

# On the Skolem Problem and the Skolem Conjecture

Richard J. Lipton

rjl@cc.gatech.edu

Department of Computer Science  
Georgia Institute of Technology  
USA

Joris Nieuwveld

Joël Ouaknine<sup>†</sup>

David Purser<sup>‡</sup>

jnieuwve@mpi-sws.org

joel@mpi-sws.org

dpurser@mpi-sws.org

Max Planck Institute for Software Systems  
Germany

Florian Luca\*

florian.luca@wits.ac.za

School of Mathematics  
University of the Witwatersrand  
South Africa

James Worrell

jbw@cs.ox.ac.uk

Department of Computer Science  
University of Oxford  
UK

## ABSTRACT

It is a longstanding open problem whether there is an algorithm to decide the Skolem Problem for linear recurrence sequences (LRS) over the integers, namely whether a given such sequence  $\langle u_n \rangle_{n=0}^\infty$  has a zero term (i.e., whether  $u_n = 0$  for some  $n$ ). A major breakthrough in the early 1980s established decidability for LRS of order 4 or less, i.e., for LRS in which every new term depends linearly on the previous four (or fewer) terms. The Skolem Problem for LRS of order 5 or more, in particular, remains a major open challenge to this day.

Our main contributions in this paper are as follows:

First, we show that the Skolem Problem is decidable for *reversible* LRS of order 7 or less. (An integer LRS  $\langle u_n \rangle_{n=0}^\infty$  is reversible if its unique extension to a bi-infinite LRS  $\langle u_n \rangle_{n=-\infty}^\infty$  also takes exclusively integer values; a typical example is the classical Fibonacci sequence, whose bi-infinite extension is  $\langle \dots, 5, -3, 2, -1, 1, 0, 1, 1, 2, 3, 5, \dots \rangle$ .)

Second, assuming the *Skolem Conjecture* (a central hypothesis in Diophantine analysis, also known as the *Exponential Local-Global Principle*), we show that the Skolem Problem for LRS of order 5 is decidable, and exhibit a concrete procedure for solving it.

## CCS CONCEPTS

• Theory of computation  $\rightarrow$  Logic and verification.

\*Also affiliated with: the Research Group in Algebraic Structures and Applications, King Abdulaziz University, Jeddah, Saudi Arabia; the Centro de Ciencias Matemáticas UNAM, Morelia, Mexico; and the Max Planck Institute for Software Systems, Saarland Informatics Campus, Germany.

<sup>†</sup>Also affiliated with Keble College, Oxford as emmy.network Fellow.

<sup>‡</sup>Also affiliated with University of Warsaw, Poland.



This work is licensed under a Creative Commons Attribution International 4.0 License.

LICS '22, August 2–5, 2022, Haifa, Israel

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9351-5/22/08.

<https://doi.org/10.1145/3531130.3533328>

## KEYWORDS

Linear recurrence sequences, Skolem Problem, Skolem Conjecture, Exponential local-global principle, decidability

### ACM Reference Format:

Richard J. Lipton, Florian Luca, Joris Nieuwveld, Joël Ouaknine, David Purser, and James Worrell. 2022. On the Skolem Problem and the Skolem Conjecture. In *37th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS) (LICS '22)*, August 2–5, 2022, Haifa, Israel. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3531130.3533328>

## 1 INTRODUCTION

### 1.1 Linear Recurrence Sequences and Bi-Sequences

A *linear recurrence relation* over a ring  $R$  is an equation of the following form:

$$u_{n+k} = a_1 u_{n+k-1} + \dots + a_k u_n, \quad (1)$$

where  $a_1, \dots, a_k \in R$ . We shall in addition require that  $a_k \neq 0$ .

Let  $S = \{1, a_k, a_k^2, a_k^3, \dots\}$ , and let  $S^{-1}R$  be the localisation of  $R$  by  $S$ .<sup>1</sup> Given initial values  $u_0, \dots, u_{k-1} \in R$ , Eqn. (1) uniquely defines an  $R$ -sequence  $\vec{u} = \langle u_n \rangle_{n=0}^\infty$ , as well as an  $(S^{-1}R)$ -bi-sequence  $\overleftarrow{u} = \langle u_n \rangle_{n=-\infty}^\infty$ . We refer to the former as a *linear recurrence sequence (LRS)*, and to the latter as a *linear recurrence bi-sequence (LRBS)*. The smallest  $k$  for which the sequence obeys a relation of the form (1) is the *order* of the sequence.

Note that if  $a_k$  is a unit of  $R$ , then  $S^{-1}R = R$  and the LRBS  $\overleftarrow{u}$  is entirely contained in  $R$ . An old result of Fatou [16] (see also [8, Chapter 7]) straightforwardly entails that when  $R$  is the ring of integers  $\mathbb{Z}$  and  $k$  is the order of  $\vec{u}$ , then  $\overleftarrow{u}$  is contained in  $\mathbb{Z}$  if and only if  $a_k = \pm 1$ . Such  $\mathbb{Z}$ -LRS are said to be *reversible*, and in turn the corresponding recurrence relation is likewise termed *reversible*. A classical example is the Fibonacci sequence which satisfies the recurrence  $u_{n+2} = u_{n+1} + u_n$  with initial values  $u_0 = 0, u_1 = 1$ , and which extends to the integer bi-sequence  $\langle \dots, 5, -3, 2, -1, 1, 0, 1, 1, 2, 3, 5, \dots \rangle$ .

<sup>1</sup> $S^{-1}R$  is a ring consisting of elements of the form  $r/a_k^m$ , with  $r \in R$  and  $m \in \mathbb{N}$ .

We shall also be interested in instances in which  $R$  is of the form  $\mathbb{Z}/m\mathbb{Z}$ , for integer  $m \geq 2$ . In such cases we will at times require that  $\gcd(m, a_k) = 1$ , which guarantees that the resulting LRBS  $\overleftarrow{u}$  is entirely contained in  $\mathbb{Z}/m\mathbb{Z}$ . Note that LRS in  $\mathbb{Z}/m\mathbb{Z}$  are necessarily always eventually periodic,<sup>2</sup> and therefore (among others) it is always possible to determine whether they contain a zero or not.

Equation (1) gives rise to a polynomial

$$g(X) := X^k - a_1 X^{k-1} - \dots - a_k. \quad (2)$$

The *characteristic polynomial* of an LRS is the polynomial associated with its minimal recurrence relation (i.e.,  $k$  is the order of the LRS). A  $\mathbb{Q}$ -LRS is *simple* if its characteristic polynomial is simple, i.e., has no repeated roots.

We refer to the roots of the characteristic polynomial of a given LRS as the *characteristic roots* of the LRS. We then say that an LRS is *non-degenerate* if it has no two distinct characteristic roots whose ratio is a root of unity. As noted in the next section, the study of the set of zeros of a given LRS can be reduced in an effective manner to the study of the set of zeros of its non-degenerate subsequences.

## 1.2 The Skolem Problem and Bi-Variants

The decidability of the following question has been open for nearly a century [15, 30]:

**Skolem Problem.** Let  $C$  be a class of linear recurrence sequences. Given an LRS  $\overrightarrow{u} \in C$ , does  $\overrightarrow{u}$  contain a zero?

(When the class  $C$  is not explicitly mentioned, it is assumed to refer to integer linear recurrence sequences.) The Skolem Problem is arguably, by some distance, the most prominent problem whose decidability status is currently unknown. It remains open even if restricting to simple LRS of order 5 [22], but is famously known to be decidable for sequences of order 4 or below [21, 31]. It is also straightforward that the Skolem Problems over  $\mathbb{Z}$  and over  $\mathbb{Q}$  are interreducible.

The Skolem Problem, along with closely related questions such as the Positivity Problem, is intimately connected to various fundamental topics in program analysis and automated verification, such as the termination and model checking of simple while loops [3, 19, 23] or the algorithmic analysis of stochastic systems [1, 2, 6, 13, 24]. It also appears in a variety of other contexts, such as formal power series [25, 29] and control theory [11, 17]. The Skolem Problem is often used as a reference to establish hardness of other open decision problems; in addition to some of the previously cited papers, the articles [4, 14], for example, specifically invoke hardness for the Skolem Problem at order 5. Thus far, the only known complexity bound for the Skolem Problem is NP-hardness [12].

Instead of sequences, one might shift one's attention to bi-sequences, leading to the following two natural problems:

**Definition 1.1** ( $\mathbb{Q}$ -Bi-Skolem Problem). Given a  $\mathbb{Q}$ -LRBS  $\overleftarrow{u}$ , does  $\overleftarrow{u}$  contain a zero?

**Definition 1.2** ( $\mathbb{Z}$ -Bi-Skolem Problem). Given a  $\mathbb{Z}$ -LRBS  $\overleftarrow{u}$ , does  $\overleftarrow{u}$  contain a zero?

*Example 1.3.* The Fibonacci sequence  $\langle 0, 1, 1, 2, 3, 5, \dots \rangle$  has a single zero, whereas its shifted sibling  $\langle 1, 1, 2, 3, 5, 8, \dots \rangle$  has none. Of course, their common canonical 'completion' as a bi-sequence,  $\langle \dots, 5, -3, 2, -1, 1, 0, 1, 1, 2, 3, 5, \dots \rangle$ , does have a zero. On the other hand, the bi-sequence

$$\langle \dots, 7, -4, 3, -1, 2, 1, 3, 4, 7, \dots \rangle,$$

which also obeys the Fibonacci recurrence, never vanishes. The latter can be ascertained either via a growth argument (in absolute value), or by observing that the corresponding bi-sequence of residues modulo 5 (i.e., viewing the LRBS as lying in  $\mathbb{Z}/5\mathbb{Z}$ ) has no zeros. It is worth pointing out that the classical Fibonacci sequence has infinitely many zeros modulo any integer  $m \geq 2$ .

It is immediate that the  $\mathbb{Z}$ -Bi-Skolem Problem reduces to the  $\mathbb{Q}$ -Bi-Skolem Problem, and that in turn the  $\mathbb{Q}$ -Bi-Skolem Problem reduces to the Skolem Problem. However, the following questions do not seem to have obvious answers:

**Problem 1.4.** *Is the  $\mathbb{Z}$ -Bi-Skolem Problem decidable?*

**Problem 1.5.** *Does the  $\mathbb{Q}$ -Bi-Skolem Problem reduce to the  $\mathbb{Z}$ -Bi-Skolem Problem?*

**Problem 1.6.** *Does the Skolem Problem reduce to the  $\mathbb{Q}$ -Bi-Skolem Problem?*

The celebrated theorem of Skolem, Mahler, and Lech (see [15]) describes the structure of the set  $\{n \in \mathbb{N} : u_n = 0\}$  of zeros of an LRS as follows:

**Theorem 1.7** (Skolem-Mahler-Lech). *Given a  $\mathbb{Q}$ -LRS  $\overrightarrow{u}$ , its set of zeros is a union of finitely many arithmetic progressions, together with a finite set.*

The statement of Theorem 1.7 can in fact be refined further. Any  $\mathbb{Q}$ -LRS can be effectively decomposed as an interleaving of finitely many non-degenerate sequences, some of which may be identically zero. The core of the Skolem-Mahler-Lech Theorem is the fact that a non-zero non-degenerate linear recurrence sequence has finitely many zeros. Unfortunately, all known proofs of this last assertion are ineffective: it is not known how to compute the finite set of zeros of a given non-degenerate LRS. It is readily seen that the existence of a procedure to do so is equivalent to solving the Skolem Problem.

## 1.3 The Skolem Conjecture

In 1937, Thoralf Skolem formulated a conjecture (also known today as the *Exponential Local-Global Principle*) on purely exponential Diophantine equations [28]. Specialised to  $\mathbb{Q}$ -LRBS, the conjecture reads as follows:

**Skolem Conjecture.** *Let  $\overleftarrow{u}$  be a simple  $\mathbb{Q}$ -LRBS satisfying Eqn. (1), with  $a_1, \dots, a_k$  and  $u_0, \dots, u_{k-1}$  in  $\mathbb{Z}$ . Then  $\overleftarrow{u}$  has no zero iff, for some integer  $m \geq 2$  such that  $\gcd(m, a_k) = 1$ , we have that for all  $n \in \mathbb{Z}$ ,  $u_n \not\equiv 0 \pmod{m}$ .*

In other words, the Skolem Conjecture asserts that if a simple LRBS fails to have a zero, then this is witnessed modulo *some* integer  $m$ . If true, this conjecture would therefore immediately entail the existence of an algorithm to solve the  $\mathbb{Q}$ -Bi-Skolem Problem for simple LRBS: simply search in parallel either for a zero of the LRBS, or for a number  $m$  substantiating the absence of zeros. If the Skolem

<sup>2</sup>Under the additional assumption that  $\gcd(m, a_k) = 1$ , such LRS (and LRBS) are even fully periodic.

Conjecture holds, then the search must necessarily terminate in finite time.

*Remark 1.8.* Note that the Skolem Conjecture only applies to *bi*-sequences, as the simple example of the shifted Fibonacci sequence beginning with  $(1, 1, \dots)$  demonstrates.

The assumption of simplicity can also not be lifted: consider the LRBS  $\overleftarrow{u}$  given by  $u_{n+2} = 4u_{n+1} - 4u_n$ , with  $u_0 = 1$  and  $u_1 = 6$ , having closed form  $u_n = (2n+1)2^n$  with the single repeated characteristic root 2.  $u_n$  is clearly never 0 for  $n \in \mathbb{Z}$ , however given any integer  $m \geq 2$ , there are infinitely many values of  $n \in \mathbb{N}$  such that  $u_n \equiv 0 \pmod{m}$ .<sup>3</sup>

There exists a substantial body of literature on the Skolem Conjecture, including proofs of a variety of special cases. In particular, as pertains to its specialisation to  $\mathbb{Q}$ -LRBS, the Skolem Conjecture has been shown to hold for simple LRBS of order 2 [7], and for certain families of LRBS of order 3 [26, 27]. In a different but related vein, Bertók and Hajdu have shown that, in some sense, the Skolem Conjecture is valid in “almost all” instances [9, 10].

It is worth noting that, in spite of the similarity in nomenclature between the Skolem Problem and the Skolem Conjecture, the truth of the latter is not known (or even believed) to imply the decidability of the former (and nor, for that matter, conversely). Indeed, as pointed out in Remark 1.8, the Skolem Conjecture only applies to *simple* LRBS. Moreover—and perhaps more significantly—the Skolem Conjecture only differentiates between LRBS having *at least one zero*, and LRBS having *none*. The Skolem Conjecture would therefore appear to be of little utility for any LRS whose bi-completion happens to harbour a zero. As a stark illustration of this state of affairs, note that an algorithm for the Skolem Problem would enable one to produce in finite time the set of all zeros of a given non-degenerate LRS, whereas there is no obvious way to achieve the same merely by virtue of the Skolem Conjecture holding.

## 1.4 Main Contributions

Recall that the Skolem Problem is presently not known to be decidable for LRS of order 5 and above [22]. It is interesting to note that, whilst most researchers in the area likely expect decidability to hold at all orders, to the best of our knowledge there is currently not a single proposed candidate procedure that might conjecturally serve to decide outstanding open cases of the Skolem Problem.

Our contributions in the present paper are threefold:

- (1) We show that the Skolem Problem is decidable for *reversible*  $\mathbb{Z}$ -LRS of order 7 or less (Theorem 4.1).
- (2) Assuming the Skolem Conjecture, we show that the Skolem Problem is decidable for all  $\mathbb{Z}$ -LRS of order 5 (Theorem 5.1).<sup>4</sup>
- (3) We exhibit concrete LRS, respectively of order 8 (Section 4.2) and of order 6 (Example 3.5), showing that the above techniques and results do not extend (at least in any obvious manner) to higher orders.

<sup>3</sup>Strictly speaking, the Skolem Conjecture only requires one to consider positive integers  $m$  such that  $\gcd(m, 4) = 1$ , which in turn ensures that the sequence of residues modulo  $m$  remains well defined for negative values of  $n$ .

<sup>4</sup>In fact, it is sufficient merely to assume that the Skolem Conjecture holds for LRBS of order 5.

Our conditional decidability result for LRS of order 5 is obtained by producing an explicit (and easily implementable) decision procedure. Correctness is unconditional and straightforward, whereas termination is guaranteed via the Skolem Conjecture (for LRBS of order 5). To the best of our knowledge, this is the first concrete plausible proposal of an algorithm to solve the Skolem Problem at order 5.

## 2 PRELIMINARIES

In this section we briefly summarise some basic notions about algebraic numbers and linear recurrences. For more details see [5, Chapters 10–12] and [15, Section 1.1.6].

Let  $K$  be a finite Galois extension of  $\mathbb{Q}$ . Write  $\mathcal{O}$  for the ring of algebraic integers in  $K$ . Every ideal of  $\mathcal{O}$  can be written uniquely (up to reordering) as a product of prime ideals of  $\mathcal{O}$ . In particular, for every rational prime  $p \in \mathbb{Z}$  we have  $p\mathcal{O} = \mathfrak{p}_1^e \cdots \mathfrak{p}_g^e$  for some prime ideals  $\mathfrak{p}_1, \dots, \mathfrak{p}_g$  and a positive integer  $e$  the *ramification index* of  $p$ . In this situation we say that the prime ideals  $\mathfrak{p}_1, \dots, \mathfrak{p}_g$  lie *above*  $p$ . Here we have that  $e$  divides  $[K : \mathbb{Q}]$  (the degree of  $K$  over  $\mathbb{Q}$ ).

Given a prime ideal  $\mathfrak{p}$  of  $\mathcal{O}$ , we define the *p-adic valuation*  $v_{\mathfrak{p}} : \mathcal{O} \rightarrow \mathbb{N} \cup \{\infty\}$  by writing  $v_{\mathfrak{p}}(0) = \infty$  and, for  $\alpha \neq 0$ , defining  $v_{\mathfrak{p}}(\alpha)$  to be the exponent of  $\mathfrak{p}$  in the prime decomposition of  $\alpha\mathcal{O}$ . We then have  $v_{\mathfrak{p}}(\alpha_1\alpha_2) = v_{\mathfrak{p}}(\alpha_1) + v_{\mathfrak{p}}(\alpha_2)$  and  $v_{\mathfrak{p}}(\alpha_1 + \alpha_2) \geq \min(v_{\mathfrak{p}}(\alpha_1), v_{\mathfrak{p}}(\alpha_2))$  for all  $\alpha_1, \alpha_2 \in \mathcal{O}$ , i.e.,  $v_{\mathfrak{p}}$  is indeed a valuation. A further useful fact is that if  $\alpha_1, \alpha_2 \in \mathcal{O}$  are such that  $\alpha_1/\alpha_2$  is not a root of unity and all Galois conjugates of  $\alpha_1/\alpha_2$  have modulus one, then there exists a prime ideal  $\mathfrak{p}$  with  $v_{\mathfrak{p}}(\alpha_1) \neq v_{\mathfrak{p}}(\alpha_2)$ .

Consider a recurrence of the form (1) with integer coefficients. Let  $K$  be the splitting field of the characteristic polynomial. Every LRS  $\overrightarrow{u} = \langle u_n \rangle_{n=0}^{\infty}$  satisfying the recurrence admits an *exponential-polynomial* representation  $u_n = \sum_{i=1}^k A_i(n)\lambda_i^n$ , where  $\lambda_1, \dots, \lambda_k$  are the distinct characteristic roots (which are algebraic integers) and  $A_1, \dots, A_k \in K[X]$  are polynomials such that the degree of  $A_i$  is less than the multiplicity of  $\lambda_i$ . We say that a characteristic root  $\lambda_i$  is *dominant in modulus* if  $|\lambda_i| \geq |\lambda_j|$  for any other characteristic root  $\lambda_j$ . Given a prime ideal  $\mathfrak{p}$  of  $\mathcal{O}$ , we say that  $\lambda_i$  is *dominant with respect to the valuation*  $v_{\mathfrak{p}}$  if  $v_{\mathfrak{p}}(\lambda_i) \leq v_{\mathfrak{p}}(\lambda_j)$  for any other characteristic root  $\lambda_j$ .

## 3 THE MSTV CLASS AND LRS MODULARITY

We introduce a class of LRS for which the Skolem Problem is decidable, named after the authors of the seminal papers [21, 31], in which decidability is established:

**Definition 3.1.** The *Mignotte-Shorey-Tijdeman-Vereshchagin (MSTV) class* consists of those  $\mathbb{Z}$ -LRS that either have at most three dominant roots in modulus or at most two dominant roots with respect to some  $p$ -adic valuation.

The MSTV class includes all integer LRS of order at most 4 and all integer LRS of order 5 with repeated characteristic roots. It also includes all order-5 LRS that have 5 characteristic roots of equal modulus.<sup>5</sup>

<sup>5</sup>The last assertion can be seen as follows. Consider a non-degenerate order-5 LRS with characteristic roots  $\lambda_1, \overline{\lambda_1}, \lambda_2, \overline{\lambda_2}, \lambda_3$ , all having the same modulus. By rescaling,

The following notion plays an instrumental role in our paper:

**Definition 3.2.** A  $\mathbb{Z}$ -LRS  $\langle u_n \rangle_{n=0}^\infty$  is *modular* if there exists an integer  $m \geq 2$  such that there are only finitely many values of  $n \in \mathbb{N}$  with  $u_n \equiv 0 \pmod{m}$ .

*Remark 3.3.* Note that, contrary to membership in the MSTV class, in general we do not know how to decide whether a given LRS is modular or not. However, if a non-degenerate LRS  $\vec{u}$  is *known* (or *assumed*) to be modular, then it is straightforward to compute its set of zeros: first find a relevant value of  $m$  through enumeration (bearing in mind that the sequence of residues modulo  $m$  is always eventually periodic), from which one can elicit an upper bound  $N$  beyond which  $\vec{u}$  is guaranteed not to vanish, and then evaluate  $\vec{u}$  at all indices less than or equal to  $N$ .

*Example 3.4.* Consider the  $\mathbb{Z}$ -LRS  $\vec{u} = \langle u_n \rangle_{n=0}^\infty$  given by

$$u_n = \frac{1}{6}((-4 + 7i)^n + (-4 - 7i)^n) + \frac{1}{3}((8 + i)^n + (8 - i)^n) - 31,$$

or equivalently

$$u_{n+5} = 9u_{n+4} - 10u_{n+3} + 522u_{n+2} - 4745u_{n+1} + 4225u_n$$

with initial values (for  $n = 0, 1, 2, 3, 4$ ) of  $\langle -30, -27, 0, 469, 1762 \rangle$ .

The five characteristic roots of  $\vec{u}$  are

$$\lambda_1 = -4 + 7i, \bar{\lambda}_1 = -4 - 7i, \lambda_2 = 8 + i, \bar{\lambda}_2 = 8 - i, \lambda_3 = 1.$$

One can verify that  $\vec{u}$  is non-degenerate. Moreover, the two Gaussian primes  $p_1 = -2 + i$  and  $p_2 = 3 - 2i$  serve as building blocks for the non-real characteristic roots:  $\lambda_1 = p_1 p_2$  and  $\lambda_2 = -p_1 \bar{p}_2$ . From this one can show that, for any prime ideal  $\mathfrak{p}$  of the ring of integers of  $\mathbb{Q}(\lambda_1, \bar{\lambda}_1, \lambda_2, \bar{\lambda}_2, \lambda_3)$ , there are always at least three dominant characteristic roots in  $\mathfrak{p}$ -adic absolute value. Finally, we have

$$|\lambda_1| = |\bar{\lambda}_1| = |\lambda_2| = |\bar{\lambda}_2| = \sqrt{65} > \lambda_3 = 1.$$

In other words,  $\vec{u}$  does not belong to the MSTV class.

Now given any  $m \geq 2$  with  $\gcd(m, 4225) = 1$ , the fact that  $u_2 = 0$  entails that there are infinitely many values of  $n$  such that  $u_n \equiv 0 \pmod{m}$ . Nevertheless,  $\vec{u}$  is modular, and one can check (preferably using a computer!) that the smallest relevant value is  $m = 12625 = 101 \cdot 5^3$ . Finally, one can then in turn verify that the only zero of  $\vec{u}$  occurs at index  $n = 2$ .

It might be tempting to speculate that any order-5  $\mathbb{Z}$ -LRS not in the MSTV class is necessarily modular, but unfortunately that is not the case: consider the LRS  $\vec{v}$  defined by  $v_n := 5^n 13^n u_n$  (with  $\langle u_n \rangle_{n=0}^\infty$  as above). One easily verifies that  $v_n$  is not in the MSTV class. Moreover, since  $4225 = 5^2 13^2$ , and making use of our earlier observations about  $u_n$ , it follows straightforwardly that  $\vec{v}$  fails to be modular.

Observe that  $\vec{v}$  is, however, *reducible*: it comprises  $\vec{u}$  as a factor, and moreover such factors can always be effectively extracted.<sup>6</sup>

we may assume that  $|\lambda_1| = |\lambda_2| = \lambda_3 = 1$ . From the fact that all the Galois conjugates of  $\lambda_1/\lambda_2$  have modulus one and that  $\lambda_1/\lambda_2$  is not a root of unity, there exists a prime ideal  $\mathfrak{p}$  such that  $v_{\mathfrak{p}}(\lambda_1) \neq v_{\mathfrak{p}}(\lambda_2)$ . Since  $v_{\mathfrak{p}}(\lambda_1) = -v_{\mathfrak{p}}(\bar{\lambda}_1)$  and  $v_{\mathfrak{p}}(\lambda_2) = -v_{\mathfrak{p}}(\bar{\lambda}_2)$ , there are at most two dominant roots among  $\{\lambda_1, \lambda_2, \bar{\lambda}_1, \bar{\lambda}_2\}$  with respect to  $v_{\mathfrak{p}}(\cdot)$  and the dominant root(s) have a non-zero valuation. Hence there are still at most two dominant roots among all five characteristic roots.

<sup>6</sup>Technically speaking, *irreducibility* is the assertion that no rational prime divides all of the characteristic roots in the relevant subring of algebraic integers—see Section 5 for details.

The crux of our main conditional decidability result, presented in Section 5, proceeds as follows: assuming the Skolem Conjecture, we show that any irreducible order-5  $\mathbb{Z}$ -LRS that does not belong to the MSTV class is necessarily modular.

Our main unconditional decidability result, according to which reversible  $\mathbb{Z}$ -LRS of order 7 or below have a decidable Skolem Problem, is obtained as follows: we show that any such non-degenerate LRS must necessarily have at most three characteristic roots of maximum modulus, whence the LRS automatically belongs to the MSTV class. This is achieved in Section 4.1 mainly via Galois-theoretic combinatorial arguments.

Finally, we present below an instance of a non-degenerate, irreducible, order-6 LRS that does not belong to the MSTV class and is not modular. This suggests that the techniques deployed in Section 5 do not immediately appear to extend to LRS of orders 6 and beyond.

*Example 3.5.* Write

$$\lambda_1 = 1 + 2i, \lambda_2 = \frac{3}{2} + \frac{1}{2}\sqrt{-11}, \lambda_3 = \frac{1}{2} + \frac{1}{2}\sqrt{-19},$$

and let  $u_n = \lambda_1^n + \bar{\lambda}_1^n + \lambda_2^n + \bar{\lambda}_2^n - 2\lambda_3^n - 2\bar{\lambda}_3^n$ ; or equivalently,

$$u_{n+6} = 6u_{n+5} - 26u_{n+4} + 66u_{n+3} - 130u_{n+2} + 150u_{n+1} - 125u_n$$

with initial values (for  $n = 0, 1, 2, 3, 4, 5$ ) of  $\langle 0, 3, 11, -12, -125, -177 \rangle$ .

The proofs that  $\vec{u}$  does not belong to the MSTV class and fails to be modular are somewhat technical, and are deferred to Appendix A.

## 4 REVERSIBLE LINEAR RECURRENCE SEQUENCES

Recall that an integer  $\mathbb{Z}$ -LRS  $\vec{u} = \langle u_n \rangle_{n=0}^\infty$  is *reversible* if its bi-completion  $\overleftarrow{\vec{u}} = \langle u_n \rangle_{n=-\infty}^\infty$  lies entirely in  $\mathbb{Z}$ . As noted in Section 1.1,  $\vec{u}$  is reversible if and only if its characteristic polynomial

$$g(X) = X^k - a_1 X^{k-1} - \dots - a_k$$

has the property that  $a_k = \pm 1$ . This in turn entails, by Vieta's formula, that the product of the characteristic roots (including repetitions) is equal to  $\pm 1$ , and therefore all characteristic roots are units in the ring of algebraic integers. Conversely, if all characteristic roots of a given  $\mathbb{Z}$ -LRS are units, then so is their product, which is equal (up to a sign) to the constant term of the characteristic polynomial of the  $\mathbb{Z}$ -LRS. The latter being a rational integer, it is  $\pm 1$ , and the original  $\mathbb{Z}$ -LRS is therefore reversible.

Let  $\vec{u}$  be a degenerate reversible  $\mathbb{Z}$ -LRS, and let  $L$  be the least common multiple of the orders of the roots of unity appearing among quotients of distinct characteristic roots. Then for any  $m \in \{0, \dots, L-1\}$ , the subsequence  $\langle u_{Ln+m} \rangle_{n=0}^\infty$  is non-degenerate, of order at most that of  $\vec{u}$ , and moreover its characteristic roots are  $L$ -th powers of the characteristic roots of  $\vec{u}$ . In particular, the characteristic roots of each such derived subsequence are units, and the subsequence is therefore also reversible.

### 4.1 The Skolem Problem for Reversible LRS

In this section, we establish the following:

**Theorem 4.1.** *The Skolem Problem for reversible  $\mathbb{Z}$ -LRS of order 7 or less is decidable.*

As noted earlier, it is sufficient to restrict our attention to non-degenerate reversible LRS.<sup>7</sup> The proof of Theorem 4.1 proceeds by showing that any non-degenerate reversible  $\mathbb{Z}$ -LRS of order 7 or less has at most three characteristic roots of maximum modulus, and therefore belongs to the MSTV class.

In order to do so, we show in Propositions 4.2, 4.3, and 4.4, that no monic polynomial  $g \in \mathbb{Z}[X]$  that has degree at most 7 and constant term  $\pm 1$  satisfies the following two properties:

- (H1)  $g$  has at least four distinct roots of maximum modulus;
- (H2) no quotient of two distinct roots of  $g$  is a root of unity.

**Proposition 4.2.** *No monic polynomial  $g \in \mathbb{Z}[X]$  of degree at most 5 and constant term  $\pm 1$  satisfies both (H1) and (H2).*

**PROOF.** We suppose that  $g$  exists with the given properties and derive a contradiction. By (H1),  $g$  has at least four distinct dominant roots<sup>8</sup> and, by (H2), at most one of the dominant roots is real. Hence the dominant roots include two complex-conjugate pairs  $\lambda_1, \bar{\lambda}_1$  and  $\lambda_2, \bar{\lambda}_2$ . The dominant roots of  $g$  must have modulus strictly greater than 1 for otherwise, by an old result of Kronecker [20], the roots of  $g$  would all be roots of unity, which is precluded by (H2).

Each root of  $g$  is a unit in the ring of algebraic integers and hence has norm<sup>9</sup> equal to  $\pm 1$ . In particular, the dominant root  $\lambda_1$  must have a Galois conjugate of modulus less than 1, that is, there is a non-dominant root  $\lambda_3$  of  $g$  and an automorphism  $\sigma$  of the splitting field of  $g$  such that  $\sigma(\lambda_1) = \lambda_3$ . Thus  $g$  has degree 5 and

$$|\sigma(\lambda_1)| < |\sigma(\bar{\lambda}_1)| = |\sigma(\lambda_2)| = |\sigma(\bar{\lambda}_2)|. \quad (3)$$

Now the dominant roots of  $g$  satisfy the equation  $\lambda_1 \bar{\lambda}_1 = \lambda_2 \bar{\lambda}_2$ . For the automorphism  $\sigma$ , mentioned above, we thus have  $\sigma(\lambda_1)\sigma(\bar{\lambda}_1) = \sigma(\lambda_2)\sigma(\bar{\lambda}_2)$ . But this clearly contradicts Equation (3).  $\square$

**Proposition 4.3.** *No monic degree-6 polynomial  $g \in \mathbb{Z}[X]$  with constant term  $\pm 1$  satisfies both (H1) and (H2).*

**PROOF.** Once again we suppose that  $g$  exists with the given properties and derive a contradiction. Let  $G$  denote the group of automorphisms of the splitting field of  $g$ .

Arguing exactly as in the proof of Proposition 4.2, the dominant roots of  $g$  include two complex-conjugate pairs  $\lambda_1, \bar{\lambda}_1$  and  $\lambda_2, \bar{\lambda}_2$ , such that

$$\lambda_1 \bar{\lambda}_1 = \lambda_2 \bar{\lambda}_2. \quad (4)$$

A fact which we will use repeatedly is that, by Equation (4), for all  $\sigma \in G$ ,

$$\sigma(\lambda_1)\sigma(\bar{\lambda}_1) = \sigma(\lambda_2)\sigma(\bar{\lambda}_2). \quad (5)$$

Since  $\lambda_1$  has norm  $\pm 1$ , being a unit, it must have a Galois conjugate of modulus less than 1, that is, there exists a non-dominant root  $\lambda_3$  of  $g$  and  $\sigma \in G$  with  $\sigma(\lambda_1) = \lambda_3$ . Now we cannot have

<sup>7</sup>In decomposing a given LRS into a collection of non-degenerate LRS, the order of the non-degenerate LRS thus obtained is always at most the order of the original LRS.

<sup>8</sup>Throughout Section 4.1, *dominant root* refers to roots of maximal modulus.

<sup>9</sup>The *norm* of an algebraic integer is the product of its Galois conjugates (including itself).

both  $\sigma(\lambda_1)$  and  $\sigma(\bar{\lambda}_1)$  non-dominant, since then  $\sigma(\lambda_2)$  and  $\sigma(\bar{\lambda}_2)$  would have to lie among the four dominant roots, and Equation (5) would fail. We thus must have that  $\sigma(\bar{\lambda}_1)$  is dominant, and exactly one element of  $\{\sigma(\lambda_2), \sigma(\bar{\lambda}_2)\}$  is dominant, with other elements having the same modulus as  $\lambda_3$ . In particular, there is a second non-dominant root, having the same modulus as  $\lambda_3$ . By Condition (H2) the two non-dominant roots cannot both be real; thus we have two non-dominant complex roots  $\lambda_3$  and  $\bar{\lambda}_3$ . Taking into account  $\sigma, \sigma^{-1}$  and complex conjugation, it is now clear that  $G$  acts transitively on the roots of  $g$ ; hence  $g$  is irreducible.

Since  $g$  is irreducible, its roots all have degree 6. It follows that the order of  $G$  (which is equal to the degree of the splitting field of  $g$  over  $\mathbb{Q}$ ) is divisible by 3. In particular, by Cauchy's theorem,  $G$  contains an element  $\sigma$  of order 3. Such an element is either a 3-cycle or is the product of two disjoint 3-cycles. We claim that in fact  $\sigma$  is a product of two disjoint 3-cycles  $(D_1 D_2 \lambda_3)$  and  $(D_3 D_4 \bar{\lambda}_3)$ , where  $D_1, \dots, D_4$  is a list of the dominant roots.

To prove the claim, suppose for a contradiction that  $\sigma$  contains a 3-cycle of dominant roots. By renaming roots we can assume without loss of generality that the cycle is  $(\lambda_1 \lambda_2 \bar{\lambda}_1)$ . Since  $\sigma$  maps each of  $\lambda_1, \lambda_2, \bar{\lambda}_1$  to a dominant root, by Equation (5) we have that  $\sigma(\bar{\lambda}_2)$  is also dominant, i.e.,  $\sigma$  fixes  $\bar{\lambda}_2$ . Now Equation (5) gives  $\lambda_2 \lambda_1 = \bar{\lambda}_1 \bar{\lambda}_2$ . Multiplying the previous equation by  $\bar{\lambda}_1 \lambda_2$  and dividing by Equation (4) gives  $\lambda_2^2 = \bar{\lambda}_1^2$ , contradicting Condition (H2). We now have that a cycle decomposition of  $\sigma$  must contain a 3-cycle of the form  $(D_1 D_2 \lambda_3)$ , with  $D_1, D_2$  dominant roots. Again, by Equation (5),  $\sigma$  must map one of the remaining two dominant roots to  $\bar{\lambda}_3$ . Thus  $\sigma$  be the product of two 3-cycles, necessarily of the form indicated in the claim.

We now consider two cases according to the cycle decomposition of  $\sigma$ .

*Case 1.* Suppose that some 3-cycle in  $\sigma$  contains a pair of complex-conjugate roots of  $g$ . By renaming dominant roots, we can assume without loss of generality that  $\sigma = (\lambda_1 \bar{\lambda}_1 \lambda_3)(\lambda_2 \bar{\lambda}_2 \bar{\lambda}_3)$ . Applying  $\sigma$  twice to Equation (4) we successively get that  $\bar{\lambda}_1 \lambda_3 = \bar{\lambda}_2 \bar{\lambda}_3$  and  $\lambda_3 \lambda_1 = \bar{\lambda}_3 \lambda_2$ . Next, multiplying the two previous equations and dividing the result by Equation (4), we get  $\lambda_3^2 = \bar{\lambda}_3^2$ . This contradicts Condition (H2). Thus Case 1 leads to a contradiction.

*Case 2.* Suppose that neither 3-cycle in the decomposition of  $\sigma$  contains a pair of complex-conjugate roots. By renaming dominant roots we can write  $\sigma = (\lambda_1 \lambda_2 \lambda_3)(\bar{\lambda}_2 \bar{\lambda}_1 \bar{\lambda}_3)$  without loss of generality (since it is not possible for both  $\lambda_2$  and  $\bar{\lambda}_2$  both to map to non-dominant roots, otherwise Equation (5) would be violated). By Equation (5) we have

$$\lambda_2 \bar{\lambda}_3 = \lambda_3 \bar{\lambda}_1 \quad (6)$$

We now consider two sub-cases:

**Sub-case 2.1.** Suppose that  $\lambda_1 \lambda_2 \lambda_3$  is a root of unity. Multiplying Equation (6) by  $\lambda_1 \lambda_3^2$  we have

$$\lambda_1 \lambda_2 \lambda_3 \bar{\lambda}_3 \lambda_3 = \lambda_1 \bar{\lambda}_1 \lambda_3^3. \quad (7)$$

Dividing Equation (7) by its complex-conjugate equation gives

$$(\lambda_1 \lambda_2 \lambda_3) / (\bar{\lambda}_1 \bar{\lambda}_2 \bar{\lambda}_3) = (\lambda_3 / \bar{\lambda}_3)^3,$$

which implies that  $\lambda_3/\overline{\lambda_3}$  is a root of unity, contradicting Condition (H2).

**Sub-case 2.2.** Suppose that  $\lambda_1\lambda_2\lambda_3$  is not a root of unity. Then  $\lambda_1\lambda_2\lambda_3$  has a Galois conjugate of modulus strictly greater than 1. But for a product of three roots of  $g$  to have modulus greater than 1 they must all be dominant, since the product of two dominant and one non-dominant root has modulus 1. Hence there exists  $\tau \in G$  such that  $\tau(\lambda_1)$ ,  $\tau(\lambda_2)$ , and  $\tau(\lambda_3)$  are all dominant. Then, since at most one of  $\tau(\lambda_1)$  and  $\tau(\lambda_2)$  can be dominant, in order for  $\tau$  to preserve Equation (4) we would need that both  $\tau(\lambda_1)$  and  $\tau(\lambda_2)$  be non-dominant, and  $\tau(\lambda_3)$  be dominant. But then  $\tau$  does not preserve Equation (6). Again we arrive at a contradiction.

We have thus given an exhaustive case analysis, with all cases leading to a contradiction. We conclude that there does not exist a polynomial  $g$  with the properties stated in the proposition.  $\square$

**Proposition 4.4.** *No monic degree-7 polynomial  $g \in \mathbb{Z}[X]$  with constant term  $\pm 1$  satisfies both (H1) and (H2).*

**PROOF.** Once more we suppose that  $g$  exists with the given properties and derive a contradiction. Let  $G$  denote the Galois group of  $g$ .

Exactly as in the proof of Proposition 4.3, one shows that there are two pairs of complex-conjugate dominant roots  $\lambda_1, \overline{\lambda_1}$  and  $\lambda_2, \overline{\lambda_2}$  and a pair of non-dominant complex conjugate roots  $\lambda_3, \overline{\lambda_3}$  such that the six mentioned roots are contained in a single orbit of the Galois group of  $g$  (and hence are all roots of the same irreducible factor of  $g$ ). But the six mentioned roots cannot be the roots of a degree-6 factor of  $g$ , since this factor would violate Proposition 4.3. Hence  $g$  must be irreducible.

Let  $\lambda_4 \in \mathbb{R}$  be the remaining root of  $g$ . We note that  $\lambda_4$  is not dominant. Indeed, if  $\lambda_4$  were dominant then we would have the equation

$$\lambda_1\overline{\lambda_1} = \lambda_2\overline{\lambda_2} = \lambda_4^2.$$

But consider the image of this equation under  $\sigma \in G$  such that  $\sigma(\lambda_1) = \lambda_3$ . In order to preserve the equation we would need one of  $\sigma(\lambda_2)$  and  $\sigma(\overline{\lambda_2})$  to be non-dominant and  $\sigma(\lambda_4)$  to also be non-dominant, that is, we would need the image of  $\sigma$  to contain three non-dominant roots—which is impossible. We conclude that  $\lambda_4$  is non-dominant.

Since  $g$  is irreducible, the order of  $G$  is divisible by 7 and hence, by Cauchy's theorem,  $G$  contains an element  $\sigma$  of order 7. Now  $\sigma$  induces a 7-cycle of the roots of  $g$ . Let  $D_1, \dots, D_4$  be a list of the dominant roots of  $g$  in some arbitrary order and likewise  $N_1, N_2, N_3$  a list of the non-dominant roots. To preserve the equation  $\lambda_1\overline{\lambda_1} = \lambda_2\overline{\lambda_2}$ ,  $\sigma$  must map exactly two dominant roots to non-dominant roots. This yields (up to cyclic symmetry) three possible patterns for  $\sigma$ . We show that each case leads to a contradiction.

**Case 1.** Suppose  $\sigma = (D_1 D_2 N_1 N_2 D_3 D_4 N_3)$ . Then  $\sigma^2$  maps three dominant roots (namely  $D_1, D_2, D_3$ ) to non-dominant roots and hence does not preserve the equation  $\lambda_1\overline{\lambda_1} = \lambda_2\overline{\lambda_2}$ . Thus there cannot be an automorphism of this form.

**Case 2.** Let  $\sigma$  be either  $(D_1 D_2 D_3 N_1 D_4 N_2 N_3)$  or  $(D_1 D_2 D_3 N_1 N_2 D_4 N_3)$ . In both of these cases (as can be seen by enumerating all powers of  $\sigma$ ), every one of the six pairs of

dominant roots is mapped by some power of  $\sigma$  to a pair of non-dominant roots. In particular, some power of  $\sigma$  maps both  $\lambda_1$  and  $\overline{\lambda_1}$  to non-dominant roots and hence does not preserve the equation  $\lambda_1\overline{\lambda_1} = \lambda_2\overline{\lambda_2}$  (for only one of  $\lambda_2$  and  $\overline{\lambda_2}$  can be sent to the remaining non-dominant root).  $\square$

## 4.2 Hard Instances at Order 8

We conclude our discussion of reversible linear recurrence sequences by showing that the preceding string of propositions cannot be extended further, that is, there exists a family of degree-8 polynomials all having constant term 1 and satisfying properties (H1) and (H2). In turn, this enables us to exhibit a family of non-degenerate reversible  $\mathbb{Z}$ -LRS that do not belong to the MSTV class.

Fix non-zero integers  $a, b$  with  $a \neq \pm b$  and let  $\rho = \sqrt{2} + 1$  (more generally, the construction below works for  $\rho$  any real quadratic unit greater than 1). Let  $k$  be an even positive integer parameter. Writing

$$g_1(X) := (X^2 - aX + \rho^k)(X^2 - bX + \rho^k),$$

the roots of  $g_1$  are

$$\frac{a \pm \sqrt{a^2 - 4\rho^k}}{2} \quad \text{and} \quad \frac{b \pm \sqrt{b^2 - 4\rho^k}}{2}.$$

For  $k$  sufficiently large,  $g_1$  has four complex roots, all with modulus  $\rho^{k/2}$ .

Now write  $g(X) := g_1(X)g_2(X)$ , where

$$g_2(X) := (X^2 - aX + \rho^{-k})(X^2 - bX + \rho^{-k}).$$

Noting that the Galois conjugate of  $\rho$  is  $-\rho^{-1}$ , since  $k$  is even we have that  $g$  is an integer polynomial of degree 8 and constant term 1. The roots of  $g_2$  are

$$\frac{a \pm \sqrt{a^2 - 4\rho^{-k}}}{2} \quad \text{and} \quad \frac{b \pm \sqrt{b^2 - 4\rho^{-k}}}{2}.$$

Thus, for suitably large  $k$ , the roots of  $g_2$  have modulus at most  $\max(|a|, |b|) < \rho^{k/2}$  and so  $g$  has four roots of maximal modulus, i.e.,  $g$  satisfies (H1).

It remains to observe that  $g$  satisfies (H2) for all but finitely many choices of  $k$ . Indeed, for sufficiently large  $k$  the non-dominant roots (in modulus) are real and all pairwise distinct. Meanwhile, the arguments of the roots of maximal modulus converge to  $\pm\pi$  as  $k$  tends to infinity. Since for every  $k$  the roots of maximal modulus lie in a field of degree at most 8 and since there are only finitely many roots of unity of degree at most 8, the desired result follows.

Let us furthermore observe that, for any prime ideal  $\mathfrak{p}$  the valuation function  $v_{\mathfrak{p}}(\cdot)$  evaluates to 0 on every root of  $g$  (as the latter are all units). Consequently,  $g$  always has eight dominant roots in  $\mathfrak{p}$ -adic absolute value. It is now straightforward to manufacture a family of non-degenerate reversible order-8  $\mathbb{Z}$ -LRS that do not belong to the MSTV class, with such LRS having instantiations of  $g$  as characteristic polynomials. We construct one such simple instance below.

**Example 4.5.** Let  $\rho = \sqrt{2} + 1$  as earlier, and write:

$$\lambda_1 = \frac{1 + \sqrt{1 - 4\rho^2}}{2} \quad \text{and} \quad \lambda_2 = \frac{2 + \sqrt{4 - 4\rho^2}}{2}.$$

The characteristic roots of maximum modulus will be  $\lambda_1, \overline{\lambda_1}, \lambda_2$ , and  $\overline{\lambda_2}$ . The other four (real) roots are

$$r_1 = \frac{1 + \sqrt{1 - 4\rho^{-2}}}{2}, \tilde{r}_1 = \frac{1 - \sqrt{1 - 4\rho^{-2}}}{2},$$

$$r_2 = \frac{2 + \sqrt{4 - 4\rho^{-2}}}{2}, \text{ and } \tilde{r}_2 = \frac{2 - \sqrt{4 - 4\rho^{-2}}}{2}.$$

Let

$$u_n = \sqrt{2} \left( \lambda_1^n + \overline{\lambda_1}^n + 2\lambda_2^n + 2\overline{\lambda_2}^n - r_1^n - \tilde{r}_1^n - 2r_2^n - 2\tilde{r}_2^n \right).$$

Equivalently, write:

$$u_{n+8} = 6u_{n+7} - 25u_{n+6} + 66u_{n+5} - 120u_{n+4} + 150u_{n+3} - 89u_{n+2} + 18u_{n+1} - u_n,$$

with initial values (for  $n = 0, \dots, 7$ ) of  $\langle 0, 0, -48, -120, 0, 520, 624, -2016 \rangle$ .

$\vec{u} = \langle u_n \rangle_{n=0}^\infty$  has zeros at indices 0, 1, and 4. It does not belong to the MSTV class, it is irreducible (as all roots are units), and it is not modular.<sup>10</sup>

## 5 THE SKOLEM PROBLEM AT ORDER 5

In this section, we prove the following result:

**Theorem 5.1.** *The Skolem Problem for  $\mathbb{Z}$ -LRS of order 5 is decidable, assuming the Skolem Conjecture.*

*Remark 5.2.* In fact, the proof of Theorem 5.1 only requires that the Skolem Conjecture hold for  $\mathbb{Z}$ -LRBS of order 5. Let us also record an immediate corollary, to the effect that the Skolem Problem for  $\mathbb{Q}$ -LRS of order 5 is also decidable subject to the Skolem Conjecture.

**PROOF.** As noted in Section 3, it suffices to consider non-degenerate order-5  $\mathbb{Z}$ -LRS that do not belong to the MSTV class. Let  $\vec{u} = \langle u_n \rangle_{n=0}^\infty$  be such an LRS. Then  $\vec{u}$  has five distinct non-zero characteristic roots  $\lambda_1, \dots, \lambda_5$ , with  $\lambda_2 = \overline{\lambda_1}, \lambda_4 = \overline{\lambda_3}, \lambda_5$  real, and such that

$$\lambda_1 \lambda_2 = \lambda_3 \lambda_4 > \lambda_5^2. \quad (8)$$

Define  $K := \mathbb{Q}(\lambda_1, \dots, \lambda_5)$  to be the field generated by the characteristic roots and write  $\mathcal{O}$  for the ring of integers of  $K$ . By rescaling, we may assume that the coefficients  $\alpha_1, \dots, \alpha_5$  in the exponential-polynomial representation  $u_n = \sum_{i=1}^5 \alpha_i \lambda_i^n$  all lie in  $\mathcal{O}$ .

Write  $d = [K : \mathbb{Q}]$  for the degree of  $K$  over  $\mathbb{Q}$ . By dividing  $\vec{u}$  into  $d$  subsequences  $\langle u_{dn+r} \rangle_{n=0}^\infty$ , for  $r = 0, \dots, d-1$ , and separately considering each subsequence, we can assume without loss of generality that for each characteristic root  $\lambda$  of  $\vec{u}$ , every valuation  $v_p(\lambda)$  is an integer multiple of  $d$ . Furthermore, by rescaling we may assume that there is no rational prime that divides all the characteristic roots (in  $\mathcal{O}$ ).

We now show, assuming the Skolem Conjecture, that  $\vec{u}$  is necessarily modular, from which decidability follows. We proceed by contradiction and suppose that  $\vec{u}$  is *not* modular.

As established in the proof of Proposition 4.2, the set  $\{\lambda_1, \dots, \lambda_4\}$  of roots that are dominant in modulus is invariant under every automorphism of  $K$ . It follows that all Galois conjugates of  $\lambda_1/\lambda_2$

<sup>10</sup> Any reversible LRS that has a zero (or whose bi-completion has a zero) will necessarily fail to be modular, since for any given integer  $m \geq 2$ , the sequence of residues modulo  $m$  is always periodic.

have modulus one. Since  $\lambda_1/\lambda_2$  is not a root of unity there is a prime ideal  $\mathfrak{p}$  of  $\mathcal{O}$  such that  $v_{\mathfrak{p}}(\lambda_1) \neq v_{\mathfrak{p}}(\lambda_2)$ . Without loss of generality we may suppose that  $\mathfrak{p}$  divides  $\lambda_1$ . We further claim that  $\mathfrak{p}$  does not divide every characteristic root. Indeed, suppose for a contradiction that  $\mathfrak{p}$  does divide every characteristic root. Let  $p$  lie above the rational prime  $p$  and denote by  $e$  its ramification index. Since the Galois group of  $K$  acts transitively on the prime ideals above  $p$  we get that every prime ideal above  $p$  divides every characteristic root. In turn this implies that every prime ideal above  $p$  divides every characteristic root to order at least  $d \geq e$ . But this implies that the rational integer  $p$  divides every characteristic root, contrary to our assumption. Thus the claim is proved.

By the previous claim we have that every  $p$ -adically dominant characteristic root  $\lambda_i$  has  $v_{\mathfrak{p}}(\lambda_i) = 0$ . We also know that  $v_{\mathfrak{p}}(\lambda_1) > 0$ . Furthermore, by the assumption that  $\vec{u}$  is not in the MSTV class we know that at least three characteristic roots have zero valuation with respect to  $v_{\mathfrak{p}}$ . By Equation (8), we may thus suppose that  $v_{\mathfrak{p}}(\lambda_1) = v_{\mathfrak{p}}(\lambda_3) > 0$  and  $v_{\mathfrak{p}}(\lambda_2) = v_{\mathfrak{p}}(\lambda_4) = v_{\mathfrak{p}}(\lambda_5) = 0$ .

Write  $f$  for the residual degree of  $\mathfrak{p}$  (i.e., such that  $\mathcal{O}/\mathfrak{p}$  has order  $p^f$ ). For some residue class  $r \in \{0, 1, \dots, p^f - 2\}$  we have that the subsequence  $\langle u_{(p^f-1)n+r} \rangle_{n=0}^\infty$  fails to be modular (otherwise it is easily seen that  $\langle u_n \rangle_{n=0}^\infty$  itself would be modular). For this value of  $r$ , define  $\vec{w} = \langle w_n \rangle_{n=0}^\infty$  by  $w_n := u_{(p^f-1)n+r}$ . Since  $\vec{w}$  is not modular, it follows that for every  $k \in \mathbb{N}$  there exists  $r_k \in \{0, 1, \dots, p^k - 1\}$  such that

$$\forall m \geq 2 \exists^\infty n \in r_k + p^k \mathbb{N} : w_n \equiv 0 \pmod{m}. \quad (9)$$

(To see this, fix  $k \in \mathbb{N}$ , and consider the infinite sequence of indices corresponding to null residues modulo  $m = \ell!$ . By the pigeonhole principle, some  $r_k$  will emerge for infinitely many instances of  $\ell!$ , and will thus satisfy Equation 9 for all  $m \geq 2$ .)

Setting  $m = p^k$  in (9), we see that for all  $k$  there exists  $n_k \in r_k + p^k \mathbb{N}$  such that  $n_k \geq k$  and  $w_{n_k} \equiv 0 \pmod{p^k}$ .

By the Skolem Conjecture and (9), for all  $k$  there exists  $n \in r_k + p^k \mathbb{Z}$  such that  $w_n = 0$ ; that is, for all  $k$ ,  $n_k$  is congruent modulo  $p^k$  to some integer zero of  $\vec{w}$ . Since  $\vec{w}$  has only finitely many integer zeros (by the Skolem-Mahler-Lech theorem), there must exist  $x \in \mathbb{Z}$  such that  $w_x = 0$  and  $x \equiv n_k \pmod{p^k}$  for infinitely many  $k$ .

Now for all  $n \in \mathbb{Z}$  we have

$$w_n = \beta_1 \lambda_1^{(p^f-1)n} + \beta_3 \lambda_3^{(p^f-1)n} + \beta_2 \lambda_2^{(p^f-1)n} + \beta_4 \lambda_4^{(p^f-1)n} + \beta_5 \lambda_5^{(p^f-1)n},$$

for some non-zero  $\beta_1, \dots, \beta_5 \in \mathcal{O}$ . In particular, for  $x$  as above, we have

$$w_x = \beta_1 \lambda_1^{(p^f-1)x} + \beta_2 \lambda_2^{(p^f-1)x} + \dots + \beta_5 \lambda_5^{(p^f-1)x} = 0. \quad (10)$$

On the other hand, since  $w_{n_k} \equiv 0 \pmod{p^k}$ ,  $n_k \geq k$ , and  $v_{\mathfrak{p}}(\lambda_1), v_{\mathfrak{p}}(\lambda_3) > 0$ , we have

$$\beta_2 \lambda_2^{(p^f-1)n_k} + \beta_4 \lambda_4^{(p^f-1)n_k} + \beta_5 \lambda_5^{(p^f-1)n_k} \equiv 0 \pmod{p^k}. \quad (11)$$

Since  $\mathcal{O}/\mathfrak{p}$  is a finite field of order  $p^f$ , we have that  $\lambda_2^{p^f-1}, \lambda_4^{p^f-1}, \lambda_5^{p^f-1} \in 1 + \mathfrak{p}$ . Now for all  $z \in 1 + \mathfrak{p}$  and  $n_1, n_2 \in \mathbb{Z}$  by the Binomial theorem we have that if  $n_1 \equiv n_2 \pmod{p^k}$  then

$z^{n_1} \equiv z^{n_2} \pmod{p^{k+1}}$  (see, e.g. [18, Section 4.6]). From this fact and Equation (11) we deduce that

$$\beta_2 \lambda_2^{(p^f-1)x} + \beta_4 \lambda_4^{(p^f-1)x} + \beta_5 \lambda_5^{(p^f-1)x} \equiv 0 \pmod{p^k}.$$

Since the above holds for all  $k \geq 0$ , we have

$$\beta_2 \lambda_2^{(p^f-1)x} + \beta_4 \lambda_4^{(p^f-1)x} + \beta_5 \lambda_5^{(p^f-1)x} = 0. \quad (12)$$

Comparing (10) and (12) we have

$$\beta_1 \lambda_1^{(p^f-1)x} + \beta_3 \lambda_3^{(p^f-1)x} = 0.$$

Applying complex conjugation to this equation gives

$$\beta_2 \lambda_2^{(p^f-1)x} + \beta_4 \lambda_4^{(p^f-1)x} = 0.$$

The previous equation together with Equation (12) entail that  $\beta_5 \lambda_5^{(p^f-1)x} = 0$ , which is impossible, since  $\beta_5, \lambda_5 \neq 0$ .  $\square$

*Remark 5.3.* It is worth noting in the above proof that the Skolem Conjecture is invoked infinitely many times, although only ever on order-5 LRBS that do not belong to the MSTV class. It would therefore suffice to prove this particular special case of the Skolem Conjecture to obtain the unconditional decidability of the Skolem Problem at order 5.

We now outline a simple procedure for computing the set of all zeros of a given non-degenerate order-5  $\mathbb{Z}$ -LRS  $\vec{u} = \langle u_n \rangle_{n=0}^\infty$ . We assume that  $\vec{u}$  does not belong to the MSTV class, otherwise one simply proceeds using one of the known existing algorithms [21, 31].

Write  $u_n = \sum_{i=1}^5 \alpha_i \lambda_i^n$ , where each  $\lambda_i$  is an algebraic integer. Let  $C \in \mathbb{N}$  be the largest rational integer such, for each  $i$ , one can write  $\lambda_i = C \mu_i$ , with  $\mu_i$  an algebraic integer. And let  $D \geq 1$  be the smallest rational integer such that, for each  $i$ ,  $\gamma_i := D \alpha_i$  is an algebraic integer.

Letting

$$v_n := \frac{D}{C^n} u_n = \sum_{i=1}^5 \gamma_i \mu_i^n,$$

we see that the zero sets of the  $\mathbb{Z}$ -LRS  $\vec{u}$  and  $\vec{v}$  are the same.

The proof of Theorem 5.1 shows that, assuming the Skolem Conjecture,  $\vec{v}$  is modular. Therefore search (by enumeration) for a suitable value of  $m$ , from which one can compute a bound  $N \in \mathbb{N}$  such that, for all  $n > N$ ,  $v_n \neq 0$ .

Finally, compute the set of zeros of  $\vec{u}$  by inspecting each  $u_n$ , for  $0 \leq n \leq N$ .

Note that whenever this procedure halts, it has correctly produced the zero set of  $\vec{u}$ . The Skolem Conjecture merely ensures that the procedure always does terminate.

## ACKNOWLEDGMENTS

This work was partially funded by DFG grant 389792660 as part of TRR 248 (see <https://perspicuous-computing.science>).

## A PROOFS FOR EXAMPLE 3.5

Recall the order-6  $\mathbb{Z}$ -LRS from Example 3.5; we have

$$\lambda_1 = 1 + 2i, \lambda_2 = \frac{3}{2} + \frac{1}{2}\sqrt{-11}, \lambda_3 = \frac{1}{2} + \frac{1}{2}\sqrt{-19},$$

and  $u_n = \lambda_1^n + \overline{\lambda_1}^n + \lambda_2^n + \overline{\lambda_2}^n - 2\lambda_3^n - 2\overline{\lambda_3}^n$ ; equivalently

$$u_n = 6u_{n-1} - 26u_{n-2} + 66u_{n-3} - 130u_{n-4} + 150u_{n-5} - 125u_{n-6}$$

with initial values (for  $n = 0, 1, 2, 3, 4, 5$ ) of  $\langle 0, 3, 11, -12, -125, -177 \rangle$ .

We prove that  $\vec{u}$  is not modular and does not belong to the MSTV class. Here we assume familiarity with basic notions concerning the  $p$ -adic numbers  $\mathbb{Q}_p$ . In particular, we refer to Hensel's Lemma [5, Theorem 12.16] and the following simple fact about  $p$ -adic exponentiation (see, e.g., [18, Section 4.6]): for  $\alpha \in \mathbb{Z}_p$  and  $k \geq 0$ , the congruence  $(1 + p\alpha)^{p^k} \equiv 1 \pmod{p^{k+1}}$  holds.

Let us first show that the sequence  $\vec{u}$  fails to be modular. This will result from the following two conditions:

(M1) For every  $m$  that is not divisible by 5, the sequence  $\langle u_n \pmod{m} \rangle_{n=0}^\infty$  is periodic.

(M2) For all  $k \geq 0$  and  $\ell \geq 1$ , we have  $u_{\ell \cdot 4 \cdot 5^k} \equiv 0 \pmod{5^k}$ .

Indeed, given any  $m$  not divisible by 5, by (M1) the sequence  $\langle u_n \pmod{m} \rangle_{n=0}^\infty$  is periodic—say with period  $\ell$ . Since  $u_0 = 0$  we have that  $u_{\ell \cdot 4 \cdot 5^k} \equiv 0 \pmod{m}$  for all  $k \geq 0$ . But by (M2) we also have  $u_{\ell \cdot 4 \cdot 5^k} \equiv 0 \pmod{5^k}$ . Hence  $u_{\ell \cdot 4 \cdot 5^k} \equiv 0 \pmod{m5^k}$  for all  $k \geq 0$ . Since every positive integer has the form  $m5^k$  for some  $k \geq 0$  and  $m$  not divisible by 5, we conclude that  $\vec{u}$  is not modular.

It remains to establish (M1) and (M2). In fact, (M1) was already noted in Section 1.1. For (M2) it is convenient to work in the field  $\mathbb{Q}_5$  of 5-adic numbers and its subring  $\mathbb{Z}_5$  of 5-adic integers.

Let  $K := \mathbb{Q}(\sqrt{-1}, \sqrt{19}, \sqrt{11})$  be the field generated over  $\mathbb{Q}$  by the characteristic roots of  $\vec{u}$ . Since  $-1, 19, 11$  are all squares in  $\mathbb{Q}_5$ , we may regard  $K$  as a subfield of  $\mathbb{Q}_5$ . Indeed, by Hensel's Lemma, we can choose  $\sqrt{-1}, \sqrt{19}, \sqrt{11}$  to be elements of  $\mathbb{Z}_5$  satisfying the following congruences in  $\mathbb{Z}_5$ :  $\sqrt{-1} \equiv 3 \pmod{5}$ ,  $\sqrt{19} \equiv 2 \pmod{5}$ , and  $\sqrt{11} \equiv 1 \pmod{5}$ . This in turn leads us to identify the characteristic roots of  $\vec{u}$  as elements of  $\mathbb{Z}_5$  such that, respectively,  $\lambda_1, \lambda_2, \lambda_3 \not\equiv 0 \pmod{5}$  and  $\overline{\lambda_1}, \overline{\lambda_2}, \overline{\lambda_3} \equiv 0 \pmod{5}$ .

Since the residue field  $\mathbb{Z}_5/5\mathbb{Z}_5$  is the finite field  $\mathbb{Z}/5\mathbb{Z}$ , we have  $\lambda_1^4, \lambda_2^4, \lambda_3^4 \equiv 1 \pmod{5}$ . Hence, applying the above-mentioned fact about exponentiation, we have that for all  $\ell, k \geq 0$ ,  $\lambda_1^{\ell \cdot 4 \cdot 5^k}, \lambda_2^{\ell \cdot 4 \cdot 5^k}, \lambda_3^{\ell \cdot 4 \cdot 5^k} \equiv 1 \pmod{5^{k+1}}$ . But we also have  $\overline{\lambda_1}^{\ell \cdot 4 \cdot 5^k}, \overline{\lambda_2}^{\ell \cdot 4 \cdot 5^k}, \overline{\lambda_3}^{\ell \cdot 4 \cdot 5^k} \equiv 0 \pmod{5^k}$ . Thus, from the formula  $u_n = \lambda_1^n + \overline{\lambda_1}^n + \lambda_2^n + \overline{\lambda_2}^n - 2\lambda_3^n - 2\overline{\lambda_3}^n$ , we have that  $u_{\ell \cdot 4 \cdot 5^k} \equiv 0 \pmod{5^k}$ . This establishes (M2) and completes the proof that  $\vec{u}$  is not modular.

We now proceed to show that the sequence  $\vec{u}$  does not lie in the MSTV class. Note that if we regard the characteristic roots of  $\vec{u}$  as lying in  $\mathbb{Q}_5$ , then, as explained above, there are three dominant roots with respect to the 5-adic absolute value. Specifically we have that  $v_5(\lambda_1) = v_5(\lambda_2) = v_5(\overline{\lambda_3}) = 0$ , with the remaining roots having positive 5-adic valuation. Let us now take some prime-ideal divisor  $\mathfrak{p}$  of a characteristic root in  $K$ . Then  $\mathfrak{p}$  divides 5 since the product of every complex-conjugate pair of characteristic roots is 5. Moreover, since the characteristic polynomial of  $\vec{u}$  splits into distinct linear factors over  $\mathbb{Q}_5$  we have that  $(K, v_{\mathfrak{p}})$  can be embedded in  $(\mathbb{Q}_5, v_5)$  as a valued field (see the discussion in [5, Section 12.8]). It follows that there are also three dominant characteristic roots with respect to  $v_{\mathfrak{p}}$ . Moreover, since each characteristic root has modulus  $\sqrt{5}$ , there



are six dominant root with respect to modulus. We conclude that  $\vec{u}$  does not lie in the MSTV class.

## REFERENCES

- [1] M. Agrawal, S. Akshay, B. Genest, and P. S. Thiagarajan. 2015. Approximate Verification of the Symbolic Dynamics of Markov Chains. *J. ACM* 62, 1 (2015), 2:1–2:34.
- [2] S. Akshay, T. Antonopoulos, J. Ouaknine, and J. Worrell. 2015. Reachability problems for Markov chains. *Inf. Process. Lett.* 115, 2 (2015), 155–158.
- [3] S. Almagor, T. Karimov, E. Kelmendi, J. Ouaknine, and J. Worrell. 2021. Deciding  $\omega$ -regular properties on linear recurrence sequences. *Proc. ACM Program. Lang.* 5, POPL (2021).
- [4] C. Baier, F. Funke, S. Jantsch, T. Karimov, E. Lefauchaux, F. Luca, J. Ouaknine, D. Purser, M. A. Whiteland, and J. Worrell. 2021. The Orbit Problem for Parametric Linear Dynamical Systems. In *32nd International Conference on Concurrency Theory, CONCUR 2021, August 24–27, 2021, Virtual Conference (LIPIcs, Vol. 203)*, Serge Haddad and Daniele Varacca (Eds.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 28:1–28:17.
- [5] A. Baker. 2012. *A Comprehensive Course in Number Theory*. Cambridge University Press.
- [6] G. Barthe, C. Jacomme, and S. Kremer. 2020. Universal equivalence and majority of probabilistic programs over finite fields. In *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8–11, 2020*, Holger Hermanns, Lijun Zhang, Naoki Kobayashi, and Dale Miller (Eds.). ACM, 155–166.
- [7] B. Bartolome, Y. Bilu, and F. Luca. 2013. On the exponential local-global principle. *Acta Arith.* 159, 2 (2013), 101–111.
- [8] J. Berstel and C. Reutenauer. 2011. *Noncommutative Rational Series with Applications*. Cambridge University Press.
- [9] Cs. Bertók and L. Hadju. 2016. A Hasse-type principle for exponential Diophantine equations and its applications. *Math. Comput.* 85 (2016), 849–860.
- [10] Cs. Bertók and L. Hadju. 2018. A Hasse-type principle for exponential Diophantine equations over number fields and its applications. *Monatshefte Math.* 187 (2018), 425–436.
- [11] V. Blondel and J. Tsitsiklis. 2000. A survey of computational complexity results in systems and control. *Automatica* 36, 9 (2000), 1249–1274.
- [12] V. D. Blondel and N. Portier. 2002. The presence of a zero in an integer linear recurrent sequence is NP-hard to decide. *Linear Algebra and Its Applications* 351–352 (2002).
- [13] K. Chatterjee and L. Doyen. 2021. Stochastic Processes with Expected Stopping Time. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*. IEEE, 1–13.
- [14] V. Chonev, J. Ouaknine, and J. Worrell. 2015. The Polyhedron-Hitting Problem. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4–6, 2015*, Piotr Indyk (Ed.). SIAM, 940–956.
- [15] G. Everest, A. van der Poorten, I. Shparlinski, and T. Ward. 2003. *Recurrence Sequences*. American Mathematical Society.
- [16] P. Fatou. 1904. Sur les séries entières à coefficients entiers. *Comptes Rendus Acad. Sci. Paris* 138, 130 (1904), 342–344.
- [17] N. Fijalkow, J. Ouaknine, A. Pouly, J. Sousa Pinto, and J. Worrell. 2019. On the decidability of reachability in linear time-invariant systems. In *Proceedings of the 22nd ACM International Conference on Hybrid Systems: Computation and Control, HSCC 2019, Montreal, QC, Canada, April 16–18, 2019*, Necmiye Ozay and Pavithra Prabhakar (Eds.). ACM, 77–86.
- [18] F. Gouvea. 1997. *p-adic Numbers: An Introduction*. Springer.
- [19] T. Karimov, E. Lefauchaux, J. Ouaknine, D. Purser, A. Varonka, M. A. Whiteland, and J. Worrell. 2022. What’s decidable about linear loops? *Proc. ACM Program. Lang.* 6, POPL (2022).
- [20] L. Kronecker. 1857. Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten. *Reine Angew. Math* 53 (1857), 173–175.
- [21] M. Mignotte, T. N. Shorey, and R. Tijdeman. 1984. The distance between terms of an algebraic recurrence sequence. *Journal für die reine und angewandte Mathematik* 349 (1984).
- [22] J. Ouaknine and J. Worrell. 2012. Decision Problems for Linear Recurrence Sequences. In *Proc. Intern. Workshop on Reachability Problems (RP) (LNCS, Vol. 7550)*. Springer.
- [23] J. Ouaknine and J. Worrell. 2015. On linear recurrence sequences and loop termination. *ACM SIGLOG News* 2, 2 (2015), 4–13.
- [24] J. Piribauer and C. Baier. 2020. On Skolem-Hardness and Saturation Points in Markov Decision Processes. In *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8–11, 2020, Saarbrücken, Germany (Virtual Conference) (LIPIcs, Vol. 168)*, Artur Czumaj, Anuj Dawar, and Emanuela Merelli (Eds.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 138:1–138:17.
- [25] G. Rozenberg and A. Salomaa. 1994. *Cornerstones of Undecidability*. Prentice Hall.
- [26] A. Schinzel. 1977. Abelian binomials, power residues and exponential congruences. *Acta Arith.* 32, 3 (1977), 245–274.
- [27] A. Schinzel. 2003. On the congruence  $u_n \equiv c \pmod{p}$  where  $u_n$  is a recurring sequence of the second order. *Acta Acad. Paedagog. Agriensis Sect. Math.* 30 (2003), 147–165.
- [28] Th. Skolem. 1937. Anwendung exponentieller Kongruenzen zum Beweis der Unlösbarkeit gewisser diophantischer Gleichungen. *Avhdl. Norske Vid. Akad. Oslo I* 12 (1937), 1–16.
- [29] M. Soittola. 1976. On D0L Synthesis Problem. In *Automata, Languages, Development*, A. Lindenmayer and G. Rozenberg (Eds.). North-Holland.
- [30] T. Tao. 2008. *Structure and Randomness*. American Mathematical Society.
- [31] N. K. Vereshchagin. 1985. The problem of appearance of a zero in a linear recurrence sequence (in Russian). *Mat. Zametki* 38, 2 (1985).