THE BULLETIN OF SYMBOLIC LOGIC
Volume 27, Number 4, December 2021

VARIANTS OF KREISEL'S CONJECTURE ON A NEW NOTION OF PROVABILITY

PAULO GUILHERME SANTOS AND REINHARD KAHLE

Abstract. Kreisel's conjecture is the statement: if, for all $n \in \mathbb{N}$, $\mathrm{PA} \vdash_{k \text{ steps}} \varphi(\overline{n})$, then $\mathrm{PA} \vdash \forall x.\varphi(x)$. For a theory of arithmetic T, given a recursive function h, $T \vdash_{\leq h} \varphi$ holds if there is a proof of φ in T whose code is at most $h(\#\varphi)$. This notion depends on the underlying coding. $P^h_T(x)$ is a predicate for $\vdash_{\leq h}$ in T. It is shown that there exist a sentence φ and a total recursive function h such that $T \vdash_{\leq h} \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner\varphi\urcorner) \rightarrow \varphi\urcorner)$, but $T \nvdash_{\leq h} \varphi$, where Pr_T stands for the standard provability predicate in T. This statement is related to a conjecture by Montagna. Also variants and weakenings of Kreisel's conjecture are studied. By the use of reflexion principles, one can obtain a theory T^h_Γ that extends T such that a version of Kreisel's conjecture holds: given a recursive function h and $\varphi(x)$ a Γ-formula (where Γ is an arbitrarily fixed class of formulas) such that, for all $n \in \mathbb{N}$, $T \vdash_{\leq h} \varphi(\overline{n})$, then $T^h_\Gamma \vdash \forall x.\varphi(x)$. Derivability conditions are studied for a theory to satisfy the following implication: if $T \vdash \forall x.P^h_T(\ulcorner\varphi(\dot{x})\urcorner)$, then $T \vdash \forall x.\varphi(x)$. This corresponds to an arithmetization of Kreisel's conjecture. It is shown that, for certain theories, there exists a function h such that $\vdash_{k \text{ steps}} \subseteq \vdash_{\leq h}$.

§1. Preliminaries. Let T be a fixed theory of arithmetic which is a consistent primitive-recursive extension of Robinson's Q.

Following [14], we say that a partial function $f : \mathbb{N}^k \rightharpoonup \mathbb{N}$ is *strongly representable* in T if there is a formula $\varphi(x_0, \dots, x_{k-1}, y)$ such that[1]

(i) For all $m_0, \dots, m_{k-1}, n \in \mathbb{N}$, $f(m_0, \dots, m_{k-1}) \simeq n \iff T \vdash \varphi(\overline{m_0}, \dots, \overline{m_{k-1}}, \overline{n})$;
(ii) $T \vdash \forall x_0 \cdots \forall x_{k-1}.\exists! y.\varphi(x_0, \dots, x_{k-1}, y)$.

If a function f is strongly representable by a Σ_1-formula $\varphi(x, y)$ in T, then f is provably recursive in T (see [13] for a definition of this last concept). Every theory of arithmetic which is a consistent primitive-recursive extension of Q can strongly represent every (partial) recursive function (not necessarily by a Σ_1-formula; see [14]).

For a standard Gödelization $\ulcorner\cdot\urcorner$ of the underlying language, we use Feferman's dot notation [16, p. 837]: Let $\mathrm{sub}(x, y)$ be a function-symbol such that, for every term t, $T \vdash \mathrm{sub}(\ulcorner\varphi\urcorner, \ulcorner t\urcorner) = \ulcorner\varphi(t)\urcorner$, and let $\mathrm{num}(x)$

Received March 6, 2020.
2020 *Mathematics Subject Classification*. 03F30, 03F03.
Key words and phrases. Kreisel's conjecture, provability predicate, formal arithmetic.

[1] $\bar{n}$ is the standard representation of the number n in the theory T.

1079-8986/21/2704-0001
DOI :10.1017/bsl.2021.68

be a function-symbol which represents the numerals in T. With $\mathrm{s}(x, y) := \mathrm{sub}(x, \mathrm{num}(y))$ we denote $\mathrm{s}(\ulcorner\varphi(x)\urcorner, y)$ by $\ulcorner\varphi(\dot{y})\urcorner$.

$P(x)$ is a *provability predicate* in T if for all formulas φ, $T \vdash P(\ulcorner\varphi\urcorner)$ if, and only if, φ is a theorem of T. Furthermore, $S(x)$ is a *notion of provability* in T if, for every formula φ, $T \vdash S(\ulcorner\varphi\urcorner)$ implies that φ is provable in T. For a provability predicate $P(x)$, we define $\mathrm{Con}_P := \neg P(\ulcorner\bot\urcorner)$.

We say a formula is Σ_n when it is equivalent in T to a Σ_n-formula.

For a (partial) recursive function h, the notation $T \vdash_{\leq h} \varphi$ expresses that $h(\#\varphi)$ is defined and φ is provable in T with a proof whose code is at most $h(\#\varphi)$. This notion generalises the approach followed in [9, pp. 33–35]. $\vdash_{\leq h}$ depends heavily on the chosen Gödelization: different codings give rise to different notions. For the rest the concrete Gödelization is assumed to be a fixed one.

Given T, the theory K_T extends T by the following axiom schema:

Axiom K. If f is a total recursive function such that, for all $n \in \mathbb{N}$, $f(n) \neq 0$, and $\mathrm{R}(x, y)$ is a formula that strongly represents f in T, then $K_T \vdash \forall x.\neg\mathrm{R}(x, 0)$.

This schema can be restricted to a smaller class of functions in such a way that K_T is recursively enumerable (for instance by considering the class of all primitive recursive functions [10]).

§2. Introduction.

According to [5], Kreisel's conjecture is the statement:

$$\text{If, for all } n \in \mathbb{N}, \mathrm{PA} \vdash_{k \text{ steps}} \varphi(\overline{n}), \text{then } \mathrm{PA} \vdash \forall x.\varphi(x).$$

Kreisel's conjecture has been studied for different systems, with partial solutions for specific theories of arithmetic other than PA—see, for instance, [1, 11, 12]; for a detailed account on the conjecture we refer to [2].

In this paper we present results similar to Kreisel's conjecture for $\vdash_{\leq h}$, which are not restricted to PA. It is an interesting feature of our approach that it does not depend so heavily on the particular axiomatisation of T that one chooses. In some sense, it can be seen as a uniform approach, since it applies to any consistent, primitive-recursive extension of Q.

For h be a total recursive function, the *adapted Kreisel's conjecture for* $\vdash_{\leq h}$ is:

$$\text{For all } n \in \mathbb{N}, T \vdash_{\leq h} \varphi(\overline{n}), \text{ then } T \vdash \forall x.\varphi(x).$$

Theorem 2.1. *The adapted Kreisel's conjecture for* $\vdash_{\leq h}$ *is false.*

Proof. Let $\mathrm{Proof}(x, y)$ be a standard proof predicate of T that expresses that x is a proof of y (see [16, p. 838] for details). Let h be the function defined by:

$$h(m) := \begin{cases} \mu k[k \text{ is the code of a proof of the formula coded by } m \text{ in } T], \\ \qquad\qquad\qquad\qquad \text{if } m = \neg\,\mathrm{Proof}(\overline{n}, \ulcorner\bot\urcorner), \text{for some } n, \\ 0, \qquad\qquad\qquad\quad \text{otherwise,} \end{cases}$$

where μ denotes the minimisation function (see [16, p. 833] for further details on minimisation). It is clear that h is a (total) recursive function. By construction, for all $n \in \mathbb{N}$, $T \vdash_{\leq h} \neg\,\mathrm{Proof}(\overline{n}, \ulcorner \bot \urcorner)$. If the adapted Kreisel's conjecture for $\vdash_{\leq h}$ was true, it would follow $T \vdash \forall x. \neg\,\mathrm{Proof}(x, \ulcorner \bot \urcorner)$, contradicting the second incompleteness theorem (see [16, p. 828] for further information). ⊣

It is not known whether the theorem still holds if one restricts oneself to primitive-recursive functions.

The result is in accordance with [6], where several reasons are given to believe that Kreisel's conjecture is, in fact, false.

Even though the adapted Kreisel's conjecture for $\vdash_{\leq h}$ is false, it is worth studying variants and weakenings of it. For example, one could ask for an extension T^h of T such that Kreisel's conjecture holds adapted to T^h: given a total recursive function h, if, for all $n \in \mathbb{N}$, $T \vdash_{\leq h} \varphi(\overline{n})$, then $T^h \vdash \forall x.\varphi(x)$.

One immediate solution would be to add the true sentence $\forall x.\varphi(x)$ as an axiom to T. We will, however, construct a theory T^h, avoiding the trivial *a priori* addition of $\forall x.\varphi(x)$ as an axiom. The approach is of interest, because it allows to establish relations between different concepts.

We will also study versions of the conjecture for theories that satisfy certain derivability conditions. We exhibit conditions for a theory to satisfy the following implication:

$$\text{If } T \vdash \forall x. P^h_T(\ulcorner \varphi(\dot{x}) \urcorner), \text{ then } T \vdash \forall x.\varphi(x).$$

This corresponds to an arithmetization of Kreisel's conjecture.

Finally, we prove, for certain theories, the existence of a total recursive function h such that $\vdash_{k \text{ steps}} \subseteq \vdash_{\leq h}$.

§3. On the notion of provability $\vdash_{\leq h}$. In this section, we study the notion $\vdash_{\leq h}$ and some of its properties. We start with a result that guarantees that $\vdash_{\leq h}$ is representable in T.

Theorem 3.1. *Given a total recursive function h, there is a notion of provability $P^h_T(x)$ that represents $\vdash_{\leq h}$ in T such that if, for all $n \in \mathbb{N}$, $T \vdash_{\leq h} \alpha(\overline{n})$, then $K_T \vdash \forall x. P^h_T(\ulcorner \alpha(\dot{x}) \urcorner)$.*

Proof. Let h be an arbitrary, but fixed total recursive function. We define f_h by:

$$f_h(n) := \begin{cases} \mu m \leq h(n)[m \text{ is the code of a proof of the formula coded by } n], \\ \qquad\qquad \text{if } n \text{ is a code of a formula and there is such an } m, \\ 0, \qquad\qquad \text{otherwise.} \end{cases}$$

f_h is a total recursive function, thus f_h can be strongly representable by a formula $\mathrm{R}_h(x, y)$ in T. Given $n, m \in \mathbb{N}$, it is clear that $m \leq h(n)$ is the smallest code of a proof of the formula coded by n if, and only if, $T \vdash \mathrm{R}_h(\overline{n}, \overline{m}) \wedge \overline{m} \neq 0$. Thus, we can define the provability predicate $P^h_T(x) := \exists y \neq 0.\, \mathrm{R}_h(x, y)$.

https://doi.org/10.1017/bsl.2021.68 Published online by Cambridge University Press

Assume that, for all $n \in \mathbb{N}$, $T \vdash_{\leq h} \alpha(\overline{n})$. Let g_h be the function defined by: $g_h(n) := f_h(\#\alpha(\overline{n}))$. g_h is a total recursive function. Furthermore, g_h is strongly representable by the formula $\mathrm{S}_h(x, y) := \mathrm{R}_h(\ulcorner\alpha(\dot{x})\urcorner, y)$ since:

(i) If $g_h(n) = m$, then $f_h(\#\alpha(\overline{n})) = m$, and thus $T \vdash \mathrm{R}_h(\ulcorner\alpha(\overline{n})\urcorner, \overline{m})$, i.e., $T \vdash \mathrm{S}_h(\overline{n}, \overline{m})$;

(ii) As $T \vdash \forall x.\exists!\, y.\, \mathrm{R}_h(x, y)$ it follows that $T \vdash \forall x.\exists!\, y.\, \mathrm{R}_h(\ulcorner\alpha(\dot{x})\urcorner, y)$, i.e., $T \vdash \forall x.\exists!\, y.\mathrm{S}_h(x, y)$.

By hypothesis, for all $n \in \mathbb{N}$, there is $m \leq h(\#\alpha(\overline{n}))$ such that m is the code of a proof of $\alpha(\overline{n})$ in T. Hence, for all $n \in \mathbb{N}$, $g_h(n) \neq 0$. As $\mathrm{S}_h(x, y)$ strongly represents g_h, we have by hypothesis that $K_T \vdash \forall x.\neg\mathrm{S}_h(x, 0)$. From $T \vdash \forall x.\exists!\, y.\mathrm{S}_h(x, y)$ follows that $T \vdash \forall x.\exists y.\mathrm{S}_h(x, y)$. Together with $K_T \vdash \forall x.\neg\mathrm{S}_h(x, 0)$, it follows that $K_T \vdash \forall x.\exists y \neq 0.\mathrm{S}_h(x, y)$, i.e., $K_T \vdash \forall x.\exists y \neq 0.\, \mathrm{R}_h(\ulcorner\alpha(\dot{x})\urcorner, y)$. So, $K_T \vdash \forall x.P^h_T(\ulcorner\alpha(\dot{x})\urcorner)$.

We show that, for all formulas φ, $T \vdash_{\leq h} \varphi \iff T \vdash P^h_T(\ulcorner\varphi\urcorner)$. Suppose that $T \vdash_{\leq h} \varphi$. For $m := f_h(\#\varphi)$, we have that $m \neq 0$. Thus, $T \vdash \mathrm{R}_h(\ulcorner\varphi\urcorner, \overline{m}) \wedge \overline{m} \neq 0$, and so $T \vdash \exists y \neq 0.\, \mathrm{R}_h(\ulcorner\varphi\urcorner, y)$. Hence, $T \vdash P^h_T(\ulcorner\varphi\urcorner)$. Now suppose that $T \vdash P^h_T(\ulcorner\varphi\urcorner)$. Let $m := f_h(\#\varphi)$. If $m \neq 0$, then $T \vdash_{\leq h} \varphi$. Suppose, towards a contradiction, that $m = 0$. We have that $T \vdash \exists y \neq 0.\, \mathrm{R}_h(\ulcorner\varphi\urcorner, y)$ and $T \vdash \mathrm{R}_h(\ulcorner\varphi\urcorner, 0)$. As $T \vdash \forall x.\exists!\, y.\, \mathrm{R}_h(x, y)$ we arrive at a contradiction. So, $m \neq 0$, as desired. ⊣

The next result follows immediately from the proof of the last theorem.

Corollary 3.1. *Given a total recursive function h, and a formula φ, we have that $T \vdash_{\leq h} \varphi \iff T \vdash P^h_T(\ulcorner\varphi\urcorner)$.*

The provability predicate P^h_T is provably decidable, in the following sense:

Theorem 3.2. *Given a total recursive function h, for every formula φ, we have that $T \vdash P^h_T(\ulcorner\varphi\urcorner)$ or $T \vdash \neg P^h_T(\ulcorner\varphi\urcorner)$.*

Proof. Suppose that $T \nvdash P^h_T(\ulcorner\varphi\urcorner)$. By the previous result, $T \nvdash_{\leq h} \varphi$. This means that $f_h(\#\varphi) = 0$. As $\mathrm{R}_h(x, y)$ strongly represents the function f_h, it follows that $T \vdash \mathrm{R}_h(\ulcorner\varphi\urcorner, 0)$. Since $T \vdash \forall x.\exists!\, y.\, \mathrm{R}_h(x, y)$, we can conclude that $T \vdash \forall y.(\mathrm{R}_h(\ulcorner\varphi\urcorner, y) \rightarrow y = 0)$, and so, $T \vdash \neg\exists y.(y \neq 0 \wedge \mathrm{R}_h(\ulcorner\varphi\urcorner, y))$, i.e., $T \vdash \neg P^h_T(\ulcorner\varphi\urcorner)$. ⊣

The next result corresponds to an arithmetization of the previous statement.

Theorem 3.3. *Given a provability predicate $P(x)$ and a total recursive function h, we have for every formula φ that $T \vdash P(\ulcorner P^h_T(\ulcorner\varphi\urcorner)\urcorner) \vee P(\ulcorner\neg P^h_T(\ulcorner\varphi\urcorner)\urcorner)$.*

Proof. From Theorem 3.2, it follows $T \vdash P^h_T(\ulcorner\varphi\urcorner)$ or $T \vdash \neg P^h_T(\ulcorner\varphi\urcorner)$. If $T \vdash P^h_T(\ulcorner\varphi\urcorner)$, then $T \vdash P(\ulcorner P^h_T(\ulcorner\varphi\urcorner)\urcorner)$, and so $T \vdash P(\ulcorner P^h_T(\ulcorner\varphi\urcorner)\urcorner) \vee P(\ulcorner\neg P^h_T(\ulcorner\varphi\urcorner)\urcorner)$. If $T \vdash \neg P^h_T(\ulcorner\varphi\urcorner)$, then $T \vdash P(\ulcorner\neg P^h_T(\ulcorner\varphi\urcorner)\urcorner)$; hence $T \vdash P(\ulcorner P^h_T(\ulcorner\varphi\urcorner)\urcorner) \vee P(\ulcorner\neg P^h_T(\ulcorner\varphi\urcorner)\urcorner)$. In sum, $T \vdash P(\ulcorner P^h_T(\ulcorner\varphi\urcorner)\urcorner) \vee P(\ulcorner\neg P^h_T(\ulcorner\varphi\urcorner)\urcorner)$. ⊣

We now prove that there is no h such that $\vdash_{\leq h}$ coincides with $\vdash$.

THEOREM 3.4. *For every total recursive function* h, *there is a formula* φ *such that* $T \vdash \varphi$, *but* $T \nvdash_{\leq h} \varphi$.

PROOF. Let h be a fixed total recursive function. Let φ be the sentence obtained from the application of the diagonalisation lemma [15, p. 169] to the formula $\neg P^h_T(x)$. Then,

$$T \vdash \varphi \leftrightarrow \neg P^h_T(\ulcorner \varphi \urcorner). \tag{I}$$

Suppose, towards a contradiction, that $T \vdash_{\leq h} \varphi$. Thus, $T \vdash \varphi$. By Corollary 3.1 we have that $T \vdash P^h_T(\ulcorner \varphi \urcorner)$, and so, by (I), $T \vdash \neg\varphi$, which contradicts $T \vdash \varphi$. Hence, $T \nvdash_{\leq h} \varphi$. From Theorem 3.2 it follows that $T \vdash \neg P^h_T(\ulcorner \varphi \urcorner)$, i.e., $T \vdash \varphi$. ⊣

The next fact will play a major role for the discussion of Kreisel's conjecture.

THEOREM 3.5. *Given a total recursive function* h, *for every formula* φ, $T \vdash P^h_T(\ulcorner \varphi \urcorner) \to \varphi$.

PROOF. Let φ be an arbitrary formula. From Theorem 3.2, $T \vdash P^h_T(\ulcorner \varphi \urcorner)$ or $T \vdash \neg P^h_T(\ulcorner \varphi \urcorner)$. First, suppose that $T \vdash P^h_T(\ulcorner \varphi \urcorner)$. Then, by Corollary 3.1, we conclude $T \vdash_{\leq h} \varphi$, from where we get $T \vdash \varphi$. Thus, $T \vdash P^h_T(\ulcorner \varphi \urcorner) \to \varphi$. Second, suppose that $T \vdash \neg P^h_T(\ulcorner \varphi \urcorner)$. Then, by logic, $T \vdash P^h_T(\ulcorner \varphi \urcorner) \to \varphi$. In all, $T \vdash P^h_T(\ulcorner \varphi \urcorner) \to \varphi$. ⊣

THEOREM 3.6. *Let* h *be a primitive-recursive function and* $P(x)$ *be a provability predicate such that*:

C1 *For all* Σ_1*-formulas* φ, $T \vdash \varphi \to P(\ulcorner \varphi \urcorner)$;
C2 *For all formulas* φ *and* ψ, $T \vdash \varphi \to \psi \implies T \vdash P(\ulcorner \varphi \urcorner) \to P(\ulcorner \psi \urcorner)$.

Then, for every formula φ, $T \vdash \neg P(\ulcorner \neg P^h_T(\ulcorner \varphi \urcorner) \urcorner) \to P^h_T(\ulcorner \varphi \urcorner)$.

PROOF. If h is primitive-recursive, then $\mathrm{R}_h(x, y)$ can be picked as being a Σ_1-formula. Clearly, $T \vdash \neg P^h_T(\ulcorner \varphi \urcorner) \leftrightarrow \mathrm{R}_h(\ulcorner \varphi \urcorner, 0)$. From **C2** we get that $T \vdash P(\ulcorner \neg P^h_T(\ulcorner \varphi \urcorner) \urcorner) \leftrightarrow P(\ulcorner \mathrm{R}_h(\ulcorner \varphi \urcorner, 0) \urcorner)$. From **C1**, $T \vdash \mathrm{R}_h(\ulcorner \varphi \urcorner, 0) \to P(\ulcorner \mathrm{R}_h(\ulcorner \varphi \urcorner, 0) \urcorner)$, so $T \vdash \neg P^h_T(\ulcorner \varphi \urcorner) \to P(\ulcorner \neg P^h_T(\ulcorner \varphi \urcorner) \urcorner)$, as wanted. ⊣

Let $\mathrm{Pr}_T(x) := \exists y.\, \mathrm{Proof}(x, y)$ denote the standard provability predicate in T [16, p. 826] and $\mathrm{Con}_T := \neg\, \mathrm{Pr}_T(\ulcorner \bot \urcorner)$.

THEOREM 3.7. *Given a primitive-recursive function* h, *for every formula* φ, $T + Con_T \vdash Pr_T(\ulcorner P^h_T(\ulcorner \varphi \urcorner) \urcorner) \to P^h_T(\ulcorner \varphi \urcorner)$.

PROOF. It is clear that $T \vdash \mathrm{Pr}_T(\ulcorner P^h_T(\ulcorner \varphi \urcorner) \urcorner) \wedge \mathrm{Pr}_T(\ulcorner \neg P^h_T(\ulcorner \varphi \urcorner) \urcorner) \to \mathrm{Pr}_T(\ulcorner \bot \urcorner)$. Thus, $T + \mathrm{Con}_T \vdash \mathrm{Pr}_T(\ulcorner P^h_T(\ulcorner \varphi \urcorner) \urcorner) \wedge \mathrm{Pr}_T(\ulcorner \neg P^h_T(\ulcorner \varphi \urcorner) \urcorner) \to \bot$. Hence, $T + \mathrm{Con}_T \vdash \neg\, \mathrm{Pr}_T(\ulcorner P^h_T(\ulcorner \varphi \urcorner) \urcorner) \vee \neg\, \mathrm{Pr}_T(\ulcorner \neg P^h_T(\ulcorner \varphi \urcorner) \urcorner)$, i.e., $T + \mathrm{Con}_T \vdash \mathrm{Pr}_T(\ulcorner P^h_T(\ulcorner \varphi \urcorner) \urcorner) \to \neg\, \mathrm{Pr}_T(\ulcorner \neg P^h_T(\ulcorner \varphi \urcorner) \urcorner)$. By the previous result we conclude that $T + \mathrm{Con}_T \vdash \mathrm{Pr}_T(\ulcorner P^h_T(\ulcorner \varphi \urcorner) \urcorner) \to P^h_T(\ulcorner \varphi \urcorner)$. ⊣

https://doi.org/10.1017/bsl.2021.68 Published online by Cambridge University Press

§4. Montagna's conjecture. Löb's theorem [9, pp. 28–29] expresses that for all formulas φ, if $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi$, then $T \vdash \varphi$. More generally, for all formulas φ,

$$T \vdash \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner) \leftrightarrow \mathrm{Pr}_T(\ulcorner \varphi \urcorner).$$

If one analyses the proof of Löb's theorem, it indicates that one can prove $\mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi$ only if one has already proved φ. It indicates, moreover, that $\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$ is only provable, if φ is proven in the first place. This intuition can be related to a problem proposed by Montagna in [3, p. 9]: "Does PA $\vdash_{k \text{ steps}} \mathrm{Pr}_{\mathrm{PA}}(\ulcorner \varphi \urcorner) \rightarrow \varphi$ imply PA $\vdash_{k \text{ steps}} \varphi$?" Let us consider a variant of this question: "Does $T \vdash_{k \text{ steps}} \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$ imply $T \vdash_{k \text{ steps}} \varphi$?" We prove that the adapted question to the provability notion $\vdash_{\leq h}$ is false: " $T \vdash_{\leq h} \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$ does not imply $T \vdash_{\leq h} \varphi$."

THEOREM 4.1. *For every primitive-recursive function $g(x, y)$ with $g(x, y) > y$, there are a sentence φ and a number n_0 such that $T \vdash_{\leq h} Pr_T(\ulcorner Pr_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$, but $T \nvdash_{\leq h} \varphi$, where $h := \lambda x.g(x, n_0)$.*

PROOF. We follow closely the proof of Theorem 14 from [9, p. 34]. Let g be a function-symbol that represents the primitive-recursive function g. By the diagonalization lemma, there is a sentence φ such that

$$T \vdash \varphi \leftrightarrow \exists y.(\,\mathrm{Proof}(\ulcorner \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)\urcorner, y) \\ \wedge \forall z \leq \mathrm{g}(\ulcorner \varphi \urcorner, y).\neg \mathrm{Proof}(\ulcorner \varphi \urcorner, z)).$$

By construction, $T + \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)\urcorner) + \neg \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \vdash \varphi$. As T is Σ_1-complete, $T + \varphi \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$. Thus, we can conclude that $T + \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)\urcorner) \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$. By Löb's theorem, it follows that $T \vdash \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)\urcorner) \leftrightarrow \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner)\urcorner)$. Hence, $T + \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner)\urcorner) \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$. Again by Löb's theorem, it follows that $T \vdash \mathrm{Pr}_T(\ulcorner \varphi \urcorner)$, and consequently $T \vdash \varphi$ and $T \vdash \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$. This means that φ is true. Let n_0 satisfy the true existential property of φ. Then, n_0 is the code of a proof of $\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$. By hypothesis on g, it follows that $n_0 < g(\#\,\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner), n_0) = h(\#\,\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner))$, *ergo* $T \vdash_{\leq h} \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$. From the fact that φ is true one can conclude that for all $z \leq g(\#\varphi, n_0)$, z is not the code of a proof of φ. This means that $T \nvdash_{\leq h} \varphi$. ⊣

If a formula φ is provable in T, we define

$$\|\varphi\|_T := \min\{n | n \text{ is the code of a proof of } \varphi \text{ in } T\}.$$

Moreover, if φ and ψ are formulas, we stipulate that $\varphi <_T \psi$ if $T \vdash \varphi \wedge \psi$ and $\|\varphi\|_T < \|\psi\|_T$. The following result confirms that the mentioned intuition that a proof of $\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner)$ should encompass, in a way, a proof of φ fails.

THEOREM 4.2. *There is a formula φ such that*

$$Pr_T(\ulcorner Pr_T(\ulcorner \varphi \urcorner) \rightarrow \varphi \urcorner) <_T \varphi.$$

PROOF. By the diagonalization lemma, there is a sentence φ such that

$$T \vdash \varphi \leftrightarrow \exists y.(\,\mathrm{Proof}(\ulcorner \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \varphi \urcorner)\urcorner, y) \\ \wedge \forall z \leq y.\neg\,\mathrm{Proof}(\ulcorner \varphi \urcorner, z)).$$

Applying the same reasoning as in the previous proof, it follows that $T \vdash \mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \varphi \urcorner) \wedge \varphi$; in particular φ is true. Take n_0 as being the natural number that is guaranteed to exist from the true formula φ. It is straightforward that $\|\,\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \varphi \urcorner)\|_T \leq n_0$. As φ is true, it follows that for all $z \leq n_0$, z is not the code of a proof of φ. Hence, $n_0 < \|\varphi\|_T$, and so $\|\,\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \varphi \urcorner)\|_T < \|\varphi\|_T$. In all, $\mathrm{Pr}_T(\ulcorner \mathrm{Pr}_T(\ulcorner \varphi \urcorner) \to \varphi \urcorner) <_T \varphi$. ⊣

§5. Variants of Kreisel's conjecture.

In this section we present some partial results related to Kreisel's conjecture, namely variants of the conjecture for provability predicates in the present of different derivability conditions. In this section, the theory T does not need to be r.e.

THEOREM 5.1. *Let h be a primitive-recursive function and T be such that there is a provability predicate $P(x)$ satisfying*:

- **C1** *If $\varphi(x)$ is a Σ_n-formula, then $T \vdash \varphi(x) \to P(\ulcorner \varphi(\dot{x})\urcorner)$;*
- **C2** $T \vdash P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \to P(\ulcorner \varphi(\dot{x})\urcorner)$;
- **C3** *For all formulas $\varphi(x)$ and $\psi(x)$, $T \vdash P(\ulcorner \varphi(\dot{x}) \to \psi(\dot{x})\urcorner) \to (P(\ulcorner \varphi(\dot{x})\urcorner) \to P(\ulcorner \psi(\dot{x})\urcorner))$.*

If $\varphi(x)$ is a Π_n-formula such that $T \vdash \forall x.P^h_T(\ulcorner \varphi(\dot{x})\urcorner)$, then $T + Con_P \vdash \forall x.\varphi(x)$.

PROOF. As $\varphi(x)$ is Π_n, by **C1**, we have $T \vdash \neg\varphi(x) \to P(\ulcorner \neg\varphi(\dot{x})\urcorner)$. Thus, $T \vdash \exists x.\neg\varphi(x) \to \exists x.P(\ulcorner \neg\varphi(\dot{x})\urcorner)$. By **C2**, $T \vdash P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \to P(\ulcorner \varphi(\dot{x})\urcorner)$. Hence, $T \vdash \forall x.P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \to \forall x.P(\ulcorner \varphi(\dot{x})\urcorner)$. So, $T + \forall x.P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \wedge \exists x.\neg\varphi(x) \vdash \exists x.P(\ulcorner \neg\varphi(\dot{x})\urcorner) \wedge \forall x.P(\ulcorner \varphi(\dot{x})\urcorner)$. By **C3**, $T + \forall x.P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \wedge \exists x.\neg\varphi(x) \vdash P(\ulcorner \bot \urcorner)$, i.e., $T + \mathrm{Con}_P \vdash \forall x.P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \to \forall x.\varphi(x)$. ⊣

The condition $T \vdash \forall x.P^h_T(\ulcorner \varphi(\dot{x})\urcorner)$ corresponds to an arithmetization of the antecedent of a version of Kreisel's conjecture. Thus, the result is weaker than Kreisel's conjecture. If $T \vdash \mathrm{Con}_P$, then the previous result can be proved inside T. The next result is a particular case of the previous theorem.

COROLLARY 5.1. *Let h be a primitive-recursive function. If $\varphi(x)$ is a Π_1-formula such that $T \vdash \forall x.P^h_{PA}(\ulcorner \varphi(\dot{x})\urcorner)$, then $PA + Con_{PA} \vdash \forall x.\varphi(x)$.*

PROOF. The corollary follows immediately from the fact that $\mathrm{Pr}_{\mathrm{PA}}$ satisfies **C1** and **C2** of the previous theorem [7]. ⊣

THEOREM 5.2. *Let h be a primitive-recursive function and T be such that there is a provability predicate $P(x)$ satisfying*:

- **C1** *For all formulas $\varphi(x)$, $T \vdash P(\ulcorner \neg\varphi(\dot{x})\urcorner) \to P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{x})\urcorner)\urcorner)$;*

https://doi.org/10.1017/bsl.2021.68 Published online by Cambridge University Press

C2 *If* $\varphi(x)$ *is a* Σ_n*-formula, then* $T \vdash \varphi(x) \to P(\ulcorner\varphi(\dot{x})\urcorner)$;
C3 *For all formulas* $\varphi(x)$ *and* $\psi(x)$, $T \vdash P(\ulcorner\varphi(\dot{x}) \to \psi(\dot{x})\urcorner) \to (P(\ulcorner\varphi(\dot{x})\urcorner) \to P(\ulcorner\psi(\dot{x})\urcorner))$.

If $\varphi(x)$ *is a* Π_n*-formula such that* $T \vdash \forall x.P^h_T(\ulcorner\varphi(\dot{x})\urcorner)$, *then* $T + Con_P \vdash \forall x.\varphi(x)$.

Proof. As $\neg\varphi(x)$ is Σ_n, by **C2** $T \vdash \neg\varphi(x) \to P(\ulcorner\neg\varphi(\dot{x})\urcorner)$. Thus, $T \vdash \neg P(\ulcorner\neg\varphi(\dot{x})\urcorner) \to \varphi(x)$. By **C3**, we have that $T + P(\ulcorner P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner) \wedge P(\ulcorner\neg P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner) \vdash P(\ulcorner\bot\urcorner)$, since $\neg\varphi := \varphi \to \bot$. Hence, $T + \mathrm{Con}_P \vdash P(\ulcorner P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner) \to \neg P(\ulcorner\neg P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner)$. Together with **C1** we get that $T + \mathrm{Con}_P \vdash P(\ulcorner P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner) \to \neg P(\ulcorner\neg\varphi(\dot{x})\urcorner)$, but, by what was previously concluded, one gets that $T + \mathrm{Con}_P \vdash P(\ulcorner P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner) \to \varphi(x)$. Suppose that $T \vdash \forall x.P^h_T(\ulcorner\varphi(\dot{x})\urcorner)$. As h is primitive-recursive, we have that $P^h_T(x)$ is Σ_1. *Ergo*, by **C2**, $T \vdash P^h_T(\ulcorner\varphi(\dot{x})\urcorner) \to P(\ulcorner P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner)$. By assumption, it follows that $T \vdash \forall x.P(\ulcorner P^h_T(\ulcorner\varphi(\dot{x})\urcorner)\urcorner)$; therefore, $T \vdash \forall x.\varphi(x)$. ⊣

In the next result, we drop the assumption that h is primitive-recursive, but we need to strengthen condition **C1**.

Theorem 5.3. *Let* T *be such that there is a provability predicate* $P(x)$ *satisfying*:

C1 *For all formulas* $\varphi(x)$, $T \vdash P^h_T(\ulcorner\varphi(\dot{x})\urcorner) \to P(\ulcorner\varphi(\dot{x})\urcorner)$;
C2 *If* $\varphi(x)$ *is a* Σ_n*-formula, then* $T \vdash \varphi(x) \to P(\ulcorner\varphi(\dot{x})\urcorner)$;
C3 *For all formulas* $\varphi(x)$ *and* $\psi(x)$, $T \vdash P(\ulcorner\varphi(\dot{x}) \to \psi(\dot{x})\urcorner) \to (P(\ulcorner\varphi(\dot{x})\urcorner) \to P(\ulcorner\psi(\dot{x})\urcorner))$.

If $\varphi(x)$ *is a* Π_n*-formula such that* $T \vdash \forall x.P^h_T(\ulcorner\varphi(\dot{x})\urcorner)$, *then* $T + Con_P \vdash \forall x.\varphi(x)$.

Proof. As $\varphi(x)$ is Π_n, by **C2**, $T \vdash \exists x.\neg\varphi(x) \to \exists x.P(\ulcorner\neg\varphi(\dot{x})\urcorner)$. As $T \vdash \forall x.P^h_T(\ulcorner\varphi(\dot{x})\urcorner)$, it follows, by **C1**, that $T \vdash \forall x.P(\ulcorner\varphi(\dot{x})\urcorner)$. This, together with the fact that $T + \exists x.\neg\varphi(x) \vdash \exists x.P(\ulcorner\neg\varphi(\dot{x})\urcorner)$, yields $T + \exists x.\neg\varphi(x) \vdash \exists x.P(\ulcorner\neg\varphi(\dot{x})\urcorner) \wedge P(\ulcorner\varphi(\dot{x})\urcorner)$. As $\neg\varphi := \varphi \to \bot$, it follows by **C3** that $T + \exists x.\neg\varphi(x) \vdash \exists x.P(\ulcorner\bot\urcorner)$, i.e., $T + \exists x.\neg\varphi(x) \vdash P(\ulcorner\bot\urcorner)$. Hence, $T + \mathrm{Con}_P \vdash \forall x.\varphi(x)$. ⊣

Feferman, in [4], requires an intensionally correct arithmetization of provability to satisfy several conditions including **C1**, **C2**, and **C3**.

Corollary 5.2. *Let* T *be such that there is a provability predicate* $P(x)$ *satisfying*:

C1 *For all formulas* $\varphi(x)$, $T \vdash P^h_T(\ulcorner\varphi(\dot{x})\urcorner) \to P(\ulcorner\varphi(\dot{x})\urcorner)$;

C2 *If* $\varphi(x)$ *is a* Σ_1*-formula, then* $T \vdash \varphi(x) \rightarrow P(\ulcorner \varphi(\dot{x}) \urcorner)$;
C3 *For all formulas* $\varphi(x)$ *and* $\psi(x)$, $T \vdash P(\ulcorner \varphi(\dot{x}) \rightarrow \psi(\dot{x}) \urcorner) \rightarrow (P(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow P(\ulcorner \psi(\dot{x}) \urcorner))$;
C4 $T \vdash Con_P$.

If $\varphi(x)$ *is a* Π_1*-formula such that* $T \vdash \forall x.P^h_T(\ulcorner \varphi(\dot{x}) \urcorner)$, *then* $T \vdash \forall x.\varphi(x)$.

PROOF. Follows immediately from the previous theorem. ⊣

By [4, 8], there is a provability predicate that satisfies **C2**, **C3**, and **C4**. Furthermore, if $P(x)$ is a provability predicate that satisfies **C2** and **C4**, then $P'(x) := P^h_T(x) \vee P(x)$ is a provability predicate that satisfies **C1**, **C2**, and **C4**. For this reason, we believe that any sufficiently strong theory T satisfies all the previous conditions.

Using the theory K_T we can go even further:

COROLLARY 5.3. *Let* T *be a theory in the conditions of the previous result. If* $\varphi(x)$ *is a* Π_1*-formula such that, for all* $n \in \mathbb{N}$, $T \vdash_{\leq h} \varphi(\overline{n})$, *then* $K_T \vdash \forall x.\varphi(x)$.

PROOF. By the proof of Corollary 5.2, it can be concluded that $T \vdash \forall x.P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow \forall x.\varphi(x)$. Thus, the result follows from Theorem 3.1. ⊣

A result similar to Theorem 5.3, for some Σ-formulas, holds in the presence of the stronger schema $\omega\text{–}\mathrm{Con}^n_P$:

$$P(\ulcorner \exists x.\varphi(x, \dot{y}) \urcorner) \rightarrow \exists x.\neg P(\ulcorner \neg\varphi(\dot{x}, \dot{y}) \urcorner), \qquad \varphi(x) \text{ is a } \Pi_n\text{-formula.}$$

THEOREM 5.4. *Let* T *be such that there is a provability predicate* $P(x)$ *satisfying*:

C1 *For all formulas* $\varphi(x)$, $T \vdash P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow P(\ulcorner \varphi(\dot{x}) \urcorner)$;
C2 *For all* Σ_n*-formulas* $\varphi(x, y)$, $T \vdash \varphi(x, y) \rightarrow P(\ulcorner \varphi(\dot{x}, \dot{y}) \urcorner)$.

Suppose that $\varphi(x)$ *is a* Π_n*-formula. If* $T \vdash \forall y.P^h_T(\ulcorner \exists x.\varphi(x, \dot{y}) \urcorner)$, *then* $T + \omega\text{–}Con^n_P \vdash \forall y.\exists x.\varphi(x, y)$.

PROOF. Suppose that $T \vdash \forall y.P^h_T(\ulcorner \exists x.\varphi(x, \dot{y}) \urcorner)$. By **C1**, we have $T \vdash \forall y.P(\ulcorner \exists x.\varphi(x, \dot{y}) \urcorner)$. Hence, $T + \omega\text{–}\mathrm{Con}^n_P \vdash \forall y.\exists x.\neg P(\ulcorner \neg\varphi(\dot{x}, \dot{y}) \urcorner)$, i.e., $T + \omega\text{–}\mathrm{Con}^n_P \vdash \neg\exists y.\forall x.P(\ulcorner \neg\varphi(\dot{x}, \dot{y}) \urcorner)$. Furthermore, by **C2**, we have $T + \exists y.\forall x.\neg\varphi(x, y) \vdash \exists y.\forall x.P(\ulcorner \neg\varphi(\dot{x}, \dot{y}) \urcorner)$. Therefore, $T + \omega\text{–}\mathrm{Con}^n_P \vdash \neg\exists y.\forall x.\neg\varphi(x, y)$, and so, $T + \omega\text{–}\mathrm{Con}^n_P \vdash \forall y.\exists x.\varphi(x, y)$. ⊣

If $T \vdash \omega\text{–}\mathrm{Con}^n_P$, then everything is provable in T. We can yet get a stronger result, but, like before, we need a stronger schema. Let $\omega\text{–}\mathrm{Con}^{3,n}_P$ be the following schema:

$$P(\ulcorner \exists y.\varphi(\dot{x}, y, \dot{z}) \urcorner) \rightarrow \exists y.\neg P(\ulcorner \neg\varphi(\dot{x}, \dot{y}, \dot{z}) \urcorner), \quad \varphi(x) \text{ is a } \Pi_n\text{-formula.}$$

THEOREM 5.5. *Let* T *be such that there is a provability predicate* $P(x)$ *satisfying*:

C1 *For all formulas* $\varphi(x)$, $T \vdash P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow P(\ulcorner \varphi(\dot{x}) \urcorner)$;

https://doi.org/10.1017/bsl.2021.68 Published online by Cambridge University Press

C2 *For all* Π_n*-formulas* $\varphi(x, y, z)$, $T \vdash P(\ulcorner \forall x.\exists y.\varphi(x, y, \dot{z})\urcorner) \rightarrow \forall x.P(\ulcorner \exists y.\varphi(\dot{x}, y, \dot{z})\urcorner)$;

C3 *For all* Σ_n*-formulas* $\varphi(x, y, z)$, $T \vdash \varphi(x, y, z) \rightarrow P(\ulcorner \varphi(\dot{x}, \dot{y}, \dot{z})\urcorner)$.

Suppose that $\varphi(x)$ *is a* Π_n*-formula. If* $T \vdash \forall z.P^h_T(\ulcorner \forall x.\exists y.\varphi(x, y, \dot{z})\urcorner)$, *then* $T + \omega\text{–}Con^{3,n}_P \vdash \forall z.\forall x.\exists y.\varphi(x, y, z)$.

Proof. $T + \forall z.P^h_T(\ulcorner \forall x.\exists y.\varphi(x, y, \dot{z})\urcorner) \vdash \forall z.P(\ulcorner \forall x.\exists y.\varphi(x, y, \dot{z})\urcorner)$, by **C1**. From **C2**, $T + \forall z.P^h_T(\ulcorner \forall x.\exists y.\varphi(x, y, \dot{z})\urcorner) \vdash \forall z.\forall x.P(\ulcorner \exists y.\varphi(\dot{x}, y, \dot{z})\urcorner)$. This means that $T + \omega\text{–}\mathrm{Con}^{3,n}_P \vdash \forall z.\forall x.\exists y.\neg P(\ulcorner \neg\varphi(\dot{x}, \dot{y}, \dot{z})\urcorner)$, i.e., $T + \omega\text{–}\mathrm{Con}^{3,n}_P \vdash \neg\exists z.\exists x.\forall y.P(\ulcorner \neg\varphi(\dot{x}, \dot{y}, \dot{z})\urcorner)$. As $\varphi(x, y, z)$ is Π_n, by **C3**, $T + \exists z.\exists x.\forall y.\neg\varphi(x, y, z) \vdash \exists z.\exists x.\forall y.P(\ulcorner \neg\varphi(\dot{x}, \dot{y}, \dot{z})\urcorner)$. Altogether, $T + \omega\text{–}\mathrm{Con}^{3,n}_P \vdash \neg\exists z.\exists x.\forall y.\neg\varphi(x, y, z)$. ⊣

By Theorem 3.5, we have that the Local Reflection Principle (see [16, p. 845]) of $P^h_T(x)$ is provable in T, i.e., $T \vdash P^h_T(\ulcorner \varphi \urcorner) \rightarrow \varphi$. In fact, we have the following result.

Theorem 5.6. *Suppose that h is primitive-recursive. Let T be such that there is a provability predicate* $P(x)$ *satisfying*:

C1 *For all formulas* $\varphi(x)$, $T \vdash P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \rightarrow P(\ulcorner \varphi(\dot{x})\urcorner)$;

C2 *For all* Σ_1*-formulas* $\varphi(x)$, $T \vdash \varphi(x) \rightarrow P(\ulcorner \varphi(\dot{x})\urcorner)$;

C3 *For all formulas* $\varphi(x)$ *and* $\psi(x)$, $T \vdash P(\ulcorner \varphi(\dot{x}) \rightarrow \psi(\dot{x})\urcorner) \rightarrow (P(\ulcorner \varphi(\dot{x})\urcorner) \rightarrow P(\ulcorner \psi(\dot{x})\urcorner))$;

C4 $T \vdash Con_P$;

C5 *For all formulas* $\varphi(x)$, $T \vdash \varphi(x) \implies T \vdash P(\ulcorner \varphi(\dot{x})\urcorner)$.

Then, $T \vdash \forall x.P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner) \rightarrow \varphi(\dot{x})\urcorner)$.

Proof. As h is primitive-recursive, we know that $P^h_T(x)$ and $\neg P^h_T(x)$ are Σ_1-formulas. By **C2**, $T \vdash P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \rightarrow P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner)$, so $T \vdash \neg P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) \rightarrow \neg P^h_T(\ulcorner \varphi(\dot{x})\urcorner)$. It holds that $T \vdash \neg P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \rightarrow P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner)$. So, $T \vdash \neg P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) \rightarrow P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner)$, i.e., $T \vdash P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) \vee P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner)$.

From logic, $T \vdash \neg P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \rightarrow (P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \rightarrow \varphi(x))$. Hence, by **C5**, $T \vdash P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner) \rightarrow (P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner) \rightarrow \varphi(\dot{x}))\urcorner)$; thus, by **C3**, $T \vdash P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) \rightarrow P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner) \rightarrow \varphi(\dot{x})\urcorner)$.

From **C1**, $T + \neg P(\ulcorner \varphi(\dot{x})\urcorner) \vdash \neg P^h_T(\ulcorner \varphi(\dot{x})\urcorner)$. By **C2**, $T + \neg P(\ulcorner \varphi(\dot{x})\urcorner) \vdash P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner)$. *Ergo* we have $T + P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) + \neg P(\ulcorner \varphi(\dot{x})\urcorner) \vdash P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) \wedge P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner)$. From **C3** and **C4**, it follows that $T + P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) + \neg P(\ulcorner \varphi(\dot{x})\urcorner) \vdash \bot$, i.e., $T \vdash P(\ulcorner P^h_T(\ulcorner \varphi(\dot{\dot{x}})\urcorner)\urcorner) \rightarrow P(\ulcorner \varphi(\dot{x})\urcorner)$. From logic, $T \vdash \varphi(x) \rightarrow (P^h_T(\ulcorner \varphi(\dot{x})\urcorner) \rightarrow \varphi(x))$; thus, by

https://doi.org/10.1017/bsl.2021.68 Published online by Cambridge University Press

C5 and **C3**, $T \vdash P(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow P(\ulcorner P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow \varphi(\dot{x}) \urcorner)$. Hence, $T \vdash P(\ulcorner P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \urcorner) \rightarrow P(\ulcorner P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow \varphi(\dot{x}) \urcorner)$.

So we have $T \vdash P(\ulcorner P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \urcorner) \vee P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \urcorner)$ and also $T \vdash P(\ulcorner \neg P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \urcorner) \rightarrow P(\ulcorner P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow \varphi(\dot{x}) \urcorner)$. From before, we have $T \vdash P(\ulcorner P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \urcorner) \rightarrow P(\ulcorner P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow \varphi(\dot{x}) \urcorner)$, and thus the result follows. ⊣

Inspired by the previous fact, one can consider the *uniform reflection principle* schema, $\mathrm{RFN}^h(T)$, for the provability notion $P^h_T(x)$ (see [16, p. 845]):

$$\forall x.(P^h_T(\ulcorner \varphi(\dot{x}) \urcorner) \rightarrow \varphi(x)), \qquad \varphi(x) \text{ has only } x \text{ free.}$$

With Γ being an arbitrary class of formulas (for example, Σ_n, Π_n, or even Δ_n), we denote by $\mathrm{RFN}^h_\Gamma(T)$ the previous schema restricted to Γ-formulas and define $T^h_\Gamma := K_T + \mathrm{RFN}^h_\Gamma(T)$. There is a deep relation between ω-consistency and reflection principles [16, p. 853]: Restrictions to Π-formulas of the uniform reflexion principle for the standard provability predicate are equivalent to restrictions of the schema $\omega\text{–}\mathrm{Con}^n_{\mathrm{Pr}}$ from above to Σ-formulas. Note that we are adding ω-consistency and not ω-completeness; hence Kreisel's conjecture—which follows immediately from ω-completeness—is not being trivialised.

Now we presented another adapted version of Kreisel's conjecture.

THEOREM 5.7. *Let h be a total recursive function and $\varphi(x)$ be a Γ-formula such t hat, for all $n \in \mathbb{N}$, $T \vdash_{\leq h} \varphi(\overline{n})$. Then, $T^h_\Gamma \vdash \forall x.\varphi(x)$.*

PROOF. Let h be a total recursive function and $\varphi(x)$ be a Γ-formula such that, for all $n \in \mathbb{N}$, $\mathrm{PA} \vdash_{\leq h} \varphi(\overline{n})$. By Theorem 3.1, we have that $K_T \vdash \forall x.P^h_T(\ulcorner \varphi(\dot{x}) \urcorner)$. Thus, by $\mathrm{RFN}^h_\Gamma(T)$, it follows that $T^h_\Gamma \vdash \forall x.\varphi(x)$. ⊣

Note that there are no particular reasons to believe that the theory K_T is effectively axiomatisable. This is something worth studying.

Furthermore, one could consider a modal logic with modalities $\Box$ (interpreted by $\mathrm{Pr}_{\mathrm{PA}}(\cdot)$) and $\Box_{\leq h}$ (with $P^h_T(\cdot)$ as an intended interpretation) and, at least, the usual axioms of $\Box$ and the properties of $P^h_T(\cdot)$. For example, as modal versions of Theorems 3.3, 3.5, and 3.6, one could add the following axioms:

Ax.1 $(\Box\Box_{\leq h} A) \vee (\Box\neg\Box_{\leq h} A)$;
Ax.2 $\Box_{\leq h} A \rightarrow A$;
Ax.3 $\neg\Box\neg\Box_{\leq h} A \rightarrow \Box_{\leq h} A$.

§6. On $\vdash_{k \text{ steps}}$ and $\vdash_{\leq h}$.

From [3, p. 8], we know the following fact:

THEOREM 6.1. *If T is a finitely axiomatised theory, then there is a total recursive function $f(k, \#\varphi)$ such that*

$$T \vdash_{k\ steps} \varphi \implies T \vdash_{f(k,\#\varphi)\ symbols} \varphi.$$

With this theorem, one can establish a relation between $\vdash_{k \text{ steps}}$ and $\vdash_{\leq h}$.

https://doi.org/10.1017/bsl.2021.68 Published online by Cambridge University Press

THEOREM 6.2. *Given k, if T is a finitely axiomatised theory, then the function*

$$g_k(\#\varphi) := \begin{cases} 1, & T \vdash_{k\ steps} \varphi, \\ 0, & otherwise \end{cases}$$

is recursive.

PROOF. Let k be fixed. We will intuitively describe the algorithm that computes the function g_k. Consider the input $\#\varphi$. Compute, by Theorem 6.2, $f(k, \#\varphi)$. If φ is provable with at most k steps, then it must be provable using at most $f(k, \#\varphi)$ symbols. In such a hypothetical proof, clearly there are, at most, $f(k, \#\varphi)$ different variables. Furthermore, the variables, besides the ones that occur in φ, can be arbitrarily chosen, i.e., if one performs a change of variables in the proof without changing the variables occurring in φ, one maintains the soundness of the proof and the number of steps in it. This means that it is enough to consider a finite set of variables consisting of: the variables in φ and $f(k, \#\varphi)$ other variables. Then, the algorithm considers all possible finite strings constructed using the finite set consisting of: the logical connectives, quantifiers, parenthesis, a blank symbol (to separate the steps in a proof), and the variables of the finite set that was mentioned. By vanishing over all the (finite) possible strings, the algorithm tests if any of them is a proof of φ with at most k steps. If there is any, it outputs 1; otherwise it ought to output 0. Thus, the algorithm outputs 1 exactly when φ is provable with at most k steps. ⊣

THEOREM 6.3. *Given k, if T is a finitely axiomatised theory, then there is a total recursive function h_k such that*

$$T \vdash_{k\ steps} \varphi \implies T \vdash_{\leq h_k} \varphi.$$

PROOF. Let g_k be as in Theorem 6.2. It is immediate that the function

$$h_k(n) := \begin{cases} m, & \text{if } g_k(n) = 1 \text{ and } m \text{ is the smallest code of a proof of the} \\ & \text{formula coded by } n \text{ with at most } k \text{ steps,} \\ 0, & \text{otherwise} \end{cases}$$

is total recursive. We show that $T \vdash_{k\text{ steps}} \varphi \implies T \vdash_{\leq h_k} \varphi$. If $T \vdash_{k\text{ steps}}$, then $g_k(n) = 1$ and so $h_k(\#\varphi)$ is the code of a proof of φ with at most k steps; by definition, $T \vdash_{\leq h_k} \varphi$. ⊣

There are two immediate consequences of the previous result.

COROLLARY 6.1. *Suppose that T is a finitely axiomatised theory satisfying the conditions of Corollary 5.3 for the function h_k and that $\varphi(x)$ is a Π_1-formula. If for all $n \in \mathbb{N}$, $T \vdash_{k\ steps} \varphi(\overline{n})$, then $K_T \vdash \forall x.\varphi(x)$.*

PROOF. Follows from the previous theorem and from Corollary 5.3. ⊣

COROLLARY 6.2. *Suppose that T is a finitely axiomatised theory and that $\varphi(x)$ be a Γ-formula. If, for all $n \in \mathbb{N}$, $T \vdash_{k\ steps} \varphi(\overline{n})$, then $T_\Gamma^{h_k} \vdash \forall x.\varphi(x)$.*

PROOF. Follows from Theorems 5.7 and 6.3. ⊣

We finish with an open problem.

PROBLEM. Is there a total recursive function h such that, for all formulas φ, $\mathrm{PA} \vdash_{k \text{ steps}} \varphi \implies \mathrm{PA} \vdash_{\leq h} \varphi$? ⊣

§7. Conclusion.

Kreisel's conjecture is the fundamental problem of k-steps-provability. As mentioned in the introduction, there are some solutions under specific conditions. Usually they rely on properties of the considered formulas or properties of the theory. In this paper, we presented a novel approach to the conjecture, where we abstracted from the concrete formalization.

We introduced a notion of provability $\vdash_{\leq h}$ expressing that $T \vdash_{\leq h} \varphi$ holds if there is a proof of φ in T whose code is at most $h(\#\varphi)$. This is clearly an intensional notion. We studied the representation of $\vdash_{\leq h}$ inside the theory T using the formula $P_T^h(x)$ and several of its properties. Montagna's conjecture ("Does $\mathrm{PA} \vdash_{k \text{ steps}} \mathrm{Pr}_{\mathrm{PA}}(\ulcorner \varphi \urcorner) \to \varphi$ imply $\mathrm{PA} \vdash_{k \text{ steps}} \varphi$?") was analysed for the notion $\vdash_{\leq h}$.

We also considered variants of Kreisel's conjecture for provability predicates with different derivability conditions. From the results, we like to highlight Theorem 5.4 that, using a form of ω-consistency ($\omega\text{–}\mathrm{Con}_P^n$) and under certain derivability conditions, allows to conclude $T + \omega\text{–}\mathrm{Con}_P^n \vdash \forall y.\exists x.\varphi(x, y)$ from $T \vdash \forall y.P_T^h(\ulcorner \exists x.\varphi(x, \dot{y})\urcorner)$.

The paper finishes with connections between $\vdash_{k \text{ steps}}$ and $\vdash_{\leq h}$, in particular, two forms of Kreisel's conjecture for $\vdash_{\leq h}$ (Corollaries 6.1 and 6.2).

Acknowledgments. This work was funded by the following FCT-projects: Centro de Matemática e Aplicações (UIDB/00297/2020) and Bolsa de Doutoramento (SFRH/BD/143756/2019). The second author was also supported by the Udo Keller Foundation.

REFERENCES

[1] M. BAAZ and P. PUDLÁK, *Kreisel's conjecture for L∃1*, ***Arithmetic, Proof Theory and Computational Complexity*** (P. Clote and J. Krajićek, editors), Oxford University Press, Oxford, 1993, pp. 29–59.

[2] S. CAVAGNETTO, *The lengths of proofs: Kreisel's conjecture and Gödel's speed-up theorem*. ***Journal of Mathematical Sciences***, vol. 158 (2009), no. 5, pp. 689–707.

[3] P. CLOTE and J. KRAJIĆEK, *Open problems*, ***Arithmetic, Proof Theory and Computational Complexity*** (P. Clote and J. Krajićek, editors), Oxford University Press, Oxford, 1993, pp. 1–19.

[4] S. FEFERMAN, *Arithmetization of metamathematics in a general setting*. ***Fundamenta Mathematicae***, vol. 49 (1960), no. 1, pp. 35–92.

[5] H. FRIEDMAN, *One hundred and two problems in mathematical logic*. ***The Journal of Symbolic Logic***, vol. 40 (1975), no. 2, pp. 113–129.

[6] P. HRUBEŠ, *Theories very close to PA where Kreisel's conjecture is false*. ***The Journal of Symbolic Logic***, vol. 72 (2007), no. 1, pp. 123–137.

[7] T. KURAHASHI, *Arithmetical soundness and completeness for Σ_2 numerations*. ***Studia Logica***, vol. 106 (2018), no. 6, pp. 1181–1196.

[8] ———, *A note on derivability conditions*. ***The Journal of Symbolic Logic***, vol. 85 (2020), no. 3, pp. 1224–1253.

[9] P. Lindström, ***Aspects of Incompleteness***, Lecture Notes in Logic 10, Cambridge University Press, Cambridge, 2017.

[10] S. Liu, *An enumeration of the primitive recursive functions without repetition*. ***Tohoku Mathematical Journal, Second Series***, vol. 12 (1960), no. 3, pp. 400–402.

[11] T. Miyatake, *On the length of proofs in a formal systems*. ***Tsukuba Journal of Mathematics***, vol. 4 (1980), no. 1, pp. 115–125.

[12] R. Parikh, *Some results on the length of proofs*. ***Transactions of the American Mathematical Society***, vol. 177 (1973), pp. 29–36.

[13] M. Rathjen and W. Sieg, *Proof theory*, ***The Stanford Encyclopedia of Philosophy*** (E. N. Zalta, editor), Metaphysics Research Lab, Stanford University, Stanford, 2018, Appendix F. Provably Recursive Functions.

[14] R. Ritchie and P. Young, *Strong representability of partial functions in arithmetic theories*. ***Information Sciences***, vol. 1 (1969), no. 2, pp. 189–204.

[15] P. Smith, ***An Introduction to Gödel's Theorems***, second ed., Cambridge Introductions to Philosophy, Cambridge University Press, Cambridge, 2013.

[16] C. Smorynski, *The incompleteness theorems*, ***Handbook of Mathematical Logic***, eighth ed., North-Holland, Amsterdam, 1993, pp. 821–865.

CENTRO DE MATEMÁTICA E APLICAÇÕES
NOVA SCHOOL OF SCIENCE AND TECHNOLOGY
P-2829-516 CAPARICA, PORTUGAL
E-mail: pgd.santos@campus.fct.unl.pt

THEORIE UND GESCHICHTE DER WISSENSCHAFTEN
UNIVERSITÄT TÜBINGEN
DOBLERSTR. 33, D-72074 TÜBINGEN, GERMANY
and
CENTRO DE MATEMÁTICA E APLICAÇÕES
NOVA SCHOOL OF SCIENCE AND TECHNOLOGY
P-2829-516 CAPARICA, PORTUGAL
E-mail: reinhard.kahle@uni-tuebingen.de