulas $[n_i]\varphi_i$, for $i < s$, enumerate all the subformulas of φ of the form $[k]\psi$.

Theorem 4 *Suppose a theory T and all the provability predicates π are sound. The following statements are equivalent:*

(i) $\mathbf{GLPS} \vdash \varphi$;

(ii) $\mathbf{GLP} \vdash H(\varphi) \to \varphi$;

(iii) $\mathbb{N} \models f_\pi(\varphi)$, *for every arithmetical realization f.*

Proof. We will follow the standard method coming from the second arithmetical completeness theorem of Solovay. The implications (i)⇒(iii) and (ii)⇒(i) are immediate. We prove (iii)⇒(ii) reasoning by contraposition.

Assume $\mathbf{GLP} \nvdash H(\varphi) \to \varphi$. Then there is a finite **J**-model $\mathcal{W}$ with the root 0 such that $\mathcal{W}, 0 \models M^+(\varphi), H(\varphi)$ and $\mathcal{W}, 0 \nvDash \varphi$. As before, we assume that $W = \{0, \dots, N\}$ and apply the construction from the proof of Theorem 2. Lemmas 2.1, 2.2 and 2.3 hold without any change.

Since we are now interested in the validity relation at 0 we augment Lemma 2.1 by the following statement.

Lemma 3.2 *For each subformula θ of the formula φ,*

(a) $\mathcal{W}, 0 \models \theta$ *implies* $T \vdash S_0 \to f_\pi(\theta)$;

(b) $\mathcal{W}, 0 \nvDash \theta$ *implies* $T \vdash S_0 \to \neg f_\pi(\theta)$.

Proof. We jointly prove (a) and (b) by induction on the build-up of θ using Lemma 2.1. All the cases are treated exactly as in Lemma 2.1 except for (a) when $\theta = [k]\psi$. In this case we have $\mathcal{W}, 0 \models \psi$, since $\mathcal{W}, 0 \models H(\varphi)$ and $\mathcal{W}, 0 \models [j]\psi$, for all $k \leqslant j \leqslant m$, because of $\mathcal{W}, 0 \models M(\varphi)$. Thus, $\mathcal{W}, x \models \psi$, for all $x \in R_k^*(0) \cup \{0\}$. Hence, by the induction hypothesis together with Lemma 2.1,

$$T \vdash \ell_m \in R_k^*(0) \cup \{0\} \to f_\pi(\psi),$$

and therefore

$$T \vdash \mathsf{Prov}_k[\ell_m \in R_k^*(0) \cup \{0\}] \to \mathsf{Prov}_k[f_\pi(\psi)]. \tag{5}$$

Since 0 is a root of $\mathcal{W}$, from S_0 one can infer $\exists n h_k(n) = 0$, whence $\mathsf{Prov}_k[\exists n h_k(n) = 0]$ and $\mathsf{Prov}_k[\ell_m \in R_k^*(0) \cup \{0\}]$ by Lemma 2.2 (iv). This implies by (5) that $\mathsf{Prov}_k[f_\pi(\psi)]$, as required. ⊠

Since $\mathbb{N} \models S_0$, this lemma yields $\mathbb{N} \nvDash f_\pi(\varphi)$, which completes the proof of Theorem 4. ⊠

References

[1] L.D. Beklemishev. Reflection principles and provability algebras in formal arithmetic. *Uspekhi Matematicheskikh Nauk*, 60(2):3–78, 2005. In Russian. English translation in: *Russian Mathematical Surveys*, 60(2): 197–268, 2005.

[2] L.D. Beklemishev. Kripke semantics for provability logic GLP. *Annals of Pure and Applied Logic*, 161:756–774, 2010. Preprint: Logic Group Preprint Series 260, University of Utrecht, November 2007. `http://preprints.phil.uu.nl/lgps/`.

[3] G. Boolos. *The Logic of Provability.* Cambridge University Press, Cambridge, 1993.

[4] K.N. Ignatiev. On strong provability predicates and the associated modal logics. *The Journal of Symbolic Logic*, 58:249–290, 1993.

[5] G.K. Japaridze. The modal logical means of investigation of provability. Thesis in Philosophy, in Russian, Moscow, 1986.

[6] C. Smoryński. The incompleteness theorems. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 821–865. North Holland, Amsterdam, 1977.

[7] R.M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 28:33–71, 1976.