Rich Sequences and Decidability of Logical Theories

Toghrul Karimov
Max Planck Institute for Software Systems
Germany

Joris Nieuwveld
Oxford University
UK

Joël Ouaknine
Max Planck Institute for Software Systems
Germany

Abstract

We show that for a large class of integer linear recurrence sequences $(u_n)_{n\in\mathbb{N}}$, the first-order theories of $\langle \mathbb{N}; <, n \mapsto \max\{0, u_n\}\rangle$ and $\langle \mathbb{N}; +, \{u_n : n \in \mathbb{N}\} \cap \mathbb{N}\rangle$ are undecidable. Our approach is to show that $(u_n)_{n\in\mathbb{N}}$ contains, in a specific sense, all finite sequences over $\mathbb{N}$, an idea that we borrow from the proof of Hieronymi and Schulz that the first-order theory of $\langle \mathbb{N}; +, \{2^n : n \in \mathbb{N}\}, \{3^n : n \in \mathbb{N}\}\rangle$ is undecidable. In a similar way, we harness a contemporary result about quasi-randomness in the values of the Ramanujan tau function to show that the first-order theory of $\langle \mathbb{N}; <, n \mapsto |\tau(n)|\rangle$ is undecidable.

ACM Reference Format:
Toghrul Karimov, Joris Nieuwveld, and Joël Ouaknine. 2026. Rich Sequences and Decidability of Logical Theories. In . ACM, New York, NY, USA, 15 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Decidability of various logical theories connected to arithmetic has been a central topic in mathematics since the formulation of *Hilbert's program* [1] in the 1920s, arguably leading to the birth of modern computer science through the works of Turing in the 1930s [2]. Hilbert believed that every true mathematical statement must be provable in some formal system using "finitary methods" [3]. His program, however, was proven unattainable with Gödel's proofs of the *incompleteness theorems* [4], which established the following: there does not exist an algorithm (in particular, an algorithm that operates on axioms using deduction rules) that takes a first-order statement in the language of $\langle \mathbb{N}; <, +, \cdot\rangle$, and decides whether it is true. Around forty years after Gödel, an even stronger result was shown by Matiyasevich, Robinson, Davis, and Putnam [5]: it is not possible to algorithmically determine whether a given multivariate polynomial $p \in \mathbb{Z}[x_1, \ldots, x_d]$ has a zero in $\mathbb{Z}^d$, famously resolving *Hilbert's tenth problem* in the negative.

On the decidability front, Presburger showed (in 1929, using quantifier elimination) that the first-order theory of $\langle \mathbb{N}; <, +\rangle$, now called *Presburger arithmetic*, is decidable [6]. This spurred almost 100 years of fruitful research aimed at accurately delineating the boundaries of decidability for expansions of $\langle \mathbb{N}; <, +\rangle$, uncovering deep connections to computer science (automata theory [7], linear programming [8]), symbolic dynamics (numeration systems [9, Chapter 2]), number theory (rational points on varieties [10]), etc.[1] A central problem in the area, which is also the starting point of this paper, is the following. Recall that an integer *linear recurrence sequence* (LRS) is a sequence $(u_n)_{n\in\mathbb{N}}$ over $\mathbb{Z}$ such that $u_{n+d} = a_1 u_{n+d-1} + \cdots + a_d u_n$ for some $d \geq 0$ and $a_1, \ldots, a_d \in \mathbb{Z}$, the Fibonacci sequence being perhaps the most emblematic example.

Problem 1. *For which integer LRS $(u_n)_{n\in\mathbb{N}}$ is the first-order theory of $\langle \mathbb{N}; +, \{u_n : n \in \mathbb{N}\} \cap \mathbb{N}\rangle$ decidable?*

Note that in any structure that has $\mathbb{N}$ as the domain and is equipped with $+$, we can define the usual order $<$ as well as any constant $c \in \mathbb{N}$; the same does not hold for structures with domain $\mathbb{Z}$. Let us next recount the main approaches to proving decidability and undecidability of the first-order theory of an expansion of $\langle \mathbb{N}; +\rangle$, and what they tell us in terms of Problem 1.

Quantifier elimination. Semënov [12] gave a sufficient (growth-type) condition for the structures $\langle \mathbb{N}; +, P\rangle$ and $\langle \mathbb{N}; +, f(\cdot)\rangle$, where $P \subseteq \mathbb{N}$ and $f : \mathbb{N} \to \mathbb{N}$, to admit quantifier elimination, which implies decidability of the first-order theory. His criterion yields decidability for $\langle \mathbb{N}; +, n \mapsto 2^n\rangle$, $\langle \mathbb{N}; +, n \mapsto n!\rangle$, $\langle \mathbb{N}; +, n \mapsto F_n\rangle$ (where F_n is the nth Fibonacci number), and $\langle \mathbb{N}; +, n \mapsto \lfloor e^n \rfloor\rangle$. In the context of integer LRS, it yields decidability of the first-order theory of

$$\langle \mathbb{N}; +, n \mapsto \max\{0, u_n\}\rangle$$

for *irreducible* $(u_n)_{n\in\mathbb{N}}$ with a *single dominant root* ρ (which will necessarily be real); see Section 2.6 for the exact definitions.[2] Note that the structure with the function $n \mapsto \max\{0, u_n\}$ is strictly more expressive than the one with the predicate $\{u_n : n \in \mathbb{N}\} \cap \mathbb{N}$. Hence Problem 1 is decidable for the aforementioned class of integer LRS. Intuitively, the restriction that $(u_n)_{n\in\mathbb{N}}$ be irreducible ensures that $(u_n)_{n\in\mathbb{N}}$ is "atomic" and does not hide multiple different integer LRS inside it. The *number of dominant roots* of an LRS, on the other hand, is a natural measure of complexity of an LRS. Finally, we mention that a much faster and practical variant of Semënov's algorithm for quantifier elimination was recently given in [13].

Axiomatisation. Point [14] showed that for any $(u_n)_{n\in\mathbb{N}}$ as above, the first-order theory of $\langle \mathbb{N}; +, \{u_n : n \in \mathbb{N}\} \cap \mathbb{N}\rangle$ admits a countable and decidable axiomatisation. Therefore, for such structures, an alternative decision procedure can be given as follows: given a sentence φ, enumerate all possible proofs (using the countable axiomatisation and the standard rules of deduction) until either a proof of φ or a proof of $\neg\varphi$ is found.

Automaticity. For any integer LRS $(u_n)_{n\in\mathbb{N}}$ as above such that the dominant root ρ is a *Pisot number*, the predicate $\{u_n : n \in \mathbb{N}\} \cap \mathbb{N}$

[1] The first-order theory of $\langle \mathbb{N}; \cdot\rangle$, known as *Skolem arithmetic*, is also decidable. However, decidability of first-order theories of various expansions of $\langle \mathbb{N}; \cdot\rangle$ has not received much attention. We mention that $<$ is definable in $\langle \mathbb{N}; +\rangle$ but not in $\langle \mathbb{N}; \cdot\rangle$: in fact, $\langle \mathbb{N}; <, \cdot\rangle$ defines addition and hence has an undecidable first-order theory [11].
[2] Semënov's criterion provably does not apply to reducible LRS, as well as any LRS with two or more dominant roots.

Conference'17, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

and the function + can be implemented using automata over finite words that operate on greedy representations of $x \in \mathbb{N}$ in a number system constructed from ρ [9]. Well-known examples include the powers of 2 and the binary representation, as well as the Fibonacci numbers and the Zeckendorf representation [15]. In these cases, whether a given formula φ is true can be determined using tools from automata theory [16].

Definability of multiplication. Undecidability, on the other hand, has almost exclusively been proven by first showing that the structure in question defines multiplication, and then invoking undecidability of the (existential fragment of the) first-order theory of $\langle \mathbb{N}; +, \cdot \rangle$. For example, Büchi observed in 1960 that for any non-linear polynomial $p \in \mathbb{Z}[x]$ such that $p(\mathbb{N}) \cap \mathbb{N}$ is infinite, the structure

$$\langle \mathbb{N}; +, \{p(n) \colon n \in \mathbb{N}\} \cap \mathbb{N} \rangle$$

defines multiplication and hence has an undecidable first-order theory. To link this result to Problem 1, we mention that for any $p \in \mathbb{Z}[x]$ of degree at least two, the sequence $(p(n))_{n \in \mathbb{N}}$ is an integer LRS with the single *repeated* dominant root $\rho = 1$. Recently, Xiao [10] showed that $\langle \mathbb{N}; 0, 1, <, +, \{n^2 \colon n \in \mathbb{N}\} \rangle$, in fact, *existentially defines* multiplication, and hence the *existential fragment* of the first-order theory of the aforementioned structure is undecidable; this settled a long-standing open problem in arithmetic geometry ([17, Sec. 8], [18]).

There are many more results that show undecidability of an expansion of Presburger arithmetic via definability of multiplication: Bès [19] showed in 1997 that, for any *multiplicatively independent* $k, l \geq 2$, the structure $\langle \mathbb{N}; +, \{k^n \colon n \in \mathbb{N}\}, V_l(\cdot) \rangle$ (where $V_l(x)$ returns the largest power of l dividing x) defines multiplication. Woods [20] proved that, assuming *Dickson's conjecture*, the structure $\langle \mathbb{N}; +, \{n \colon n \text{ is prime}\} \rangle$ defines multiplication. For more examples in this vein, see the survey [21].

The Hieronymi-Schulz method. This paper, however, is about another powerful, yet under-explored technique for proving undecidability of first-order theories, originating in the recent work [22]. There, Hieronymi and Schulz prove that the first-order theory of

$$\mathbb{M} \coloneqq \langle \mathbb{N}; +, \{2^n \colon n \in \mathbb{N}\}, \{3^n \colon n \in \mathbb{N}\} \rangle$$

is undecidable, thus resolving a problem that had been open at least since 1992 [23].[3] At the same time Schulz [25] has shown that the structure $\mathbb{M}$ *does not* define multiplication. The approach of [22] is, therefore, radically different: the idea is to show that we can extract from $\mathbb{M}$ arbitrary finite sequences over $\mathbb{N}$, in a manner that satisfies some specific conditions that allows directly simulating Turing machines. For $x \geq 1$, define $\alpha(x)$ by $2^{\alpha(x)} \leq x < 2^{\alpha(x)+1}$, and for $x \geq 1$, x not a power of 2, let $\beta(x) = \alpha(x - 2^{\alpha(x)})$. That is, β calculates the (index of the) second-largest bit in the binary expansion of x.

Theorem 1 (Lemma 3.4 in [22]). *For every finite sequence $(t_i)_{i=1}^N$ over $\mathbb{N}$ there exist $a, b \in \mathbb{N}$ such that*

$$(\beta(3^n) - \beta(3^a))_{n=a+1}^{b-1} \cap [0, \beta(3^b) - \beta(3^a)] = (t_i)_{i=1}^N.$$

Here the $\cap$ operation takes a sequence and an interval, and returns the sequence obtained by only keeping the elements that belong to the specified interval. Theorem 1 can be interpreted as follows: there exists a function with a fixed number of inputs (two in Theorem 1, namely $3^a, 3^b \in \mathbb{N}$) that can be implemented (in a very specific sense, as we will see in Section 3) in $\mathbb{M}$ that outputs all possible finite sequences over $\mathbb{N}$. In this paper, we

- define what it means for a structure $\mathbb{M}$ to *simulate counter machines* (Section 3), which abstracts the Hieronymi-Schulz approach to proving undecidability, and
- use number-theoretic tools to show that certain classes of structures simulate counter machines and hence have undecidable first-order theories.

We focus on non-degenerate integer LRS $(u_n)_{n \in \mathbb{N}}$ with exactly two non-repeated dominant roots.[4] (In a measure-theoretic sense, almost all integer LRS are non-degenerate and have either one or two dominant roots: see, e.g., [27].) Such sequences are of the form

$$u_n = a\lambda^n + \overline{a}\,\overline{\lambda}^n + v_n$$

where λ is non-real, $|\lambda| > 1$, and $|a\lambda^n + \overline{a}\,\overline{\lambda}^n|$ grows much faster than $|v_n|$. A concrete example is $u_n = (2 + \boldsymbol{i})^n + (2 - \boldsymbol{i})^n + 2^n$, which satisfies the recurrence relation $u_{n+3} = 6u_{n+2} - 13u_{n+1} + 10u_n$.

When $(u_n)_{n \in \mathbb{N}}$ satisfying the conditions above is clear from the context, we write $U = \{u_n \colon n \in \mathbb{N}\} \cap \mathbb{N}$ (which will necessarily be infinite), $(p_n)_{n \in \mathbb{N}}$ for the ordering of U, and $u(\cdot)$ for the function $n \mapsto \max\{0, u_n\}$. Our main results are the following.

Theorem 2. *Let $(u_n)_{n \in \mathbb{N}}$ be a non-degenerate integer LRS $(u_n)_{n \in \mathbb{N}}$ with exactly two non-repeated dominant roots. Then there exists a computable $\zeta > 1$ with the following property. For all $\ell \geq 1$ and $1 < \gamma_1 < \delta_1 \leq \gamma_2 < \delta_2 \leq \cdots \leq \gamma_\ell < \delta_\ell < \zeta$, there exist infinitely many n such that*

$$\frac{p_{n+j}}{p_n} \in (\gamma_j, \delta_j)$$

for all $1 \leq j \leq \ell$.

That is, we can control the ratios of consecutive elements of U: this is the "everything that can happen, will happen" result analogous to Theorem 1 that we need. Using the framework of structures that simulate counter machines, we deduce the following from Theorem 2.

Theorem 3. *Let $(u_n)_{n \in \mathbb{N}}$ be a non-degenerate integer LRS $(u_n)_{n \in \mathbb{N}}$ with exactly two non-repeated dominant roots. The structure $\langle \mathbb{N}; +, U \rangle$ simulates counter machines and therefore has an undecidable first-order theory.*

Note that since U is definable in $\langle \mathbb{Z}; <, +, \{u_n \colon n \in \mathbb{N}\} \rangle$, the first-order theory of the latter is also undecidable. At the time of writing, decidability of the first-order theory of $\langle \mathbb{Z}; +, \{u_n \colon n \in \mathbb{N}\} \rangle$ remains open.

The quasi-randomness of the integer LRS that we consider was already hinted at in the recent paper [28], where the *decidability* of

[3] More precisely, the $\exists^*\forall^*\exists^*$ fragment of the first-order theory of $\langle \mathbb{N}; 0, 1, <, +, \{2^n \colon n \in \mathbb{N}\}, \{3^n \colon n \in \mathbb{N}\} \rangle$ is undecidable, and the existential fragment of the same theory is decidable [24]. Decidability thus remains open for the $\exists^*\forall^*$ fragment.

[4] See Section 9 for a discussion on first-order theories of integer LRS with exactly one dominant root. We also note that the following problem subsumes the *Skolem problem* [26] for integer LRS of order 5, which is currently open: given $x \in \mathbb{N}$ and an integer LRS $(u_n)_{n \in \mathbb{N}}$ with four (non-repeated) dominant roots, decide whether $x \in U$, where $U = \{u_n \colon n \in \mathbb{N}\} \cap \mathbb{N}$. That is, integer LRS with four or more dominant roots are, at the time of writing, considered *mathematically intractable*.

the *monadic second-order* (MSO) theory of $\langle \mathbb{N}; <, U\rangle$ was established via the following theorem.

Theorem 4 ([28, Theorem 4]). *Let $(u_n)_{n\in\mathbb{N}}$ be a non-degenerate integer LRS $(u_n)_{n\in\mathbb{N}}$ with exactly two non-repeated dominant roots, $m \geq 1$, and*

$$\Sigma_m = \{0 \leq r < m \colon u_n \equiv r \ (\mathrm{mod}\ m) \text{ for infinitely many } n\}.$$

Then for any $t_0, \ldots, t_\ell \in \Sigma_m$ we can compute infinitely many n such that for all $0 \leq i \leq \ell$,

$$p_{n+i} \equiv t_i \ (\mathrm{mod}\ m).$$

Even though our original motivation was to attack Problem 1, when proving Theorem 3, we discovered some other, very surprising undecidability results. The first one is about what happens if, instead of the predicate U, we have the (more expressive) function $u(\cdot)$. It turns out that, in this case, we do not even need the + operator to obtain undecidability.

Theorem 5. *Let $(u_n)_{n\in\mathbb{N}}$ be a non-degenerate integer LRS $(u_n)_{n\in\mathbb{N}}$ with exactly two non-repeated dominant roots. Then the structure $\langle \mathbb{N}; <, u(\cdot)\rangle$ simulates counter machines and therefore has an undecidable first-order theory.*

Theorem 5 is proven via the following quasi-randomness result analogous to Theorem 2.

Theorem 6. *Let $(u_n)_{n\in\mathbb{N}}$ be a non-degenerate integer LRS $(u_n)_{n\in\mathbb{N}}$ with exactly two non-repeated dominant roots. Then there exists computable $\zeta > 1$ with the following property. For every $\ell \geq 1$ and $1 < \gamma_j < \delta_j < \zeta$ with $1 \leq j \leq \ell$, there exist infinitely many $n, \widetilde{n}$ such that, writing $(n_j)_{j=0}^{k}$ for the ordering of $\{m \colon n \leq m < \widetilde{n}$ and $u_n \leq u_m < \zeta u_n\}$, we have that $u_n > 0$, $k = \ell$ and*

$$\frac{u_{n_j}}{u_n} \in (\gamma_j, \delta_j)$$

for all $1 \leq j \leq \ell$.

Motivated by theorems 3 and 5 above, we say that a sequence $(u_n)_{n\in\mathbb{N}}$ is *rich* if $\langle \mathbb{N}; +, \{u_n \colon n \in \mathbb{N}\} \cap \mathbb{N}\rangle$ or $\langle \mathbb{N}; <, n \mapsto \max\{0, u_n\}\rangle$ simulates counter machines. It turns out that quasi-randomness properties of various special sequences in mathematics have already been extensively studied, and we can deduce richness results for the corresponding structures relatively easily. In this paper, we do this for $\tau \colon \mathbb{N} \to \mathbb{Z}$, the *Ramanujan tau function*, using a recent quasi-randomness result of Bilu, Deshouillers, Gun, and Luca [29]; see Section 4 for the details.

Theorem 7. *Assuming Lehmer's conjecture, $\langle \mathbb{N}; <, |\tau(\cdot)|\rangle$ simulates counter machines and hence has an undecidable first-order theory.*

Note that, once again, we do not need the + operator to prove undecidability.

2 Preliminaries

We denote the cardinality of a set X by $\#X$. For $x \in \mathbb{R}$ and $y \in \mathbb{R}\backslash\{0\}$, let $[[x]]_y$ be the smallest distance from x to an integer multiple of y. We denote by $\boldsymbol{i}$ the imaginary number and by Log the principal branch of the complex logarithm, which satisfies $\mathrm{Log}(x + y\boldsymbol{i}) = \boldsymbol{i}\theta + \log\sqrt{x^2 + y^2}$, $\theta \in (-\pi, \pi]$ for all $x, y \in \mathbb{R}$. We write $\mathbb{T}$ for the unit circle in $\mathbb{C}$, and $\mathbb{T}_+$ for $\{z \in \mathbb{T} \colon \mathrm{Re}(z) > 0\}$. For $z_1, z_2 \in \mathbb{T}$, we write $\Delta(z_1, z_2)$ for the length of the shorter arc of $\mathbb{T}$ connecting z_1 and z_2.

2.1 Structures and their theories

A *structure* $\mathbb{M}$ consists of a domain D, constants $c_1, \ldots, c_k \in D$, predicates $P_1, \ldots, P_l$ where each $P_i \subseteq D^{\mu(i)}$ for some $\mu(i) \geq 1$, and functions $f_1, \ldots, f_m$ where each f_i has the type $f_i \colon D^{\delta(i)} \to D$ for some $\delta(i) \geq 1$. We denote such $\mathbb{M}$ by $\langle D; c_1, \ldots, c_k, P_1, \ldots, P_l, f_1, \ldots, f_m\rangle$. We do note explicitly mention = as a predicate, but assume that every structure has it. A *theory* is simply a set of sentences, i.e. first-order formulas without free variables. The (first-order) theory of a structure $\mathbb{M}$ is the set of all well-formed sentences (i.e. formulas without free variables) constructed from the symbols $c_1, \ldots, c_k, P_1, \ldots, P_l, f_1, \ldots, f_m$ as well as $\wedge, \vee, \neg, \exists, \forall$ that are true in $\mathbb{M}$. We write $\mathbb{M} \models \varphi$ to mean that φ holds in $\mathbb{M}$. A formula is *existential* if it is of the form $\exists x_1 \cdots \exists x_m \colon \varphi(x_1, \ldots, x_m)$ for φ quantifier-free. A theory $\mathcal{T}$ is *decidable* if there exists an algorithm that takes a sentence φ and decides whether $\varphi \in \mathcal{T}$ and *undecidable* otherwise.

2.2 Counter machines

A *k-counter machine* $\mathcal{M}$ consists of counters $c_1, \ldots, c_k$ taking positive integer values and instructions numbered $1, \ldots, H$ for some H. Without loss of generality, we assume that $H > 1$. The instructions are of the form INC c_i, IF $c_i > 1$ THEN DEC c_i, GOTO l, and HALT, where c_i is a counter and l is an instruction number. That is, the counters can be incremented and decremented, but they cannot go below 1. We additionally assume, without loss of generality, that the machine starts with the instruction numbered 1, and has a single HALT instruction, numbered H. The initial values of the counters are all 1. We write $\delta_M \colon \{1, \ldots, H\} \times \mathbb{N}_{>0}^k \to \{1, \ldots, H\} \times \mathbb{N}_{>0}^k$ for the (partial) transition function of $\mathcal{M}$ that describes how a configuration consisting of an instruction number $l \neq H$ and the values of the k counters is updated in one step. By the *trace* of $\mathcal{M}$ we mean the (finite or infinite) sequence

$$(s_n)_n = \langle 0, \iota_0, c_{1,0}, \ldots, c_{k,0}, 0, \iota_1, c_{1,1}, \ldots, c_{k,1}, \ldots\rangle$$

such that $s_n = 0 \Leftrightarrow n \equiv 0 \ (\mathrm{mod}\ k+2)$, $\iota_0 = 1$, $c_{j,0} = 0$ for all j, and

$$(\iota_{j+1}, c_{1,j+1}, \ldots, c_{k,j+1}) = \delta_{\mathcal{M}}(\iota_j, c_{1,j}, \ldots, c_{k,j})$$

for all j such that $\iota_j \neq H$. Note that we use 0 as a delimiter between consecutive configurations of $\mathcal{M}$. The Halting Problem asks to decide whether the execution of a given machine $\mathcal{M}$ ever reaches the HALT instruction, and is undecidable already for two-counter machines [30].

2.3 Algebraic numbers

An algebraic number α is a complex number that is a root of a polynomial $p \in \mathbb{Q}[x]$. The unique monic polynomial p of the smallest degree that has α as a root is called the *minimal polynomial* α. The set of algebraic numbers forms a field, written $\overline{\mathbb{Q}}$. An algebraic number α can be represented in computer memory, for example, by its minimal polynomial p as well as sufficiently close rational approximations to $\mathrm{Re}(\alpha)$ and $\mathrm{Im}(\alpha)$. In this representation, all usual arithmetic operations can be effectively performed on algebraic numbers [31, Chapter 4].

2.4 Baker's theorem

Baker's theorem and its p-adic analogue are among the most important mathematical tools in the study of linear recurrence sequences. Let $\Lambda = b_1 \operatorname{Log}(\alpha_1) + \cdots + b_m \operatorname{Log}(\alpha_m)$, where $b_i \in \mathbb{Z} \setminus \{0\}$, $\alpha \in \overline{\mathbb{Q}} \setminus \{0\}$. Such a Λ is called a linear form in logarithms.

Theorem 8. *Let $B \geq 3$ be such that $B > |b_i|$ for all i. There exists a computable constant $C > 0$ (that only depends on $\alpha_1, \ldots, \alpha_m$) such that*

$$|\Lambda| > B^{-C}$$

whenever $\Lambda \neq 0$.

Lemma 9. *Let $\alpha, \beta \in \mathbb{T}$, $n \in \mathbb{N}$ be such that $\alpha^n \neq \beta$. There exists a computable constant $C_0 > 0$ such that*

$$\Delta(\alpha^n, \beta) > \frac{1}{(\max\{2, n\})^{C_0}}.$$

Proof. We have that

$$\begin{aligned}\Delta(\alpha^n, \beta) &= \pi[[n \operatorname{Log}(\alpha) - \operatorname{Log}(\beta)]]_{2\pi} \\ &= \pi(n \operatorname{Log}(\alpha) - \operatorname{Log}(\beta) - k\pi) \\ &= \pi(n \operatorname{Log}(\alpha) - \operatorname{Log}(\beta) - k \operatorname{Log}(-1))\end{aligned}$$

where $-n-1 \leq k \leq n+1$. Applying Baker's theorem with $B = \max\{3, n+1\}$, there exists a constant $C > 0$ such that

$$\Delta(\alpha^n, \beta) > \frac{1}{(\max\{3, n+1\})^C}.$$

Finally, choose any $C_0 > 0$ such that $3^{-C} > 2^{-C_0}$ and $(n+1)^{-C} > n^{-C_0}$ for all $n \geq 3$. □

2.5 Linear recurrence sequences

A sequence $(u_n)_{n\in\mathbb{N}}$ over a ring R is a *linear recurrence sequence* (LRS) over R if there exist $d \geq 0$ and $a_1, \ldots, a_d \in R$ such that

$$u_{n+d} = a_1 u_{n+d-1} + \cdots + a_d u_n \quad (1)$$

for all $n \in \mathbb{N}$. The smallest such d is called the *order* of $(u_n)_{n\in\mathbb{N}}$. In this paper, we work with LRS over $\mathbb{Z}$, which we also call *integer LRS*. For example, the Fibonacci sequence satisfies $u_{n+2} = u_{n+1} + u_n$ for all $n \in \mathbb{N}$ and is an integer LRS of order two. We refer the reader to the book [32] for a detailed account of LRS.

Let $R \subseteq \overline{\mathbb{Q}}$ and $(u_n)_{n\in\mathbb{N}}$ be an LRS over R of order d. Then there exist unique $a_1, \ldots, a_d \in R$ (with $a_d \neq 0$) such that $(u_n)_{n\in\mathbb{N}}$ satisfies the recurrence relation (1). The *minimal polynomial* of $(u_n)_{n\in\mathbb{N}}$ is $p(x) = x^d - \sum_{i=1}^d c_i x^{d-i}$. Suppose p has the (distinct) roots $\lambda_1, \ldots, \lambda_m \in \overline{\mathbb{Q}}$, called the *characteristic roots* of $(u_n)_{n\in\mathbb{N}}$. Then there exist unique non-zero polynomials $q_1, \ldots, q_m \in \overline{\mathbb{Q}}[x]$ such that

$$u_n = q_1(n)\lambda_1^n + \cdots + q_m(n)\lambda_m^n \quad (2)$$

for all $n \in \mathbb{N}$. Equation (2) is known as the *exponential-polynomial form* of $(u_n)_{n\in\mathbb{N}}$. A characteristic root λ_i is called *non-repeated* (alternatively, *simple*) if q_i is constant. The sequence $(u_n)_{n\in\mathbb{N}}$ is called *diagonalisable* (alternatively, *simple*) if every λ_i is non-repeated. A characteristic root λ_i is called *dominant* if $|\lambda_i| \geq |\lambda_j|$ for all $1 \leq j \leq m$. We say that $(u_n)_{n\in\mathbb{N}}$ is *non-degenerate* if $z \coloneqq \lambda_i/\lambda_j$ is not a root of unity for all $i \neq j$, i.e., $z^k \neq 1$ for all $k \neq 0$. For any LRS $(u_n)_{n\in\mathbb{N}}$, there exists L (that is effectively computable for integer LRS) such that the subsequences $(u_{nL+r})_{n\in\mathbb{N}}$ are non-degenerate for all $0 \leq r < L$. By the Skolem-Mahler-Lech theorem, every non-zero and non-degenerate LRS over R (in fact, over any ring of characteristic zero) has finitely many zeros. We say that an integer LRS $(u_n)_{n\in\mathbb{N}}$ is *irreducible* if its characteristic polynomial is irreducible over $\mathbb{Q}$.

Decision problems of linear recurrence sequences, despite being of central interest in algebraic number theory, largely remain open. The most famous example is the *Skolem Problem*, which asks to decide whether a given integer LRS contains zero. It is known to be decidable for non-degenerate LRS with at most 3 distinct dominant roots [33].

2.6 LRS with two dominant roots

In this section, fix a non-degenerate integer LRS $(u_n)_{n\in\mathbb{N}}$ with exactly two dominant roots. Because the dominant roots are closed under complex conjugation, we either have two real dominant roots $\lambda, -\lambda$, or two non-real dominant roots $\lambda, \overline{\lambda}$. We will adopt the following notation throughout this paper: $u_n = v_n + r_n$, $v_n = a\lambda^n + \overline{a}\,\overline{\lambda}^n$ where $a \in \overline{\mathbb{Q}} \setminus \{0\}$, $\rho = |\lambda| > 0$, $\lambda = \rho\mu$, $\mu = e^{i\theta} \in \mathbb{T}$, $a = |a|\xi$, and $\xi = e^{i\varphi} \in \mathbb{T}$ We have that

$$a\lambda^n + \overline{a}\,\overline{\lambda}^n = |a|\rho^n(\xi\mu^n + \overline{\xi}\,\overline{\mu}^n) = 2|a|\rho^n\cos(n\theta + \varphi)$$

and, by the assumption on dominant roots, $|r_n| = o((\rho - \varepsilon)^n)$ for all sufficiently small $\varepsilon > 0$ (where the implied constant is effective). By non-degeneracy, θ is not a rational multiple of π. Finally, we mention that both $(v_n)_{n\in\mathbb{N}}$ and $(r_n)_{n\in\mathbb{N}}$ are themselves LRS over $\mathbb{R} \cap \overline{\mathbb{Q}}$.

By Kronecker's theorem in Diophantine approximation [34], because θ is not a rational multiple of π, we have that $(\cos(n\theta + \varphi))_{n\in\mathbb{N}}$ is dense in $[-1, 1]$; equivalently, $(\xi\mu^n)_{n\in\mathbb{N}}$ is dense in $\mathbb{T}$. We next argue that $\rho > 1$. Suppose $\rho \leq 1$. Then, as ρ is dominant, $|u_n|$ is bounded, and so as u_n takes integer values, $u_n = \cos(n\theta + \varphi)\rho^n + o((\rho - \varepsilon)^n)$ only takes finitely many values for small enough $\varepsilon > 0$. Hence, as $(\cos(n\theta + \varphi))_{n\in\mathbb{N}}$ is dense in $[-1, 1]$, $\rho < 1$ and thus $(u_n)_{n\in\mathbb{N}}$ is eventually constantly 0. Combined with the non-degeneracy assumption and the Skolem-Mahler-Lech theorem, this implies that u_n must be identically zero, which is a contradiction.

We next give a few lemmas for our class of LRS. The following two results are proven using Baker's theorem.

Lemma 10 ([35, Lemma 3])**.** *There exist computable constants $N, C > 0$ such that for all $n \geq N$, $|v_n| > \frac{\rho^n}{n^C}$.*

Theorem 11 ([33, Theorem 3]). *There exist computable constants $N, C > 0$ such that for all $n \geq N$ and $m < n$,*

$$|u_n - u_m| > \rho^n n^{-C(\log(n+1))^2}.$$

In the following lemmas, we show that, for our purposes, we can work with v_n instead of u_n, provided that n is sufficiently large.

Lemma 12. *There exists computable M with the following property. For all distinct $n_1, n_2 \in \mathbb{N}$ with $n_1 \geq M$ we have that $u_{n_1} \neq u_{n_2}$, $u_{n_1} \neq 0$, $\operatorname{sign}(u_{n_1}) = \operatorname{sign}(v_{n_1})$, and $\operatorname{sign}(u_{n_1} - u_{n_2}) = \operatorname{sign}(v_{n_1} - v_{n_2})$.*

Proof. Let N, C be as in Theorem 11. Then for all $n_1 \neq n_2$ such that at least one $n_i \geq N$, we have that $u_{n_1} \neq u_{n_2}$. Since $|r_n| < o((\rho - \varepsilon)^n)$ for all sufficiently small $\varepsilon > 0$, applying Lemma 10 we can compute $N' \geq N$ such that $|v_n| > |r_n|$ and hence $u_n \neq 0$ for all $n \geq N'$.

Consider $n_1 \geq N'$ and $n_1 \neq n_2$. Exchanging n_1 and n_2 if necessary, we can assume that $n_1 > n_2$. Then

$$v_{n_1} - v_{n_2} = u_{n_1} - u_{n_2} + (r_{n_2} - r_{n_1}). \tag{3}$$

Since $(r_n)_{n\in\mathbb{N}} = o((\rho - \varepsilon)^n)$ for all sufficiently small $\varepsilon \geq 0$, we can compute $\widetilde{N}$ and $\varepsilon > 0$ such that

$$|r_{n_2} - r_{n_1}| < (\rho - \varepsilon)^{n_1}$$

for all n_1, n_2 with $n_1 \geq \widetilde{N}$. Applying Theorem 11, for all sufficiently large n_1 and (any) $n_1 > n_2$, we have that

$$\frac{|u_{n_1} - u_{n_2}|}{|r_{n_2} - r_{n_1}|} > \frac{\rho^{n_1} n_1^{-C(\log(n_1+1))^2}}{(\rho - \varepsilon)^{n_1}} > 1$$

which implies (together with Equation (3)) that

$$\operatorname{sign}(v_{n_1} - v_{n_2}) = \operatorname{sign}(u_{n_1} - u_{n_2}) \in \{-, +\}. \qquad \square$$

Lemma 13. *For every $\varepsilon \in \mathbb{Q}_{>0}$ there exists computable M_ε such that for all $n \geq M_\varepsilon$*

$$1 - \varepsilon < \frac{|u_n|}{|v_n|} < 1 + \varepsilon.$$

Proof. We have

$$\frac{u_n}{v_n} = 1 + \frac{r_n}{v_n}.$$

It remains to observe that by Lemma 10,

$$\lim_{n\to\infty} \frac{r_n}{v_n} = 0$$

effectively. $\square$

2.7 The Hieronymi-Schulz interval stacking lemma

The following lemma plays a crucial role in the proof of Theorem 1, as well as our theorems 2 and 6.

Lemma 14 ([22, Lemma 2.1]). *Let $I \subseteq \mathbb{T}$ be an interval and $(I_n)_{n\in\mathbb{N}}$ be a sequence of intervals such that $\sum_{n=0}^{\infty} |I_n \cap I| < |I|$ and $(I_n)_{n\in\mathbb{N}}$ is dense in I. That is, for every open $J \subseteq I$ there exist infinitely many n such that I_n intersects J. Then there exist infinitely many k such that $I_k \subseteq I$ and $I_k \cap I_n = \varnothing$ for all $n < k$.*

Proof. See Appendix A. $\square$

3 How to prove undecidability

We say that a structure $\mathbb{M}$ with domain D *simulates counter machines* if there exist $l, m \geq 1$, maps $\mathsf{Seq}\colon D^l \to \mathbb{N}^*$, $\mathsf{Rep}\colon D^l \to (D^m)^*$, and formulas rep, cnst_k (where $k \in \mathbb{N}$ and we additionally require that cnst_k be effectively computable given k), succ, inc, eq in the language of $\mathbb{M}$ with $l+m$, $l+m$, $l+2m$, $l+2m$, $l+2m$ free variables, respectively, that satisfy the following.

(1) The map Seq is surjective. Intuitively, Seq is a black box with l inputs that is "implementable" in $\mathbb{M}$ that outputs all possible finite sequences over $\mathbb{N}$.
(2) For all $x \in D^l$, $|\mathsf{Seq}(x)| = |\mathsf{Rep}(x)|$, and the terms of $\mathsf{Rep}(x)$ are distinct. Intuitively, $\mathsf{Rep}(x)$ is a finite sequence of distinct terms that index the sequence $\mathsf{Seq}(x)$.
3) For all $x \in D^l$ and $y \in D^m$, $\mathbb{M} \models \mathsf{rep}(x, y)$ if and only if $y = \mathsf{Rep}(x)_i$ for some i. That is, y appears in $\mathsf{Rep}(x)$; equivalently, y is the index of some term in $\mathsf{Seq}(x)$.
4) For all $k \in \mathbb{N}$, $x \in D^l$ and $y \in D^m$, $\mathbb{M} \models \mathsf{cnst}_k(x, y)$ if and only if there exists i such that $y = \mathsf{Rep}(x)_i$ and $\mathsf{Seq}(x)_i = k$: that is, y is the index of a term in $\mathsf{Seq}(x)$ that is equal to k.
5) For all $x \in D^l$ and $y, z \in D^m$, $\mathbb{M} \models \mathsf{succ}(x, y, z)$ if and only if y, z are two consecutive terms appearing in $\mathsf{Rep}(x)$: that is, y, z index two consecutive terms of $\mathsf{Seq}(x)$. Note that $\mathsf{succ}(x, y, z)$ implies $\mathsf{rep}(x, y)$ and $\mathsf{rep}(x, z)$.
6) For all $x \in D^l$ and $y, z \in D^m$, $\mathbb{M} \models \mathsf{inc}(x, y, z)$ if and only if there exist i, j such that that $y = \mathsf{Rep}(x)_i$, $z = \mathsf{Rep}(x)_j$, and $\mathsf{Seq}(x)_j = \mathsf{Seq}(x)_i + 1$: that is, y, z index two terms t_1, t_2, respectively, of $\mathsf{Seq}(x)$ that satisfy $t_2 = t_1 + 1$.
7) For all $x \in D^l$ and $y, z \in D^m$, $\mathbb{M} \models \mathsf{eq}(x, y, z)$ if and only if there exist i, j such that $y = \mathsf{Rep}(x)_i$, $z = \mathsf{Rep}(x)_j$ and $\mathsf{Seq}(x)_j = \mathsf{Seq}(x)_i$: that is, y, z index two terms of $\mathsf{Seq}(x)$ that are equal.

Note that "$\mathbb{M}$ simulates counter machines" is not the same at all as "arbitrary finite sequences over $\mathbb{N}$ are definable in $\mathbb{M}$" in the classical sense (of model theory): the latter applies to structures $\mathbb{M}$ with domain $D \supseteq \mathbb{N}$, and means[5] that there exists $l \geq 1$ and a single formula φ with $l+2$ free variables such that when we plug $x \in D^l$ into φ, the resulting formulas (with two free variables) define all finite sequences over $\mathbb{N}$, where $(t_i)_{i=1}^N$ is viewed as the set $\{(1, t_1), \ldots, (N, t_N)\}$.

Theorem 15. *Let $\mathbb{M}$ be a structure with domain D that simulates counter machines. The first-order theory of $\mathbb{M}$ is undecidable.*

Proof. We will describe a procedure that takes as input a two-counter machine $\mathcal{M}$, and outputs a formula Φ in the language of $\mathbb{M}$ that is true (in $\mathbb{M}$) if and only if $\mathcal{M}$ halts. Recall that $\mathcal{M}$ has counters c_1, c_2 (initialised to 1) that take positive integer values and instructions $1, \ldots, H$ for $H > 1$. The execution starts at line $l = 1$, and $l = H$ is the unique halting instruction. The trace of $\mathcal{M}$ is $\langle 0, \iota_0, c_{1,0}, c_{2,0}, 0, \iota_1, c_{1,1}, c_{2,1}, \ldots\rangle$ where $\iota_n, c_{1,n}, c_{2,n}$ are, respectively, the (next) instruction to be executed, the value of c_1, and the value of c_2 at time n. We will construct a formula Φ that states "there exists x such that $\mathsf{Seq}(x)$ is a finite sequence that is the trace of $\mathcal{M}$ ending in the halting state", which is true if and only if $\mathcal{M}$ halts.

[5]There is some room for interpretation of what definability of arbitrary finite sequences in $\mathbb{M}$ should mean in the setting of model theory.

Define $\Phi \coloneqq \exists x \in D^l \colon \big(\Phi_{\mathsf{init}}(x) \wedge \Phi_{\mathsf{final}}(x) \wedge \Phi_\delta(x)\big)$ where

$$\Phi_{\mathsf{init}}(x) \coloneqq \exists y_1, y_2, y_3, y_4 \in D^m \colon \Bigg(\bigwedge_{i=1}^{3} \mathsf{succ}(x, y_i, y_{i+1}) \wedge \mathsf{cnst}_0(x, y_1) \wedge \bigwedge_{i=2}^{4} \mathsf{cnst}_1(x, y_i) \Bigg)$$

$$\Phi_{\mathsf{fin}}(x) \coloneqq \exists y_1, y_2, y_3, y_4 \in D^m \colon \Bigg(\bigwedge_{i=1}^{3} \mathsf{succ}(x, y_i, y_{i+1}) \wedge \mathsf{cnst}_0(x, y_1) \wedge \mathsf{cnst}_H(x, y_2) \Bigg)$$

$$\Phi_\delta(x) \coloneqq \forall y_1, \dots, y_8 \in D^m \colon \Bigg(\mathsf{cnst}_0(x, y_1) \wedge \bigwedge_{i=1}^{7} \mathsf{succ}(x, y_i, y_{i+1}) \Rightarrow \mathsf{cnst}_0(x, y_5) \wedge \Psi(x, y_2, y_3, y_4, y_6, y_7, y_8) \Bigg)$$

and Ψ is a positive Boolean combination of the formulas implementing the transition function of $\mathcal{M}$, i.e. a positive Boolean combination of formulas $\mathsf{inc}(x, y_i, y_j)$, $\mathsf{cnst}_k(x, y_i)$, and $\neg\mathsf{cnst}_k(x, y_i)$ for $i, j \in \{2, 3, 4, 6, 7, 8\}$ and $1 \leq k \leq H$. The formula $\Phi_{\mathsf{init}}(x)$ states that the sequence $\mathsf{Seq}(x)$ starts with four consecutive terms that are $0, 1, 1, 1$, respectively. The formula $\Phi_{\mathsf{fin}}(x)$ states that $\mathsf{Seq}(x)$ contains four consecutive terms $s_1, \dots, s_4$ such that $s_1 = 0$ and $s_2 = H$. Together, the two formulas imply that $\mathsf{Seq}(x)$ must contain at least eight terms: in particular, two blocks of four consecutive terms that start with the delimiter 0. Finally, $\Phi_\delta(x)$ states that whenever $s_1, \dots, s_8$ are consecutive terms of $\mathsf{Seq}(x)$ such that $s_1 = 0$, then $s_5 = 0$ and the transition function of $\mathcal{M}$ satisfies $\delta_{\mathcal{M}}(s_2, s_3, s_4) = (s_6, s_7, s_8)$. □

4 The first-order theory of $\langle \mathbb{N}; <, |\tau(\cdot)| \rangle$

We start with Theorem 7, as it is the easiest one among our undecidability results, in the sense that the hard mathematical work has already been done in [29]. Let

$$\mathcal{H} = \{z \in \mathbb{C} \colon \operatorname{Im}(z) > 0\}.$$

A *modular form* of weight k is an analytic function $f \colon \mathcal{H} \to \mathbb{C}$ satisfying the following conditions.

- For any
$$\Gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$
and $z \in \mathcal{H}$,
$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z).$$
The Möbius transformations $z \mapsto \frac{az+b}{cz+d}$ for a, b, c, d as above are precisely the automorphisms of $\mathcal{H}$; thus the condition above states that f behaves well with respect to the symmetries of $\mathcal{H}$.
- We can write
$$f(z) = \sum_{n=0}^{\infty} a_n q^n$$
where $a_n \in \mathbb{C}$ for all n and $q = e^{i2\pi z}$. That is, the Fourier expansion of f in terms of q does not have any negative powers. In the most interesting cases, a_n are often real numbers, and sometimes even integers.

We say that f is a *cusp modular form* if additionally $a_0 = 0$. Modular forms (as well as their Fourier coefficients) play a fundamental role in contemporary mathematics, e.g. in the theory of elliptic curves and studying the solutions of Diophantine equations, Wiles' proof of Fermat's Last Theorem [36] being a prominent example of both. In this section, we focus on the modular form called the *quadratic discriminant*, denoted Δ, and its Fourier coefficients, which are the values of the famous *Ramanujan tau function*; however, our undecidability result can be easily generalised to (coefficients of) a large class of modular forms called *primitive forms.*

The quadratic discriminant is a cusp modular form of weight 12 (no such forms of weight $2, \dots, 11$ exist) defined by

$$\Delta(z) = q \prod_{n=1}^{\infty} (1 - q^n)^{24}$$

where $q = e^{i2\pi z}$. The Ramanujan tau function returns the value of the nth Fourier coefficient of Δ, i.e.,

$$\Delta(z) = \sum_{n=0}^{\infty} \tau(n) q^n, \quad \tau(0) = 0, \quad \tau \colon \mathbb{N} \to \mathbb{Z}.$$

The tau function makes fascinating appearances in a diverse range of fields of mathematics; see, e.g., [37] for an exposition. It is known to be non-zero infinitely often and multiplicative: $\tau(mn) = \tau(m)\tau(n)$ for any coprime m, n. *Lehmer's conjecture* states that $\tau(n) \neq 0$ for all $n \geq 1$; this has been empirically verified [38] at least for all $n \leq 10^{20}$. Given n in binary, $\tau(n)$ can be computed in polynomial time, thanks to deep connections between the τ function and the number of rational points on elliptic curves [39, Chapter 15]. The only technical lemma about the tau function that we will need is the following, which is a specialisation of the main result of [29] concerning coefficients of primitive modular forms. For the remainder of this section, we write $f(n) = |\tau(n)|$.

Theorem 16. *Let $\sigma \colon \{1, \dots, k\} \to \{1, \dots, k\}$ be a permutation. Assuming Lehmer's conjecture, there exist infinitely many $c \in \mathbb{N}$ such that*

$$f(c + \sigma(1)) < \cdots < f(c + \sigma(k)),$$

i.e. for all $c < n_1, n_2 \leq c + k$,

$$f(n_1) < f(n_2) \Leftrightarrow \sigma(n_1 - c) < \sigma(n_2 - c).$$

Phrased differently, from τ we can extract, in a very simple manner, arbitrary finite permutations.[6] It remains to show how to simulate finite sequences over $\mathbb{N}$ using finite permutations.

Proof of Theorem 7. For $c, d, e \in \mathbb{N}$ we define

$$\mathsf{Rep}(c, d, e) = (d + 1, d + 2, \dots, e)$$

and $\mathsf{Seq}(c, d, e)$ to be the finite sequence $(t_i)_{i=1}^{e-d}$ over $\{0, \dots, d - c\}$ where

$$t_i = \#\{c < n \leq d \colon f(n) < f(d + i)\}.$$

[6] By slightly modifying the arguments of [29], we can show that one can, in fact, extract arbitrary permutations from $(|\tau(n)|)_{n \in \mathbb{N}}$ even if the tau function were to have zeros. That is, with more work we can prove undecidability without assuming Lehmer's conjecture.

Thus $d+i$ indexes the ith term of $\mathrm{Seq}(c,d,e)$, in the sense of Section 3. We claim that, assuming Lehmer's conjecture, $\mathrm{Seq}\colon \mathbb{N}^3 \to \mathbb{N}^*$ is onto. To see this, consider a non-empty sequence $(t_i)_{i=1}^N$ over $\{0,\ldots,R\}$; note that for the empty sequence, we can take, e.g., $c=d=e$. Let σ be a permutation of $\{1,\ldots,N+R\}$ such that $\sigma(1)<\cdots<\sigma(R)$ and

$$\#\{1\le j\le R\colon \sigma(j)<\sigma(R+i)\}=t_i$$

for all $1\le i\le N$. Applying Theorem 16, there exist infinitely many $c\in\mathbb{N}$ such that for all $c<n_1,n_2\le N+R$,

$$f(n_1)<f(n_2)\Leftrightarrow \sigma(n_1-c)<\sigma(n_2-c).$$

We can choose any such c, and set $d=c+R$, $e=c+R+N$. Then for all $1\le i\le N$,

$$\begin{aligned}\{c<n\le d\colon f(n)<f(d+i)\}&=\{c<n\le d\colon \sigma(n-c)<\sigma(b+i)\}\\&=\{1\le j\le b\colon \sigma(j)<\sigma(b+i)\}.\end{aligned}$$

Therefore, $\mathrm{Seq}(c,d,e)=(t_i)_{i=1}^N$. Next, we define

$$\begin{aligned}
&\mathrm{rep}(c,d,e,n)\coloneqq d+1\le n\le e\\
&\mathrm{cnst}_0(c,d,e,n)\coloneqq \mathrm{rep}(c,d,e,n)\wedge\\
&\quad\neg\exists m\in(c,d]\colon f(m)<f(n)\\
&\mathrm{cnst}_k(c,d,e,n)\coloneqq \mathrm{rep}(c,d,e,n)\wedge\\
&\quad\exists!\{m_1,\ldots,m_k\}\subset(c,d]\colon f(m_1),\ldots,f(m_k)<f(n)\\
&\mathrm{succ}(c,d,e,n_1,n_2)\coloneqq \mathrm{rep}(c,d,e,n_1)\ \wedge\ \mathrm{rep}(c,d,e,n_2)\wedge\\
&\quad\forall m\in(n_1,n_2)\colon \neg\mathrm{rep}(c,d,e,m)\\
&\mathrm{inc}(c,d,e,n_1,n_2)\coloneqq \mathrm{rep}(c,d,e,n_1)\ \wedge\ \mathrm{rep}(c,d,e,n_2)\wedge\\
&\quad\exists!m\in(c,d]\colon f(m)\in[f(n_1),f(n_2))\\
&\mathrm{eq}(c,d,e,n_1,n_2)\coloneqq \mathrm{rep}(c,d,e,n_1)\ \wedge\ \mathrm{rep}(c,d,e,n_2)\wedge\\
&\quad\neg\exists m\in(c,d]\colon f(m)\in[f(n_1),f(n_2))\cup[f(n_2),f(n_1))
\end{aligned}$$

where $k\in\mathbb{N}_{\ge1}$ and $\exists!\{m_1,\ldots,m_l\}\subset(c,d]$ means "there exists a unique set $\{m_1,\ldots,m_k\}$ of k numbers from $(c,d]$". (Similarly, $\exists!m\in(c,d]$ means "there exists unique $m\in(c,d]$".) Observe that for valid representatives n_1,n_2 (i.e., when $d+1\le n_1,n_2\le e$), $\mathrm{eq}(c,d,e,n_1,n_2)$ should evaluate to false if and only if there exists $m\in(c,d]$ such that either

- $f(m)<f(n_2)$ but $f(m)\not<f(n_1)$, or
- $f(m)<f(n_1)$ but $f(m)\not<f(n_2)$.

That is, $f(m)\in[f(n_1),f(n_2))\cup[f(n_2),f(n_1))$. Similarly, for valid representatives n_1,n_2, $\mathrm{inc}(c,d,e,n_1,n_2)$ should evaluate to true if and only if $f(n_2)>f(n_1)$ and there exists unique $c<m\le d$ such that $f(m)<f(n_2)$ but $f(m)\not<f(n_1)$. The conditions (2-7) in the definition of simulating counter machines are thus satisfied, and we can apply Theorem 15. □

5 The first-order theory of $\langle\mathbb{N};<,u(\cdot)\rangle$

Let $(u_n)_{n\in\mathbb{N}}$ be an integer LRS with exactly two non-repeated dominant roots as in the Introduction. We prove Theorem 5 in the remainder of this section by explicitly giving the maps $\mathrm{Seq}\colon\mathbb{N}^3\to\mathbb{N}^*$, $\mathrm{Rep}\colon\mathbb{N}^3\to\mathbb{N}^*$ and the formulas rep, cnst_k, succ, inc, eq. Our main tool for proving surjectivity of Seq is Theorem 6, which will be proved in Section 7. Let $\zeta>1$ be as in the statement of Theorem 6.

Let $c,d,e\in\mathbb{N}$, and suppose that $c<d<e$ and

$$0<u(c)<u(d)<u(e).$$

(For c,d,e that do not satisfy these conditions, we define $\mathrm{Seq}(c,d,e)$ and $\mathrm{Rep}(c,d,e)$ to be the empty sequence.) Let

$$X=\{c<n\le d\colon u(c)<u(n)<u(e)\}$$

and $n_1<\cdots<n_k$ be the ordering of all $d<n<e$ such that $u(c)<u(n)<u(e)$. We extract from (c,d,e) a finite sequence $\mathrm{Seq}(c,d,e)$ of length k by defining

$$\mathrm{Rep}(c,d,e)\coloneqq\langle n_1,n_2,\ldots,n_k\rangle$$

and

$$\mathrm{Seq}(c,d,e)_i\coloneqq\#\{x\in X\colon u(x)<u(n_i)\}\in\{0,\ldots,\#X\}$$

for $1\le i\le k$. Note that this construction is slightly different from the one we used in Section 4.

Example 1. Consider the sequence defined by $u_{n+3}=-u_{n+2}+u_n$, $u_0=0$, $u_1=1$, and $u_2=2$. We have that

$$u_n=a\lambda^n+\overline{a}\,\overline{\lambda}^n+yr^n$$

where $\lambda\approx-0.88+0.74\boldsymbol{i}$, $r\approx0.75$, $a\approx-0.68-0.61\boldsymbol{i}$ and $y\approx1.17$. Let us compute $\mathrm{Seq}(c,d,e)$ and $\mathrm{Rep}(c,d,e)$ for $(c,d,e)=(81,92,100)$. (We mention that $u_{81},u_{92},u_{100}>0$.) First, determine all $n\in(81,100)$ such that $u_{81}<u_n<u_{100}$. This leaves us with $n=82$, 84, 87, 89, 92 , 94, 95, 99. Since $d=92$, $X=\{82,84,87,89,92\}$. Next, we observe that

$$\begin{aligned}
&u_{81}<u_{82}<u_{84}<u_{87}<u_{89}<u_{94}<u_{92}<u_{100}\\
&u_{81}<u_{82}<u_{95}<u_{84}<u_{87}<u_{89}<u_{92}<u_{100}\\
&u_{81}<u_{82}<u_{84}<u_{87}<u_{99}<u_{89}<u_{92}<u_{100}.
\end{aligned}$$

For each red term, we count the number of black terms that are smaller. (The blue terms are just delimiters, and are not counted.) Hence the triple (c,d,e) defines the sequence $\mathrm{Seq}(81,92,100)=\langle4,1,3\rangle$ with the corresponding representatives $\mathrm{Rep}(81,92,100)=\langle94,95,99\rangle$.

We next show how the maps Seq and Rep can be implemented using first-order formulas. We proceed similarly to Section 4, mostly just replacing $f(\cdot)$ with $u(\cdot)$. Define

$$\begin{aligned}
&\mathrm{rep}(c,d,e,n)\coloneqq c<d<n<e\wedge\\
&\quad0<u(c)<u(d)<u(e)\ \wedge\ u(c)<u(n)<u(e)\\
&\mathrm{cnst}_0(c,d,e,n)\coloneqq \mathrm{rep}(c,d,e,n)\wedge\\
&\quad\neg\exists m\in(c,d]\colon u(m)\in(u(c),u(n))\\
&\mathrm{cnst}_k(c,d,e,n)\coloneqq \mathrm{rep}(c,d,e,n)\wedge\\
&\quad\exists!\{m_1,\ldots,m_k\}\subset(c,d]\colon u(m_1),\ldots,u(m_k)\in(u(c),u(n))\\
&\mathrm{succ}(c,d,e,n_1,n_2)\coloneqq \mathrm{rep}(c,d,e,n_1)\ \wedge\ \mathrm{rep}(c,d,e,n_2)\wedge\\
&\quad\forall m\in(n_1,n_2)\colon\neg\mathrm{rep}(c,d,e,m)\\
&\mathrm{inc}(c,d,e,n_1,n_2)\coloneqq \mathrm{rep}(c,d,e,n_1)\ \wedge\ \mathrm{rep}(c,d,e,n_2)\wedge\\
&\quad\exists!m\in(c,d]\colon u(m)\in[u(n_1),u(n_2))\\
&\mathrm{eq}(c,d,e,n_1,n_2)\coloneqq \mathrm{rep}(c,d,e,n_1)\ \wedge\ \mathrm{rep}(c,d,e,n_2)\wedge\\
&\quad\neg\exists m\in(c,d]\colon u(m)\in[u(n_1),u(n_2))\cup[u(n_2),u(n_1)).
\end{aligned}$$

where $k\in\mathbb{N}_{\ge1}$, and $\exists!\{m_1,\ldots,m_k\}$ and $\exists!$ are interpreted as in the proof of Theorem 7. Note that the definitions of succ, inc, and eq are identical (up to replacing $u(\cdot)$ with $f(\cdot)$) with those given in Section 4. Items (2-7) in the definition of simulating counter machines are satisfied by construction. To prove Theorem 5 (by applying Theorem 15) it remains to show the following.

Lemma 17. *The map* $\mathrm{Seq}\colon \mathbb{N}^3 \to \mathbb{N}^*$ *above is surjective.*

Proof. To define the empty sequence, we can choose, for example, $(c, d, e) = (0, 0, 0)$. Now consider a non-empty finite sequence $(t_i)_{i=1}^N$ over $\{0, \ldots, R\}$. We will construct, using Theorem 6,

$$c = n_0 < n_1 < \cdots < n_R = d < n_{R+1} < \cdots < n_{R+N+1} = e$$

such that

1) $0 < u(c) < u(n + i) < u(e)$ for all $1 \le i \le R + N$,
2) $u(n) \notin (u(c), u(e))$ for any $n \in (c, e) \setminus \{n_1, \ldots, n_{R+N}\}$, and
3) writing $X = \{n_1, \ldots, n_R\}$,
$$\#\{x \in X\colon u(x) < u(n_{R+i})\} = t_i$$
for all $1 \le i \le N$.

Then $\mathrm{Seq}(c, d, e) = (t_i)_{i=1}^N$.

Example 2. Figure 1 illustrates our construction for the finite sequence $\langle 1, 3, 0, 1\rangle$. In this case $N = 4$ and $R = 3$. We partition $(1, \zeta)$ into 10 intervals of equal length η. The intervals $(1, 1 + \eta)$ and $(\zeta - \eta, \zeta)$ are just buffers: they are needed because in the statement of Theorem 6, γ_j and δ_j must be strictly between 1 and ζ for all k. We define

- $(\gamma_1, \delta_1) = (1+2\eta, 1+3\eta)$, $(\gamma_2, \delta_2) = (1+4\eta, 1+5\eta)$, $(\gamma_3, \delta_3) = (1 + 6\eta, 1 + 7\eta)$,
- $(\gamma_4, \delta_4) = (1+3\eta, 1+4\eta)$, $(\gamma_5, \delta_5) = (1+7\eta, 1+8\eta)$, $(\gamma_6, \delta_6) = (1 + \eta, 1 + 2\eta)$, $(\gamma_7, \delta_7) = (1 + 3\eta, 1 + 4\eta)$, and
- $(\gamma_8, \delta_8) = (\zeta - 2\eta, \zeta - \eta)$.

Then $1 < \gamma_j < \delta_j < \zeta$ for all j. By Theorem 6 there exist infinitely many $n_0 < n_1 < \cdots < n_8$ such that $u_{n_j} > 0$ and,

$$\frac{u(n_j)}{u(n_0)} \in (\gamma_j, \delta_j)$$

for all j, and for all integers $n \in [n_0, n_8] \setminus \{n_0, \ldots, n_8\}$, either $u(n) < u(n_j)$ for all j, or $u(n) > u(n_j)$ for all j. Note that because $\gamma_8 \ge \delta_j$ for $1 \le j \le 7$, we have that $u(n_0) < u(n_j) < u(n_8)$ for all $1 \le j \le 7$. From our construction of (γ_j, δ_j), $1 \le j \le \ell$ it then follows that $\mathrm{Seq}(n_0, n_3, n_8) = \langle 1, 3, 0, 1\rangle$.

We now proceed with the proof. Let $\ell = R + N + 1$ and $\eta = (\zeta - 1)/(2R + 4)$. We define

$$(\gamma_k, \delta_k) = (1 + 2k\eta, 1 + (2k + 1)\eta)$$

for $1 \le k \le R$. For $k = 1, \ldots, N$, we define

$$(\gamma_{R+k}, \delta_{R+k}) = (1 + (2t_k + 1)\eta, 1 + (2t_k + 2)\eta)$$

Finally, we set $(\gamma_\ell, \delta_\ell) = (1 + (2R + 2)\eta, 1 + (2R + 3)\eta)$. Note that $1 < \gamma_j < \delta_j < \zeta$ for all j. Applying Theorem 6 we obtain infinitely many $n_0 < \cdots < n_\ell$ such that

- $u(n_j) > 0$ for all j,
- $\frac{u(n_j)}{u(n_0)} \in (\gamma_j, \delta_j)$ for all j, which implies that $u(n_0) < u(n_j) < u(n_\ell)$ for all $0 \le j < \ell$, and
- $u(n) \notin (u(n_0), u(n_\ell))$ for all integers $n_0 < n < n_\ell$ not equal to $\{n_1, \ldots, n_{\ell-1}\}$.

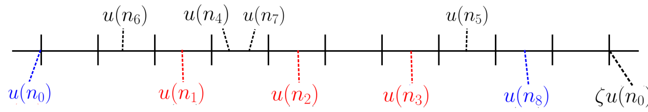

Figure 1: The construction of the proof of Lemma 17 for the finite sequence $\langle 1, 3, 0, 1\rangle$. The horizontal line is $\mathbb{R}$, and the vertical ticks are the points $u(n_0) + \eta k$ for $0 \le k \le 10$ and $\eta = u(n_0)(\zeta - 1)/10$. The dashed lines indicate the locations of $u(n_0), \ldots, u(n_8), \zeta u(n_0)$.

Thus $n_0, \ldots, n_\ell$ satisfy conditions (1-2) above. It remains to verify (3). For all $1 \le i \le N$ we have that

$$\{x \in X\colon u(x) < u(n_{R+i})\} = \left\{1 \le k \le R\colon \frac{u(n_k)}{u(n_0)} < \frac{u(n_{R+i})}{u(n_0)}\right\}.$$

Moreover,

$$\frac{u(n_{R+i})}{u(n_0)} \in (1 + (2t_i + 1)\eta, 1 + 2(t_i + 1)\eta)$$

and

$$\frac{u(n_k)}{u(n_0)} \in (1 + 2k\eta, 1 + (2k + 1)\eta)$$

for all $1 \le k \le R$. Therefore,

$$\left\{1 \le k \le R\colon \frac{u(n_k)}{u(n_0)} < \frac{u(n_{R+i})}{u(n_0)}\right\} = \{1, \ldots, t_i\}$$

and hence $\#\{x \in X\colon u(x) < u(n_{R+i})\} = t_i$. □

6 The first-order theory of $\langle \mathbb{N}; +, U\rangle$

We prove Theorem 3 in this section. We first define the maps $\mathrm{Seq}\colon \mathbb{N}^3 \to \mathbb{N}^*$ and $\mathrm{Rep}\colon \mathbb{N}^3 \to (\mathbb{N}^2)^*$. Recall that $<$ can be defined in $\langle \mathbb{N}; +, U\rangle$ by $x < y \Leftrightarrow x \ne y \wedge \exists z\colon x + z = y$. Consider $(p_c, p_d, p_e) \in U^3$ with $c < d < e$ (where $(p_n)_{n\in\mathbb{N}}$ enumerates U as in the Introduction); for all other triples, both Seq and Rep return the empty sequence. Let $R = d - c - 1$ and $N = e - d$. For $1 \le i \le N$, we define

$$\mathrm{Rep}(p_c, p_d, p_e)_i = (p_{d+i-1}, p_{d+i})$$

and $\mathrm{Seq}(p_c, p_d, p_e)_i$ as

$$\#\{c < n < d\colon p_n < p_c + p_{d+i} - p_{d+i-1}\} \in \{0, \ldots, R\}.$$

We next show how to implement Seq and Rep using first-order formulas; these are modifications of the formulas given in Section 5. We write $\mathbf{r}$ for the pair of variables r_1, r_2, and $\tilde{\mathbf{r}}$ for the pair $\tilde{r}_1, \tilde{r}_2$. Define

$$\begin{aligned}
&\mathrm{rep}(y_1, y_2, y_3, \mathbf{r}) \coloneqq y_1, y_2, y_3 \in U \wedge \\
&\quad y_1 < y_2 \le r_1 < r_2 \le y_3 \ \wedge\ \forall r \in (r_1, r_2)\colon r \notin U \\
&\mathrm{const}_0(y_1, y_2, y_3, \mathbf{r}) \coloneqq \mathrm{rep}(y_1, y_2, y_3, \mathbf{r}) \wedge \\
&\quad \neg\exists x \in (y_1, y_2) \cap U\colon x < y_1 + r_2 - r_1 \\
&\mathrm{const}_k(y_1, y_2, y_3, \mathbf{r}) \coloneqq \mathrm{rep}(y_1, y_2, y_3, \mathbf{r}) \wedge \\
&\quad \exists!\{x_1, \ldots, x_k\} \subset (y_1, y_2) \cap U\colon x_1, \ldots, x_k < y_1 + r_2 - r_1 \\
&\mathrm{succ}(y_1, y_2, y_3, \mathbf{r}, \tilde{\mathbf{r}}) \coloneqq \mathrm{rep}(y_1, y_2, y_3, \mathbf{r}) \wedge \\
&\quad \mathrm{rep}(y_1, y_2, y_3, \tilde{\mathbf{r}}) \ \wedge\ r_2 = \tilde{r}_1
\end{aligned}$$

where $k \in \mathbb{N}_{\ge 1}$ and $\exists!\{m_1, \ldots, m_k\}$ (as well as $\exists!$ used below) are interpreted as in the proof of Theorem 7. It remains to define eq and inc. Write t and $\tilde{t}$ for $y_1 + r_2 - r_1$ and $y_2 + \tilde{r}_2 - \tilde{r}_1$, respectively. Then we can define

$$\begin{aligned}&\mathrm{eq}(y_1, y_2, y_3, \mathbf{r}, \tilde{\mathbf{r}}) \coloneqq \mathrm{rep}(y_1, y_2, y_3, \mathbf{r}) \ \wedge \ \mathrm{rep}(y_1, y_2, y_3, \tilde{\mathbf{r}}) \ \wedge \\ &\quad \forall x \in (y_1, y_2) \cap U \colon x \notin [t, \tilde{t}\,) \cup [\tilde{t}, t) \\ &\mathrm{inc}(y_1, y_2, y_3, \mathbf{r}, \tilde{\mathbf{r}}) \coloneqq \mathrm{rep}(y_1, y_2, y_3, \mathbf{r}) \ \wedge \ \mathrm{rep}(y_1, y_2, y_3, \tilde{\mathbf{r}}) \ \wedge \\ &\quad \exists! x \in (y_1, y_2) \cap U \colon x \in [t, \tilde{t}\,).\end{aligned}$$

Items (2-6) in the definition of simulating counter machines are satisfied by construction. It remains to prove the following.

Lemma 18. *The map* $\mathrm{Seq}\colon \mathbb{N}^3 \to \mathbb{N}^*$ *above is surjective.*

Proof. Take a non-empty finite sequence $(t_i)_{i=1}^N$ over $\{0, \ldots, R\}$. We will construct $c < d < e$ such that $\mathrm{Seq}(p_c, p_d, p_e) = (t_i)_{i=1}^N$. (For the empty sequence, we can take any $c = d = e$.) It suffices to find n such that

$$\#\{n < m \leq n + R \colon p_m < p_n + p_{n+R+i+1} - p_{n+R+i}\} = t_i \tag{4}$$

for all $1 \leq i \leq N$. We can then choose $c = n$, $d = n + R + 1$ and $e = n + R + N + 1$. For $1 \leq k \leq R$, we define

$$(\gamma_k, \delta_k) = (1 + (2k - 1)\eta, 1 + 2k\eta)$$

where $\eta > 0$ is to be determined. Further define $(\gamma_{R+1}, \delta_{R+1}) = (1 + (2R + 1 - 1/8)\eta, 1 + (2R + 1 + 1/8)\eta)$. So far we have that

$$\gamma_1 < \delta_1 = \gamma_2 < \delta_2 = \cdots = \gamma_R < \delta_R < \gamma_{R+1} < \delta_{R+1}.$$

For $1 \leq i \leq N$, we write $k = R + i + 1$ and define

$$\gamma_k = 1 + \left(-\frac{1}{8} + 2R + 1 + \sum_{j=1}^{i} \left(2t_j + \frac{1}{2}\right)\right)\eta$$

$$\delta_k = 1 + \left(\frac{1}{8} + 2R + 1 + \sum_{j=1}^{i} \left(2t_j + \frac{1}{2}\right)\right)\eta.$$

Observe that $\gamma_{R+1} < \delta_{R+1} < \cdots < \gamma_{R+1+N} < \delta_{R+1+N}$, and for any $z \in (\gamma_{R+i+1}, \delta_{R+i+1})$ and $x \in (\gamma_{R+i}, \delta_{R+i})$ we have that

$$\frac{z - x}{\eta} \in \left(2t_i + \frac{1}{2} - \frac{1}{4}, 2t_i + \frac{1}{2} + \frac{1}{4}\right),$$

i.e., $z - x \approx (2t_i + 1/2)\eta$. Since

$$1 < \gamma_j, \delta_j < 1 + 2\left(R + N + \sum_{i=1}^{N} t_i\right)\eta,$$

for all j, we choose

$$\eta = \frac{\zeta - 1}{2\left(R + N + \sum_{j=1}^{N} t_j\right)}.$$

Then $1 < \gamma_j < \delta_j < \zeta$ for all j.

Applying Theorem 2, construct n such that

$$\frac{p_{n+j}}{p_n} \in (\gamma_j, \delta_j)$$

for all j. It remains to show that (4) holds for all $1 \leq i \leq N$. Observe that

$$\begin{aligned}&\{n < m \leq n + R \colon p_m < p_n + p_{n+R+i+1} - p_{n+R+i}\} \\ &\qquad = \left\{1 < k \leq R \colon \frac{p_{n+k}}{p_n} - 1 < \frac{p_{n+R+i+1} - p_{n+R+i}}{p_n}\right\}.\end{aligned}$$

Then it follows that we have that $p_{n+R+i+1}/p_n \in (\gamma_{R+i+1}, \delta_{R+i+1})$ and $p_{n+R+i}/p_n \in (\gamma_{R+i}, \delta_{R+i})$. Therefore, as discussed earlier,

$$\frac{p_{n+R+i+1} - p_{n+R+i}}{p_n} \in \left(\left(2t_i + \frac{1}{4}\right)\eta, \left(2t_i + \frac{3}{4}\right)\eta\right).$$

On the other hand,

$$\frac{p_{n+k}}{p_n} - 1 \in (1 + (2k - 1)\eta, 1 + 2k\eta)$$

for all $1 \leq k \leq R$. Therefore,

$$\left\{1 < k \leq R \colon \frac{p_{n+k}}{p_n} - 1 < \frac{p_{n+R+i+1} - p_{n+R+i}}{p_n}\right\} = \{1, \ldots, t_i\}. \qquad \square$$

7 Proof of Theorem 6

Assume the notation of Section 2.6. Recall that we have

$$\begin{aligned}u_n &= v_n + r_n = a\lambda^n + \overline{a}\,\overline{\lambda}^n + r_n \\ r_n &= o((\rho - \varepsilon)^n) \text{ for all sufficiently small } \varepsilon > 0 \\ v_n &= |a|\rho^n(\xi\mu^n + \overline{\xi}\,\overline{\mu}^n) = 2|a|\rho^n \cos(n\theta + \varphi).\end{aligned}$$

Define

$$\mathcal{I} = \left\{z \in \mathbb{T} \colon \ \mathrm{Re}(z) > \frac{1}{\rho}\right\}.$$

Then, by elementary geometry,

$$2 \cdot \sqrt{1 - \frac{1}{\rho^2}} < |\mathcal{I}| < \pi. \tag{5}$$

The significance of $\mathcal{I}$ is as follows.

Lemma 19. *For all n, if $\xi\mu^n \in \mathcal{I}$ then $v_n > 0, v_0, \ldots, v_{n-1}$.*

Proof. Suppose $\xi\mu^n \in \mathcal{I}$. Then $\xi\mu^n \in \mathbb{T}_+$ and hence $v_n > 0$. Next, take $0 \leq m < n$ such that $v_m > 0$; otherwise it is immediate that $v_n > v_m$. We have that

$$\frac{v_n}{v_m} = \frac{\rho^n \,\mathrm{Re}(\xi\mu^n)}{\rho^m \,\mathrm{Re}(\xi\mu^m)} > \frac{\rho^{n-m}}{\mathrm{Re}(\xi\mu^m)} \cdot \frac{1}{\rho} \geq \rho^{n-m-1} \geq 1. \qquad \square$$

For $z \in \mathbb{T}_+$, let

$$g_d(z) = \frac{z\lambda^d + \overline{z}\,\overline{\lambda}^d}{z + \overline{z}}.$$

Note that

$$g_d(z) = \eta \Leftrightarrow z^2 = -\frac{\overline{\lambda}^d - \eta}{\lambda^d - \eta}. \tag{6}$$

For $\gamma, \delta \in \mathbb{R}$ and $d \geq 1$, we define

$$\mathcal{J}_d(\gamma, \delta) = \{z \in \mathbb{T}_+ \colon g_d(z) \in (\gamma, \delta)\}.$$

We next argue that each $\mathcal{J}_d(\gamma, \delta)$ is an interval.

Lemma 20. *For all $d \geq 1$, $g_d \colon \mathbb{T}_+ \to \mathbb{R}$ is a homeomorphism.*

Proof. Define $f_d \colon (-\pi/2, \pi/2) \to \mathbb{R}$ by $f_d(x) = g_d(e^{ix})$. We have that

$$f_d(x) = \frac{\rho^d \cos(x + d\theta)}{\cos(x)} = \rho^d\big(\cos(d\theta) - \tan(x)\sin(d\theta)\big)$$

where the last equality follows from the usual trigonometric relations. As $\mu = e^{i\theta}$ is not a root of unity, $\sin(d\theta) \neq 0$. Because $\tan(x)$ is a homeomorphism from $(-\pi/2, \pi/2)$ to $\mathbb{R}$, so is f_d. It remains to write $g_d(z) = f_d(\mathrm{Log}(z))$, which is a composition of two homeomorphisms. $\square$

Corollary 21. *For every $\gamma, \delta \in \mathbb{R}$ with $\gamma < \delta$ and $d \geq 1$, $\mathcal{J}_d(\gamma, \delta)$ is a non-empty and open interval.*

Proof. We have that
$$\mathcal{J}_d(\gamma, \delta) = \{z \in \mathbb{T}_+ \colon f_d(z) \in (\delta, \gamma)\} = f_d^{-1}((\gamma, \delta)). \qquad \square$$

The intervals $\mathcal{J}_d(\gamma, \delta)$ play a key role in the proof of Theorem 4 [28]. Their idea behind their definition is that for all $n \in \mathbb{N}$ and $d \geq 1$,
$$\xi\mu^n \in \mathcal{J}_d(\gamma, \delta) \Leftrightarrow v_n > 0 \ \wedge \ \frac{v_{n+d}}{v_n} \in (\gamma, \delta). \tag{7}$$
Recall from Section 2.6 that $(\xi\mu^n)_{n\in\mathbb{N}}$ is dense in $\mathbb{T}$. Hence Corollary 21 tells us that for any $d \geq 1$ and $\gamma < \delta$, we can find infinitely many n such that $v_n > 0$ and
$$\frac{v_{n+d}}{v_n} \in (\gamma, \delta).$$
Proving Theorem 6 amounts to proving a version of the preceding statement that involves arbitrarily many terms of $(v_n)_{n\in\mathbb{N}}$ as opposed to only v_n and v_{n+d}.

We next estimate the length of $\mathcal{J}_d(\delta, \gamma)$. Lemma 22 is similar to a result proven in [28], whereas Lemma 23 is much stronger than the analogous result from [28]. Recall that we denote by Δ the arc distance function on the unit circle $\mathbb{T} \subseteq \mathbb{C}$.

Lemma 22. *There exists computable $C_1 > 0$ with the following property. For any $0 \leq \gamma < \delta \leq \frac{\rho+1}{2}$ and $d \geq 1$,*
$$|\mathcal{J}_d(\gamma, \delta)| < \frac{C_1(\delta - \gamma)}{\rho^d}.$$

Proof. See Appendix A. $\square$

For $d \geq 1$ let α_d be the unique $z \in \mathbb{T}_+$ such that $g_d(z) = 0$. For any d, γ, and δ, we refer to α_d as the *anchor point* of $\mathcal{J}_d(\gamma, \delta)$. The reason for this is that for any fixed $\gamma < \delta$, as $d \to \infty$, the intervals $\mathcal{J}_d(\gamma, \delta)$ become arbitrarily small while getting arbitrarily close to the point α_d. Note that α_d satisfies $\alpha_d\lambda^d + \overline{\alpha_d}\overline{\lambda}^d = 0$ and $|\alpha_d| = 1$. Hence $\alpha_d \in \{\boldsymbol{i}\mu^{-d}, -\boldsymbol{i}\mu^{-d}\}$.

Lemma 23. *There exists computable $C_2 > 0$ with the following property. Let $0 \leq \gamma < \delta \leq \frac{\rho+1}{2}$, $d \geq 1$, and suppose $\mathcal{J}_d(\gamma, \delta) \subseteq \mathcal{I}$. Then*
$$|\mathcal{J}_d(\gamma, \delta)| > \frac{C_2(\delta - \gamma)}{\rho^d}.$$

Proof. See Appendix A. $\square$

We will also need the following lemma, which is analogous to the density of $\xi\mu^n \in \mathbb{T}$ proven via Kronecker's theorem.

Lemma 24. *For any $0 \leq \gamma < \delta$, the intervals $(\mathcal{J}_d(\gamma, \delta))_{d=1}^{\infty}$ are dense in $\mathbb{T}_+$.*

Proof. Recall that α_d is an endpoint of $\mathcal{J}_d(0, \delta) \subseteq \mathbb{T}_+$, and that $\lim_{d\to\infty} |\mathcal{J}_d(0, \delta)| = 0$ by Lemma 22. Therefore, it suffices to prove that $(\alpha_d)_{d=1}^{\infty}$ is dense in $\mathbb{T}_+$. Recall that α_d is either $\boldsymbol{i}\mu^{-d}$ or $-\boldsymbol{i}\mu^{-d}$, whichever is in $\mathbb{T}_+$. (Exactly one of these is always the case, since by the non-degeneracy assumption, μ is not a root of unity.) By Kronecker's theorem, $(\boldsymbol{i}\mu^{-d})_{d=1}^{\infty}$ is dense in $\mathbb{T}$. Since $(\alpha_d)_{d=1}^{\infty}$ contains all terms of $(\boldsymbol{i}\mu^{-d})_{d=1}^{\infty}$ that lie in $\mathbb{T}_+$, it is dense in $\mathbb{T}_+$. $\square$

Henceforth fix $C_1, C_2 > 0$ as in the two lemmas above. Choose ζ such that
$$\zeta > 1, \qquad \zeta < \frac{\rho+1}{2}, \qquad \frac{C_1(\zeta - 1)}{\rho(\rho - 1)} < |\mathcal{I}|. \tag{8}$$
We will only work with subintervals of $\mathcal{J}_d(1, \zeta)$ for $d \geq 1$.

The next lemma tells us that as $d_1 \to \infty$, it becomes exponentially harder for $\mathcal{J}_{d_2}(1, \zeta)$ with $d_2 > d_1$ to intersect $\mathcal{J}_{d_1}(1, \zeta)$.

Lemma 25. *Suppose $d_1 < d_2$ and*
$$\mathcal{J}_{d_1}(1, \zeta) \cap \mathcal{J}_{d_2}(1, \zeta) \neq \emptyset. \tag{9}$$
Then
$$d_2 - d_1 > \left(\frac{\rho^{d_1}}{2C_1\zeta}\right)^{1/C_0} - 1 \tag{10}$$
where C_0, C_1 are the constants of lemmas 9 and 22, respectively.

Proof. See Appendix A. $\square$

Next, we further study when $\mathcal{J}_{d_1}(1, \zeta) \cap \mathcal{J}_{d_2}(1, \zeta) \neq \emptyset$.

Definition 26. *An interval $I \subseteq \mathcal{I}$ is protected from time $d \geq 0$ onwards if*
$$|I| > \sum_{k=d+1}^{\infty} |I \cap \mathcal{J}_k(1, \zeta)|.$$

Lemma 27. *The interval $\mathcal{I}$ is protected from time 0 onwards.*

Proof. By Lemma 22 and the construction of ζ,
$$\sum_{d=1}^{\infty} |\mathcal{I} \cap \mathcal{J}_d(1, \zeta)| \leq \sum_{d=1}^{\infty} |\mathcal{J}_d(1, \zeta)| \leq \frac{C_1(\zeta - 1)}{\rho(\rho - 1)} < |\mathcal{I}|. \qquad \square$$

We are now ready to prove a version of Theorem 6 for $(v_n)_{n\in\mathbb{N}}$; soon thereafter we will move to $(u_n)_{n\in\mathbb{N}}$. First, a helpful lemma.

Lemma 28. *Let $1 < \gamma < \delta < \zeta$. There exists $D \geq 0$ such that for all $d \geq D$, if $\mathcal{J}_d(\gamma, \delta) \subseteq \mathcal{I}$ then $\mathcal{J}_d(\gamma, \delta)$ is protected from time d onwards.*

Proof. See Appendix A. $\square$

Lemma 29. *Suppose we are given $\ell \geq 1$ and for $1 \leq j \leq \ell$, $1 < \gamma_j < \delta_j < \zeta$. We can construct $0 < d_1 < \cdots < d_\ell$ with the following properties.*

(a) $\mathcal{I} \supseteq \mathcal{J}_{d_1}(\gamma_1, \delta_1) \supseteq \cdots \supseteq \mathcal{J}_{d_\ell}(\gamma_\ell, \delta_\ell)$.
(b) For all $1 \leq d \leq d_\ell$, if $d \notin \{d_1, \ldots, d_\ell\}$ then
$$\mathcal{J}_d(1, \zeta) \cap \mathcal{J}_{d_\ell}(\gamma_\ell, \delta_\ell) = \emptyset.$$
(c) Every $\mathcal{J}_{d_j}(\gamma_j, \delta_j)$ is protected from time d_i onwards.

Proof. We proceed by induction on ℓ. For $\ell = 1$, the conclusion follows immediately from Lemma 27, Lemma 24, and Lemma 14.

Next, consider $\ell = m + 1 \geq 2$. Apply the induction hypothesis with $\gamma_1, \delta_1, \ldots, \gamma_m, \delta_m$ to construct $d_1, \ldots, d_m$. By Lemma 24, $(\mathcal{J}_d(1, \zeta))_{d > d_m}$ are dense in $\mathcal{J}_{d_m}(\gamma_m, \delta_m)$. Applying (c) of the induction hypothesis and Lemma 14, there exist infinitely many $d > d_m$ such that $\mathcal{J}_d(1, \zeta) \subseteq \mathcal{J}_{d_m}(\gamma_m, \delta_m) \subseteq \mathcal{I}$ and for all $1 \leq d' < d$, $\mathcal{J}_d(1, \zeta) \cap \mathcal{J}_{d'}(1, \zeta) = \emptyset$. Applying Lemma 28, we can construct infinitely many d for which we additionally have that $\mathcal{J}_d(\gamma_\ell, \delta_\ell)$ is protected from time d onwards. We can then pick d_ℓ to be equal to any such d. $\square$

Before finally proving Theorem 6, let us have a look at what Lemma 29 immediately gives us. Consider $1 < \gamma_j < \delta_j < \zeta$ for $1 \le j \le \ell$. Construct $d_1, \dots, d_\ell > 0$ using Lemma 29. By the density of $(\xi\mu^n)_{n\in\mathbb{N}}$ in $\mathbb{T}$, there exist infinitely many $n \in \mathcal{J}_{d_\ell}(\gamma_\ell, \delta_\ell) \subseteq \mathcal{I}$. Then $v_n > 0$, and by (a), $n \in \mathcal{J}_{d_j}(\gamma_j, \delta_j)$ for all i. Write $n_j = n + d_j$ for $1 \le j \le \ell$. By Equation (7),

$$\frac{v_{n_j}}{v_n} \in (\gamma_j, \delta_j) \subset (1, \zeta)$$

for all $1 \le j \le \ell$. On the other hand, from (c) it follows that for all $n \le m \le n_\ell$,

$$m \notin \{n_1, \dots, n_\ell\} \Rightarrow \frac{v_m}{v_n} \notin (1, \zeta), \tag{11}$$

Therefore, the pattern $v_n, v_{n_1}, \dots, v_{n_\ell}$ appears in the ordering of $(v_m)_{m=n}^{n_\ell}$, and has the ratios of consecutive elements controlled by (γ_j, δ_j), $1 \le j \le \ell$ for all j.

Proof of Theorem 6. Let ζ be as in Equation (8) and M be as in Lemma 12. Suppose we are given $1 < \gamma_j < \delta_j < \zeta$ for $1 \le i \le \ell$. First construct $\widetilde{\gamma}_j, \widetilde{\delta}_j$ satisfying $\gamma_j < \widetilde{\gamma}_j < \widetilde{\delta}_j < \delta_j$ for $1 \le i \le \ell$, and sufficiently small $\varepsilon > 0$ such that

$$(1+\varepsilon) \cdot \widetilde{\delta}_j \cdot \frac{1}{1-\varepsilon} < \delta_j$$
$$(1-\varepsilon) \cdot \widetilde{\gamma}_j \cdot \frac{1}{1+\varepsilon} > \gamma_j.$$

Apply Lemma 13 with $\varepsilon > 0$ to construct M_ε. Then let $\widetilde{M} = \max\{M, M_\varepsilon\}$. Apply Lemma 29 with $\widetilde{\gamma}_j, \widetilde{\delta}_j$ to construct $d_1, \dots, d_\ell$. Let $I = \mathcal{J}_{d_\ell}(\widetilde{\delta}_\ell, \widetilde{\gamma}_\ell)$. By the density of $\xi\mu^n$ in $\mathbb{T}$, there exist infinitely many $n \ge \widetilde{M}$ such that $\xi\mu^n \in I$. Pick such n, and let $\widetilde{n} = n + d_\ell + 1$. Then for all $n \le m < \widetilde{n}$,

- by Lemma 29 (a), if $m = n + d_j$ for some j, then $v_m/v_j \in (\gamma_j, \delta_j)$, and
- by Lemma 29 (b) and Lemma 13, if $m \ne n + d_j$ for all j, then $v_m/v_j \notin [1, \zeta)$.

Order $\{m : n \le m < \widetilde{n} \text{ and } u_n \le u_m < \zeta u_n\}$ as $(n_j)_{j=0}^k$. Then $k = \ell$ and $n_j = n + d_j$ for all $1 \le j \le \ell$. Because $\xi\mu^n \in I \subseteq \mathbb{T}_+$, we have that $v_n > 0$, and by Lemma 12, $u_n > 0$. By the application of Lemma 13,

$$\frac{u_{n_j}}{u_n} = \frac{u_{n_j}}{v_{n_j}} \cdot \frac{v_{n_j}}{v_n} \cdot \frac{v_n}{u_n} < (1+\varepsilon) \cdot \widetilde{\delta}_j \cdot \frac{1}{1-\varepsilon} < \delta_j$$

and similarly $\frac{u_{n_j}}{u_n} > (1-\varepsilon) \cdot \widetilde{\gamma}_j \cdot \frac{1}{1+\varepsilon} > \gamma_j$ for all j. ◻

8 Proof of Theorem 2

We will proceed similarly to the proof of Theorem 6, but we need much stronger technical machinery. Let $C_1, C_2 > 0$ and $\zeta > 1$ be as in the previous section, and C_0 be the constant of Lemma 9.

Lemma 30. *There exists computable $C_3 > 0$ such that for all $1 \le \gamma < \delta \le \zeta$ and $d, n \ge 1$, if $\xi\mu^n \in \mathcal{J}_d(\gamma, \delta)$ then*

$$d < C_3 \log(n+1).$$

Proof. See Appendix A. ◻

Next, we use Baker's theorem to prove a bound on how long it takes for $(\mu^n)_{n\in\mathbb{N}}$ to fall into a given sub-interval of $\mathbb{T}$.

Lemma 31. *There exists computable $C_4 > 0$ such that for any sub-interval $J \subseteq \mathbb{T}$, the following holds. For any $N \in \mathbb{N}$, there exists*

$$N \le n < N + \left(\frac{4\pi}{|J|}\right)^{C_4}$$

such that $\mu^n \in J$.

Proof. See Appendix A. ◻

Recall that our interval theory applies to $(v_n)_{n\in\mathbb{N}}$, and we translate results to $(u_n)_{n\in\mathbb{N}}$ using lemmas 12 and 13. We next prove a modification of Theorem 2 where we replace u_n with v_n. Let $(p'_n)_{n\in\mathbb{N}}$ be the ordering of $\{v_n \ge 0 : n \in \mathbb{N}\}$.

Lemma 32. *Let $\ell \ge 1$ and $1 < \gamma_1 < \delta_1 < \cdots < \gamma_\ell < \delta_\ell < \zeta$. There exist infinitely many n such that for all $1 \le j \le \ell$,*

$$\frac{p'_{n+j}}{p'_n} \in (\gamma_j, \delta_j).$$

Proof. From all γ_j and δ_j construct $d_1, \dots, d_\ell$ as in Lemma 29. Let $I = \mathcal{J}_{d_\ell}(\gamma_\ell, \delta_\ell)$ and $\mathcal{D} = \{0, d_1, \dots, d_\ell\}$. Suppose $n \in \mathbb{N}$ is such that $\xi\mu^n \in I$. Then we have the following.

- Because $I \subseteq \mathcal{I}$, $v_n > 0$ and $v_n > v_k$ for all $k < n$.
- For all $1 \le j \le \ell$, because $I \subseteq \mathcal{J}_{d_j}(\gamma_j, \delta_j)$, we have that
$$\frac{v_{n+d_j}}{v_n} \in (\gamma_j, \delta_j).$$
- Consider $1 \le k \le d_\ell$ with $k \notin \mathcal{D}$. Because $I \cap \mathcal{J}_k(1, \zeta)$ is empty (by Lemma 29 (b)) and $1 < \gamma_j < \delta_j < \zeta$ for all j, we have that either $v_{n+k} < v_{n+d}$ for all $d \in \mathcal{D}$, or $v_{n+k} > v_{n+d}$ for all $d \in \mathcal{D}$.
- Therefore, if additionally $\xi\mu^n \notin \mathcal{J}_k(1, \zeta)$ for all $k > d_\ell$, then the pattern $(v_n, v_{n+d_1}, \dots, v_{n+d_\ell})$ appears in $(p'_m)_{m\in\mathbb{N}}$. In particular, if $v_n = p'_m$, then $\gamma_j < p'_{m+j}/p'_m < \delta_j$ for all $1 \le j \le \ell$.

Hence it suffices to construct infinitely many $n \in \mathbb{N}$ such that $\xi\mu^n \in I$, but $\xi\mu^n \notin \mathcal{J}_k(1, \zeta)$ for all $k > d_\ell$. Let

$$Y_m = I \setminus \bigcup_{k=d_\ell+1}^{m} \mathcal{J}_k(1, \zeta).$$

Then by Lemma 30 it suffices to construct infinitely many n such that $\xi\mu^n \in Y_{\lfloor C_3 \log(n+1) \rfloor}$.

Let $N \in \mathbb{N}$. We will construct $n \ge N$ with the desired property. By construction of I (Lemma 29 (c)), there exists $\tau > 0$ such that $|Y_m| < |I| - \tau$ for all m. By the pigeonhole principle, each Y_m contains an open interval of size

$$\frac{\tau}{1 + C_3 \log(n+1)}.$$

With C_4 as in Lemma 31, pick $M \in \mathbb{N}$ such that

$$M - N \ge \left(\frac{4\pi(1 + C_3 \log(M))}{\tau}\right)^{C_4}.$$

Consider $Y_N, \dots, Y_{M-1}$. We have that

$$Y_N \supseteq Y_{N+1} \supseteq \cdots \supseteq Y_{M-1}$$

Therefore, there exists an open interval $Z \subseteq I$ of size

$$\frac{\tau}{1 + C_3 \log(M)}$$

such that $Z \subseteq Y_N, \ldots, Y_{M-1}$. Applying Lemma 31, there exists n such that

$$N \leq n < N + \left(\frac{4\pi(1 + C_3 \log(M))}{\tau}\right)^{C_4} \leq M$$

and $\mu\xi^n \in Z$. That is, $\mu\xi^n \in Y_n$ for some $N \leq n < M$. □

Proof of Theorem 2. Let M be as in Lemma 12. Suppose we are given $1 < \gamma_j < \delta_j < \zeta$ for $1 \leq j \leq \ell$. First construct $\widetilde{\gamma}_j, \widetilde{\delta}_j$ satisfying $\gamma_j < \widetilde{\gamma}_j < \widetilde{\delta}_j < \delta_j$ for $1 \leq j \leq \ell$, and sufficiently small $\varepsilon > 0$ such that

$$(1 + \varepsilon) \cdot \widetilde{\delta}_j \cdot \frac{1}{1 - \varepsilon} < \delta_j$$

$$(1 - \varepsilon) \cdot \widetilde{\gamma}_j \cdot \frac{1}{1 + \varepsilon} > \gamma_j.$$

Apply Lemma 13 with $\varepsilon > 0$ to construct M_ε. Then let $\widetilde{M} = \max\{M, M_\varepsilon\}$. By Lemma 32, there exist infinitely many n such that for all j

$$\frac{p'_{n+j}}{p'_n} \in (\widetilde{\gamma}_j, \widetilde{\delta}_j).$$

Therefore, there exist infinitely many $n_0, \ldots, n_\ell \geq \widetilde{M}$ such that $(v_{n_0}, \ldots, v_{n_\ell})$ appears in $(q'_n)_{n\in\mathbb{N}}$ and $v_{n_j} > 0$ for all j. By the application of Lemma 12, we have that $(u_{n_0}, \ldots, u_{n_\ell})$ appears in $(p_n)_{n\in\mathbb{N}}$, and $u_{n_j} > 0$ for all j. Consider such $n_0, \ldots, n_\ell$.

By the application of Lemma 13, for all $0 \leq j \leq \ell$ we have that $1 - \varepsilon < u_{n_j}/v_{n_j} < 1 + \varepsilon$. Therefore,

$$\frac{u_{n_j}}{u_{n_0}} = \frac{u_{n_j}}{v_{n_j}} \cdot \frac{v_{n_j}}{v_{n_0}} \cdot \frac{v_{n_0}}{u_{n_0}} \in \left(\frac{(1-\varepsilon)\widetilde{\gamma}_j}{1+\varepsilon}, \frac{(1+\varepsilon)\widetilde{\delta}_j}{1-\varepsilon}\right).$$

The latter interval is contained in (γ_j, δ_j) by the construction of ε, $\widetilde{\gamma}_j$, and $\widetilde{\delta}_j$. □

9 Discussion

Let us briefly discuss the broader implications of our results for the study of decidability of logical theories. Firstly, we believe that our approach for proving undecidability of the first-order theory of $\langle \mathbb{N}; <, |\tau(\cdot)| \rangle$, with some modifications, should work for many special functions, e.g. Euler's totient function and the Möbius function [40]. As mentioned earlier, the expansion of $\langle \mathbb{N}; + \rangle$ with powers of 2 and powers of 3 does not define multiplication, and hence the Hieronymi-Schulz approach is necessary in this case. This leads us to the following.

Problem 2. *Is addition (resp. multiplication) first-order definable in $\langle \mathbb{N}; <, |\tau(\cdot)| \rangle$? What about expansions of $\langle \mathbb{N}; < \rangle$ with other special functions?*

Given that we now have undecidability results for expansions of $\langle \mathbb{N}; < \rangle$ with "natural" functions (see for an example of a predicate P, that includes information about Turing machines, such that the first-order theory of $\langle \mathbb{N}; <, P \rangle$ is undecidable), we also ask the following.

Problem 3. *Is the first-order theory of $\langle \mathbb{N}; <, n \mapsto 2^n, n \mapsto 3^n \rangle$ is decidable?*

We move on to our results about integer LRS. We studied *non-degenerate* (but possibly reducible) LRS of the form $u_n = a\lambda^n + \overline{a}\,\overline{\lambda}^n + r_n$, where r_n is the non-dominant part. This restriction can be weakened to only the dominant part $v_n = a\lambda^n + \overline{a}\,\overline{\lambda}^n$ being non-degenerate, but not further. To see this, let $(w_n)_{n\in\mathbb{N}}$ be any integer LRS, $W = \{w_n : n \in \mathbb{N}\} \cap \mathbb{N}$, and $z_n = \langle 0, w_1, 0, w_2, \ldots \rangle$, which itself is an integer LRS. We have that $z_{2n+1} = w_n$ for all n. Next, take sufficiently large $a \in \mathbb{N}$ such that $a \equiv 1 \pmod 3$ and $|z_n| = O(a^n)$, and consider $u_n = (a\boldsymbol{i})^n + (-a\boldsymbol{i})^n + 3z_n$. The sequence $(u_n)_{n\in\mathbb{N}}$ has exactly two non-repeated, non-real dominant roots, but is degenerate: $(a\boldsymbol{i})/(-a\boldsymbol{i}) = -1$ is a root of unity. In particular, for all $n \in \mathbb{N}$ we have that $u_{4n} = 2a^n + 3z_{4n}$, $u_{4n+1} = 3z_{4n+1}$, $u_{4n+2} = -2a^n + z_{4n+2}$, and $u_{4n+3} = z_{4n+3}$. Therefore, we can define W in $\langle \mathbb{N}; +, U \cap \mathbb{N} \rangle$, where $U = \{u_n : n \in \mathbb{N}\}$: we have that for all sufficiently large x (where the implied threshold is effective), $x \in W \Leftrightarrow 3x \in U$. Since $(w_n)_{n\in\mathbb{N}}$ can be any integer LRS, we see that lifting the non-degeneracy assumption has the effect of voiding any restriction on the number of dominant roots of $(u_n)_{n\in\mathbb{N}}$. We believe that, nevertheless, our techniques can be extended to at least integer LRS whose dominant part is non-degenerate and of the form $\sum_{i=1}^m a_i\lambda_i^n + \overline{a_i}\,\overline{\lambda_i}^n$, where $a_1, \lambda_1, \ldots, a_m, \lambda_m$ are non-zero algebraic numbers and $|\lambda_1| = \cdots = |\lambda_m| > 1$.

Let us now revisit Problem 1 for integer LRS $(u_n)_{n\in\mathbb{N}}$ with exactly one non-repeated dominant root. As discussed in the Introduction, if the characteristic polynomial p of $(u_n)_{n\in\mathbb{N}}$ is irreducible (over $\mathbb{Z}$), then the first-order theory of $\langle \mathbb{N}; +, U \rangle$ is decidable, as shown by Semënov [12] and Point [14]. What happens if $(u_n)_{n\in\mathbb{N}}$ is reducible? We give an example.

Let $u_n = (2+\boldsymbol{i})^n + (2-\boldsymbol{i})^n$, which satisfies the recurrence relation $u_{n+2} = 4u_{n+1} - 5u_n$. Consider $v_n = 5^n + u_n$, which satisfies $v_n > 0$ for all n and has the characteristic polynomial $p(x) = (x-5)(x^2-4x+5)$ with the single, non-repeated dominant root $\lambda = 5$. What can we say about decidability of the first-order theory of $\langle \mathbb{N}; +, V \rangle$, where $V = \{v_n : n \in \mathbb{N}\}$? Define, in $\langle \mathbb{N}; +, V \rangle$, the predicate $W \subseteq \mathbb{N}$ by

$$x \in W \Leftrightarrow x \geq 0 \ \wedge \ \exists y_1, y_2 \in V : y_1 < y_2 \ \wedge \\ x = y_2 - 5y_1 \ \wedge \ \forall y_3 \in (y_1, y_2) : y_3 \notin V.$$

Then, using that $(v_n)_{n\in\mathbb{N}}$ is increasing, $W = \{w_n : n \in \mathbb{N}\} \cap \mathbb{N}$ for

$$w_n = u_{n+1} - 5u_n = (\boldsymbol{i} - 3)(2 + \boldsymbol{i})^n + (-\boldsymbol{i} - 3)(2 - \boldsymbol{i})^n$$

which is a non-degenerate integer LRS with exactly two dominant roots. Therefore, by Theorem 3, the first-order theory of $\langle \mathbb{N}; +, W \rangle$ is undecidable, which implies the same for $\langle \mathbb{N}; +, V \rangle$. More generally, the trick above can be used more to construct, given a reducible LRS with characteristic roots $\lambda_1, \ldots, \lambda_m$, new LRS whose characteristic roots form a subset of $\{\lambda_1, \ldots, \lambda_m\}$. We conjecture that, at least for irreducible LRS, by combining our approach and Semënov's result it should be possible to completely classify $(u_n)_{n\in\mathbb{N}}$ such that the first-order theory of $\langle \mathbb{N}; +, \{u_n : n \in \mathbb{N}\} \cap \mathbb{N} \rangle$ is decidable.

References

[1] R. Zach, "Hilbert's program then and now," in *Philosophy of logic*. Elsevier, 2007, pp. 411–447.

[2] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society*, vol. 58, no. 345-363, p. 5, 1936.

[3] D. Hilbert, "Die logischen grundlagen der Mathematik," *Mathematische Annalen*, vol. 88, no. 1, pp. 151–165, 1922.

[4] K. Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I," *Monatshefte für Mathematik und Physik*, vol. 38, pp. 173–198, 1931.

[5] Y. V. Matiyasevich, *Hilbert's Tenth Problem*. MIT Press, 1993.

[6] M. Presburger, "Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt," *Comptes Rendus du I congres de Mathématiciens des Pays Slaves*, pp. 92–102, 1929.
[7] A. Blumensath and E. Gradel, "Automatic structures," in *Proceedings of the 15th Annual IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, 2000, pp. 51–62.
[8] S. Hitarth, A. Mansutti, and G. Shabadi, "Optimization modulo integer linear-exponential programs," in *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 2026, pp. 5270–5343.
[9] W. Steiner, "Numeration systems: automata, combinatorics, dynamical systems, number theory," *Habilitationa Diriger des Recherches en Informatique. Université de Paris*, 2021.
[10] S. Y. Xiao, "Hilbert's tenth problem for systems of diagonal quadratic forms, and Büchi's problem," *arXiv preprint arXiv:2412.16740*, 2024.
[11] A. Bès and D. Richard, "Undecidable extensions of Skolem arithmetic," *The Journal of Symbolic Logic*, vol. 63, no. 2, pp. 379–401, 1998.
[12] A. L. Semenov, "Logical theories of one-place functions on the set of natural numbers," *Mathematics of the USSR-Izvestiya*, vol. 22, no. 3, p. 587, 1984.
[13] D. Chistikov, A. Mansutti, and M. R. Starchak, "Integer linear-exponential programming in NP by quantifier elimination," in *51st International Colloquium on Automata, Languages, and Programming (ICALP 2024)*, 2024, pp. 132:1–132:20.
[14] F. Point, "On decidable extensions of Presburger arithmetic: from A. Bertrand numeration sytems to Pisot numbers," *The Journal of Symbolic Logic*, vol. 65, no. 3, pp. 1347–1374, 2000.
[15] H. Mousavi, L. Schaeffer, and J. Shallit, "Decision algorithms for Fibonacci-automatic words, I: Basic results," *RAIRO-Theoretical Informatics and Applications-Informatique Théorique et Applications*, vol. 50, no. 1, pp. 39–66, 2016.
[16] B. Khoussainov and A. Nerode, "Automatic presentations of structures," in *International Workshop on Logic and Computational Complexity*. Springer, 1994, pp. 367–392.
[17] J. R. Büchi, "Weak second-order arithmetic and finite automata," in *The Collected Works of J. Richard Büchi*. Springer, 1990, pp. 398–424.
[18] P. Vojta, "Diagonal quadratic forms and Hilbert's tenth problem." *Contemporary Mathematics*, vol. 270, 2000.
[19] A. Bès, "Undecidable extensions of Büchi arithmetic and Cobham-Semënov theorem," *The Journal of Symbolic Logic*, vol. 62, no. 4, pp. 1280–1296, 1997.
[20] A. Woods, "Some problems in logic and number theory and their connections," Ph.D. dissertation, Univ. of Manchester, 1981.
[21] A. Bès, "A survey of arithmetical definability," *A tribute to Maurice Boffa*, pp. 1–54, 2002.
[22] P. Hieronymi and C. Schulz, "A strong version of Cobham's theorem," in *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, 2022, pp. 1172–1179.
[23] R. Villemaire, "The theory of $(\mathbb{N}, +, V_k, V_l)$ is undecidable," *Theoretical Computer Science*, vol. 106, no. 2, pp. 337–349, 1992.
[24] T. Karimov, F. Luca, J. Nieuwveld, J. Ouaknine, and J. Worrell, "On the decidability of Presburger arithmetic expanded with powers," in *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 2025, pp. 2755–2778.
[25] C. Schulz, "Undefinability of multiplication in Presburger arithmetic with sets of powers," *The Journal of Symbolic Logic*, pp. 1–15, 2023.
[26] R. Lipton, F. Luca, J. Nieuwveld, J. Ouaknine, D. Purser, and J. Worrell, "On the Skolem problem and the Skolem conjecture," in *Proceedings of the 37th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, 2022, pp. 1–9.
[27] A. Dubickas and M. Sha, "Positive density of integer polynomials with some prescribed properties," *Journal of Number Theory*, vol. 159, pp. 27–44, 2016.
[28] J. Nieuwveld and J. Ouaknine, "On Expansions of Monadic Second-Order Logic with Dynamical Predicates," in *50th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, vol. 345, 2025, pp. 80:1–80:17.
[29] Y. F. Bilu, J.-M. Deshouillers, S. Gun, and F. Luca, "Random ordering in modulus of consecutive Hecke eigenvalues of primitive forms," *Compositio Mathematica*, vol. 154, no. 11, pp. 2441–2461, 2018.
[30] M. L. Minsky, *Computation*. Prentice-Hall Englewood Cliffs, 1967.
[31] H. Cohen, *A course in computational algebraic number theory*. Springer Science & Business Media, 2013, vol. 138.
[32] G. Everest, A. J. Van Der Poorten, I. Shparlinski, and T. Ward, *Recurrence sequences*. American Mathematical Society, 2003, vol. 104.
[33] R. Tijdeman, M. Mignotte, and T. N. Shorey, "The distance between terms of an algebraic recurrence sequence." *Journal für die reine und angewandte Mathematik*, vol. 346, pp. 63–76, 1984.
[34] S. M. Gonek and H. L. Montgomery, "Kronecker's approximation theorem," *Indagationes Mathematicae*, vol. 27, no. 2, pp. 506–523, 2016.
[35] N. K. Vereshchagin, "Occurrence of zero in a linear recursive sequence," *Mathematical notes of the Academy of Sciences of the USSR*, vol. 38, no. 2, pp. 609–615, 1985.
[36] A. Wiles, "Modular elliptic curves and Fermat's Last Theorem," *Annals of Mathematics*, vol. 142, pp. 4–13, 1995.
[37] B. C. Berndt and P. Moree, "Sums of two squares and the tau-function: Ramanujan's trail," 2025. [Online]. Available: https://arxiv.org/abs/2409.03428
[38] J. Zeng and L. Yin, "On the computation of coefficients of modular forms: The reduction modulo p approach," *Mathematics of Computation*, vol. 84, no. 293, pp. 1469–1488, 2015.
[39] J.-M. Couveignes and B. Edixhoven, *Computational aspects of modular forms and Galois representations*. Princeton University Press, 2011.
[40] N. Frantzikinakis and B. Host, "The logarithmic Sarnak conjecture for ergodic weights," *Annals of Mathematics*, vol. 187, no. 3, pp. 869–931, 2018.

A Omitted proofs

Proof of Lemma 14. Let $N \in \mathbb{N}$. We will construct $k > N$ with the required property. Let J be a maximal interval component of $I \setminus \bigcup_{n=0}^{N} I_n$ such that $|J| > \sum_{n=N+1}^{\infty} |I_n \cap J|$; such J must exist by the assumption that $\sum_{n=0}^{\infty} |I_n \cap I| < |I|$, and must satisfy $|J| > 0$. Let z_1, z_2 be two distinct points in J outside $\bigcup_{n=0}^{\infty} I_n$. Further let $J' \subseteq J$ be an arc spanned by z_1 and z_2. We choose k to be the smallest n such that $I_n \cap J' \neq \emptyset$, which must exist by the density assumption. Such I_k must also satisfy $I_k \subset J'$, as $z_1, z_2 \notin I_n$ for all n, and the conclusion follows. □

Proof of Lemma 22. Let $z_1, z_2 \in \mathbb{T}_+$ be such that $g_d(z_1) = \gamma$ and $g_d(z_2) = \delta$. Then $|\mathcal{J}_d(\gamma, \delta)| = \Delta(z_1, z_2)$. By (6),

$$z_1^2 = -\frac{\overline{\lambda}^d - \gamma}{\lambda^d - \gamma}$$
$$z_2^2 = -\frac{\overline{\lambda}^d - \delta}{\lambda^d - \delta}.$$

By the geometry of the unit circle, we have that

$$|z_1 - z_2| < \Delta(z_1, z_2) < \frac{\pi}{2}|z_1 - z_2|. \tag{12}$$

Hence it suffices to estimate

$$|z_1 - z_2| = \frac{|z_1^2 - z_2^2|}{|z_1 + z_2|}. \tag{13}$$

We first argue that $\Delta(z_1, z_2) < \pi/2$, which implies that

$$\sqrt{2} < |z_1 + z_2| < 2. \tag{14}$$

Suppose that $\Delta(z_1, z_2) \geq \pi/2$. Then, by the continuity of g_d, there exists $\gamma < \beta \leq \delta$ and $z_3 \in \mathbb{T}_+$ such that $g_d(z) = \beta$ and $\Delta(z_1, z_3) = \pi/2$. Then either $z_3 = \boldsymbol{i} z_1$ or $z_3 = -\boldsymbol{i} z_1$, and hence $z_1^2 = -z_3^2$. Applying Equation (6), we obtain that

$$\frac{\overline{\lambda}^d - \gamma}{\lambda^d - \gamma} = -\frac{\overline{\lambda}^d - \beta}{\lambda^d - \beta}$$

which simplifies to

$$2\lambda^d\overline{\lambda}^d + (\beta - \gamma)(\overline{\lambda}^d - \lambda^d) + 2\gamma\beta = 0.$$

Observe that $\lambda^d\overline{\lambda}^d = \rho^{2d}$, β and γ are all real. Hence $\overline{\lambda}^d - \lambda^d$ is real. In particular, λ^d is real and thus $\lambda/\overline{\lambda}$ is a root of unity, which contradicts the non-degeneracy assumption. We conclude that $\Delta(z_1, z_2) < \pi/2$.

Therefore, to prove an upper bound on $\Delta(z_1, z_2)$ it suffices to consider $|z_1^2 - z_2^2|$. We have

$$z_1^2 - z_2^2 = \frac{(\delta - \gamma)\rho^d}{|\lambda^d - \gamma||\lambda^d - \delta|}\left|\mu^{2d} - 1\right|. \tag{15}$$

Consider $|\lambda^d - \gamma|$. By assumption, $|\lambda^d| = \rho^d > \gamma$. Moreover, λ^d is non-real. Hence $|\lambda^d - \gamma| > \rho^d - \gamma \geq \rho^d - \frac{\rho+1}{2} = \rho^d(1 - \frac{\rho+1}{2\rho^d}) \geq \rho^d \cdot c$ where $c = 1 - \frac{\rho+1}{2\rho} > 0$. Similarly, $|\lambda^d - \delta| > \rho^d \cdot c$. Therefore, we only need to bound $|\mu^{2d} - 1|$ from above by a constant. Since $|\mu^{2d} - 1| \leq 2$, we obtain

$$|z_1^2 - z_2^2| < \frac{2(\delta - \gamma)}{c^2 \rho^d}.$$

and hence

$$\Delta(z_1, z_2) < \frac{\pi}{2} \cdot \frac{2(\delta - \gamma)}{c^2 \rho^d} \cdot \frac{1}{\sqrt{2}}. \qquad \square$$

Proof of Lemma 23. Define z_1 and z_2 as was done in the proof of Lemma 22; then $|\mathcal{J}_d(\gamma, \delta)| > |z_1 - z_2|$. From Equations (12) to (15) we obtain

$$|z_1 - z_2| > \frac{(\delta - \gamma)\rho^d}{2|\lambda^d - \gamma||\lambda^d - \delta|} |\mu^{2d} - 1|.$$

As argued in the proof of Lemma 22, $|\lambda^d - \gamma|, |\lambda^d - \delta| < \rho^d \cdot c$ for $c = 1 - \frac{\rho+1}{2\rho}$. Hence it suffices to give a constant lower bound for $|\mu^{2d} - 1|$, which we do below.

Let $\widetilde{\mathcal{I}} = \{z \in \mathbb{T}_+ \colon \operatorname{Re}(z) > 1/(2\rho)\} \supset \mathcal{I}$ and $D \geq 1$ be such that $|\mathcal{J}_d(0, \frac{\rho+1}{2})| < \frac{1}{2\rho}$ for all $d \geq D$. Then for all $d \geq D$, if $\mathcal{J}_d(\gamma, \delta) \subseteq \mathcal{I}$ then $\mathcal{J}_d(0, \frac{\rho+1}{2}) \subseteq \widetilde{\mathcal{I}}$ and, as α_d is an endpoint of $\mathcal{J}_d(0, \frac{\rho+1}{2})$, $\alpha_d \in \widetilde{\mathcal{I}}$. We define

$$a = \min_{1 \leq d \leq D} |\mu^{2d} - 1|$$

which is positive by the assumption that μ is not a root of unity. We will show that

$$|\mu^{2d} - 1| \geq \min\{a, (2/\rho)^2\}.$$

If $d \leq D$, this is immediate. Now suppose $d \geq D$. Recall that $\mu^d = \frac{\boldsymbol{i}}{\pm\alpha_d}$ where $\pm\alpha_d$ is one of α_d and $-\alpha_d$. Then, as $\alpha_d \in \widetilde{\mathcal{I}}$, we have that $\operatorname{Re}(\alpha_d) > 2/\rho$ and hence

$$|\mu^{2d} - 1| = |\mu^d - 1| \cdot |\mu^d + 1| = |\pm\alpha_d - \boldsymbol{i}| \cdot |\pm\alpha_d + \boldsymbol{i}| \geq (2/\rho)^2. \quad \square$$

Proof of Lemma 25. By the triangle inequality,

$$\Delta(\alpha_{d_1}, \alpha_{d_2}) \leq \Delta(\alpha_{d_1}, z) + \Delta(z, \alpha_{d_2}) \tag{16}$$

for $z \in \mathcal{J}_{d_1}(1, \zeta) \cap \mathcal{J}_{d_2}(1, \zeta)$. Recall that α_d is an endpoint of $\mathcal{J}_d(0, \zeta)$. Hence we have

$$\Delta(\alpha_{d_1}, \alpha_{d_2}) \leq |\mathcal{J}_{d_1}(0, \zeta)| + |\mathcal{J}_{d_2}(0, \zeta)| < \frac{2C_1\zeta}{\rho^{d_1}} \tag{17}$$

where the last inequality is deduced from Lemma 22. Write $\alpha_{d_i} = \chi_i \mu^{-d_i}$ for $i = 1, 2$ where $\chi_i \in \{\boldsymbol{i}, -\boldsymbol{i}\}$. Then

$$\Delta(\alpha_{d_1}, \alpha_{d_2}) = \Delta(\chi_1 \chi_2^{-1}, \mu^{-d_2 + d_1}).$$

Note that $\alpha_{d_1} \neq \alpha_{d_2}$ as otherwise we would have $\mu^{-d_1} = \mu^{-d_2}$ or $\mu^{-d_1} = -\mu^{-d_2}$, contradicting the assumption that μ is not a root of unity. Applying Baker's theorem on the left-hand side (note that $\max\{2, d_2 - d_1\} \leq 1 + d_2 - d_1$),

$$\Delta(\alpha_{d_1}, \alpha_{d_2}) > \frac{1}{(1 + d_2 - d_1)^{C_0}}.$$

Therefore,

$$\frac{1}{(1 + d_2 - d_1)^{C_0}} < \frac{2C_1\zeta}{\rho^{d_1}}.$$

Rearranging gives the desired conclusion. $\square$

Proof of Lemma 28. Applying Lemma 25, we have that for all $d > 0$,

$$\sum_{k=d+1}^{\infty} |\mathcal{J}_d(\gamma, \delta) \cap \mathcal{J}_k(1, \zeta)| \leq \sum_{k=\varphi(d)}^{\infty} |\mathcal{J}_k(1, \zeta)| \leq \frac{C_1(\zeta - 1)}{\rho^{\varphi(d)} \cdot (\rho - 1)}$$

where

$$\varphi(d) = \left\lceil d + \left(\frac{\rho^d}{2C_1\zeta}\right)^{1/C_0} - 1 \right\rceil.$$

Because $\varphi(d)$ grows much faster than d, we can compute D such that for all $d \geq D$,

$$\frac{C_1(\zeta - 1)}{\rho^{\varphi(d)} \cdot (\rho - 1)} < \frac{C_2(\delta - \gamma)}{\rho^d}.$$

Therefore, for all $d \geq D$, if $\mathcal{J}_d(\gamma, \delta) \subseteq \mathcal{I}$ then by Lemma 23,

$$|\mathcal{J}_d(\delta, \gamma)| > \frac{C_2(\delta - \gamma)}{\rho^d} > \sum_{k=d+1}^{\infty} |\mathcal{J}_d(\gamma, \delta) \cap \mathcal{J}_k(1, \zeta)|. \qquad \square$$

Proof of Lemma 30. It suffices to prove the statement for $\gamma = 1$ and $\delta = \zeta$. By Lemma 22, whenever $\xi\mu^n \in \mathcal{J}_d(1, \zeta)$ we have that

$$0 < \Delta(\xi\mu^n, \alpha_d) < \frac{C_1\zeta}{\rho^d}.$$

Since $|\mu| = 1$ and α_d is one of $\pm\boldsymbol{i}\mu^{-d}$,

$$\Delta(\xi\mu^n, \alpha_d) = \Delta(\mu^{n+d}, z)$$

where z is one of $\pm\boldsymbol{i}/\xi$. Applying Lemma 9 to the right-hand side gives

$$\frac{C_1\zeta}{\rho^d} > \Delta(\xi\mu^n, \alpha_d) > \frac{1}{(n+d)^{C_0}}$$

for all $d, n \geq 1$ such that $\xi\mu^n \in \mathcal{J}_d(1, \zeta)$. Hence

$$n > \frac{(1+\varepsilon)^d}{(C_1\zeta)^{1/C_0}} - d$$

where $\varepsilon \coloneqq \rho^{1/C_0} - 1 > 0$. Let $D \geq 1$ and $\widetilde{\varepsilon} \in (0, \varepsilon)$ be such that for all $d \geq D$,

$$\frac{(1+\varepsilon)^d}{(C_1\zeta)^{1/C_0}} - d > (1 + \widetilde{\varepsilon})^d.$$

Then for all $n \geq 1$ and $d \geq D$ such that $\xi\mu^n \in \mathcal{J}_d(1, \zeta)$,

$$n + 1 > n > (1 + \widetilde{\varepsilon})^d$$

which implies that

$$\frac{1}{\log(1 + \widetilde{\varepsilon})} \cdot \log(n+1) > d.$$

Finally, note that for $d < D$ and $n \geq 1$, regardless of whether $\xi\mu^n \in \mathcal{J}_d(1, \zeta)$ we have that $d < \frac{D}{\log(2)} \log(n+1)$ as $\log(n+1)/\log(2) \geq 1$. We can therefore take

$$C_3 = \max\left\{\frac{D}{\log(2)}, \frac{1}{\log(1 + \widetilde{\varepsilon})}\right\}. \qquad \square$$

Proof of Lemma 31. It suffices to prove the claim in case J is open. Let $l = \left\lfloor \frac{2\pi}{|J|} \right\rfloor$ and consider the intervals $\{J, \ldots, \mu^l J\}$ on $\mathbb{T}$. As $(l+1)|J| > 2\pi$, there exist $0 \leq m < s \leq l$ such that $\mu^m J$ intersects $\mu^s J$. Let $k = s - m$ and $\theta = |\operatorname{Log}(\mu^k)|$. We have that $0 \leq k \leq l$ and $\theta < |J|$. Since μ is not a root of unity, $\mu^m \neq \mu^s$ and hence $\theta > 0$.

We next compute a lower bound on θ. Observe that $\theta > |\mu^k - 1|$. Applying Lemma 9,

$$\theta > |\mu^k - 1| > (\max\{2, k\})^{-B}$$

for a (computable) constant $B > 0$. Since $k \leq l < 4\pi/|J|$ and $2 \leq 4\pi/|J|$, we have that

$$\theta > (4\pi/|J|)^{-B}.$$

Let $L = \lceil 2\pi/\theta \rceil$. By the lower bound on θ above, $L < (4\pi/|J|)^C$ for a constant $C > 0$.

Consider the sequence $(z_n)_{n\in\mathbb{N}}$ of points on $\mathbb{T}$ that is defined by $z_n = \mu^{N+kn}$. We have that $z_{n+1} = \mu^k z_n$ and hence $|z_{n+1} - z_n| < |J|$ for all n. Moreover, the finite sequence $\langle z_0, \ldots, z_L \rangle$ winds around $\mathbb{T}$ at least once. Hence there exists

$$0 \leq r < L$$

such that $z_r \in J$. That is, $\mu^n \in J$ for $n = N + kr$. It remains to observe that $N \leq N + kr < N + kL$, and recall the bounds on k and L. □