

SELF-REFERENCE UPFRONT: A STUDY OF SELF-REFERENTIAL GÖDEL NUMBERINGS

BALTHASAR GRABMAYR

Humboldt University of Berlin
and

ALBERT VISSER

Utrecht University

Abstract. In this paper we examine various requirements on the formalisation choices under which self-reference can be adequately formalised in arithmetic. In particular, we study self-referential numberings, which immediately provide a strong notion of self-reference even for expressively weak languages. The results of this paper suggest that the question whether truly self-referential reasoning can be formalised in arithmetic is more sensitive to the underlying coding apparatus than usually believed. As a case study, we show how this sensitivity affects the formal study of certain principles of self-referential truth.

§1. Introduction. Notions of self-reference feature prominently in the philosophical literature, yet they are notoriously elusive and mostly left imprecise. The study of certain logical and philosophical issues, however, such as semantic paradoxes and the semantics of both natural and formal languages, requires a precise understanding of self-reference. For instance, an adequate evaluation of the contentious hypothesis that circularity lies at the root of all semantic paradoxes requires a satisfactory and precise explication of self-reference (see [17]). Following the wide-spread custom of *arithmetisation*, the explication and investigation of self-reference is therefore typically carried out in a formal arithmetical framework.

There are essentially two ways an arithmetical sentence can be taken to (directly) refer to itself, namely, in virtue of containing a term denoting (the code of) itself, or by means of quantification [17, 21].¹ In this paper we are only concerned with the first kind of self-reference, also called “self-reference by mention” or in short “m-self-reference” [21, p. 581]. Halbach and Visser [10] trace this notion of self-reference back to the

Received: August 12, 2020.

2020 *Mathematics Subject Classification*: 03A05, 03F30, 03F40.

Key words and phrases: Self-reference, Gödel numberings, arithmetic, truth-theories.

¹ In his Princeton seminar on truth, September 29, 1982, Saul Kripke distinguishes these two notions of self-reference from *demonstrative* self-reference, which is obtained by indexicals, such as the word “this”, and therefore is unattainable in standard formalisms of arithmetic. For proposals how such indexicals can be added to systems of arithmetic see [29, 31]. We are grateful to Allen Hazen for sharing with us his transcript of Kripke’s seminar and to Saul Kripke for his permission to use them.

© The Author(s), 2021. Published by Cambridge University Press on behalf of The Association for Symbolic Logic. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

famous exchange of Kreisel and Henkin regarding Henkin's question as to whether sentences that state their own PA-provability are provable in PA. Heck [13], for instance, considers m-self-reference the only legitimate way to formalise truly self-referential reasoning in arithmetic. A precise notion of m-self-referentiality in an arithmetical framework rests on three formalisation choices:

- i. a formal language
- ii. a Gödel numbering
- iii. a naming device

Given the central importance of m-self-reference in the literature, the primary concern of this paper is to examine the formalisation choices under which m-self-reference is attainable. That is, we ask for which choices of (i)–(iii) we can find for each formula $A(x)$, with free variable x , a closed term t which denotes the Gödel code of $A(t)$.

Let \mathcal{L}^0 be the arithmetical language which has 0, S, + and \times as its non-logical vocabulary. Let ξ be a standard numbering of \mathcal{L}^0 and consider the canonical naming device which maps each number n to its standard numeral \underline{n} . Taken together, these canonical formalisation choices do not, generally, permit the construction of m-self-referential sentences.² There are two well-known routes to overcome this.

The first route consists in enriching the arithmetical language by function symbols for primitive recursive functions, such that the resulting language \mathcal{L}^+ contains a term $\delta(x)$ which represents a diagonal function D . Given a standard numbering ξ^+ of \mathcal{L}^+ , a diagonal function D , adequate for these choices, maps (the ξ^+ -code of) an \mathcal{L}^+ -formula $A(x)$ with x free to (the ξ^+ -code of) its diagonalisation $A(\delta(\xi^+(A)))$. This language extension can be achieved in different ways. Since D is primitive recursive, one may simply add a function symbol for D to \mathcal{L}^0 , or one can add function symbols for other p.r. functions such as the substitution function and the numeral function and then represent D as a complex term, etc. Let $R_{\mathcal{L}^+}$ be the result of adding all true equations of the form $t = \underline{n}$ to the Tarski–Mostowski–Robinson theory R, where t is a closed \mathcal{L}^+ -term. The Strong Diagonal Lemma then provides the existence of m-self-referential sentences:

LEMMA 1.1. (Strong Diagonal Lemma for \mathcal{L}^+). *For every \mathcal{L}^+ -formula $A(x)$ with x as a free variable, there exists a closed \mathcal{L}^+ -term t such that $R_{\mathcal{L}^+} \vdash t = \underline{\xi(A(t))}$.*

The second route is based on so-called *self-referential* Gödel numberings. A numbering ξ is called self-referential if, for any formula $A(x)$, we can find a number n such that $n = \xi(A(\underline{n}))$. Self-referential numberings thus immediately provide m-self-referential sentences, without the need of extending the language (and the theory). The idea of self-referential Gödel numberings can be traced back to [15, footnote 6] and [5, p. 80]. Constructions of self-referential numberings are given in Kripke's 1982 Princeton seminar (see footnote 1) and in his recent [16], in [32, 33], in [13] and in the Appendix of [11]. The last construction is for efficient numerals: see below.

The second route is typically taken to be contrived and unsatisfactory since the self-referential numberings existing in the literature to date do not satisfy certain desirable properties. One such property is *monotonicity*, which requires the Gödel number of an expression to be larger than the Gödel numbers of its sub-expressions. Halbach [9], for instance, does not consider non-monotonic Gödel numberings as

² See Lemma 8.3 for a generalisation of this fact.

adequate formalisation choices.³ Indeed, in the study of self-reference, monotonicity is prevalently required to hold for any *reasonable* or *adequate* numbering. It is, thus, widely believed that for adequate choices of Gödel numberings m-self-reference is not attainable in arithmetic formulated in \mathcal{L}^0 (see [13]). Accordingly, the natural setting to formalise self-referential reasoning is typically based on the first approach, by increasing the expressiveness of the language \mathcal{L}^0 .

The aim of the present paper is to further investigate this trade-off between the (term-)expressiveness of the language and the naturalness of the underlying coding. In particular, we will show that the received view regarding the unattainability of m-self-reference in \mathcal{L}^0 is committed to much stronger assumptions on the coding apparatus than usually assumed in the literature.

We will start by investigating which requirements on Gödel numberings are sufficient to rule out self-referential numberings as admissible formalisation choices. We will first observe that when employing standard numerals, self-referential Gödel numberings indeed cannot be monotonic (see Section 3.2). Thus, by Halbach's standards, self-referential Gödel numberings would be disqualified as adequate formalisation choices. They could be, at most, a technical tool providing an alternative proof of the Gödel Fixed Point Lemma and, thus, of the Incompleteness Theorems. (See for instance [6, lemma 4.12].)

This is not the end of the story, however. While the attainability of m-self-referential sentences is not affected by the underlying naming device (as will become clear from inspection of Definition 3.2), we will show that the (in-)compatibility of monotonicity and self-referentiality of numberings depends crucially on this aspect of formalisation. In fact, we will show in this paper that we can produce effective self-referential monotonic numberings by basing the definition of *self-referential* on efficient instead of standard numerals (see Section 4 and Appendix A). That is, for any formula $A(x)$, we can find a number n such that $n = \xi(A(\bar{n}))$, where \bar{n} denotes the efficient numeral of n (see Section 2.4). Setting $t := \bar{n}$ hence yields $R \vdash t = \xi(A(t))$. The adequacy constraints on numberings have thus to be strictly more restrictive than effectiveness and monotonicity in order to rule out self-referential numberings and, in particular, m-self-referentiality in \mathcal{L}^0 .

In Section 6, we introduce a strengthened notion of monotonicity put forward by Halbach [9], which captures the idea that (efficient) numerals are arithmetical proxies for quotations. A monotonic coding is called strongly monotonic if the code of the Gödel numeral of an expression is larger than the code of the expression itself. This constraint on numberings is sufficiently restrictive to exclude self-referential numberings for any numeral system. However, we will show that even strong monotonicity is not restrictive enough to exclude m-self-referentiality, i.e., strong diagonalisation. In fact, we will construct an effective and strongly monotonic numbering which gives rise to the Strong Diagonal Lemma for \mathcal{L}^0 , thus providing m-self-referential sentences formulated in \mathcal{L}^0 .

In Section 7, we introduce computational constraints which are more restrictive than effectiveness and which may serve as additional adequacy constraints for numberings. A Gödel numbering is called \mathfrak{E} I-adequate if it represents a large portion of syntactic

³ We do think more argument is needed to philosophically justify constraints on Gödel numberings. However, our present interest is not in that discussion, but in the further implications of such constraints.

relations and operations by *elementary* relations and operations on ω . We show that the numberings constructed in this paper are $\mathfrak{E}\mathfrak{I}$ -adequate in this sense. Hence, strong monotonicity and $\mathfrak{E}\mathfrak{I}$ -adequacy taken together, are once again not restrictive enough to exclude m-self-referentiality in \mathcal{L}^0 .

In Section 8, we briefly discuss the constraint of *regularity* on numberings for languages $\mathcal{L} \supseteq \mathcal{L}^0$ due to Heck [13]. Indeed, this constraint is sufficiently restrictive to rule out the existence of m-self-referential sentences formulated in \mathcal{L} . However, we construct a decent numbering which is *not* regular for \mathcal{L}^0 . Regularity thus hardly serves as a necessary constraint for admissible numberings and m-self-referentiality in \mathcal{L}^0 is in the clear.

The obtained results suggest the following disjunctive conclusion: when formalising m-self-reference in arithmetic, the adequacy constraints on reasonable numberings are more restrictive than widely assumed, or m-self-reference can be adequately formalised in a less expressive language than usually believed.⁴ More specifically, we conclude that either the constraints on reasonable numberings are more restrictive than $\mathfrak{E}\mathfrak{I}$ -adequacy and strong monotonicity taken together, or m-self-reference is already attainable in \mathcal{L}^0 .

We close, in Section 9, by showing how these results bear on the study of axiomatic truth theories. In particular, we show that the constraints of $\mathfrak{E}\mathfrak{I}$ -adequacy and (strong) monotonicity taken together are not sufficient to determine the consistency of certain type-free truth theories. Thus, the formalisation of certain informal principles of truth in an arithmetical setting is highly sensitive to the underlying formalisation choices. These results raise doubts as to what extent such axiomatic theories can be taken to faithfully reflect informal reasoning regarding the underlying principles of truth.

Finally, we hope to provide entertaining examples of Gödel numberings which bring to light some surprising subtleties regarding the interaction between self-reference and the employed formalisation devices.

§2. Technical Preliminaries. In this section, we introduce the necessary basic notions concerning syntax, theories and numberings.

2.1. Languages. Languages can be represented in many ways: as free algebras, as sets of strings, as labeled directed acyclic graphs, as ... These choices are both philosophically and technically important. In our context of the study of Gödel numberings, the choice of a format for the language will often suggest a particular choice of numerical representation. E.g., labeled directed acyclic graphs can be modeled in the hereditarily finite sets and we can map these sets into numbers using the Ackermann coding. It would be very natural to base a Gödel numbering on this idea. In this paper, we will mainly employ the algebraic perspective and on the string perspective. This does not reflect a philosophical standpoint, but is just a limitation dictated by length. In Section 8.2.2, we will briefly consider the idea of sharing material in syntax.

Let \mathcal{L}^0 be the language of first-order arithmetic, which contains $\perp, \top, =, \neg, \wedge, \vee, \rightarrow, \forall$ and \exists as logical constants, as well as the non-logical symbols $0, S, +$ and \times . The infix expressions of \mathcal{L}^0 are given as follows.

- $x ::= v \mid x'$

⁴ In this paper, we do not take a stand on which of the disjuncts obtains.

- $t ::= x \mid 0 \mid St \mid (t + t) \mid (t \times t)$
- $A ::= \perp \mid \top \mid t = t \mid \neg A \mid (A \wedge A) \mid (A \vee A) \mid (A \rightarrow A) \mid \forall x A \mid \exists x A$

Alternatively, we will sometimes consider \mathcal{L}^0 to be given in Polish notation:

- $x ::= v \mid x'$
- $t ::= x \mid 0 \mid St \mid Att \mid Mtt$
- $A ::= \perp \mid \top \mid =tt \mid \neg A \mid \wedge AA \mid \vee AA \mid \rightarrow AA \mid \forall x A \mid \exists x A$

Let \mathbb{N} be the standard interpretation of \mathcal{L}^0 , with ω as its domain. In this paper, we consider arbitrary languages $\mathcal{L} \supseteq \mathcal{L}^0$ with finite signature such that each constant symbol c and function symbol f of \mathcal{L} has an intended interpretation $c^{\mathbb{N}}$ and $f^{\mathbb{N}}$ in \mathbb{N} . Thus, in particular, the evaluation function ev is well-defined on closed terms of \mathcal{L} . For any such language \mathcal{L} , both definitions above extend in the obvious way. When we do not specify the employed notation system, we assume \mathcal{L} to be given in infix notation.

2.2. Theories. We will mainly consider theories in the language of arithmetic. Our basic theory is be the Tarski–Mostowski–Robinson theory R (see [30, p. 53]). We will extend R in a standard way to $R_{\mathcal{L}}$ by adding all true equations of the form $t = \underline{n}$ to R , where t is a closed \mathcal{L} -term. Of course, this requires the intended interpretations of the new function symbols in the background.

2.3. Gödel numberings. Let S be a domain which permits a robust notion of effectiveness for functions $\xi: S \rightarrow \omega$.⁵ We say that a function $\xi: S \rightarrow \omega$ is a *Gödel numbering* or *coding* of S , if ξ is injective and effective.⁶ We also call $\xi(A)$ the (ξ -)code of A (for terms and formulæ A). In this paper we consider numberings of languages \mathcal{L} given in infix or Polish notation. We note that choice of infix language versus Polish language is, in a sense, immaterial, since the two languages/representations of the language are connected by a standard bijection.

We occasionally also consider a language as embedded in a set of strings. Then, a numbering of the strings over the given finite alphabet will induce a numbering of the language. For instance, the infix expressions of \mathcal{L}^0 can be conceived of as strings over an alphabet containing 17 symbols, while its Polish notations can be formulated as strings over an alphabet with 15 symbols. (So, when we consider the language as embedded in strings the difference between infix and Polish suddenly does have some role.)

We note that, e.g., if we would think of formulæ as labeled directed acyclic graphs (dags), the numbering’s domain S could be the totality of all finite labeled dags with

⁵ Traditionally, the notion of effectiveness for functions with domain S is reduced to Turing computability or recursiveness by coding the elements of S as strings or numbers respectively. However, different coding devices in general yield different extensions of effective functions. Here, we require that S is a domain for which the notion of effectiveness does not depend on such choices. For instance, S may be taken to be a space of finite objects, in the sense of Shoenfield [28]. In this paper, we will be solely concerned with domains of strings and other syntactic expressions which permit robust notions of effectiveness. Due to limitations of space we leave a discussion of this important matter for another occasion.

⁶ There are metamathematical contexts in which Gödel numberings are not required to be effective. See, for example, Theorem 1 in [30, p. 46]. Effectiveness is a minimal constraint on numberings to ensure *weak* diagonalisation [6, sec. 2.3 and lemma 4.12). Since here we are concerned with the attainability of *strong* diagonalisation, we simply assume the effectiveness of Gödel numberings at the outset.

labels in a fixed alphabet. Such choices often reflect a syntax theory. If we view syntax as *sui generis*, S will usually be the set of expressions itself. If we view the syntactic objects as specima of a wider variety X , we will usually take $S := X$.

REMARK 2.1. *We could allow one extra degree of freedom for Gödel numberings: we could drop functionality. For example, suppose we want to an existing coding of syntax in the finite sets to define our Gödel numbering via coding the finite sets as numbers. There are various ways to code the finite sets. Suppose we do it by interpreting sequences first and then ignoring the order of the components. This gives us a non-functional Gödel numbering. Note that it would still be injective. For the purposes of the present paper we will not need the extra flexibility of non-functionality.*

REMARK 2.2. *The Gödel numbering employed in [4] is a nice example of a Gödel numbering that looks directly at the language-as-algebra without considering it as embedded in strings. For Feferman's coding, the choice between infix and Polish is irrelevant.*

An important example of a numbering of strings is the *length-first ordering*. Let \mathcal{A}^* be a finite alphabet and suppose some ordering of \mathcal{A} is given. We order the strings of \mathcal{A} using the *length-first ordering* $(\alpha_n)_{n \in \omega}$ in which we enumerate the strings according to increasing length, where the strings of same length are ordered alphabetically. We set $g(\alpha) = n$ if $\alpha = \alpha_n$.⁷ We write $|\alpha|$ for the length of α .

Throughout this paper, we will use the following basic fact about the length-first ordering.

LEMMA 2.3. *Suppose the alphabet \mathcal{A} has $N \geq 2$ letters. We then have:*

$$\frac{N^{|\alpha|} - 1}{N - 1} \leq g(\alpha) < \frac{N^{|\alpha|+1} - 1}{N - 1}.$$

It follows that $|\alpha| \leq g(\alpha)$. Moreover, whenever $|\alpha| < |\beta|$, then $g(\alpha) < g(\beta)$.

Proof. Clearly, for any string α of \mathcal{A} we have

$$\underbrace{1 \cdots 1}_{|\alpha| \times} \leq g(\alpha) \leq \underbrace{1 \cdots 1}_{|\alpha|+1 \times},$$

where, for any m , $\overbrace{1 \cdots 1}^{m \times}$ is considered as an N -adic notation. The claim follows immediately from a well-known property of geometric series:

$$\overbrace{1 \cdots 1}^{m \times} = \sum_{i=0}^{m-1} N^i = \frac{N^m - 1}{N - 1}. \quad \square$$

2.4. Naming devices. Let $\text{CTerm}_{\mathcal{L}}$ denote the set of closed \mathcal{L} -terms. We call a function $v: \omega \rightarrow \text{CTerm}_{\mathcal{L}}$ a *numeral function for \mathcal{L}* , if v is injective and effective and the closed term $v(n)$ has value n , for each $n \in \omega$. We call $v(n)$ the *v -numeral of n* . Standard numerals form a canonical choice of a numeral function and are defined as follows.

- $\alpha ::= 0 \mid S\alpha$.

⁷ We note that g has the alphabet and the ordering on the alphabet as hidden parameters. To make our notation not too heavy we will always suppress these data.

We write \underline{n} for the ordinary, or standard, numeral with value n , and also write \cdot for the standard numeral function. These numerals provide unique normal forms for the natural numbers. These normal forms reflect the fact that we naturally read the axioms for addition and multiplication as directed. We have $(x + 0) \rightsquigarrow x$, $(x + Sy) \rightsquigarrow S(x + y)$, $(x \times 0) \rightsquigarrow 0$, $(x \times Sy) \rightsquigarrow ((x \times y) + x)$. The numerals are the normal forms for the closed terms in this reduction system.

In the literature on weak systems, however, another kind of numeral is used that corresponds to binary or dyadic notations. The main reason for this is simply that the arithmetisation of the numeral function for standard numerals and the usual Gödel numberings exhibits exponential growth. This holds for example for the Gödel numbering \mathfrak{g} based on the length-first ordering. The numeral function for the numerals in the other style is more efficient. Thus, these alternative numerals are also called *efficient numerals*. We define efficient numerals for dyadic representations as follows.

- $\alpha ::= 0 \mid S(SS0 \times \alpha) \mid SS(SS0 \times \alpha)$.

We write \bar{n} for the efficient numeral with value n and write $\bar{\cdot}$ for the efficient numeral function. We note that efficient numerals also correspond more naturally to reduction systems in a different signature, e.g., a theory of strings formulated with two successors.

REMARK 2.4. Consider the Gödel numbering \mathfrak{g} for the language \mathcal{L}^0 based on the length-first ordering, where we assume we used infix notations. We note that the cardinality of the alphabet is 17. Consider any number n .

Clearly, we have $|\underline{n}| = n + 1$. So, $\frac{17^{n+1}-1}{16} \leq \mathfrak{g}(\underline{n})$. So $\mathfrak{g}(\underline{n})$ has an exponential lower bound in n .

We give an upper bound for $\mathfrak{g}(\bar{n})$. Let the length of the dyadic notation of n be k . We have: $2^k - 1 \leq n < 2^{k+1} - 1$. We can estimate the length of the dyadic numeral \bar{n} by considering the worst case where the dyadic notation of n is a string of 2's. We find that $|\bar{n}| \leq 8k + 1$. Thus, using that $17 < 2^5$, we see that:

$$\mathfrak{g}(\bar{n}) < \frac{17^{8k+2} - 1}{16} \leq 64 \cdot 2^{40k} \leq 64 \cdot (n + 1)^{40}.$$

So, we see that $\mathfrak{g}(\bar{n})$ has a polynomial upper bound in n .

§3. Formalisations of m-Self-Reference. The notion of self-reference considered here is defined by means of a reference relation on sentences. According to Leitgeb [17], “a singular sentence might [...] be defined to refer to all the referents of all of its singular terms, and only to them” (p. 4). A sentence is then said to be self-referential if and only if it refers to itself. By employing a Gödel numbering, this notion of self-reference can be formalised in an arithmetical framework according to the following definition.⁸

⁸ Picollo distinguishes the formal notions of m-reference *simpliciter* [21] and alethic m-reference [22]. While the former is intended to capture a pre-theoretical notion of reference *simpliciter* as indicated by the above quotation by Leitgeb, the latter is more restrictive and specifically tailored to studying the reference patterns which underlie paradoxical expressions in the context of truth. Even though Definitions 3.1 and 3.2 capture m-self-reference *simpliciter*, the results of this paper also apply to alethic m-self-reference.

DEFINITION 3.1. *A sentence of the form $A(t)$ is called m-self-referential with respect to the formula $A(x)$, if the closed term t denotes (the code of) $A(t)$.*

This notion of self-reference is for example captured by the Kreisel–Henkin Criterion for self-reference (see [10, p. 684]). According to this criterion, given a formula $A(x)$ expressing a property P , a sentence $A(t)$ says of itself that it has property P iff $A(t)$ is m-self-referential with respect to $A(x)$ (in the sense of the above definition).

In order to make this notion of m-self-reference mathematically precise, Definition 3.1 has to be specified with respect to the formalisation choices (i)–(iii). Accordingly, we call a triple $\langle \mathcal{L}, \xi, \nu \rangle$ a *formalisation choice*, if

- \mathcal{L} is a language as specified in Section 2.1;
- ξ is a Gödel numbering of \mathcal{L} ; and
- ν is a numeral function for \mathcal{L} .

DEFINITION 3.2. *We say that m-self-reference is attainable for a formalisation choice $\langle \mathcal{L}, \xi, \nu \rangle$, if for each \mathcal{L} -formula $A(x)$ there exists a closed \mathcal{L} -term t such that $R_{\mathcal{L}} \vdash t = \nu(\xi(A(t)))$.⁹*

It is easy to see that the attainability of m-self-reference is invariant regarding the choice of the numeral function. That is, m-self-reference is attainable for $\langle \mathcal{L}, \xi, \nu_1 \rangle$ iff m-self-reference is attainable for $\langle \mathcal{L}, \xi, \nu_2 \rangle$, for any language \mathcal{L} , numbering ξ and numeral functions ν_1 and ν_2 for \mathcal{L} . Hence, instead of requiring the provability of $t = \nu(\xi(A(t)))$ in Definition 3.2, one may equivalently require that $R_{\mathcal{L}} \vdash t = \underline{\xi(A(t))}$. The reason for considering arbitrary numeral functions as parameters in formalisation choices is that certain adequacy conditions on numberings will be shown to be sensitive to this aspect of formalisation.

Moreover, instead of using provability in $R_{\mathcal{L}}$, one can equivalently require in Definition 3.2 that $\mathbb{N} \models t = \underline{\xi(A(t))}$, since $R_{\mathcal{L}}$ proves every true equation of closed \mathcal{L} -terms.

3.1. Adequacy of formalisation. In formal studies of truth, self-reference and the semantic paradoxes, it is common practice to employ natural numbers as theoretical proxies for syntactic expressions. Accordingly, certain designated arithmetical domains or theories serve as domains of expressions or syntax theories respectively. This practice of *arithmetisation* rests on Gödel numberings, which can be seen as translation devices between expressions and numbers. However, as a consequence of intensionality phenomena in metamathematics, not every numbering, i.e., injective and effective function, can be considered an equally adequate candidate.

When formalising notions such as truth or (self-)reference over an arithmetical domain or theory, it is therefore essential that the underlying Gödel numbering constitutes an *adequate* translation device between expressions and numbers. Only then are we justified in considering the given arithmetical framework a domain or theory of expressions and in taking the formalised notions to apply to linguistic expressions as intended. In order to guarantee a faithful formalisation of the given philosophical notions, *adequate choices of formalisation* are thus required.

It is notoriously difficult to precisely characterise the notion of adequacy for numberings for a given metamathematical or philosophical purpose. We will therefore

⁹ We intend, of course, the call-by-value reading here: $\nu(\xi(A(t)))$ yields some term u , such that $R_{\mathcal{L}} \vdash t = u$.

adopt a more modest approach in this paper by discussing certain *necessary* conditions for adequate numberings which can be found in the literature on self-reference and axiomatic truth theories.

3.1.1. Monotonicity. Perhaps, the most prominent adequacy constraint is *monotonicity*. Let S be a set and let \triangleleft be a strict partial order relation on S . A Gödel numbering ξ of S is called *monotonic with respect to* \triangleleft , if $\alpha \triangleleft \beta$ implies $\xi(\alpha) < \xi(\beta)$, for all $\alpha, \beta \in S$. Let \mathcal{L} be a language given in infix or Polish notation (see Section 2.1). We call a numbering ξ of \mathcal{L} *monotonic*, if it is monotonic with respect to the strict sub-expression relation \prec given on infix or Polish notations respectively. Monotonicity in this context can therefore be equivalently characterised by requiring a numbering ξ of \mathcal{L} to satisfy the following three conditions:

- M1. for all \mathcal{L} -terms s, t , if $s \prec t$ then $\xi(s) < \xi(t)$;
- M2. for all \mathcal{L} -formulae A, B , if $A \prec B$ then $\xi(A) < \xi(B)$; and
- M3. for all \mathcal{L} -terms s and formulae A , if $s \prec A$ then $\xi(s) < \xi(A)$.

Thus, when we consider numberings of well-formed expressions, we do not demand monotonicity with respect to sub-strings. E.g., in case $\xi((A \wedge B))$ would be $\langle 7, \langle \xi(A), \xi(B) \rangle \rangle$, where $\langle \cdot, \cdot \rangle$ is the Cantor Pairing, the brackets do not appear in the code, so we do not need a Gödel number for a bracket. However, when considering numberings of strings, monotonicity may be also required with respect to the (strict) sub-string relation. Clearly, any numbering which is monotonic in this sense also satisfies the conditions M1–M3.

Halbach [9] for instance does not consider non-monotonic Gödel numberings as adequate formalisation choices. Indeed, the study of self-reference is typically based on numberings which are required to be monotonic. This requirement is explicit for instance in [19, footnote 4], [8, p. 33], [21, pp. 573f.], [22, footnote 4] and [23, footnote 3] (where Milne requires monotonicity with respect to the sub-string relation). A good example of an application of (a strengthened version of) monotonicity is the proof of the falsity of the Σ_1^0 -truth teller for fixed-point operators with certain good properties. See [11, theorem 7.7].

One way to motivate this constraint as a necessary condition of adequacy for numberings proceeds as follows. Let \leq be the usual less-than (or equal) relation on numbers and let \preceq denote the sub-expression relation on the given domain \mathcal{L} of expressions. One may conceive of \preceq as a parthood relation on \mathcal{L} . One may endow \leq with such a mereological interpretation by representing numbers as strings of strokes. Another approach is to represent numbers as Von Neumann ordinals, and to understand the parthood relation on sets as the containment relation \subseteq (see e.g., [18]). Then \leq is indeed a parthood relation on numbers.¹⁰ Taking parthood as an important structural feature of the domain of expressions, adequate numberings may then be required to preserve structure in virtue of representing the parthood relation \preceq on \mathcal{L} by a relation on ω contained in the parthood relation \leq on numbers, which is

¹⁰ Note that this interpretation breaks down for other reasonable representations, such as Zermelo ordinals or equivalence classes of finite sets under the equivalence relation of equinumerosity. For a critical discussion of monotonicity, see [6, sec. 3.4].

tantamount to requiring monotonicity. We will introduce further adequacy conditions for numberings in later sections of this paper.

3.1.2. Numerals. Thus far, we have only discussed adequacy with respect to Gödel numberings. It remains to be clarified which numeral functions constitute adequate formalisation choices. One approach is to characterise a numeral function as adequate iff it corresponds to an *acceptable* or *canonical* naming system for the natural numbers (see [14, 27]). Note that on this approach, standard numerals as well as efficient numerals qualify as adequate numeral systems.

One way to view the natural numbers is as the free algebra for a unary function and one generator. Addition and multiplication are, then, defined by recursion, where the possibility of such recursions is guaranteed by the fact that we have a free algebra. From this point of view, the standard or tally numbers indeed seem to have a preferred status: they are the standard syntactic representations of the elements of the free algebra. However, we would like to point out that the point of view of this specific free algebra is not the only way to think of the natural numbers. In a sense, viewing the natural numbers as this free algebra *is* taking the tally numerals to be the basic numerals. Alternatively, we could take the dyadic representation to be primary and work with a theory with two successors. In such a context, the dyadic numerals would be the usual syntactic representations. Similarly, the notion of natural number as finite cardinal does not automatically dictate zero and successor to be the preferred signature.

We will not deal with the question of the status of numerals in this paper and simply assume that both standard numerals as well as efficient numerals are adequate choices of a naming device. For a further discussion the reader is referred to [1].

3.2. Monotonicity and self-referentiality. The notion of self-referential numberings can be generalised to arbitrary numeral functions.

DEFINITION 3.3. *A numbering ξ is called self-referential for a numeral function v if, for any formula $A(x)$, there exists $n \in \omega$ such that $n = \xi(A(v(n)))$.*

As opposed to the notion of m-self-reference for sentences (see Definition 3.1), self-reference for numberings is a mere technical concept and not intended to capture a pre-theoretical or philosophical notion. However, as we have seen, self-referential numberings immediately yield m-self-referential sentences.

In the following sections we will investigate whether adequate numberings can be self-referential, i.e., whether self-referential numberings are indeed as contrived as typically believed. We will first examine whether monotonicity serves as an adequacy condition on numberings which is sufficiently restrictive to exclude self-referential numberings and will then turn to additional more restrictive adequacy constraints for numberings.

We start by showing that a monotonic numbering cannot be self-referential for a numeral function v , if the number of subterms of $v(n)$ grows sufficiently fast. In order to make this precise, let $\text{st}: \text{ClTerm}_{\mathcal{L}} \rightarrow \omega$ be the function assigning to each closed \mathcal{L} -term the number of its proper subterms-qua-type. We then get the following incompatibility result.

LEMMA 3.4. *Let v be a numeral function and suppose that there is a constant $c \in \omega$ such that $\text{st}(v(n)) < c$, for all $n \in \omega$. Then, there is no monotonic numbering which is self-referential for v .*

Proof. Assume that ξ is monotonic and self-referential for v . Let $B(x)$ be an \mathcal{L} -formula in which x does occur freely at least once. Set

$$A(x) := \underbrace{B(x) \vee \dots \vee B(x)}_{(c+1)\times}.$$

Then there exists $n \in \omega$ such that $\xi(A(v(n))) = n$. By monotonicity we then get

$$\begin{aligned} n &= \xi(A(v(n))) \\ &= \xi(\underbrace{B(v(n)) \vee \dots \vee B(v(n))}_{(c+1)\times}) \\ &\geq \xi(\underbrace{B(v(n)) \vee \dots \vee B(v(n))}_{c \times}) + 1 \geq \dots \\ &\geq \xi(B(v(n))) + c \\ &\geq \xi(v(n)) + c \\ &\geq \text{st}(v(n)) + c, \end{aligned}$$

in contradiction to the assumption that $n - \text{st}(v(n)) < c$. □

Since $n - \text{st}(n) < 1$ for all $n \in \omega$, we conclude from the above lemma:

COROLLARY 3.5. *There is no self-referential numbering for standard numerals which is monotonic.*

After the second author’s lecture on *Cogwheels of Self-reference* at the workshop *Ouroboros, Formal Criteria of Self-Reference in Mathematics and Philosophy* on February 17, 2018, Joel Hamkins asked whether the result on the non-monotonicity of self-referential Gödel numberings also holds when we consider efficient numerals.

We first note that Lemma 3.4 does not apply to efficient numerals. We define:

- $\tilde{0} := 0$ and
- $n + 1 := S(SS0 \times \tilde{n})$.

Let ev be the evaluation function for arithmetical closed terms and let $|\alpha|$ denote the length of the string α . We find:

LEMMA 3.6. *For every $n \in \omega$*

- $\text{ev}(\tilde{n}) = 2^n - 1$;
- $|\tilde{n}| = 7n + 1$; and
- *the number of subterms-qua-type of \tilde{n} is $\leq 2n + 3$.*

From this lemma we conclude that the function $\lambda n.2^n - 1$ grows exponentially while the number of subterms of $\tilde{2^n - 1}$ is not larger than $2n + 3$ and thus only grows linearly. Hence, the assumption of Lemma 3.4 cannot be satisfied for efficient numerals, since there is no constant c such that $n - \text{st}(\tilde{n}) < c$ for all $n \in \omega$.

Indeed, in what follows we show that the answer to Hamkins’ question is *no*, by constructing a monotonic numbering which is self-referential for efficient numerals.

§4. A Monotonic Self-Referential Numbering. Let \mathcal{L} be an arithmetical language, as introduced in Section 2.1, and let $\mathcal{L}(c)$ be \mathcal{L} extended with a fresh constant c . Let

A_0, A_1, \dots be an effective enumeration of all expressions of $\mathcal{L}(c)$. We assume that if $C \preceq A_n$, then, for some $k \leq n$, we have $C = A_k$.

We define $\lceil A \rceil$ as the number of sub-expressions-qua-type of A , in other words, $\lceil A \rceil$ is the cardinality of $\{B \mid B \preceq A\}$. The following trivial observation is very useful:

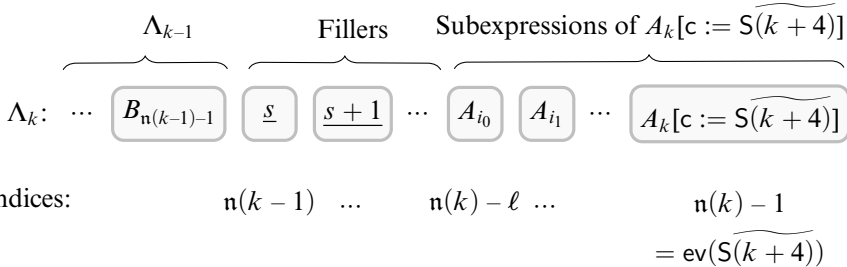
LEMMA 4.1. $\lceil A_k \rceil \leq k + 1$.

Let $A[c := t]$ denote the result of substituting t for all occurrences of c in A . Here is our construction.

We construct a list $\Lambda := (B_n)_{n \in \omega}$ in stages k . Let $n(k) := 2^{k+4} + 1$. Each stage k will result in a list $\Lambda_k = B_0, \dots, B_{n(k)-1}$. To simplify the presentation, we make Λ_{-1} the empty list and $n(-1) := 0$.

In stage k , we act as follows. Let $A_k^* := A_k[c := \widetilde{S(k+4)}]$. Let the sub-expressions of A_k^* that do not occur in Λ_{k-1} and are not of the form m be $A_{i_0}, \dots, A_{i_{\ell-1}}$, where the sequence i_j is strictly increasing. We note that ℓ could be 0. We define $B_{n(k)-\ell+j} := A_{i_j}$, for $j < \ell$. Let s be the smallest number such that \underline{s} is not in Λ_{k-1} . We set $B_{n(k-1)+p} := \underline{s+p}$, for $p < n(k) - n(k-1) - \ell$.

The construction in stage k of the list Λ_k can be illustrated as follows:



It follows from the construction that if A_k^* is not in Λ_{k-1} , then

$$B_{n(k)-1} = A_k^* = A_k[c := \widetilde{S(k+4)}]. \tag{1}$$

We will show in Lemma 4.6 that (1) holds for each A_k which contains c .

To see that our construction is well-defined it is sufficient that $n(k) - \ell \geq n(k-1)$ or, equivalently, $\ell \leq n(k) - n(k-1)$. We note that:

$$\ell \leq \lceil A_k^* \rceil \leq \lceil A_k \rceil + \lceil \widetilde{S(k+4)} \rceil - 1 \leq k + 1 + 2(k+4) + 3 = 3k + 12.$$

(The -1 in the right-hand-side of the second inequality can be seen as follows. In case c occurs in A_k it is subtracted in the substitution. If c does not occur, we have $A_k = A_k^*$ and, from this, the inequality follows immediately.)

If $k = 0$, we have $3 \cdot 0 + 12 = 12 < 17 = 2^{0+4} + 1 - 0 = n(0) - n(-1)$. Let $k > 0$. We find:

$$3k + 12 < 2^{k+3} = 2^{k+4} + 1 - 2^{k+3} - 1 = n(k) - n(k-1).$$

LEMMA 4.2. *Suppose $A_k \in \mathcal{L}$. Then, A_k is in Λ_k and, hence, in Λ .*

Proof. Consider stage k . We note that $A_k^* = A_k$. In case A_k occurs in Λ_{k-1} , we are done. In case A_k is not in Λ_{k-1} and not of the form \underline{m} , clearly, A_k will be added, so we are again done. Suppose $A_k = \underline{m}$, for some m . It follows that $m + 1 = \lceil \underline{m} \rceil \leq k + 1$. We note that all sub-expressions of A_k are all of the form \underline{m}' , for $m' \leq m$. So $\ell = 0$. Clearly, $m + 1 \leq k + 1 < n(k) - n(k - 1)$. So, all sub-expressions of A_k are either in Λ_{k-1} or will be added. \square

LEMMA 4.3. *Suppose $C \preceq B_n$. Then, for some $j \leq n$, we have $C = B_j$.*

Proof. Let $C \preceq B_n$. Suppose B_n is added in stage k . Then, B_n is either of the form \underline{m} or a sub-expression of A_k^* not of the form \underline{m} .

Suppose $B_n = \underline{m}$. Let s be the smallest number such that \underline{s} is not in Λ_{k-1} . For all $s' < s$, we have $\underline{s'} \in \Lambda_{k-1}$. By our construction, all $\underline{s''}$, for $s \leq s'' \leq m$ are added at stage k in ascending order. So, for all $s^\circ \leq m$, we find that $\underline{s^\circ}$ precedes \underline{m} in Λ_k . Finally, each sub-expression C of \underline{m} is of the form $\underline{s^\circ}$, for some $s^\circ \leq m$.

Suppose B_n is added as a sub-expression of A_k^* not of the form \underline{m} . First suppose $C = \underline{p}$. We note that $p + 1 = \lceil \underline{p} \rceil < \lceil A_k \rceil \leq k + 1$. This tells us, by Lemma 4.2, that \underline{p} is in Λ_{k-1} . Hence, it is added to Λ before B_n . Now suppose that C is not of the form \underline{p} . Then either C is in Λ_{k-1} or added to Λ_k before B_n . \square

LEMMA 4.4. *The enumeration Λ is without repetitions.*

Proof. Consider any $C \in \mathcal{L}$. If C is not of the form \underline{m} by our construction it will be only added once.

Suppose $C = \underline{m}$. We note that, by Lemma 4.3, the n such that $\underline{n} \in \Lambda_{k-1}$ are downwards closed. This means that, in our construction, there is no $m \geq s$, such that $\underline{m} \in \Lambda_{k-1}$. So all the \underline{m} that are added in any stage k are new. \square

LEMMA 4.5. *Suppose $S\tilde{n}$ occurs in Λ_k , then $n \leq k + 4$.*

Proof. We prove this by induction on k . In stage 0, we easily verify that the largest term of the form $S\tilde{n}$ can be at most $S\tilde{4}$.

Suppose, we have our desired estimate for $k - 1$, where $k > 0$. We prove our estimate for k . Clearly, the $S\tilde{n}$ occurring in Λ_{k-1} do satisfy our estimate. Moreover, all the \underline{p} added in stage k do not provide new elements of the form $S\tilde{n}$. So, the only interesting case is the case where A_k is not a standard numeral. Suppose it is not. The $S\tilde{n}$ that are added are all sub-expressions of $A_k^* = A_k[c := S(k + 4)]$. So, we should focus on the largest such subterm. Suppose $S\tilde{n} \preceq A_k^*$. There are two possibilities:

- i. $S\tilde{n} \preceq A_k$. In this case, we have $n < \lceil S\tilde{n} \rceil \leq \lceil A_k \rceil \leq k + 1 \leq k + 4$.
- ii. $n = k + 4$, since the subterm results from the substitution for c .

In each of the possible cases, we are done. \square

LEMMA 4.6. *Consider any A such that c occurs in A . Then, for some n , $B_n = A[c := \bar{n}]$.*

Proof. Suppose c occurs in A . Let $A := A_k$. The only thing we have to show is that A_k^* is not in Λ_{k-1} . This is certainly true for $k = 0$. Suppose $k > 0$. We note that $S(k + 4)$ occurs in A_k^* . So, it is sufficient to show that $S(k + 4)$ does not occur in Λ_{k-1} . By Lemma 4.5, whenever $S\tilde{n}$ occurs in Λ_{k-1} , we have $n \leq k + 3$. We may conclude that $B_{n(k)-1} = A_k[c := S(k + 4)]$. Let $n := n(k) - 1$. Since $S(k + 4) = \overline{n(k) - 1}$, we are done. \square

Setting $gn_0(A)$ to be the unique n such that $A = B_n$, we have found a monotonic self-referential Gödel numbering gn_0 , where monotonicity is defined with respect to the sub-expression relation.

§5. Strings. Let \mathcal{L} be an arithmetical language, as introduced in Section 2.1, and let \mathcal{A} be the alphabet of \mathcal{L} . Let c be a fresh letter. We define:

- \mathcal{A}^* is the set of all strings of \mathcal{A} , including the empty string ε .
- \mathcal{A}_c is \mathcal{A} extended with c and \mathcal{A}_c^* is the set of all strings of \mathcal{A}_c . (In $\mathcal{L}(c)$ viewed as a subset of \mathcal{A}_c^* , we treat c as a constant-symbol.)
- $\alpha \sqsubseteq \beta$ iff α is a sub-string of β .
- $|\alpha|$ is the length of α .
- $[\alpha]$ is the number of sub-strings-qua-type of α , in other words, $[\alpha]$ is the cardinality of $\{\beta \mid \beta \sqsubseteq \alpha\}$.

We have the following well-known fact.

LEMMA 5.1. *Suppose $|\alpha| = n$. Then, the number of non-empty-sub-string-occurrences in α is $\frac{n(n+1)}{2}$. As a consequence, $[\alpha] \leq \frac{n^2+n+2}{2}$.*

Proof. Suppose $|\alpha| = n$. Let us number the spaces before and after the letter-occurrences in α : $1, \dots, n + 1$. Each non-empty-sub-string-occurrence corresponds uniquely to the pair of the space before and the space after the occurrence. So, the number of such occurrences is $\binom{n+1}{2} = \frac{n(n+1)}{2}$. □

Since we are not striving for maximal efficiency, we will estimate $[\alpha]$ by $|\alpha|^2 + 1$. Let $\mathfrak{a} := S(SS0 \times$. Then, $S\tilde{n} = Sa^n0^n$. So, $|S\tilde{n}| = 1 + 6n + 1 + n = 7n + 2$ and, thus, $[S\tilde{n}] \leq (7n + 2)^2 + 1 = 49n^2 + 28n + 5$.

REMARK 5.2. *We note that a much better estimate of $[S\tilde{n}]$ is possible but that its growth rate will still be quadratic. In case we switch to Polish notation, this becomes linear. Such notational choices like infix versus Polish are irrelevant in our construction and its verification when we are considering sub-expressions. In the sub-string format, they become active.*

We provide the estimate in the Polish case. For the rest of this remark we work with Polish notation. As in Section 2.1, we write M for multiplication.

Let $\mathfrak{b} := SMSS0$. We have $S\tilde{n} = S\mathfrak{b}^n0$. We note that, unlike \mathfrak{a} , the string \mathfrak{b} stands for a meaningful entity, to wit $\lambda k \in \omega.2k + 1$. See [34]. We count $[S\tilde{n}]$ as follows. Clearly $[S\tilde{0}] = [S0] = 4$. Suppose $n > 0$. We have:

- *The empty string: 1.*
- *The number of non-empty sub-strings of \mathfrak{b} : $\frac{5 \cdot 6}{2} - 2 = 13$. (We subtract 2, since S occurs three times.)*
- *All strings of the form $\alpha \mathfrak{b}^i \beta$, where $i < n - 1$, α is a non-empty final string of \mathfrak{b} and β is a non-empty initial string of \mathfrak{b} : $25(n - 1)$.*
- *All sub-strings that contain the the final occurrence of 0 , except 0 itself: $5n + 1$. (We note that all these strings do not occur elsewhere, since they end with 00 .)*
- *All sub-strings that contain the initial occurrence of S , except $S\tilde{n}$ itself, S and SS : $5n - 1$. (We note that the strings considered are unique since they start with SSM . Only $S\tilde{n}$ is an exception, since we already counted it in the previous case.)*

So, in toto, we have: $[S\tilde{n}] = 1 + 13 + 25(n - 1) + 5n + 1 + 5n - 1 = 35n - 11$. A similar estimate is also possible in the infix case, of course. However, the closing brackets will force the number to be quadratic in n .

We will use the following insights.

LEMMA 5.3. Suppose $m \neq 0$ and $n \neq 0$. Suppose a string $\alpha \in \mathcal{A}_c^*$ contains overlapping occurrences o of $S\tilde{n}$ and p of $S\tilde{m}$. Then, $o = p$.

Proof. Without loss of generality we may assume that p starts at the same place as or at a later place than o . Since o does not end with either S , or SS , the initial string $SS($ of p must lie in o . But the only sub-string of o of the form $SS($ is the initial one. Hence o and p start at the same place. By balance considerations, the matching closing brackets of the initial $SS($ in o and p should coincide. So, $o = p$ \square

LEMMA 5.4. Suppose $\alpha \in \mathcal{A}_c^*$ and $\beta \in \mathcal{A}^*$, where c occurs in α and β is non-empty. Let $[\alpha] = p$, $[\beta] = q$ and $|\beta| = r$. Then, $[\alpha[c := \beta]] \leq (p - 1)r^2 + q$. Since $q \leq r^2 + 1$, it follows that $[\alpha[c := \beta]] \leq pr^2 + 1$.

Proof. Let α^* be $\alpha[c := \beta]$ where the original string α is supposed to be black and where we colored each sub-string occurrence of β that resulted from the substitution with a unique color. We will consider black not to be a color here. Consider any non-empty sub-string γ^* of α^* . We replace all maximal sub-string occurrences of the same color by a single occurrence of c . This corresponds to a unique sub-string occurrence of γ in α . If γ is c , it has $q - 1$ originals-qua-type. If $\gamma \neq c$, the largest number of originals-qua-type obtains when γ both begins and ends with a c . This number of originals is $\leq r^2$. So, $[\alpha[c := \beta]] \leq (p - 1)r^2 + q - 1 + 1 = (p - 1)r^2 + q$. \square

Let $\alpha_0, \alpha_1, \dots$ be an effective enumeration of the strings in \mathcal{A}_c^* . We assume that if $\gamma \sqsubseteq \alpha_n$, then, for some $k \leq n$, we have $\gamma = \alpha_k$. The following trivial observation is very useful:

LEMMA 5.5. $[\alpha_k] \leq k + 1$.

Here is our construction.

We construct a list $\Lambda := (\beta_n)_{n \in \omega}$ in stages k . Let $n(k) := 2^{k+15} + 1$. Each stage k will result in a list $\Lambda_k = \beta_0, \dots, \beta_{n(k)-1}$. To simplify the presentation, we make Λ_{-1} the empty list and $n(-1) := 0$.

In stage k , we act as follows. Let $\alpha_k^* := \alpha_k[c := \widetilde{S(k + 15)}]$. Let the sub-strings of α_k^* that do not occur in Λ_{k-1} and are not of the form \perp^m be $\alpha_{i_0}, \dots, \alpha_{i_{\ell-1}}$, where the sequence i_j is strictly increasing. We note that ℓ could be 0. We define $\beta_{n(k)-\ell+j} := \alpha_{i_j}$, for $j < \ell$. Let s be the smallest number such that \perp^s is not in Λ_{k-1} . We set $\beta_{n(k-1)+p} := \perp^{s+p}$, for $p < n(k) - n(k - 1) - \ell$.

To see that our construction is well-defined it is sufficient that $n(k) - \ell \geq n(k - 1)$ or, equivalently, $\ell \leq n(k) - n(k - 1)$. We note that, if $k = 0$, we have $\alpha_0 = \varepsilon$, and hence $\ell = 1$. Moreover, $n(0) - n(0 - 1) = 2^{15} + 1$. So, we are done. Suppose $k > 0$. Clearly, $n(k) - n(k - 1) = 2^{k+14}$. In case c does not occur in α_k , we have $\ell \leq [\alpha_k^*] = [\alpha_k] \leq k + 1 \leq 2^{k+14}$. In case c does occur α_k , we have:

$$\ell \leq [\alpha_k^*] \leq [\alpha_k][\widetilde{S(k + 15)}]^2 + 1 \leq (k + 1)(7(k + 15) + 2)^2 + 1 \leq 2^{k+14}$$

We leave the verification of the last inequality to the diligent reader.

LEMMA 5.6. *Suppose $\alpha_k \in \mathcal{A}^*$. Then, α_k is in Λ_k and, hence, in Λ .*

Proof. Consider stage k . We note that $\alpha_k^* = \alpha_k$. In case α_k occurs in Λ_{k-1} , we are done. In case α_k is not in Λ_{k-1} and not of the form \perp^m , clearly, α_k will be added, so we are again done. Suppose $\alpha_k = \perp^m$, for some m . It follows that $m + 1 = \lfloor \perp^m \rfloor \leq k + 1$. We note that all sub-strings of α_k are all of the form $\perp^{m'}$, for $m' \leq m$. So $\ell = 0$. Clearly, $m + 1 \leq k + 1 < n(k) - n(k - 1)$. So, all sub-expressions of α_k are either in Λ_{k-1} or will be added. □

LEMMA 5.7. *Suppose $\gamma \sqsubseteq \beta_n$. Then, for some $j \leq n$, we have $\gamma = \beta_j$.*

Proof. Let $\gamma \sqsubseteq \beta_n$. Suppose β_n is added in stage k . Then, β_n is either of the form \perp^m or a sub-string of α_k^* not of the form \perp^m .

Suppose $\beta_n = \perp^m$. Let s be the smallest number such that \perp^s is not in Λ_{k-1} . For all $s' < s$, we have $\perp^{s'} \in \Lambda_{k-1}$. By our construction, all $\perp^{s''}$, for $s \leq s'' \leq m$ are added at stage k in ascending order. So, for all $s^\circ \leq m$, we find that \perp^{s° precedes \perp^m in Λ_k . Finally, γ as a sub-string of \perp^m is of the form \perp^{s° , for some $s^\circ \leq m$.

Suppose β_n is added as a sub-string of α_k^* not of the form \perp^m . First suppose $\gamma = \perp^p$. We note that $p + 1 = \lfloor \perp^p \rfloor < \lfloor \alpha_k \rfloor \leq k + 1$. This tells us, by Lemma 5.6, that \perp^p is in Λ_{k-1} . Hence, it is added to Λ before β_n . Now suppose that γ is not of the form \perp^p . Then, either γ is in Λ_{k-1} or added to Λ_k before β_n . □

LEMMA 5.8. *The enumeration Λ is without repetitions.*

Proof. Consider any $\gamma \in \mathcal{A}^*$. If γ is not of the form \perp^m by our construction it will be only added once. Suppose $\gamma = \perp^m$. We note that, by Lemma 5.7, the n such that $\perp^n \in \Lambda_{k-1}$ are downwards closed. This means that, in our construction, there is no $m \geq s$, such that $\perp^m \in \Lambda_{k-1}$. So all the \perp^m that are added in any stage k are new. □

LEMMA 5.9. *Suppose $S\tilde{n}$ occurs in Λ_k , then $n \leq k + 15$.*

Proof. We prove this by induction on k . In stage 0, we easily verify that the largest term of the form $S\tilde{n}$ can be at most $S\tilde{5}$.

Suppose, we have our desired estimate for $k - 1$, where $k > 0$. We prove our estimate for k . Clearly, the $S\tilde{n}$ occurring in Λ_{k-1} do satisfy our estimate. Moreover, all the \perp^p added in stage k do not provide new elements of the form $S\tilde{n}$. So, the only interesting case is the case where α_k is not a string of \perp 's. Suppose it is not. The $S\tilde{n}$ that are added are all sub-strings of $\alpha_k^* = \alpha_k[c := S(k + 15)]$. So, we should focus on the largest such sub-string. Suppose $S\tilde{n} \sqsubseteq \alpha_k^*$. There are two possibilities:

- i. $S\tilde{n} \sqsubseteq \alpha_k$. In this case, we have $n < \lfloor S\tilde{n} \rfloor \leq \lfloor \alpha_k \rfloor \leq k + 1 \leq k + 15$.
- ii. $n = k + 15$, since the sub-string results from the substitution for c .

In each of the possible cases, we are done. □

LEMMA 5.10. *Consider any α such that c occurs in β . Then, for some n , $\beta_n = \alpha[c := \bar{n}]$.*

Proof. Suppose c occurs in α . Let $\alpha := \alpha_k$. The only thing we have to show is that α_k^* is not in Λ_{k-1} . This is certainly true for $k = 0$. Suppose $k > 0$. We note that $S(k + 15)$ occurs in α_k^* . So, it is sufficient to show that $S(k + 15)$ does not occur in Λ_{k-1} . By Lemma 5.9, whenever $S\tilde{n}$ occurs in Λ_{k-1} , we have $n \leq k + 14$. We may conclude that $\beta_{n(k)-1} = \alpha_k[c := S(k + 15)]$. Since $S(k + 15) = \overline{n(k) - 1}$, we are done. □

Setting $gn_1(\alpha)$ to be the unique n such that $\alpha = \beta_n$, we have found a monotonic self-referential Gödel numbering gn_1 . Here monotonicity is defined with respect to the sub-string relation.

§6. Gödel Numerals as Quotations. In the philosophical literature, Gödel numerals are typically conceived of as canonical names for their coded expressions. One of the most canonical naming devices is quotation, assigning to each expression A its name “ A .” Accordingly, the Tarski biconditional $\top(\underline{\xi(A)}) \leftrightarrow A$ for instance is often taken as an arithmetical proxy for the sentence “‘ A ’ is true iff A .”

The standard numeral $\xi(A)$ is *prima facie* not the only candidate for an arithmetical proxy for quotation. As there are different adequate ways to quote an expression A , such as “ A ”, ‘ A ’, « A », etc., different numerals may be reasonably taken as adequate arithmetical proxies of A ’s quotation, such as $\zeta(A)$, $\xi(A)$, etc.¹¹

Once v -numerals are taken as adequate arithmetical proxies for quotations, another criterion for numberings may be extracted from the basic idea underlying monotonicity, which has been suggested by Halbach [9]. Recall that monotonic numberings have the property that the code of an expression A is larger than the code of any expression (strictly) contained in A . Since each quotation properly contains its quoted expression, it may be argued, along the lines of Section 3.1, that adequate numberings should preserve this structural feature by requiring that the code of the Gödel v -numeral of an expression A is larger than the code of A itself:

$$M4(v). \quad \zeta(A) < \zeta(v(\zeta(A))) \text{ for all } A \in \mathcal{L},$$

We call a numbering ζ of \mathcal{L} *strongly monotonic for v-numerals*, if ζ satisfies M1–M3 as well as M4(v).

We note that M4(v) rules out that $\zeta(0) = 0$ and $v(0) = 0$, which holds for certain length-first orderings and standard numerals. These standard numberings are not excluded by its weakened version M4*(v), which is obtained by replacing $<$ by \leq in M4(v). In what follows, we will always employ the version of M4(v) which yields the stronger result.

Since the standard numeral \underline{n} has n -many proper subterms, any numbering satisfying M1 also satisfies M4(\cdot). Strong monotonicity is however not entailed by monotonicity when efficient numerals are employed. For instance, as will follow from the next lemma, the numbering gn_0 is monotonic but not strongly monotonic for efficient numerals. In fact, strong monotonicity is incompatible with the self-referentiality of numberings.

LEMMA 6.1. *Let v be a numeral function for \mathcal{L} . There is no numbering of \mathcal{L} which is self-referential for v and satisfies M3 and M4*(v).*

¹¹ While we agree that both standard and efficient Gödel numerals are canonical names for their coded expressions, we remain sceptical as to whether they should be understood as proxies for quotations. For instance, the word “snow” can be adequately named by its structural-descriptive name *ess en oh doubleyoo*, by “wons’,” by “ssnooww’,” by baptising it “(1),” etc., all of which do not contain the named expression. The justification of the constraint M4(v), introduced below, seems thus to rely on a specific feature of quotation which is not shared by several other adequate naming devices.

Proof. Suppose ξ is a self-referential numbering of \mathcal{L} . Consider any formula $A(x)$ with free variable x and $n \in \omega$ such that $\xi(A(v(n))) = n$. Since $v(n) \prec A(v(n))$, using M3 and M4*(v) yields the contradiction

$$n = \xi(A(v(n))) \leq \xi(v(\xi(A(v(n)))))) = \xi(v(n)) < \xi(A(v(n))) = n. \quad \square$$

Thus, if we strengthen the conception of monotonicity in the above manner, then, by Halbach’s standards, there are no reasonable numberings which are self-referential.

6.1. A critical remark. We are somewhat sceptical whether strong monotonicity really is a necessary constraint for reasonable numberings (see also footnote 11). For instance, we now show that the length-first numbering is not strongly monotonic for a fairly reasonable numeral function based on prime decomposition. Any philosophically stable view which excludes self-referential numberings as adequate formalisation choices along the above lines, hence appears to be committed to render the introduced numeral function below an inadequate naming device.

EXAMPLE 6.2. Let $\mathcal{L}^Q := \mathcal{L}^0 \cup \{Q\}$, where Q is a unary function symbol for taking the square. Let \mathfrak{g} be a numbering of \mathcal{L}^Q based on the length-first ordering.

Are there plausible numerals for which strong monotonicity fails w.r.t. \mathfrak{g} ? This is indeed the case. We give an example of such numerals that seems reasonably natural.

We consider the following numeral function pd . (‘pd’ stands for prime decomposition.)

- i. $\text{pd}(0) := 0$ and $\text{pd}(1) := 50$.
- ii. Suppose $n > 1$, and n is not a power of a prime. Let p be the smallest prime that divides n . Say $n = p^i \cdot m$, where p does not divide m and $i > 0$. Then, $\text{pd}(n) := \times \text{pd}(p^i) \text{pd}(m)$.
- iii. Suppose $n = p^i$, where p is prime and $i > 0$. We take $\text{pd}(p) := \underline{p}$. In case $i = 2k$, for $k > 0$, we set $\text{pd}(n) := Q \text{pd}(p^k)$. In case $i = 2k + 1$, for $k > 0$, we set $\text{pd}(n) := \times \underline{p} Q \text{pd}(p^k)$.

We find that, e.g.,

$$\text{pd}(2^{2^m}) = \overbrace{Q \cdots Q}^{m \times} \underline{2}.$$

It follows that $|\text{pd}(2^{2^m})| = m + 3$. Hence, by Lemma 2.3,

$$\mathfrak{g}(\text{pd}(2^{2^m})) \leq \frac{16^{m+4} - 1}{15}.$$

Since 2^{2^m} has double exponential growth, the failure of strong monotonicity is immediate.

We note that pd has good properties for multiplication but does not seem to be very friendly towards successor and addition.

One can imagine many variants of pd . For example, we can replace \underline{p} by \overline{p} in the definition. Alternatively, we can make \mathfrak{p} where we replace clause (iii) by:

- Suppose $n = p^i$, where p is prime and $i > 0$. We take $\mathfrak{p}(p) := S\mathfrak{p}(p - 1)$. In case $i = 2k$, for $k > 0$, we set $\mathfrak{p}(n) := Q \mathfrak{p}(p^k)$. In case $i = 2k + 1$, for $k > 0$, we set $\mathfrak{p}(n) := \times S\mathfrak{p}(p - 1) Q \mathfrak{p}(p^k)$.

We leave an in-depth discussion of this constraint’s adequacy for another occasion and, from now on, simply assume that there are good reasons to consider strong monotonicity to be a necessary constraint for reasonable numberings. In the next

section we further examine the implications of this constraint on the attainability of self-reference.

6.2. A strongly monotonic numbering with strong diagonalisation. While strong monotonicity excludes self-referential numberings (Lemma 6.1), we now show that this constraint is not sufficiently restrictive to rule out m-self-reference in \mathcal{L}^0 . In particular, we construct a strongly monotonic numbering $\widehat{\text{gn}}_2$ for efficient numerals which satisfies the Strong Diagonal Lemma for $\mathcal{L} \supseteq \mathcal{L}^0$:

LEMMA 6.3 (Strong Diagonal Lemma for \mathcal{L}). *Let C be an \mathcal{L} -expression with x as a free variable. Then there exists a closed term t in \mathcal{L} such that $\text{R}_{\mathcal{L}} \vdash t = \widehat{\text{gn}}_2(C[x := t])$.*

We note that we escape Lemma 6.1 here by allowing t to be different from the efficient numeral of the value of t .

Let \mathcal{L} be an arithmetical language, as introduced in Section 2.1, and let $\mathcal{L}(c)$ be \mathcal{L} extended with a fresh constant c . Let A_0, A_1, \dots be an effective enumeration of all expressions of $\mathcal{L}(c)$. We assume that if $C \preceq A_n$, then, for some $k \leq n$, we have $C = A_k$.

We define $[A]$ as the number of subexpressions-qua-type of A , in other words, $[A]$ is the cardinality of $\{B \mid B \preceq A\}$. The following trivial observation is very useful:

LEMMA 6.4. $[A_k] \leq k + 1$.

Let the function $\widehat{\cdot} : \omega \rightarrow \mathcal{L}$ be given by setting $\widehat{0} := v$ and $\widehat{n + 1} := \widehat{n}'$. As in the case of standard numerals, expressions of the form \widehat{m} are downwards closed with regard to the sub-expression relation and we have that $[\widehat{m}] = m + 1$. In the following construction, we use expressions of the form \widehat{m} instead of standard numerals as fillers. This ensures that the constructed enumeration does not start with $\widehat{0}$ and thus that the resulting numbering is strongly monotonic.

Here is our construction.

We construct a list $\Lambda := (B_n)_{n \in \omega}$ in stages k . Let $n(k) := 2^{2k+4} + 1$. Each stage k will result in a list $\Lambda_k = B_0, \dots, B_{n(k)-1}$. To simplify the presentation, we make Λ_{-1} the empty list and $n(-1) := 0$.

In stage k , we act as follows. Let $A_k^* := A_k[c := S(k + 2)]$. Let the sub-expressions of A_k^* that do not occur in Λ_{k-1} and are not of the form \widehat{m} be $A_{i_0}, \dots, A_{i_{\ell-1}}$, where the sequence i_j is strictly increasing. We note that ℓ could be 0. We define $B_{n(k)-\ell+j} := A_{i_j}$, for $j < \ell$. Let s be the smallest number such that \widehat{s} is not in Λ_{k-1} . We set $B_{n(k-1)+p} := \widehat{s + p}$, for $p < n(k) - n(k-1) - \ell$.

To see that our construction is well-defined it is sufficient that $n(k) - \ell \geq n(k-1)$ or, equivalently, $\ell \leq n(k) - n(k-1)$. We note that:

$$\ell \leq [A_k^*] \leq [A_k] + [S(k + 2)] - 1 \leq k + 1 + 2(k + 2) + 3 = 3k + 8. \quad (*)$$

(The -1 in the right-hand-side of the second inequality can be seen as follows. In case c occurs in A_k it is subtracted in the substitution. If c does not occur, we have $A_k = A_k^*$ and, from this, the inequality follows immediately.)

If $k = 0$, we have $3 \cdot 0 + 8 = 8 < 17 = 2^{2(0+2)} + 1 - 0 = n(0) - n(-1)$. Let $k > 0$. We find:

$$3k + 8 < 3 \cdot 2^{2k+2} = 2^{2k+4} + 1 - 2^{2k+2} - 1 = n(k) - n(k-1).$$

LEMMA 6.5. *Suppose $A_k \in \mathcal{L}$. Then, A_k is in Λ_k and, hence, in Λ .*

Proof. Consider stage k . We note that $A_k^* = A_k$. In case A_k occurs in Λ_{k-1} , we are done. In case A_k is not in Λ_{k-1} and not of the form \widehat{m} , clearly, A_k will be added, so we are again done. Suppose $A_k = \widehat{m}$, for some m . It follows that $m + 1 = \lceil \widehat{m} \rceil \leq k + 1$. We note that all sub-expressions of A_k are of the form \widehat{n} , for $n \leq m$. So $\ell = 0$. Clearly, $m + 1 \leq k + 1 < n(k) - n(k - 1)$. So, all sub-expressions of A_k are either in Λ_{k-1} or will be added. □

LEMMA 6.6. *Suppose $C \preceq B_n$. Then, for some $j \leq n$, we have $C = B_j$.*

Proof. Let $C \preceq B_n$. Suppose B_n is added in stage k . Then, B_n is either of the form \widehat{m} or a sub-expression of A_k^* not of the form \widehat{m} .

Suppose $B_n = \widehat{m}$. Let s be the smallest number such that \widehat{s} is not in Λ_{k-1} . For all $u < s$, we have $\widehat{u} \in \Lambda_{k-1}$. By our construction, all \widehat{v} , for $s \leq v \leq m$ are added at stage k in ascending order. So, for all $w \leq m$, we find that \widehat{w} precedes \widehat{m} in Λ_k . Finally, each sub-expression C of \widehat{m} is of the form \widehat{w} , for some $w \leq m$.

Suppose B_n is added as a sub-expression of A_k^* not of the form \widehat{m} . First suppose $C = \widehat{p}$. We note that $p + 1 = \lceil \widehat{p} \rceil < \lceil A_k \rceil \leq k + 1$. This tells us, by Lemma 6.5, that \widehat{p} is in Λ_{k-1} . Hence, it is added to Λ before B_n . Now suppose that C is not of the form \widehat{p} . Then, either C is in Λ_{k-1} or added to Λ_k before B_n . □

LEMMA 6.7. *The enumeration Λ is without repetitions.*

Proof. Consider any $C \in \mathcal{L}$. If C is not of the form \widehat{m} by our construction it will be only added once.

Suppose $C = \widehat{m}$. We note that, by Lemma 6.6, the n such that $\widehat{n} \in \Lambda_{k-1}$ are downwards closed. This means that, in our construction, there is no $m \geq s$, such that $\widehat{m} \in \Lambda_{k-1}$. So all the \widehat{m} that are added in any stage k are new. □

LEMMA 6.8. *Suppose $S\widetilde{n}$ occurs in Λ_k , then $n \leq k + 2$.*

Proof. We prove this by induction on k . In stage 0, we easily verify that the largest term of the form $S\widetilde{n}$ can be at most $S\widetilde{2}$.

Suppose, we have our desired estimate for $k - 1$, where $k > 0$. We prove our estimate for k . Clearly, the $S\widetilde{n}$ occurring in Λ_{k-1} do satisfy our estimate. Moreover, all the \widehat{p} added in stage k do not provide new elements of the form $S\widetilde{n}$. So, the only interesting case is the case where A_k is not of the form \widehat{p} . Suppose it is not. The $S\widetilde{n}$ that are added are all sub-expressions of $A_k^* = A_k[c := S(k + 2)]$. So, we should focus on the largest such subterm. Suppose $S\widetilde{n} \preceq A_k^*$. There are two possibilities:

- i. $S\widetilde{n} \preceq A_k$. In this case, we have $n < \lceil S\widetilde{n} \rceil \leq \lceil A_k \rceil \leq k + 1 \leq k + 2$.
- ii. $n = k + 2$, since the subterm results from the substitution for c .

In each of the possible cases, we are done. □

Setting $gn_2(A)$ to be the unique n such that $A = B_n$, we have found a monotonic Gödel numbering gn_2 .

In order to show that gn_2 satisfies $M4(\cdot)$, i.e., that gn_2 is strongly monotonic for efficient numerals, we prove the following auxiliary lemma.

LEMMA 6.9. *Suppose $\overline{p} \preceq A_k[c := \overline{r}]$, for some $p, r \in \omega$. Then, $p < 2^k(r + 2)$.*

Proof. Let $p, r \in \omega$ be given such that $\bar{p} \preceq A_k[c := \bar{r}]$. Then (i) $\bar{p} \preceq A_k$ or (ii) $\bar{p} \preceq \bar{r}$ or (iii) there exists $t \preceq A_k$ containing c such that $\bar{p} = t[c := \bar{r}]$. In case (i), we have, $p < 2^{\lceil \bar{p} \rceil} \leq 2^{\lceil A_k \rceil} \leq 2^{k+1} \leq 2^k(r+2)$. In case (ii), we have $p \leq r$. So both in Case (i) and (ii), we find $p < 2^k(r+2)$.

Suppose we are in case (iii). Since \bar{r} and $t[c := \bar{r}]$ are efficient numerals, there exist k, j such that \bar{r} and t are of the following forms:

- $\bar{r} = S_{a_0}(\dots S_{a_k}(0) \dots)$;
- $t = S_{c_0}(\dots S_{c_j}(c) \dots)$;

where $a_p, c_q \in \{1, 2\}$ for each $p \leq k$ and $q \leq j$, and $S_1(x) := S(SS0 \times x)$ and $S_2(x) := SS(SS0 \times x)$. Moreover, $ev(\bar{r}) = \sum_{i=0}^k a_i b^i$. Since

$$t(\bar{r}) = S_{c_0}(\dots S_{c_j}(S_{a_0}(\dots S_{a_k}(0) \dots)) \dots),$$

we have

$$\begin{aligned} p &= ev(t(\bar{r})) \\ &= \sum_{i=0}^j c_i 2^i + \sum_{i=0}^k a_i 2^{i+j+1} \\ &= \sum_{i=0}^j c_i 2^i + 2^{j+1} \sum_{i=0}^k a_i 2^i \\ &< 2^{j+2} + 2^{j+1}r \\ &= 2^{j+1}(r+2). \end{aligned}$$

For each $i \leq j$, the expression $S_{c_i}(\dots S_{c_j}(c) \dots)$ is a subterm of A_k . Since also c is a subterm of A_k , we have $j+2 \leq \lceil A_k \rceil \leq k+1$. We thus conclude that $p < 2^k(r+2)$. □

We now show that gn_2 satisfies M4($\bar{\cdot}$).

LEMMA 6.10. $m < gn_2(\bar{m})$ for all $m \in \omega$.

Proof. Let \bar{m} be added in stage k , i.e., $\bar{m} \in \Lambda_k \setminus \Lambda_{k-1}$. We then have $B_{n(k)-\ell+j} = \bar{m}$, for some $j < \ell$. (We use here the fact that no filler is of the form \bar{m}). We have $\ell \leq 3k+8$, as noted in (*) subsequent to presenting our construction. Hence,

$$2^{2k+4} - 3k - 7 = n(k) - (3k + 8) \leq n(k) - \ell \leq gn_2(\bar{m}).$$

We moreover have $\bar{m} \preceq A_k[c := \widetilde{S(k+2)}]$. Since $\widetilde{S(k+2)} = \overline{2^{k+2}}$ we get

$$m < 2^k(2^{k+2} + 2) = 3 \cdot 2^{k+1},$$

by Lemma 6.9. It is easy to check that $3 \cdot 2^{k+1} < 2^{2k+4} - 3k - 7$. Hence, $m < gn_2(\bar{m})$. □

Even though gn_2 cannot be self-referential for efficient numerals by Lemma 6.1, the following modification holds.

LEMMA 6.11. Consider any A such that c occurs in A . Then there exists n such that $gn_2(A[c := \bar{n}]) = n^2$.

Proof. Suppose c occurs in A . Let $A := A_k$. The only thing we have to show is that A_k^* is not in Λ_{k-1} . This is certainly true for $k = 0$. Suppose $k > 0$. We note that $S(k+2)$ occurs in A_k^* . So, it is sufficient to show that $S(k+2)$ does not occur in Λ_{k-1} . By Lemma 6.8, whenever $S\bar{n}$ occurs in Λ_{k-1} , we have $n \leq k+1$. We may conclude that $B_{n(k)-1} = A_k[c := S(k+2)]$. Let $n := 2^{k+2}$. Since $n^2 = n(k) - 1$ and $S(k+2) = \bar{n}$, we are done. \square

We now can immediately derive the Strong Diagonal Lemma.

Proof of Lemma 6.3. Let $C(x)$ be given and set $A := C[x := c \times c]$. By Lemma 6.11 there exists $n \in \omega$ such that $gn_2(A[c := \bar{n}]) = n^2$. Set $t := \bar{n} \times \bar{n}$. We then have $A[c := \bar{n}] = C[x := t]$. Moreover $R_{\mathcal{L}} \vdash t = \underline{n^2}$. Hence $R_{\mathcal{L}} \vdash t = gn_2(C[x := t])$. \square

By suitably adapting the above construction, we obtain a numbering which is monotonic with respect to the sub-string relation, satisfies $M4(\bar{\cdot})$ and provides the existence of m -self-referential sentences for \mathcal{L} (i.e., satisfies Lemma 6.3). We leave the details to the diligent reader.

§7. Computational Constraints. In addition to requiring monotonicity, other adequacy constraints for Gödel numberings can be extracted from the literature. Let $\mathfrak{P}o$, $\mathfrak{E}l$ and $\mathfrak{P}r$ denote the classes of p -time, (Kalmár) elementary and primitive recursive functions respectively. As usual, these classes can be extended to relations and partial functions on ω as follows.

DEFINITION 7.1. Let $\mathcal{C} \in \{\mathfrak{P}o, \mathfrak{E}l, \mathfrak{P}r\}$. We say that a subset of ω^k or a relation on ω is in \mathcal{C} , if its characteristic function is in \mathcal{C} .

Let $R \subseteq \omega^k$ be given. We say that a function $f : R \rightarrow \omega$ is in \mathcal{C} , if the total function $f' : \omega^k \rightarrow \omega$ given by

$$f'(m_1, \dots, m_k) := \begin{cases} f(m_1, \dots, m_k) + 1 & \text{if } m_1, \dots, m_k \in R; \\ 0 & \text{otherwise.} \end{cases}$$

is in \mathcal{C} .

Reasonable numberings are commonly required to represent certain syntactic relations and operations by primitive recursive relations and operations on ω (see [8, sec. 5.1]). The usual definitions of the syntactic relations (i.e., their extensions) $Var_{\mathcal{L}}$ (“is a variable”), $Ter_{\mathcal{L}}$ (“is a term”), $AtFml_{\mathcal{L}}$ (“is an atomic formula”), $Fml_{\mathcal{L}}$ (“is an a formula”), $Sent_{\mathcal{L}}$ (“is a sentence”), “is a free variable in,” “is a PA-derivation of” etc., are either explicit or of simple recursive structure. Adequate numberings may be required to preserve this simple algorithmic nature, in virtue of representing these relations by primitive recursive relations on ω . More precisely, for any adequate numbering ξ , the characteristic functions of the sets of ξ -codes of $Var_{\mathcal{L}}$, $Ter_{\mathcal{L}}$, etc. are required to be primitive recursive. A similar constraint can be extracted for the syntactic functions $\dot{=}$, $\dot{\wedge}$, $\dot{\vee}$, $\dot{\exists}$, \dot{S} , $\dot{+}$, $\dot{\times}$, Sub etc., where for instance, $\dot{\wedge} : Fml_{\mathcal{L}}^2 \rightarrow Fml_{\mathcal{L}}$ maps a pair $\langle A, B \rangle$ of formulæ to the conjunction $A \wedge B$, $\dot{\forall} : Fml_{\mathcal{L}} \times Var_{\mathcal{L}} \rightarrow Fml_{\mathcal{L}}$ maps a pair $\langle A, x \rangle$ to the universal formula $\forall x A$, etc., and where $Sub(A, x, t)$ is the result of substituting the term t for the variable x in the formula A . Once again, adequate numberings may be required to represent these syntactic functions

by primitive recursive functions on ω . Finally, the standard and the efficient numeral function may be required to be represented primitive recursively.

The arithmetisation of syntax in weaker theories, like Buss' S^1_2 and $IA\Delta_0 + \Omega_1$, requires representation of the syntactic relations and functions in weaker theories.¹² A closer inspection of what is going on shows that, in the presence of a reasonable Gödel numbering like the one based on the length-first ordering, the syntactic relations and functions can be made p-time *without extra effort*. See [2, 7]. The Gödel numberings we construct here fit most naturally in the intermediate function class $\mathcal{E}l$.

DEFINITION 7.2. *Let $C \in \{\mathfrak{P}o, \mathcal{E}l, \mathfrak{P}r\}$ and let v be a numeral function. We say that a numbering ξ of \mathcal{L} is C, v -adequate, if*

1. *the set of ξ -codes of each syntactic relation specified above is in C ;*
2. *the ξ -tracking function of each syntactic function specified above is in C , i.e., for each k -ary syntactic function f , the function $\xi(f)$ given by $\xi(f)(m_1, \dots, m_k) := \xi(f(\xi^{-1}(m_1), \dots, \xi^{-1}(m_k)))$ is in C ; and*
3. *the function $\xi \circ v$ is in C .*

We say that ξ is C -adequate iff ($C \in \{\mathcal{E}l, \mathfrak{P}r\}$ and ξ is both $C, (\bar{\cdot})$ -adequate and $C, (\cdot)$ -adequate) or ($C = \mathfrak{P}o$ and ξ is $C, (\bar{\cdot})$ -adequate).

The definition of C -adequate is somewhat awkward since, for a good numbering like the one based on the length-first ordering, the function $\xi \circ (\cdot)$ will be exponential. This is the reason for the use of efficient numerals in the context of weak theories. It would be interesting, and conceivably genuinely useful, to explore whether there are $\mathfrak{P}o$ -adequate numberings ξ for which $\xi \circ (\cdot)$ is p-time.

REMARK 7.3. *Feferman's Gödel numbering, say fef , in [4] is a good example to reflect on. The tracking functions are all elementary. The code for 0 is 3 and the code for successor is 7. We have $fef(St) = 2^7 \cdot 3^{fef(t)}$. So, the numeral function for standard numerals will be superexponential and not elementary. Similarly, for efficient numerals.*

Suppose we replace Feferman's coding for the terms by an efficient coding but leave his code formation for formulæ in place (taking appropriate measures to insure disjointness). Then, we obtain an $\mathcal{E}l$ -adequate Gödel numbering F . However, we still would have: $F(\neg A) := 2 \cdot 3^{F(A)}$. This means that, if we iterate negations, we run up an exponential tower.

The example of F illustrates that it is possible that a Gödel numbering is $\mathcal{E}l$ -adequate but not elementary in the length of expressions. This observation suggests that, where p-time and primitive recursive satisfy a certain equilibrium, possibly elementary does not. If the tracking function for function application is exponential, the codes of numerals, both standard and efficient, grow too fast. So, the tracking function for function application is severely constrained. But if this function is so slow, it seems a certain imbalance to make other forms of application much faster. So, why is this numbering not already $\mathfrak{P}o$ -adequate? ...

We now introduce a useful proof-theoretical characterisation of the classes $\mathfrak{P}o$, $\mathcal{E}l$ and $\mathfrak{P}r$. This will result in a specification of the theories in which C -adequate numberings permit the arithmetisation of syntax. To this end, we first introduce so-called *provably recursive* functions.

¹² See [24] for ways to deal with inefficient representations in weak theories.

DEFINITION 7.4. Let T be an \mathcal{L} -theory. A function $f : \omega^k \rightarrow \omega$ is called Σ_n -definable in T , if there exists a Σ_n -formula $A(x_1, \dots, x_k, y)$ such that

- $\mathbb{N} \models A(\underline{m_1}, \dots, \underline{m_k}, \underline{p})$ iff $f(m_1, \dots, m_k) = p$, for all $m_1, \dots, m_k, p \in \omega$ and
- $T \vdash \forall x_1, \dots, x_k \exists! y A(x_1, \dots, x_k, y)$.

The function f is called provably recursive in T if f is Σ_1 -definable in T .

REMARK 7.5. We note that there is some awkwardness to this definition since the provably recursive functions of all Σ_1 -unsound theories extending EA, like $\mathbb{I}\Sigma_1 + \text{incon}(\text{ZF})$, are precisely all recursive functions. This creates the mistaken impression that, e.g., $\mathbb{I}\Sigma_1 + \text{incon}(\text{ZF})$ is stronger than PA.

With an extra argument, one can show that this malaise persists even when we replace ' $\mathbb{N} \models A(\underline{m_1}, \dots, \underline{m_k}, \underline{p})$ ' in the definition by ' $T \vdash A(\underline{m_1}, \dots, \underline{m_k}, \underline{p})$ '.

We are unaware of attempts to address this awkwardness.

The classes $\mathfrak{E}\mathfrak{l}$ and $\mathfrak{P}\mathfrak{r}$ can now be characterised as exactly the functions which are provably recursive in a certain theory. The situation for the p-time computable functions is a bit more delicate.

THEOREM 7.6. Let f be a number-theoretic function. Then

- f is p-time iff f is Σ_1^b -definable in S_2^1 , where Σ_1^b is the special formula class introduced in [2];
- f is elementary iff f is provably recursive in EA;
- f is primitive recursive iff f is provably recursive in $\mathbb{I}\Sigma_1$.

Hence, for example, for any $\mathfrak{P}\mathfrak{r}$ -adequate numbering ξ there exists a Σ_1 -formula $\text{Conj}(x, y, z) \in \mathcal{L}$, such that

- $\mathbb{N} \models \text{Conj}(\xi(A), \xi(B), \xi(C))$ iff $C = A \wedge B$, for all formulæ $A, B, C \in \mathcal{L}$;
- $\mathbb{I}\Sigma_1 \vdash \forall x, y \exists! z \text{Conj}(x, y, z)$;

i.e., the ξ -tracking function of $\hat{\lambda}$ is provably recursive in $\mathbb{I}\Sigma_1$. Indeed, the representation of $\hat{\lambda}$ by a provably recursive function (in PA) is required as a necessary condition for reasonable numberings in [8, p. 33]. As a result of the above theorem, $\mathfrak{P}\mathfrak{r}$ -adequate numberings allow a large portion of syntax to be formalised in the theory $\mathbb{I}\Sigma_1$ and, similarly, for $\mathfrak{E}\mathfrak{l}$ and EA and for $\mathfrak{P}\mathfrak{o}$ and S_2^1 .¹³

Recall that the constructions of the numberings gn_0, gn_1 and gn_2 rely on an enumeration $(A_n)_{n \in \omega}$ of \mathcal{L} -expressions. Thus far, assuming $(A_n)_{n \in \omega}$ to be effective has sufficed to ensure the effectiveness of the resulting numberings gn_0, gn_1 and gn_2 . Since no further computational constraint has been imposed on $(A_n)_{n \in \omega}$, there is however no reason to expect these numberings to be even $\mathfrak{P}\mathfrak{r}$ -adequate. In the remainder of this paper, we will assume that the enumeration $(A_n)_{n \in \omega}$ (without repetitions) employed in the constructions of the numberings gn_0, gn_1 and gn_2 is obtained from a standard numbering which is $\mathfrak{E}\mathfrak{l}$ -adequate, such that $(A_n)_{n \in \omega}$ also satisfies the remaining assumptions imposed in the constructions of gn_0, gn_1 and gn_2 respectively. That is, $(A_n)_{n \in \omega} = (gn_*^{-1}(n))_{n \in \omega}$, for some suitable standard numbering gn_* .

Under the assumption that $(A_n)_{n \in \omega}$ is $\mathfrak{E}\mathfrak{l}$ -adequate, we now show that the numberings gn_0, gn_1 and gn_2 are $\mathfrak{E}\mathfrak{l}$ -adequate, and, thus, also $\mathfrak{P}\mathfrak{r}$ -adequate. This will

¹³ We note that Halbach's constraint loses its meaning in Σ_1 -unsound extensions of EA.

follow from the fact that the standard numbering gn_* is $\mathfrak{E}\mathfrak{I}$ -adequate and that gn_* , gn_0 , gn_1 and gn_2 are elementarily intertranslatable. Let $tr_{i,j} : gn_j(\mathcal{L}) \rightarrow \omega$ be the translation function, given by $tr_{i,j} := gn_i \circ gn_j^{-1}$, for each $i, j \in \{*, 0, 1, 2\}$. We then have:

LEMMA 7.7. *For each $i, j \in \{*, 0, 1, 2\}$, the translation functions $tr_{i,j}$ are elementary.*

Proof. We treat the case of the intertranslation of $*$ and 2 . To show that $tr_{*,i}$ and $tr_{i,*}$ are elementary for $i = 0, 1$ proceeds similarly. From these cases we can conclude that $tr_{i,j}$ is elementary for $i, j \leq 2$.

We use the fact that the coding machinery for finite sequences can be developed in $\mathfrak{E}\mathfrak{I}$ (see, for example, [26, chap. 2]). Let $\# : \omega^{<\omega} \rightarrow \omega$ be an elementary coding function of finite sequences and let lh be the length function. Let $g : \omega \rightarrow \omega$ serve as a parameter. We define the function $\sigma_g : \omega \rightarrow \omega$, by setting $\sigma_g(k) := \#(i_0, \dots, i_{\ell-1})$, where $A_{i_0}, \dots, A_{i_{\ell-1}}$ are exactly the sub-expressions of $A_k[c := S(k+2)]$ which are not of the form \widehat{m} such that $g(n) \neq i_j$ for all $n < n(k-1)$ and $j < lh(\sigma_g(k)) = \ell$. Since $n \in \mathfrak{E}\mathfrak{I}$ and gn_* is $\mathfrak{E}\mathfrak{I}$ -adequate, $g \in \mathfrak{E}\mathfrak{I}$ implies $\sigma \in \mathfrak{E}\mathfrak{I}$, by the usual closure properties of elementary functions. In case that $g := tr_{2,*}$, the last clause is equivalent to $A_{i_j} \notin \Lambda_{k-1}$ for each $j < \ell$.

Set $\sigma := \sigma_{tr_{2,*}}$. For any $n \in \omega$, we compute $tr_{2,*}(n)$ by course-of-values recursion as follows. As usual, we will take the empty sum to be 0.

Compute the smallest $k (\leq n)$ such that $n \leq 2^{2k+4}$. Is $p := 2^{2k+4} - n < lh(\sigma(k))$?

If *yes*, set $tr_{2,*}(n) := [\sigma(k)]_{lh(\sigma(k))-1-p}$.

If *no*, take

$$m := n - \sum_{i=0}^{k-1} lh(\sigma(i)).$$

Set $tr_{2,*}(n) := gn_*(\widehat{m})$.

It is left to the reader to verify that this computation really defines the function $tr_{2,*}$. Note that when computing $tr_{2,*}(n)$, we only resort to values of $tr_{2,*}$ for arguments smaller than n , since $n(k-1) \leq n$. Moreover, let $B : \omega \rightarrow \omega$ the elementary function given by $B(k) := gn_*(A_k[c := S(k+2)])$. Note that $tr_{2,*}(n)$ codes a filler expression of the form \widehat{m} , or a sub-expression of $A_k[c := S(k+2)]$. Since in each stage k , there are less than $n(k)$ -many fillers added, we obtain the estimate $tr_{2,*}(n) \leq \max(gn_*(\widehat{(k+1) \cdot n(k)}, B(k)))$. Hence, the employed recursion is *limited*. Since elementary functions are closed under limited course-of-values recursion, we conclude that $tr_{2,*} \in \mathfrak{E}\mathfrak{I}$.

In order to show that $tr_{*,2}$ is elementary, we use the fact that a function is elementary if (1) its graph is elementary and (2) it can be dominated by an elementary function. Since $tr_{2,*}$ is elementary, it is easy to check that the graph of $tr_{*,2}$ is elementary. Hence, it is sufficient to show (2). Let $A_k \in \mathcal{L}$. By Lemma 6.5, $A_k \in \Lambda_k$. Hence, $tr_{*,2}(k) < n(k)$ and we are done. □

It is easy to show that the elementary (or primitive recursive) relations and functions are invariant regarding numberings which are elementarily (or primitive recursively) intertranslatable.

LEMMA 7.8. Let $\mathcal{C} \in \{\mathfrak{Po}, \mathfrak{El}, \mathfrak{Pr}\}$. Let S_i be sets and let $\alpha_i, \beta_i: S_i \rightarrow \omega$ be injective functions such that the translation functions $\alpha_i \circ \beta_i^{-1}$ and $\beta_i \circ \alpha_i^{-1}$ are in \mathcal{C} , for $i \in \{1, \dots, k\}$ and $k \in \omega$. For every $R \subseteq S_1 \times \dots \times S_k$, $Q \subseteq S_{k+1}$ and function $f: R \rightarrow Q$ we then have:

- The relation

$$\vec{\alpha}(R) := \{\langle \alpha_1(s_1), \dots, \alpha_k(s_k) \rangle \mid \langle s_1, \dots, s_k \rangle \in R\}$$

is in \mathcal{C} iff $\vec{\beta}(R)$ is in \mathcal{C} .

- The function $f_{\vec{\alpha}}: \vec{\alpha}(R) \rightarrow \omega$ given by

$$f_{\vec{\alpha}}(m_1, \dots, m_k) := \alpha_{k+1} f(\alpha_1^{-1}(m_1), \dots, \alpha_k^{-1}(m_k))$$

is in \mathcal{C} iff the function $f_{\vec{\beta}}: \vec{\beta}(R) \rightarrow \omega$ is in \mathcal{C} .

We have seen in this section that \mathfrak{Pr} -adequacy can be extracted as a necessary condition for reasonable numberings from the literature. Since gn_* is \mathfrak{El} -adequate we conclude from Lemmas 7.7 and 7.8 that the numberings gn_0, gn_1 and gn_2 are \mathfrak{El} -adequate and, in particular, \mathfrak{Pr} -adequate. Thus, the computational constraint of \mathfrak{Pr} -adequacy (or \mathfrak{El} -adequacy) is not sufficiently restrictive to deny these numberings the status of adequate formalisation choices.

OPEN QUESTION 7.9 *It remains open whether there are \mathfrak{Po} -adequate versions of the constructions given in this paper. More specifically:*

- Can we find \mathfrak{Po} -adequate self-referential Gödel numberings that are monotonic?
- Can we find \mathfrak{Po} -adequate Gödel numberings which satisfy the Strong Diagonal Lemma 6.3 that are strongly monotonic?

§8. Domination and Regularity. We have seen in Section 6 that strong monotonicity is not sufficient to rule out m-self-referential sentences formulated in \mathcal{L} . This is due to the fact that the “fixed point” term t used in the proof of Lemma 6.3 is not an efficient numeral. Hence, conditions M3 and M4($\bar{\tau}$) do not force the value of t to be smaller than the code of the formula containing t .

For the remainder of this section, let \mathcal{L}, \mathcal{K} be fixed languages with $\mathcal{L}^0 \subseteq \mathcal{K} \subseteq \mathcal{L}$, as specified in Section 2.1. We introduce two related constraints on numberings of \mathcal{L} which prohibit the code of a closed \mathcal{K} -term to be smaller than its value. We will see that together with requiring monotonicity, these constraints, applied to \mathcal{L} , rule out m-self-referential sentences formulated in \mathcal{L} .

8.1. Domination. We start by defining the constraint of *domination*.

DEFINITION 8.1. *Let ξ be a numbering of \mathcal{L} . We say that ξ is \mathcal{K} -dominating, if it satisfies the following principle:*

M5(\mathcal{K}). *For all $t \in \text{CTerm}_{\mathcal{K}}$ whose value is in $\xi(\mathcal{L})$: $\text{ev}(t) \leq \xi(t)$.*

We say that ξ is dominating, if it is \mathcal{L} -dominating.

Monotonic dominating numberings are strongly monotonic. More precisely, M5(\mathcal{K}) implies M4*(v) for every numeral function v for \mathcal{K} (cf. Section 6):

COROLLARY 8.2. *Let v be a numeral function for \mathcal{K} and let ξ be a numbering of \mathcal{L} . If ξ satisfies M5(\mathcal{K}), then ξ satisfies M4*(v).*

Proof. We have $n = \text{ev}(v(n)) \leq \xi(v(n))$ for all $n \in \omega$. Hence, $\xi(A) \leq \xi(v(\xi(A)))$ for all $A \in \mathcal{L}$. □

We now show that when employing monotonic and \mathcal{K} -dominating numberings, m-self-reference is not attainable in \mathcal{K} . That is, taken together, these two constraints rule out the existence of strong fixed point terms in \mathcal{K} .

LEMMA 8.3. *Let ξ be a numbering of \mathcal{L} satisfying M3 and M5(\mathcal{K}). Then for all \mathcal{L} -formula $A(x)$ and closed \mathcal{K} -terms t , $\mathbb{N} \not\models t = \xi(A(t))$.*

Proof. Assume that there exists an \mathcal{L} -formula $A(x)$ and a closed \mathcal{K} -term t such that $\mathbb{N} \models t = \xi(A(t))$. Using M3 and M5(\mathcal{K}), we then derive the contradiction $\text{ev}(t) = \xi(A(t)) > \xi(t) \geq \text{ev}(t)$. □

Let ξ be a standard numbering of \mathcal{L}^+ , where \mathcal{L}^+ contains a term which represents the canonical strong diagonal function. That is, \mathcal{L}^+ represents the function which maps (the ξ -code of) an \mathcal{L}^+ -formula $A(x)$ with x free to (the ξ -code of) its diagonalisation $A(\xi(A))$. As we have seen already, there exists an \mathcal{L}^+ -formula $A(x)$ and a term $t \in \mathcal{L}^+$ such that $\mathbb{N} \models t = \xi(A[x := t])$ (Lemma 1.1).

The reason to define the constraints of domination with respect to a *sub*-language of the domain of a numbering can be illustrated as follows. While ξ cannot be dominating with respect to ξ 's domain (namely, \mathcal{L}^+) by Lemma 8.3, it is possibly dominating with respect to the sub-language \mathcal{L}^0 (for instance in case that ξ is the length-first ordering).

However, as we will show in Section 8.2.2, there are reasonable numberings which even fail to be \mathcal{L}^0 -dominating. Moreover, there is a large class of standard numberings found in the literature which is not dominating with respect to languages which contain a function symbol for exponentiation or the smash function (see Section 8.2.3). These examples put considerable pressure on the view that domination is a necessary condition for reasonable choices of numberings. This constraint thus appears to be of no use to a philosophically adequate approach which aims to block the attainability of m-self-reference in arithmetic formulated in \mathcal{L}^0 .

Exactly the same remarks apply to the constraint of regularity, introduced in the next section.

8.2. Regularity. We now introduce another constraint on numberings, along the lines of a notion put forward in [13, p. 17].¹⁴

DEFINITION 8.4. *Let ξ be a numbering of \mathcal{L} . We say that ξ is \mathcal{K} -regular if*

- $\xi(c) \geq c^{\mathbb{N}}$ for all constant symbols $c \in \mathcal{K}$
- $\xi(f(t_1, \dots, t_k)) > f^{\mathbb{N}}(\xi(t_1), \dots, \xi(t_k))$ for all k -ary function symbols $f \in \mathcal{K}$

Moreover, we call ξ regular, if it is \mathcal{L} -regular.

By induction we immediately get

¹⁴ Heck's [13] notion is an amalgam of our notion of regularity and monotonicity. In fact, Heck's notion of regularity is entailed by M2, M3 and M5(\mathcal{L}^0) (with \geq replaced by $>$). Heck employs regularity as a merely technical constraint, ensuring the termination of "quotational expansion" [13, p. 15], which consists in a process of iterated transformation steps of the form $T(t) \rightsquigarrow T[t := 'A']$, where $T(x)$ is a primitive truth-predicate and t is a closed term denoting the code of A . Heck raises the question as to how this constraint can be philosophically motivated in their closing footnote 40.

LEMMA 8.5. *Let ξ be a \mathcal{K} -regular numbering of \mathcal{L} . Then ξ satisfies $M5(\mathcal{K})$.*

Thus, \mathcal{K} -regular numberings are \mathcal{K} -dominating.

8.2.1. *Example of a regular numbering.* A large class of standard numberings found in the literature are \mathcal{L}^0 -regular and hence in particular \mathcal{L}^0 -dominating. As an example, we show that the length-first numbering of \mathcal{L} is \mathcal{L}^0 -regular, where \mathcal{L} is given in Polish notation (see Section 2.1). This follows as a corollary from a more general observation which has been suggested to us by Fedor Pakhomov.

Let S be a set of strings that is closed under the constructor operations \dot{A} and \dot{M} , given by $\langle \alpha, \beta \rangle \mapsto A\alpha\beta$ and $\langle \alpha, \beta \rangle \mapsto M\alpha\beta$ respectively.

Let $\xi: S \rightarrow \omega$ be a numbering of S and let $g: S^2 \rightarrow S$ be a binary operation on S . We call the function which maps ξ -codes $\langle m, n \rangle$ to $\xi(g(\xi^{-1}(m), \xi^{-1}(n)))$ the ξ -tracking function of g . As usual, we call a function $f: \omega^2 \rightarrow \omega$ monotonic, if $f(x_1, x_2) \leq f(y_1, y_2)$ for all $x_1 \leq y_1$ and $x_2 \leq y_2$.

LEMMA 8.6 (Pakhomov). *Let $\xi: S \rightarrow \omega$ be a bijection such that the ξ -tracking functions of \dot{A} and \dot{M} are monotonic. We have for all $\alpha, \beta \in S$:*

$$\xi(A\alpha\beta), \xi(M\alpha\beta) > \xi(\alpha) \cdot \xi(\beta).$$

Proof. We show the claim for $M\alpha\beta$. Let $\alpha, \beta \in S$. Since ξ is bijective, there are $(\xi(\alpha) + 1) \cdot (\xi(\beta) + 1)$ -many pairs of expressions $\langle \alpha', \beta' \rangle$ such that $\xi(\alpha') \leq \xi(\alpha)$ and $\xi(\beta') \leq \xi(\beta)$. For each such pair $\langle \alpha', \beta' \rangle$, we have $\xi(M\alpha'\beta') \leq \xi(M\alpha\beta)$, since the tracking function of \dot{M} is monotonic. Moreover, for any $\langle \alpha', \beta' \rangle \neq \langle \alpha'', \beta'' \rangle$ we have $\xi(M\alpha'\beta') \neq \xi(M\alpha''\beta'')$ by the injectivity of \dot{M} and ξ . We thus conclude

$$\xi(\alpha) \cdot \xi(\beta) + \xi(\alpha) + \xi(\beta) = (\xi(\alpha) + 1) \cdot (\xi(\beta) + 1) - 1 \leq \xi(M\alpha\beta).$$

In particular, $\xi(\alpha) \cdot \xi(\beta) < \xi(M\alpha\beta)$, if one of $\xi(\alpha), \xi(\beta)$ is > 0 . Moreover, if both of $\xi(\alpha), \xi(\beta)$ are 0, then, by injectivity, $M\alpha\beta$ cannot be 0, so we have $\xi(M\alpha\beta) > 0 = \xi(\alpha) \cdot \xi(\beta)$. □

Let \mathcal{L} be given in Polish notation and let \mathcal{A} be the alphabet of \mathcal{L} . Let \mathfrak{g} be the length-first numbering of \mathcal{A}^* . Note that since \mathfrak{g} is bijective and monotonic with respect to the sub-string relation, the assumptions of Lemma 8.6 are satisfied (for $S := \mathcal{A}^*$). It is then easy to derive the \mathcal{L}^0 -regularity of \mathfrak{g} .

COROLLARY 8.7. *The length-first numbering \mathfrak{g} is \mathcal{L}^0 -regular.*

For any infinite subset Y of ω we define coll_Y to be the unique order preserving bijection from Y to ω .

REMARK 8.8. *Let $X := \mathfrak{g}(\mathcal{L})$. Consider the numbering $\mathfrak{h} := \text{coll}_X \circ \mathfrak{g}$ of \mathcal{L} . The numbering \mathfrak{h} can be seen as a length-first enumeration of the well-formed expressions in \mathcal{L} (as opposed to considering arbitrary strings of \mathcal{A}). Arguably, \mathfrak{h} is a reasonable choice of a Gödel numbering.*

Is \mathfrak{h} \mathcal{L}^0 -regular? Note that Lemma 8.6 cannot be applied to \mathfrak{h} . This is because its domain \mathcal{L} is not single-sorted as it contains both terms and formulae. Hence, \mathcal{L} is not closed under the constructor operations \dot{A} and \dot{M} . However, we can use this lemma to show that the length-first enumeration of the closed-term fragment of \mathcal{L}^0 is \mathcal{L}^0 -regular. Let \mathcal{L}° be given by

- $t ::= 0 \mid St \mid Att \mid Mtt.$

Let $X^\circ := \mathfrak{g}(\mathcal{L}^\circ)$. We take $\mathfrak{h}^\circ := \text{coll}_{X^\circ} \circ \mathfrak{g}$ to be the numbering of \mathcal{L}° . Note \mathcal{L}° is closed under both \mathbb{A} and \mathbb{M} . Hence, we can conclude from Lemma 8.6 that \mathfrak{h}° is \mathcal{L}° -regular.

8.2.2. *Example of a non-regular numbering.* We define a non-regular numbering of \mathcal{L}^0 , given in Polish notation (see Section 2.1).¹⁵ Let \mathcal{L}^* denote the following extension of \mathcal{L}^0 :

- $x ::= v \mid x'$
- $t ::= 0 \mid x \mid St \mid Att \mid Ad \mid Mtt \mid Md$
- $A ::= \perp \mid \top \mid =tt \mid =d \mid \neg A \mid \wedge AA \mid \wedge D \mid \vee AA \mid \vee D \mid \rightarrow AA \mid \rightarrow D \mid \forall x A \mid \exists x A$
- $d ::= \delta t$
- $D ::= \Delta A$

The language \mathcal{L}^* can be seen to implement the sharing of certain sub-expressions. That is, \mathcal{L}^0 -expressions of the form Btt , for any binary constructor symbol B , are “contracted” to expressions of the form $B\delta t$. This contraction and its reversal is performed by the translation functions $\mathfrak{c}: \mathcal{L}^0 \rightarrow \mathcal{L}^*$ and $\mathfrak{t}: \mathcal{L}^* \rightarrow \mathcal{L}^0$, defined recursively as follows: Both translations commute with all construction steps for t and A involving the zero-ary operations, the unary operations and the quantifiers. Moreover:

- $\mathfrak{c}(Att) := \begin{cases} A\delta\mathfrak{c}(t) & \text{if } t = u; \\ A\mathfrak{c}(t)\mathfrak{c}(u) & \text{if } t \neq u; \end{cases}$ • $\mathfrak{t}(A\delta t) := A\mathfrak{t}(t)\mathfrak{t}(t);$
 similarly, for \mathbb{M} ; similarly, for \mathbb{M} ;
- $\mathfrak{c}(=tu) := \begin{cases} =\delta\mathfrak{c}(t) & \text{if } t = u; \\ =\mathfrak{c}(t)\mathfrak{c}(u) & \text{if } t \neq u; \end{cases}$ • $\mathfrak{t}(=\delta t) := =\mathfrak{t}(t)\mathfrak{t}(t);$
- $\mathfrak{c}(\wedge AB) := \begin{cases} \wedge\Delta\mathfrak{c}(A) & \text{if } A = B; \\ \wedge\mathfrak{c}(A)\mathfrak{c}(B) & \text{if } A \neq B; \end{cases}$ • $\mathfrak{t}(\wedge\Delta A) := \wedge\mathfrak{t}(A)\mathfrak{t}(A);$
 similarly, for \vee and \rightarrow . similarly for \vee and \rightarrow .

The translation \mathfrak{t} is well-defined. This follows from unique readability, which is proved in the lemma below.

LEMMA 8.9. \mathcal{L}^* satisfies unique readability.

First Proof. Consider the language $\hat{\mathcal{L}}$ that consists of the arithmetical language extended with a new constant δ and a new (zero-ary) proposition symbol Δ . We prove by an easy induction that every syntactical string of \mathcal{L}^* is a syntactical string of $\hat{\mathcal{L}}$. Now suppose an \mathcal{L}^* syntactical object has two readings. We consider e.g., the case where one of those readings is $M\delta t$. Clearly, the other reading can only be of the form $M\delta t$ since no \mathcal{L}^* -term can begin with δ . Now consider the case where one reading is Mt_1t_2 . The other reading must be of the form Mu_1u_2 . We apply the embedding into $\hat{\mathcal{L}}$ and unique reading in $\hat{\mathcal{L}}$, to see that $t_1 = u_1$ and $t_2 = u_2$. □

Second Proof. Consider the language $\check{\mathcal{L}}$ that consists of the arithmetical language extended with new unary functions A_δ and M_δ , a new unary predicate symbol $=_\delta$ and

¹⁵ We are indebted to the MathOverflow users Matt F., Fedor Pakhomov and Konrad Zdanowski for their helpful suggestions which inspired our construction of the numbering \mathfrak{gn}_3 (see [35]).

new unary connectives \wedge_Δ , \vee_Δ and \rightarrow_Δ . We map \mathcal{L}^* into $\check{\mathcal{L}}$ by the obvious mapping \mathfrak{d} that sends, e.g., $A t_1 t_2$ to $A \mathfrak{d}(t_1) \mathfrak{d}(t_2)$ and $A \delta t$ to $A_\delta \mathfrak{d}(t)$. It is easy to see that \mathfrak{d} is well-defined (since no \mathcal{L}^* -term can start with δ and no \mathcal{L}^* -formula can start with Δ) and injective.

Now suppose an \mathcal{L}^* syntactical object has two readings. We consider, e.g., the case where one of those readings is $M \delta t$. Clearly, the other reading can only be of the form $M \delta t$ since no \mathcal{L}^* -term can begin with δ . Now consider the case where one reading is $M t_1 t_2$. The other reading must be of the form $M u_1 u_2$. We apply the embedding \mathfrak{d} into $\hat{\mathcal{L}}$ and unique reading in $\hat{\mathcal{L}}$, to see that $\mathfrak{d}(t_1) = \mathfrak{d}(u_1)$ and $\mathfrak{d}(t_2) = \mathfrak{d}(u_2)$. By the injectivity of \mathfrak{d} , we find $t_1 = u_1$ and $t_2 = u_2$. □

Let \mathcal{A} be the alphabet of \mathcal{L}^* with some fixed ordering. Let \mathfrak{g} be the length-first numbering of \mathcal{A}^* . We define our numbering \mathfrak{gn}_3 of \mathcal{L}^0 by setting $\mathfrak{gn}_3(\varphi) := \mathfrak{g}(c(\varphi))$. Clearly, \mathfrak{gn}_3 is injective, effective and (strongly) monotonic. We show that \mathfrak{gn}_3 is not regular.

THEOREM 8.10. *\mathfrak{gn}_3 is not regular.*

Proof. Since the size of \mathcal{A} is 17, by Lemma 2.3 we have for each $\alpha \in \mathcal{A}^*$

$$\frac{17^{|\alpha|} - 1}{16} \leq \mathfrak{g}(\alpha) < \frac{17^{|\alpha|+1} - 1}{16}.$$

Hence,

- $\mathfrak{gn}_3(\underline{n}) = \mathfrak{g}(\underline{n}) \geq \frac{17^{n+1}-1}{16}$ and
- $\mathfrak{gn}_3(M \underline{n} \underline{n}) = \mathfrak{g}(M \delta \underline{n}) < \frac{17^{n+4}-1}{16}$.

But then, for all $n > 4$,

$$\begin{aligned} \mathfrak{gn}_3(\underline{n}) \cdot \mathfrak{gn}_3(\underline{n}) &\geq \left(\frac{17^{n+1} - 1}{16}\right)^2 \\ &\geq \left(\frac{17^n}{16}\right)^2 \\ &\geq \frac{17^{2n-1}}{16} \\ &\geq \frac{17^{n+4} - 1}{16} \\ &> \mathfrak{gn}_3(M \underline{n} \underline{n}). \end{aligned} \quad \square$$

We note that the computation in the above proof does not depend on the specific size of the alphabet, here represented by the numbers 17 and $16 = 17 - 1$. It works as long as the alphabet has at least two symbols. So, certainly, it works for all extensions for \mathcal{L}^0 .

REMARK 8.11. *The present example of a non-regular Gödel numbering touches lightly upon a theme that we do not discuss in this paper: Gödel numberings based on alternative notions of syntax. Our treatment is mainly inspired by string approaches and algebraic approaches. In our example we look at sharing in the narrowest possible sense, but there are more principled and more encompassing approaches to sharing involving directed acyclic graphs and the like.*

Similarly, there are different approaches to variables like de Bruijn notations and the Peirce–Quine linking notation. In Peirce diagrams and in the clause-set notation, we abstract away from the order of finite conjunctions. Etcetera.

We think that gn_3 is a perfectly reasonable Gödel numbering. We strengthen this intuition by showing that gn_3 has good properties. Of course, even with the insights below, one may still have doubts about the naturality of gn_3 . Taking for a moment the standpoint of such a skeptic, one could say that we show how *not* to argue against gn_3 .

THEOREM 8.12. *gn_3 is monotonic. Moreover, it is strongly monotonic w.r.t. standard and efficient numerals.*

Proof. We note that c preserves direct sub-expressions, in other words, if τ is a direct sub-expression of ρ , then $c(\tau)$ is a direct sub-expression of $c(\rho)$. It follows that c preserves sub-expressions. Since g is monotonic, we may conclude that $gn_3 = g \circ c$ is also monotonic. Moreover, it is easy to see that c transforms standard numerals and efficient numerals identically. Hence, gn_3 inherits strong monotonicity both with respect to standard numerals and to efficient numerals from g . □

REMARK 8.13. *We can easily adapt the numerals introduced in Section 6.2 to the present situation in order to obtain numerals for which strong monotonicity fails w.r.t. gn_3 . The numeral function on \mathcal{L}^0 would look as follows.*

- i. $pd^*(0) := 0$ and $pd^*(1) := S0$.
- ii. Suppose $n > 1$, and n is not a power of a prime. Let p be the smallest prime that divides n . Say $n = p^i \cdot m$, where p does not divide m and $i > 0$. Then $pd^*(n) := \times pd^*(p^i)pd^*(m)$.
- iii. Suppose $n = p^i$, where p is prime and $i > 0$. We take $pd^*(p) := \underline{p}$. In case $i = 2k$, for $k > 0$, we set $pd^*(n) := \times pd^*(p^k)pd^*(p^k)$. In case $i = 2k + 1$, for $k > 0$, we set $pd^*(n) := \times \underline{p} \times pd^*(p^k)pd^*(p^k)$.

It is easy to see that $pd^\circ := c \circ pd^$ behaves like pd from Example 6.2 with $M\delta$ in the role of Q . It follows that $gn_3(2^{2^n})$ grows exponentially in n .*

8.2.3. Smash and exponentiation. Let $\#$ be a binary function symbol with Nelson’s [20] smash function, given by $(m, n) \mapsto 2^{|m| \cdot |n|}$, as its intended interpretation. Let \exp be a unary function symbol for exponentiation. We set $\mathcal{L}^\# := \mathcal{L}^0 \cup \{\#\}$ and $\mathcal{L}^{\exp} := \mathcal{L}^0 \cup \{\exp\}$. We now show that there are standard numberings found in the literature which are neither $\mathcal{L}^\#$ -regular nor \mathcal{L}^{\exp} -regular.

Let ξ be a numbering such that there exists a polynomial P with

$$\xi(A) < 2^{P(|A|)} \text{ for all } A. \tag{*}$$

For instance, [2, 3] contain numberings of $\mathcal{L}^\#$ with this property. Moreover, length-first numberings, such as our numbering g , satisfy $(*)$. To see this, let g be a length-first

numbering over an alphabet with size $N > 1$. Consider any $m \in \omega$ such that $N \leq 2^m$. By Lemma 2.3 we find, for all A :

$$g(A) < \frac{N^{|A|+1} - 1}{N - 1} < N^{|A|+1} \leq (2^m)^{|A|+1} = 2^{m \cdot |A| + m}.$$

We now show that any numbering ζ satisfying (*) is neither $\mathcal{L}^\#$ -regular, nor \mathcal{L}^{exp} -regular.

THEOREM 8.14. ζ is not $\mathcal{L}^\#$ -regular.

Proof. We define:

- $\widehat{1} := (\text{SS0} \# \text{SS0})$ and
- $\widehat{n + 1} := (\text{SS0} \# \widehat{n})$.

It is easy to check that $\text{ev}(\widehat{n}) = 2^{2^{n+1}}$ and $|\widehat{n}| = 6n + 3$. Since P is a polynomial there exists $n \in \omega$ such that $P(6n + 3) < 2^{n+1}$ and hence

$$\zeta(\widehat{n}) < 2^{P(6n+3)} < 2^{2^{n+1}} = \text{ev}(\widehat{n}).$$

Thus by, Lemma 8.5, ζ is not $\mathcal{L}^\#$ -regular. □

THEOREM 8.15. ζ is not \mathcal{L}^{exp} -regular.

Proof. We define:

- $\widehat{0} := \text{SS0}$ and
- $\widehat{n + 1} := \text{exp}(\widehat{n})$.

We then have $\text{ev}(\widehat{n}) = 2 \uparrow (n + 1)$ and $|\widehat{n}| = 3n + 3$.¹⁶ Since P is a polynomial there exists $n \in \omega$ such that

$$\zeta(\widehat{n}) < 2^{P(3n+3)} < 2 \uparrow (n + 1) = \text{ev}(\widehat{n}).$$

Thus, by Lemma 8.5, ζ is not \mathcal{L}^{exp} -regular. □

§9. Type-free Truth Theories. Building on work by Heck [13] and Schindler [25], we now apply the results obtained in this paper to the study of type-free axiomatic truth theories. More specifically, we show that the consistency of certain type-free truth theories depends on the employed formalisation choices. Consider for instance the following two important principles of truth:¹⁷

- NOT: The negation of a sentence A is true iff A is not true;
- DISQ: A sentence of the form $\ulcorner t$ is true \urcorner is true iff t denotes a true sentence.

Taken together, these principles can intuitively be shown to be inconsistent. In order to show this, consider the Liar sentence

(L): L is not true.

¹⁶ We use here Knuth’s up-arrow notation.
¹⁷ For the philosophical significance of these principles see [12].

Since L is the sentence “ L is not true”, L is true iff “ L is not true” is true. By (NOT), “ L is not true” is true iff “ L is true” is not true. By (DISQ), “ L is true” is not true iff L is not true. Hence, L is true iff L is not true. Contradiction (see [13, p. 12]).

Using a similar argument, we can intuitively also show the inconsistency of the following truth principle:

NDISQ: A sentence of the form $\ulcorner t$ is not true \urcorner is true iff t denotes a sentence which is not true.

Following Heck [13] and Schindler [25] we ask in which arithmetical framework this informal reasoning can be captured. Let \mathcal{L} be any primitive recursive extension of \mathcal{L}^0 , i.e., the result of adding function symbols for certain primitive recursive functions to \mathcal{L}^0 . We set $\mathcal{L}_\top := \mathcal{L} \cup \{\top\}$, where \top is a new unary predicate symbol. The theory $R_{\mathcal{L}}$ serves as a (weak) syntax theory in the formalisation of informal truth principles. Let moreover ν be a numeral function for \mathcal{L} and let ξ be a numbering of \mathcal{L}_\top . We now consider the following arithmetical formalisations of (NOT), (DISQ) and (NDISQ):

$$\begin{aligned} \text{Not}(\nu, \xi) &: \top(\nu(\xi(\neg A))) \leftrightarrow \neg \top(\nu(\xi(A))); \\ \text{Disq}(\nu, \xi) &: \top(\nu(\xi(\top(t)))) \leftrightarrow \top(t); \\ \text{Disq}^*(\nu, \xi) &: \top(\nu(\xi(\top(\nu(\xi(A))))) \leftrightarrow \top(\nu(\xi(A))); \\ \text{NDisq}(\nu, \xi) &: \top(\nu(\xi(\neg \top(t)))) \leftrightarrow \neg \top(t); \\ \text{NDisq}^*(\nu, \xi) &: \top(\nu(\xi(\neg \top(\nu(\xi(A))))) \leftrightarrow \neg \top(\nu(\xi(A))); \end{aligned}$$

where A is any \mathcal{L}_\top -sentence and t is any closed \mathcal{L} -term. Roughly speaking, $\text{Disq}(\nu, \xi)$ formalises (DISQ) by using all singular terms (of the given language) as names of sentences, while the formalisation $\text{Disq}^*(\nu, \xi)$ only employs certain canonical names, namely, ν -numerals. The same holds for $\text{NDisq}(\nu, \xi)$ and $\text{NDisq}^*(\nu, \xi)$. For every formalisation choice $\langle \mathcal{L}_\top, \nu, \xi \rangle$ we consider the formal truth theories

$$\begin{aligned} \mathcal{S}(\mathcal{L}_\top, \nu, \xi) &:= R_{\mathcal{L}_\top} + \text{Not}(\nu, \xi) + \text{Disq}(\nu, \xi); \\ \mathcal{S}^*(\mathcal{L}_\top, \nu, \xi) &:= R_{\mathcal{L}_\top} + \text{Not}(\nu, \xi) + \text{Disq}^*(\nu, \xi); \\ \mathcal{T}(\mathcal{L}_\top, \nu, \xi) &:= R_{\mathcal{L}_\top} + \text{NDisq}(\nu, \xi); \\ \mathcal{T}^*(\mathcal{L}_\top, \nu, \xi) &:= R_{\mathcal{L}_\top} + \text{NDisq}^*(\nu, \xi). \end{aligned}$$

In general, different formalisation choices $\langle \mathcal{L}_\top, \nu, \xi \rangle$ yield different theories.

We now address the question under which constraints on the underlying formalisation choices $\langle \mathcal{L}_\top, \nu, \xi \rangle$ the \mathcal{L}_\top -theories $\mathcal{S}(\mathcal{L}_\top, \nu, \xi)$, $\mathcal{S}^*(\mathcal{L}_\top, \nu, \xi)$, $\mathcal{T}(\mathcal{L}_\top, \nu, \xi)$ and $\mathcal{T}^*(\mathcal{L}_\top, \nu, \xi)$ are inconsistent. That is, under which conditions do these theories provide a faithful formalisation of the intuitive reasoning regarding (NOT) and (DISQ) as well as (NDISQ)?

The reader is reminded that \mathcal{L}^+ is a language which results from adding function symbols for primitive recursive functions to \mathcal{L}^0 , such that \mathcal{L}^+ contains a term which represents the canonical strong diagonal function. Building on work by Heck and Schindler, the following answer can be extracted from the results of this paper.

THEOREM 9.1. *Let ν be given such that $\nu = \cdot$ or $\nu = \bar{\cdot}$. Let γ be an \mathfrak{E} -adequate and strongly monotonic numbering of \mathcal{L}_\top^+ for efficient numerals, which is \mathcal{L}^0 -regular.*

We then have

- | | |
|--|--|
| 1. $\mathcal{S}(\mathcal{L}_T^0, v, \text{gn}_2) \vdash \perp$ | 5. $\mathcal{T}(\mathcal{L}_T^0, v, \text{gn}_2) \vdash \perp$ |
| 2. $\mathcal{S}^*(\mathcal{L}_T^+, \bar{\cdot}, \text{gn}_0) \vdash \perp$ | 6. $\mathcal{T}^*(\mathcal{L}_T^+, \bar{\cdot}, \text{gn}_0) \vdash \perp$ |
| 3. $\mathcal{S}(\mathcal{L}_T^0, v, \gamma) \not\vdash \perp$ | 7. $\mathcal{T}(\mathcal{L}_T^0, v, \gamma) \not\vdash \perp$ |
| 4. $\mathcal{S}^*(\mathcal{L}_T^+, v, \gamma) \not\vdash \perp$ | 8. $\mathcal{T}^*(\mathcal{L}_T^+, v, \gamma) \not\vdash \perp$ |

Proof. (3) follows from [13, theorem 1] and (7) is Proposition 2.2.6 in [25]. (4) and (8) can be shown similarly, using the observation that v -numerals are \mathcal{L}^0 -terms. See also [13, p. 13].

We prove (2). By Lemma 4.6 there exists $n \in \omega$ such that $(\text{Eq}) n = \text{gn}_0(\neg T(\bar{n}))$. We can then derive in $\mathcal{S}^*(\mathcal{L}_T^+, \bar{\cdot}, \text{gn}_0)$ the following contradiction

$$\begin{aligned} T(\bar{n}) &\leftrightarrow T(\overline{\text{gn}_0(\neg T(\bar{n}))}), \text{ using (Eq)} \\ &\leftrightarrow \neg T(\overline{\text{gn}_0(T(\bar{n}))}), \text{ using Not}(\bar{\cdot}, \text{gn}_0) \\ &\leftrightarrow \neg T(\overline{\text{gn}_0(T(\overline{\text{gn}_0(\neg T(\bar{n}))))}), \text{ using (Eq)} \\ &\leftrightarrow \neg T(\overline{\text{gn}_0(\neg T(\bar{n}))}), \text{ using Disq}^*(\bar{\cdot}, \text{gn}_0) \\ &\leftrightarrow \neg T(\bar{n}), \text{ using (Eq)}. \end{aligned}$$

The proof of (6) proceeds similarly. (1) and (5) can be shown by applying Lemma 6.3. □

The results of the theorem can be summarised in the following table:

Constraints on numberings	consistent	inconsistent
$\mathfrak{E}\mathfrak{I}$ -adequate and strongly monotonic	$\mathcal{S}(\mathcal{L}_T^0, v, \gamma)$ $\mathcal{T}(\mathcal{L}_T^0, v, \gamma)$	$\mathcal{S}(\mathcal{L}_T^0, v, \text{gn}_2)$ $\mathcal{T}(\mathcal{L}_T^0, v, \text{gn}_2)$
$\mathfrak{E}\mathfrak{I}$ -adequate and monotonic	$\mathcal{S}^*(\mathcal{L}_T^+, v, \gamma)$ $\mathcal{T}^*(\mathcal{L}_T^+, v, \gamma)$	$\mathcal{S}^*(\mathcal{L}_T^+, \bar{\cdot}, \text{gn}_0)$ $\mathcal{T}^*(\mathcal{L}_T^+, \bar{\cdot}, \text{gn}_0)$

Thus, the conditions of $\mathfrak{E}\mathfrak{I}$ -adequacy and monotonicity taken together are not restrictive enough to determine the consistency of $\mathcal{S}^*(\mathcal{L}_T^+, \bar{\cdot}, \xi)$ and of $\mathcal{T}^*(\mathcal{L}_T^+, \bar{\cdot}, \xi)$. Moreover, the conditions of $\mathfrak{E}\mathfrak{I}$ -adequacy and strong monotonicity for efficient numerals are together not restrictive enough to determine the consistency of $\mathcal{S}(\mathcal{L}_T^0, v, \xi)$ and of $\mathcal{T}(\mathcal{L}_T^0, v, \xi)$.

§10. Summary and Perspectives. In our paper we have studied various requirements on formalisation choices, such as the Gödel numbering, under which m -self-reference is attainable in arithmetic. In particular, we have examined the admissibility of self-referential numberings, which provide immediate means to formalise m -self-reference. We have shown that monotonicity and $\mathfrak{E}\mathfrak{I}$ -adequacy taken together do not exclude the self-referentiality of a numbering. While the constraint of strong monotonicity excludes the self-referentiality of numberings, we have shown that it does not block the attainability of m -self-reference in the basic language \mathcal{L}^0 . As a counterpoint, we have shown that the demand that the numbering is regular does indeed render m -self-reference in \mathcal{L}^0 unattainable, but that, on the other hand, some completely decent

numberings are non-regular. The obtained results show that the attainability of m-self-reference in \mathcal{L}^0 is more sensitive to the underlying formalisation choices than widely believed. Finally, we have shown that this sensitivity also impinges on the formal study of certain principles of self-referential truth. Namely, whether or not certain axiomatic theories of self-referential truth are consistent is not determined by the constraints of $\mathfrak{E}\mathfrak{I}$ -adequacy and (strong) monotonicity alone.

In the paper we formulated some open questions that we repeat here.

- A. Are there $\mathfrak{P}\mathfrak{O}$ -adequate numberings ξ for which $\xi \circ (\cdot)$ is p-time?
- B. Can we find $\mathfrak{P}\mathfrak{O}$ -adequate self-referential Gödel numberings that are monotonic?
- C. Can we find $\mathfrak{P}\mathfrak{O}$ -adequate Gödel numberings which satisfy the Strong Diagonal Lemma 6.3 that are strongly monotonic?
- D. Is the numbering \mathfrak{h} (defined in Remark 8.8) \mathcal{L}^0 -regular?

Our paper is a primarily technical contribution providing *data for philosophy*. We have only gestured at philosophical motivations of the various constraints. We do believe that the motivations given in the literature for various constraints are extremely thin. However, we submit that this does not undermine the value of our contribution too much, since our work will still provide a philosopher who sets out to carefully argue for a constraint with an impression of what such a constraint does or does not mean. Moreover, both the self-referential numberings themselves and the study of various constraints on numberings have some technical interest entirely independent of the philosophical motivation of various constraints.

Acknowledgements. We thank Arnon Avron, Volker Halbach, Joel D. Hamkins, Fedor Pakhomov and two referees for valuable comments and suggestions. We would also like to thank the participants of Tel Aviv University’s *Logic Seminar*, of Carl Posy’s research seminar and of the workshop *Ouroboros, Formal Criteria of Self-Reference in Mathematics and Philosophy*. Balthasar Grabmayr’s work was partially supported by the Israeli Council for Higher Education (CHE).

§A. Appendix: An Alternative Construction of a Monotonic Self-Referential Numbering for Strings. We provide an alternative construction of a coding for strings which is self-referential for efficient numerals and monotonic with respect to the sub-expression relation. We believe that the inclusion of this construction is instructive, since in some sense it proceeds in a “bottom-up” and more direct fashion than the constructions given in Sections 4 and 5.

A.1. The length-first ordering and gn_4 . Let \mathcal{A} be the alphabet of the arithmetical language \mathcal{L}^0 introduced in Subsection 2.1 and let \mathcal{A}^+ be \mathcal{A} extended with a fresh constant c and a fresh separator-symbol $;$. We remind the reader of the fact that \mathcal{A} has 17 letters. Thus, \mathcal{A}^+ has 19 letters. We suppose some ordering of \mathcal{A}^+ is given. Let gn_4 be the length-first ordering of strings of \mathcal{A}^+ . Let α_n denote the string of \mathcal{A}^+ with gn_4 -code n . We write $|\alpha|$ for the length of α as before.

We remind the reader that, by Lemma 2.3, we have

$$\frac{19^{|\alpha|} - 1}{18} \leq gn_4(\alpha) < \frac{19^{|\alpha|+1} - 1}{18}.$$

It follows that $|\alpha| \leq gn_4(\alpha)$. Moreover, whenever $|\alpha| < |\beta|$, then $gn_4(\alpha) < gn_4(\beta)$.

We define \tilde{n} as before and remind the reader that $\text{ev}(\tilde{n}) = 2^n - 1$ and $|\tilde{n}| = 7n + 1$ (see Lemma 3.6).

A.2. On gn_5 . We say that α is *acceptable* if it is of the form $\gamma;\delta$, where γ and δ are $;$ -free and where γ is a sub-string of δ . (Here we allow $\gamma = \delta$.)

Suppose $\alpha_n = \gamma;\delta$ is acceptable. We take $\beta_n := \gamma[c := 2^{\text{gn}_4(\delta;\delta)}]$. In other words, when $\delta;\delta = \alpha_k$, then $\beta_n := \gamma[c := \tilde{S}k]$. In all other cases we set β_n to a *don't care* value, say ε , the empty string.

LEMMA A.1. $(\beta_n)_{n \in \omega}$ enumerates the \mathcal{A} -strings.

Proof. Clearly the β_n neither contain $;$ nor c , so they are \mathcal{A} -strings.

Conversely, suppose θ is an \mathcal{A} -string. Let $\alpha_k = \theta;\theta$. Then, clearly, $\beta_k = \theta$. □

We note that the enumeration $(\beta_n)_{n \in \omega}$ has repetitions. We proceed with a technical lemma.

LEMMA A.2. *Suppose $\gamma;\delta$ is adequate and $\text{gn}_4(\delta;\delta) = n$. Then, $|\tilde{n}| > |\delta;\delta| > |\gamma|$, and, hence, $S\tilde{n}$ does not occur in γ .*

Proof. We have:

$$\begin{aligned} |\gamma| &< |\delta;\delta| \\ &\leq \frac{19^{|\delta;\delta|} - 1}{18} \\ &\leq n \\ &< 7n + 1 \\ &= |\tilde{n}|. \end{aligned} \quad \square$$

LEMMA A.3. *Consider acceptable $\zeta := \gamma;\delta$ and $\theta := \mu;\eta$. Let $\text{gn}_4(\delta;\delta) = n$ and $\text{gn}_4(\eta;\eta) = m$. Suppose that c occurs both in γ and in μ and $v := \gamma[c := S\tilde{n}] = \mu[c := S\tilde{m}]$. Then, $\zeta = \theta$.*

Proof. We assume the conditions of the lemma. Note that n and m cannot be 0. In case $n = m$, we see, by Lemmas 5.3 and A.2, that both γ and ε are the result of replacing $S\tilde{n}$ in v by c . It follows that $\gamma = \mu$. Trivially, $\delta = \eta$. So, $\zeta = \theta$.

Suppose $m \neq n$. This contradicts the fact that the largest term of the form $\tilde{S}k$ in $v = \gamma[c := S\tilde{n}]$ is $S\tilde{n}$ and that the largest term of the form $\tilde{S}k$ in $v = \mu[c := S\tilde{m}]$ is $S\tilde{m}$. □

LEMMA A.4. *Consider any \mathcal{A} -string $\theta \neq \varepsilon$. Let $\alpha_m = \theta;\theta$. We have $\beta_m = \theta$. Suppose $\alpha_n = \gamma;\delta$ is acceptable and c occurs in γ and $\beta_n = \theta$. Then $n < m$.*

Proof. The first claim is trivial. Suppose the conditions of the second claim. Let $\alpha_k = \delta;\delta$. We have, by Lemma A.2, that $|\theta;\theta| > |\theta| \geq |\tilde{S}k| > |\delta;\delta| \geq |\gamma;\delta|$. □

Consider any \mathcal{A} -string ζ . Let us say that ζ is *essential* if, for some n , we have $\zeta = \beta_n$ and $\alpha_n = \gamma;\delta$, where c occurs in γ . Otherwise, we call ζ *inessential*.

The combination of Lemmas A.3 and A.4 tells us that the enumeration of the β_n looks as follows. Suppose ζ is non-empty. If ζ is inessential it will be enumerated first as $\beta_{\text{gn}_4(\zeta;\zeta)}$. All occurrences of ζ in the enumeration will have the form $\beta_{\text{gn}_4(\zeta;\eta)}$, where ζ is a sub-string of η . If ζ is essential, its first occurrence will be $\beta_{\text{gn}_4(\gamma;\delta)}$, where γ, δ is adequate and c occurs in γ and $\zeta = \gamma[c := \widetilde{S(\text{gn}_4(\delta;\delta))}]$. After that the enumeration mirrors the inessential case.

REMARK A.5. We note that we could define a many-valued Gödel numbering $gn^*(\zeta) := \{n \mid \zeta = \beta_n\}$. There are no fundamental objections to many-valued Gödel numberings as long as they are total, in the sense that to each string a non-empty set of values is assigned, and injective, in the sense that the sets of values assigned to two different strings are disjoint. However, since in our framework, we opted for the more conventional choice of functional Gödel numberings, we will not consider a Gödel numbering like gn^* .

We define, for an \mathcal{A} -string θ : $gn_5(\theta)$ is the smallest n such that $\beta_n = \theta$.

We could compute gn_5 as follows. Let an \mathcal{A} -string ζ be given.

Is ζ the empty string? If *yes*, $gn_5(\zeta) := 0$. If *no*, does ζ have a sub-string of the form $S\tilde{k}$? If *no*, set $gn_5(\zeta) := gn_4(\zeta; \zeta)$. If *yes*, find the largest such k , say it is n . Let γ be the result of replacing all occurrences of $S\tilde{n}$ in ζ by c . Is α_n of the form $\delta; \delta$, where $\gamma; \delta$ is adequate? If *no*, set $gn_5(\zeta) := gn_4(\zeta; \zeta)$. If *yes*, put $gn_5(\zeta) := gn_4(\gamma; \delta)$.

The Gödel numbering gn_5 is not monotonic with respect to the sub-string ordering and \leq . Suppose, e.g., $n = gn_4(\delta; \delta)$. Clearly, \tilde{n} is inessential, since it does not contain any sub-string of the form $S\tilde{k}$. We remind the reader that $|\tilde{n}| = 7n + 1$. So, we have:

$$gn_5(\tilde{n}) = gn_4(\tilde{n}; \tilde{n}) \geq \frac{19^{14n+3} - 1}{18} > n.$$

$$gn_5(S\tilde{n}) = gn_4(c; \delta) < gn_4(\delta; \delta) = n.$$

So, $gn_5(S\tilde{n}) < gn_5(\tilde{n})$.

We will prove two lemmas about classes of cases where gn_5 is monotonic.

LEMMA A.6. Suppose ζ is inessential and θ is a sub-string of ζ . Then, $gn_5(\theta) \leq gn_5(\zeta)$.

Proof. Suppose ζ is inessential. Then, $gn_5(\theta) \leq gn_4(\theta; \theta) \leq gn_4(\zeta; \zeta) = gn_5(\zeta)$. \square

Let \mathcal{W} be the set of well-formed expressions of the arithmetical language. We partition \mathcal{W} in two classes \mathcal{W}_0 and \mathcal{W}_1 . Here:

- $\mathcal{W}_0 := \{SS0\} \cup \{\tilde{n} \mid n \in \omega\} \cup \{(SS0 \times \tilde{n}) \mid n \in \omega\}$
- $\mathcal{W}_1 := \mathcal{W} \setminus \mathcal{W}_0$

LEMMA A.7. Suppose $\theta \in \mathcal{W}_1$ and $\zeta \in \mathcal{W}$. Suppose further that θ is a sub-expression of ζ . Then, $gn_5(\theta) \leq gn_5(\zeta)$.

Proof. Suppose $gn_5(\zeta) = gn_4(\gamma; \delta)$. Let $n := gn_4(\delta; \delta)$ and let θ_0 be the result of replacing all occurrences of $S\tilde{n}$ in θ by c . (We include the case here where there are no such occurrences.)

We claim that θ_0 is a sub-expression of γ . To see this consider the parse-tree of ζ . Since θ is in \mathcal{W}_1 , it cannot occur strictly below a node labeled $S\tilde{n}$. So, relabeling the nodes labeled $S\tilde{n}$ with c and removing all nodes below them in the parse-tree of ζ will result in occurrences of θ_0 at the places where we originally had occurrences of θ .

It follows that $gn_5(\theta) \leq gn_4(\theta_0; \delta) \leq gn_4(\gamma; \delta) = gn_5(\zeta)$. \square

A.3. On gn_6 . Finally, we are ready and set to define the desired Gödel numbering gn_6 .

- $gn_6(SS0) = 2$

- $gn_6(\tilde{n}) := 4n + 1$
- $gn_6((SS0 \times \tilde{n})) := 4n + 3$
- $gn_6(\alpha) := 2^{gn_5(\alpha)}$, if α is in \mathscr{W}_1 .

We have:

LEMMA A.8. gn_6 is injective on \mathscr{W} .

Proof. Suppose α is well-formed. The only case where we could go wrong is where $gn_5(\alpha)$ is 0 or 1. Since α is well-formed, it is not the empty string. So, for some adequate $\gamma; \delta$, where γ is well-formed, we have $gn_5(\alpha) = gn_4(\gamma; \delta)$. We note that $|\gamma; \delta| \geq 3$, so $gn_5(\alpha) = gn_4(\gamma; \delta) \geq \frac{19^3-1}{18} > 1$. □

So gn_6 is indeed a Gödel numbering.

LEMMA A.9. Suppose α is in \mathscr{W}_0 . Then, $gn_6(\alpha) \leq |\alpha| \leq gn_4(\alpha) \leq gn_5(\alpha)$.

Proof. Suppose α is in \mathscr{W}_0 . We note that, for no k , we have $S\tilde{k}$ is a sub-string of α . So, $gn_5(\alpha) = gn_4(\alpha; \alpha)$. It follows that $|\alpha| \leq gn_4(\alpha) \leq gn_4(\alpha; \alpha) = gn_5(\alpha)$. Moreover, we have:

$$\begin{aligned} gn_6(SS0) &= 2 < |SS0| \\ gn_6(\tilde{n}) &= 4n + 1 \leq 7n + 1 = |\tilde{n}| \\ gn_6((SS0 \times \tilde{n})) &= 4n + 3 < 7n + 6 = |(SS0 \times \tilde{n})|. \end{aligned} \quad \square$$

LEMMA A.10. Suppose $S\tilde{n}$ is essential. Then, $gn_6(\tilde{n}) < gn_6(S\tilde{n})$

Proof. Suppose $S\tilde{n}$ is essential. Since $S\tilde{n}$ cannot have a proper subterm of the form $S\tilde{m}$, we find that $gn_5(S\tilde{n}) = gn_4(c; \delta)$ for some δ that contains c . Moreover, $n = gn_4(\delta; \delta)$. We have:

$$\begin{aligned} gn_6(\tilde{n}) &= 4n + 1 \\ &< 4 \frac{19^{2|\delta|+2} - 1}{18} + 1 \\ &< 2 \frac{19^{|\delta|+2} - 1}{18} \\ &\leq 2^n. \end{aligned} \quad \square$$

LEMMA A.11. gn_6 restricted to \mathscr{W} is monotonic w.r.t. the sub-expression ordering and $<$.

Proof. Suppose θ is a sub-expression of ζ .

Case 1: Suppose $\theta \in \mathscr{W}_1$. Then, we are immediately done by Lemma A.7.

Case 2: Suppose $\theta \in \mathscr{W}_0$.

Case 2.1: Suppose $\zeta \in \mathscr{W}_0$. Then, we are done, by inspection of the definition of gn_6 on \mathscr{W}_0 .

Case 2.2: Suppose $\zeta \in \mathscr{W}_1$.

Case 2.2.1: Suppose ζ is inessential. Then, by Lemma A.9, $gn_6(\theta) \leq gn_5(\theta)$ and, by Lemma A.6, $gn_5(\theta) \leq gn_5(\zeta)$. Hence,

$$gn_6(\theta) \leq gn_5(\theta) \leq gn_5(\zeta) < 2^{gn_5(\zeta)} = gn_6(\zeta).$$

Case 2.2.2: Suppose ζ is essential. It follows that $gn_5(\zeta) = gn_4(\gamma; \delta)$, where c occurs in γ . Let $n := gn_4(\delta; \delta)$. By Lemma A.2, we see that $|\tilde{n}| > |\gamma|$. It follows that any subterm of γ that is in \mathscr{W}_0 must be a subterm of \tilde{n} . Hence, any subterm of ζ that is in \mathscr{W}_0 must be in \tilde{n} . So, by Case 2.1, $gn_6(\theta) \leq gn_6(\tilde{n})$. Also, clearly, $S\tilde{n}$ is essential, so, by Lemma A.10, $gn_6(\tilde{n}) < gn_6(S\tilde{n})$. Finally, by Case 1: $gn_6(S\tilde{n}) \leq gn_6(\zeta)$. So, $gn_6(\theta) \leq gn_6(\zeta)$. □

We end with the obvious insight that gn_6 is self-referential.

LEMMA A.12. Consider any arithmetical formula $A(x)$ in which there is a free occurrence of x . Let $n = \text{gn}_4(A(c); A(c))$ and let $k = 2^n$. Then, $k = \text{gn}_6(A(\bar{k}))$.

Proof. We have $\text{gn}_5(A(\bar{k})) = \text{gn}_4(A(c); A(c))$ by the definition of gn_5 . By the definition of gn_6 , we get $\text{gn}_6(A(\bar{k})) = 2^{\text{gn}_5(A(\bar{k}))} = 2^n = k$. \square

We may conclude:

THEOREM A.13. The function gn_6 is a monotonic Gödel numbering that is self-referential for dyadic numerals.

We note that our Gödel numbering is elementary and the tracking functions for the connectives also will be elementary.

BIBLIOGRAPHY

- [1] Auerbach, D. (1994). Saying it with numerals. *Notre Dame Journal of Formal Logic*, **35**(1), 130–146.
- [2] Buss, S. (1986). *Bounded Arithmetic*. Napoli: Bibliopolis.
- [3] ———. editor. (1998). *Handbook of Proof Theory*. Amsterdam: North-Holland.
- [4] Feferman, S. (1960). Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, **49**, 35–92.
- [5] ———. (1984). Toward useful type-free theories I. *The Journal of Symbolic Logic*, **49**(1), 75–111.
- [6] Grabmayr, B. (2021). On the invariance of Gödel’s second theorem with regard to numberings. *The Review of Symbolic Logic*, **14**(1), 51–84.
- [7] Hájek, P., & Pudlák, P. (1993). *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Berlin: Springer.
- [8] Halbach, V. (2014). *Axiomatic Theories of Truth* (second edition). Cambridge: Cambridge University PressUK.
- [9] ———. (2018). Personal communication, March 20, 2018.
- [10] Halbach, V., & Visser, A. (2014a). Self-reference in arithmetic I. *The Review of Symbolic Logic*, **7**(4), 671–691.
- [11] ———. (2014b). Self-reference in arithmetic II. *The Review of Symbolic Logic*, **7**(4), 692–712.
- [12] Heck, R. K. (2004). Truth and disquotation. *Synthese*, **142**(3), 317–352 (originally published under the name “Richard G. Heck, Jr”).
- [13] ———. (2007). Self-reference and the languages of arithmetic. *Philosophia Mathematica*, **15**(1), 1–29 (originally published under the name “Richard G. Heck, Jr”).
- [14] Horsten, L. (2005). Canonical naming systems. *Minds and Machines*, **15**(2), 229–257.
- [15] Kripke, S. A. (1975). Outline of a theory of truth. *Journal of Philosophy*, **72**(19), 690–716.
- [16] ———. (2021). Gödel’s theorem and direct self-reference. Preprint, 2021, arXiv:2010.11979.
- [17] Leitgeb, H. (2002). What is a self-referential sentence? *Logique et Analyse*, **177**(178), 3–14.
- [18] Lewis, D. K. (1991). *Parts of Classes*. Oxford: Blackwell.

- [19] Milne, P. (2007). On Gödel sentences and what they say. *Philosophia Mathematica*, **15**(2), 193–226.
- [20] Nelson, E. (1986). *Predicative Arithmetic*. Princeton, NJ: Princeton University Press.
- [21] Picollo, L. (2018). Reference in arithmetic. *The Review of Symbolic Logic*, **11**(3), 573–603.
- [22] ———. (2020a). Alethic reference. *Journal of Philosophical Logic*, **49**, 417–438.
- [23] ———. (2020b). Reference and truth. *Journal of Philosophical Logic*, **49**, 439–474.
- [24] Pudlák, P. (1985). Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, **50**(2), 423–441.
- [25] Schindler, T. (2015). Type-Free Truth. Ph.D. Thesis, Ludwig Maximilians Universität München.
- [26] Schwichtenberg, H., & Wainer, S. S. (2012). *Proofs and Computations* (first edition). Cambridge: Cambridge University Press.
- [27] Shapiro, S. (1982). Acceptable notation. *Notre Dame Journal of Formal Logic*, **23**(1), 14–20.
- [28] Shoenfield, J. R. (1972). *Degrees of Unsolvability*. New York: American Elsevier.
- [29] Smullyan, R. M. (1984). Chameleonic languages. *Synthese*, **60**(2), 201–224.
- [30] Tarski, A., Mostowski, A., & Robinson, R. (1953). *Undecidable Theories*. Studies in Logic and the Foundations of Mathematics, Vol. 13. Amsterdam: North-Holland.
- [31] van Fraassen, B. C. (1970). Inference and self-reference. *Synthese*, **21**(3–4), 425–438.
- [32] Visser, A. (1989). Semantics and the liar paradox. In Gabbay, D., and Guenther, F., editors. *Handbook of Philosophical Logic, Topics in the Philosophy of Language*, Vol. IV. Dordrecht: Reidel, pp. 617–706.
- [33] ———. (2004). Semantics and the liar paradox. In Gabbay, D., and Guenther, F., editors. *Handbook of Philosophical Logic* (second edition), Vol. 11. Heidelberg: Springer, pp. 149–240.
- [34] ———. (2011). On the ambiguation of polish notation. *Theoretical Computer Science*, 412(28), 3404–3411. Festschrift in Honour of Jan Bergstra.
- [35] Zdanowski, K. (2020). Is there any reasonable non-regular Gödel numbering of the language of arithmetic? MathOverflow. Available from: <https://mathoverflow.net/q/361002> (version: 2020-05-22).

DEPARTMENT OF PHILOSOPHY
 HUMBOLDT UNIVERSITY OF BERLIN
 UNTER DEN LINDEN 6
 BERLIN 10099, GERMANY
 E-mail: balthasar.grabmayr@gmx.net

PHILOSOPHY, FACULTY OF HUMANITIES
 UTRECHT UNIVERSITY
 JANSKERKHOF 13
 3512BL UTRECHT, THE NETHERLANDS
 E-mail: a.visser@uu.nl