On the Complexity of the Skolem Problem at Low Orders

Piotr Bacik Joël Ouaknine James Worrell

Abstract

The Skolem Problem asks to determine whether a given linear recurrence sequence (LRS) $\langle u_n \rangle_{n=0}^{\infty}$ over the integers has a zero term, that is, whether there exists n such that $u_n = 0$. Decidability of the problem is open in general, with the most notable positive result being a decision procedure for LRS of order at most 4.

In this paper we consider a bounded version of the Skolem Problem, in which the input consists of an LRS $\langle u_n \rangle_{n=0}^{\infty}$ and a bound $N \in \mathbb{N}$ (with all integers written in binary), and the task is to determine whether there exists $n \in \{0, \ldots, N\}$ such that $u_n = 0$. We give a randomised algorithm for this problem that, for all $d \in \mathbb{N}$, runs in polynomial time on the class of LRS of order at most d. As a corollary we show that the (unrestricted) Skolem Problem for LRS of order at most 4 lies in coRP, improving the best previous upper bound of $\mathsf{NP}^{\mathsf{RP}}$.

The running time of our algorithm is exponential in the order of the LRS—a dependence that appears necessary in view of the NP-hardness of the Bounded Skolem Problem. However, even for LRS of a fixed order, the problem involves detecting zeros within an exponentially large range. For this, our algorithm relies on results from p-adic analysis to isolate polynomially many candidate zeros and then test in randomised polynomial time whether each candidate is an actual zero by reduction to arithmetic-circuit identity testing.

1 Introduction

A *linear recurrence sequence* (LRS) $\boldsymbol{u} = \langle u_n \rangle_{n=0}^{\infty}$ is a sequence of integers satisfying a linear recurrence relation:

$$u_{n+d} = a_{d-1}u_{n+d-1} + \cdots + a_0 u_n \qquad (n \in \mathbb{N})\,, \tag{1.1}$$

where $a_0, \ldots, a_{d-1} \in \mathbb{Z}$. The *order* of $\boldsymbol{u}$ is the smallest d such that $\boldsymbol{u}$ satisfies a relation of the form (1.1). Even though LRS are ubiquitous in computer science, mathematics, and beyond, the decidability of many basic computational problems about LRS remain open. The most famous of these is the Skolem Problem [16, 19]:

Problem 1.1 (The Skolem Problem)**.** Given an LRS $\boldsymbol{u}$, does there exist $n \in \mathbb{N}$ such that $u_n = 0$?

There is substantial interest in the decidability and complexity of the Skolem Problem, being closely connected with many topics in theoretical computer science and other areas, including loop termination [14, 3], formal power series (e.g. [6], Section 6.4), matrix semigroups [5], stochastic systems [1, 4], and control theory [8].

The most significant structural result concerning the zeros of an LRS is the celebrated Skolem-Mahler-Lech Theorem [18, 13, 12], which states that the set of zeros of an LRS is a union of a finite number of arithmetic progressions and a finite set. The statement may be refined via the concept of *non-degeneracy*. Define the *characteristic polynomial* of the recurrence (1.1) to be

$$g(X) := X^d - a_{d-1}X^{d-1} - \cdots - a_0\,. \tag{1.2}$$

Let $\lambda_1, \ldots, \lambda_s \in \overline{\mathbb{Q}}$ be the distinct roots of g; these are called the *characteristic roots* of $\boldsymbol{u}$. We say $\boldsymbol{u}$ is *non-degenerate* if no ratio λ_i/λ_j of distinct characteristic roots is a root of unity. A given LRS can be effectively decomposed as the interleaving of finitely many non-degenerate LRS [11, Theorem 1.6]. The core of the Skolem-Mahler-Lech Theorem is that a non-degenerate LRS that is not identically zero has finitely many zero terms. The Skolem Problem therefore reduces to determining whether a non-degenerate LRS has a zero term.

To date, only partial decidability results are known, either by restricting the order of the LRS or by restricting the structure of the characteristic roots. The breakthrough papers [20, 22] achieve decidability for LRS up to order 4 by exhibiting an effective bound N such that if an LRS $\boldsymbol{u}$ of order at most 4 has a zero term then it has a zero term $u_n = 0$ where $n < N$. It is shown in [10] that these methods yield an $\mathsf{NP}^{\mathsf{RP}}$ complexity upper bound for the Skolem Problem on LRS of order at most 4: specifically [10, Appendix C] shows that the threshold N has magnitude exponential in the description length of the LRS and that the existence of $n < N$ such that $u_n = 0$ can be determined in $\mathsf{NP}^{\mathsf{RP}}$. A procedure to substantiate the latter bound involves guessing n and building an arithmetic circuit representing u_n and then checking zeroness of u_n in randomised polynomial time (an instance of the problem EqSLP of testing zeroness of arithmetic circuits, which is in coRP [17]).

The above considerations lead us to define the following problem:

Problem 1.2 (The Bounded Skolem Problem)**.** Given an LRS $\boldsymbol{u}$ and an integer $N > 0$, does there exist $n \in \{0, \ldots, N\}$ such that $u_n = 0$?

We can then reformulate the contribution of [10] as showing that the Bounded Skolem Problem has complexity in $\mathsf{NP}^{\mathsf{RP}}$ and that the Skolem Problem for LRS of order 4 reduces in polynomial time to the bounded version.

In this paper we show that for each fixed $d \in \mathbb{N}$ the Bounded Skolem Problem for LRS of order at most d lies in coRP. As a corollary we show that the Skolem Problem for LRS of order at most 4 also lies in coRP, improving the previous bound of $\mathsf{NP}^{\mathsf{RP}}$. Since the threshold N in the Bounded Skolem Problem is encoded in binary, the range $\{0, \ldots, N\}$ in which a zero is sought has cardinality exponential in the size of the input. Nevertheless, we give a randomised algorithm that can detect the existence of a zero in polynomial time for LRS of a fixed order. The running time of our algorithm is exponential in the order of the LRS. The exponential dependence on the order seems unavoidable, since the NP-hardness proof of the Skolem Problem straightforwardly carries over to its bounded version, although it no longer applies if the order of the LRS is fixed [2, 9].[1]

Our algorithm for the Bounded Skolem Problem is based on a p-adic approach that was introduced by Skolem in his original proof of the Skolem-Mahler-Lech Theorem. This involves extending an integer LRS $\boldsymbol{u} = \langle u_n \rangle_{n=0}^{\infty}$ to a p-adic analytic function $F : \mathcal{O}_p \to \mathcal{O}_p$, for a suitable prime p, such that $F(n) = u_n$ for all $n \in \mathbb{N}$. Here, $\mathcal{O}_p$ is the valuation ring of the field $\mathbb{C}_p$ which is the p-adic analogue of the field of complex numbers.

As we explain below, while it is not known how to decide whether an LRS has an integer zero, one can decide in polynomial time whether its extension F has a zero in $\mathcal{O}_p$. More generally, by considering zeros of the power series $F(p^r x + m)$ for given $r \in \mathbb{N}$ and $m \in \{0, \ldots, p^r - 1\}$, one can determine whether F has a zero $a \in \mathcal{O}_p$ such that $a \equiv m \bmod p^r$.

Our procedure performs a depth-first search of all residues $m \in \{0, \ldots, p^r - 1\}$, for successively larger r, such that F has a zero in $\mathcal{O}_p$ whose residue modulo p^r is equal to m. Assume without loss of generality that the input N to the Bounded Skolem Problem has the form $N = p^n - 1$. We call each residue $m \in \{0, \ldots, N\}$ for which there exists $a \in \mathcal{O}_p$ with $F(a) = 0$ and $a \equiv m \bmod (N+1)$ a *candidate zero* of $\boldsymbol{u}$. Every integer zero of $\boldsymbol{u}$ in the interval $\{0, \ldots, N\}$ is a candidate zero and, conversely, every candidate zero is the residue modulo $N + 1$ of a zero of F in $\mathcal{O}_p$. For each candidate zero $m \in \{0, \ldots, N\}$, we determine whether u_m is zero in randomised polynomial time by constructing an arithmetic circuit representing u_m, using iterated matrix powering, and checking the circuit for zeroness using standard methods in arithmetic circuit identity testing.

The overall polynomial-time bound of the procedure (for each fixed recurrence depth d) crucially relies on a quantitative refinement of Skolem's proof, due to Van der Poorten and Schlickewei [21]. We use this result in two different ways. First it gives an upper bound on the number of zeros of F in $\mathcal{O}_p$, which translates in our setting to a polynomial bound on the number of candidate zeros that our algorithm examines. In other words, we need only

[1] The NP-hardness proof for the Skolem Problem in [2, Theorem 6] is by reduction from Subset Sum. The reduction maps an instance of Subset Sum to an LRS that has a zero iff the instance of Subset Sum has a solution. The constructed LRS has period the product $N := p_1 \cdots p_m$ of the first m primes, where m is the number of integers in the instance of Subset Sum. Evidently the same reduction works also for the Bounded Skolem Problem.

search a small fragment of the exponentially large search tree of residues modulo $N+1$. Secondly, we use the result of [21] in combination with the Weierstrass Preparation Theorem (which characterises the number of roots of a power series in terms of the absolute values of its coefficients) to decide in polynomial time whether the power series F has a zero in $\mathcal{O}_p$. The key issue in this case is to obtain a polynomial bound on the number of coefficients of a power series that must be examined to determine whether it has a root in $\mathcal{O}_p$.

Although the correctness of our algorithm is based on p-adic techniques, the implementation works with the underlying LRS and does not directly compute with p-adic numbers. This is because we work with the so-called Mahler-series representation of p-adic power series, which are based on the Mahler polynomials $\binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}$ for $k \in \mathbb{N}$. Conveniently for our purposes, the interpolation F of an LRS has a Mahler series expansion $F(x) = \sum_{k=0}^{\infty} \beta_k \binom{x}{k}$ whose coefficients β_k lie in $\mathbb{Z}$ and can, moreover, be directly calculated from the LRS by a simple formula involving the discrete difference operator. The critical information for applying the p-adic Weierstrass Preparation Theorem for locating the roots of F can be extracted from its Mahler series expansion.

2 Preliminaries

2.1 Linear recurrence sequences

We recall some basic notions about linear recurrence sequences; more details may be found in [11]. We also establish some notation. Let $\boldsymbol{u}$ be an LRS of order d satisfying the relation (1.1). The *companion matrix* of $\boldsymbol{u}$ is defined to be

$$A = \begin{pmatrix} a_{d-1} & \dots & a_1 & a_0 \\ 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & 0 \\ 0 & \dots & 1 & 0 \end{pmatrix}.$$

Writing $\alpha = \begin{pmatrix} 0 & 0 \dots & 1 \end{pmatrix}$, $\beta = \begin{pmatrix} u_{d-1} & u_{d-2} \dots & u_0 \end{pmatrix}^T$, we get $u_n = \alpha A^n \beta$. We denote by $\|\boldsymbol{u}\|$ the *size* of the representation of $\boldsymbol{u}$, defined as

$$\|\boldsymbol{u}\| = \sum_{i=0}^{d-1} (\|a_i\| + \|u_i\|)$$

where $\|y\|$ denotes the number of bits needed to represent the integer y. Throughout the paper, let $\mathsf{poly}(\|\boldsymbol{u}\|)$ denote any quantity that is bounded above by $\|\boldsymbol{u}\|^{O(1)}$.

2.2 p-adic numbers

We briefly recall relevant notions about p-adic numbers. We refer to [15] for more details. Given a prime number p, every non-zero rational number $x \in \mathbb{Q}$ may be written as $x = \frac{a}{b} p^r$ for some integers a, b coprime to p, with b non-zero, and $r \in \mathbb{Z}$. We define the valuation $v_p(x) = r$, and $v_p(0) = \infty$. From this derivation we see that v_p satisfies the ultrametric inequality $v_p(x+y) \geq \min(v_p(x), v_p(y))$ for all $x, y \in \mathbb{Q}$.

We define an absolute value on $\mathbb{Q}$ by $|x|_p = p^{-v_p(x)}$. The ultrametric inequality on v_p translates to the strong triangle inequality: $|x+y|_p \leq \max(|x|_p, |y|_p)$ for all $x, y \in \mathbb{Q}$. In the same way as one completes $\mathbb{Q}$ with respect to the standard absolute value $|\cdot|$ to get $\mathbb{R}$, one may complete $\mathbb{Q}$ with respect to $|\cdot|_p$ to get the p-adic numbers $\mathbb{Q}_p$. If one completes $\mathbb{Z}$ with respect to $|\cdot|_p$ then one gets the p-adic integers $\mathbb{Z}_p$, which may also be defined as the unit disc in $\mathbb{Q}_p$:

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\} = \{x \in \mathbb{Q}_p : v_p(x) \geq 0\}\,.$$

The absolute value $|\cdot|_p$ on $\mathbb{Q}_p$ extends uniquely to an absolute value on the algebraic closure $\overline{\mathbb{Q}_p}$. Specifically, given $\alpha \in \overline{\mathbb{Q}_p}$, we define $|\alpha|_p := |N(\alpha)|_p^{1/n}$, where n is the degree of α over $\mathbb{Q}_p$ and $N(\alpha)$ is the *norm* of α, that is the determinant of the $\mathbb{Q}$-linear transformation $\mu_\alpha : \mathbb{Q}(\alpha) \to \mathbb{Q}(\alpha)$ given by $\mu_\alpha(x) = \alpha x$. We correspondingly extend the valuation $v_p(\cdot)$ to $\overline{\mathbb{Q}_p}$ by $v_p(\alpha) := \frac{1}{n} v_p(N(\alpha))$. With these definitions the identity $|\alpha|_p = p^{-v_p(\alpha)}$ holds for all $\alpha \in \overline{\mathbb{Q}_p}$. While $\overline{\mathbb{Q}_p}$ is not complete with respect to $|\cdot|_p$, taking the Cauchy completion we obtain a field $\mathbb{C}_p$ that is both algebraically closed and complete under the extension of $|\cdot|_p$ to $\mathbb{C}_p$.

Define the closed disc $\overline{D}(z,r)$ centred at $z \in \mathbb{C}_p$ with valuation r to be

$$\overline{D}(z,r) = \{x \in \mathbb{C}_p : v_p(x-z) \geq r\} = \{x \in \mathbb{C}_p : |x-z|_p \leq p^{-r}\} \tag{2.1}$$

and denote the unit disc of $\mathbb{C}_p$ by

$$\mathcal{O}_p = \overline{D}(0,0) = \{x \in \mathbb{C}_p : v_p(x) \geq 0\} = \{x \in \mathbb{C}_p : |x-z|_p \leq 1\}\,.$$

Note that $\mathcal{O}_p$ is a subring of $\mathbb{C}_p$.

2.3 p-adic power series

Let $\langle b_j \rangle_{j=0}^{\infty}$ be a sequence in $\mathbb{Z}_p$. A sufficient and necessary condition for the power series $F(x) = \sum_{j=0}^{\infty} b_j x^j$ to converge on $\mathbb{Z}_p$ is that $v_p(b_j) \to \infty$ as $j \to \infty$. A *p-adic analytic function* $F : \mathbb{Z}_p \to \mathbb{Z}_p$ is one that is given by a convergent power series. For a p-adic analytic function F and $\alpha \in \mathcal{O}_p$ it is the case that $F(\alpha)$ converges to an element of $\mathcal{O}_p$. Thus we may refer to the zeros of F in $\mathcal{O}_p$.

Theorem 2.1 (p-adic Weierstrass Preparation Theorem)**.** *Let $F(x) = \sum_{j=0}^{\infty} b_j x^j \in \mathbb{Z}_p[\![x]\!]$ be a non-zero convergent power series. Suppose j_0 is an integer such that*

$$v_p(b_{j_0}) = \min_j v_p(b_j) \qquad \text{and} \qquad v_p(b_{j_0}) < v_p(b_j) \quad \forall j > j_0\,.$$

Then there is a polynomial $g(x)$ of degree j_0 and a power series $h(x)$ with no zeros in $\mathcal{O}_p$ such that

$$F(x) = g(x)h(x)\,.$$

In particular, F has exactly j_0 zeros in $\mathcal{O}_p$, counting multiplicity.

To facilitate the application of this theorem, we formulate a further definition and corollary.

Definition 2.2. Let F and j_0 be as in Theorem 2.1. Define $V(F) := \min_j v_p(b_j) = v_p(b_{j_0})$, and write $j(F) := j_0$.

Equivalently, $j(F)$ is the largest index of a power series coefficient of F with minimal valuation among all coefficients, and $V(F)$ is this minimal valuation.

Corollary 2.3. Let F be as in Theorem 2.1. Define $h_{z,r}(x) = F(p^r x + z)$. Then F has exactly $j(h_{z,r})$ zeros in the disc $\overline{D}(z,r)$.

Proof. By Theorem 2.1, $h_{z,r}$ has exactly $j(h_{z,r})$ zeros in $\mathcal{O}_p$. Note that $x \in \mathcal{O}_p = \overline{D}(0,0)$ if and only if $p^r x + z \in \overline{D}(z,r)$, so F has exactly $j(h_{z,r})$ zeros in the disc $\overline{D}(z,r)$. □

2.4 Mahler Series

The computation of $j(F)$ for p-adic analytic functions F arising from LRS plays a critical role in our zero-finding procedure. The following two sections show how this can be done directly from the LRS, without the need to compute the coefficients of F. For this we need some background on Mahler series. The following may be found in [15, Section IV 2.3-2.4, Section VI 4.7].

Define the difference operator Δ by $(\Delta F)(x) = F(x+1) - F(x)$ for any function $F : \mathbb{Z}_p \to \mathbb{Z}_p$. For $k \geq 0$, consider the polynomial

$$\binom{x}{k} := \frac{x(x-1)\cdots(x-k+1)}{k!}\,.$$

Since $\binom{n}{k} \in \mathbb{Z}$ for all $n \in \mathbb{Z}$ and $\mathbb{Z}$ is dense in $\mathbb{Z}_p$ with respect to the topology induced by $|\cdot|_p$, it follows that $\binom{x}{k}$ maps $\mathbb{Z}_p$ to $\mathbb{Z}_p$. A *Mahler series* is a formal series

$$F(x) = \sum_{k=0}^{\infty} \beta_k \binom{x}{k}\,, \tag{2.2}$$

where $\beta_k \in \mathbb{Z}_p$ is such that $v_p(\beta_k) \to \infty$ as $k \to \infty$. Such a series defines a continuous function $F : \mathbb{Z}_p \to \mathbb{Z}_p$ and, conversely, every continuous function $F : \mathbb{Z}_p \to \mathbb{Z}_p$ can be represented by a Mahler series of the form (2.2) in which

$$\beta_k = (\Delta^k F_\ell)(0) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} F(j)\,. \tag{2.3}$$

Under the further assumption that $v_p(\beta_k/k!) \to \infty$ as $k \to \infty$, the Mahler series (2.2) defines an analytic function $F : \mathbb{Z}_p \to \mathbb{Z}_p$, that is, F admits a power series expansion $F(x) = \sum_{j=0}^{\infty} b_j x^j$ where $b_j \in \mathbb{Z}_p$ for all j and $v_p(b_j) \to 0$.

The following proposition characterises $j(F)$ for an analytic function F in terms of its Mahler series.

Proposition 2.4. Suppose $F : \mathbb{Z}_p \to \mathbb{Z}_p$ is an analytic function whose respective power series and Mahler series representations are:

$$F(x) = \sum_{j=0}^{\infty} b_j x^j = \sum_{k=0}^{\infty} \beta_k \binom{x}{k} \tag{2.4}$$

Define $V'(F) := \min_{k \geq 0} v_p(\beta_k/k!)$ and define $k(F)$ to be the largest integer k such that $v_p(\beta_k/k!) = V'(F)$. Then $V(F) = V'(F)$ and $j(F) = k(F)$.

Proof. We have

$$k!\binom{x}{k} = \sum_{j=0}^{k} (-1)^{k-j} \begin{bmatrix} k \\ j \end{bmatrix} x^j \quad \text{and} \quad x^j = \sum_{k=0}^{j} \begin{Bmatrix} j \\ k \end{Bmatrix} k! \binom{x}{k} \tag{2.5}$$

where $\begin{bmatrix} k \\ j \end{bmatrix}$ and $\begin{Bmatrix} j \\ k \end{Bmatrix}$ are positive integers known as the Stirling numbers of the first and second kind respectively. By combining (2.5) with (2.4) we see that

$$b_j = \sum_{k=j}^{\infty} (-1)^{k-j} \frac{\beta_k}{k!} \begin{bmatrix} k \\ j \end{bmatrix} \quad \text{and} \quad \frac{\beta_k}{k!} = \sum_{j=k}^{\infty} b_j \begin{Bmatrix} j \\ k \end{Bmatrix}$$

which gives

$$V(F) = v_p(b_{j(F)}) = v_p\left(\frac{\beta_{j(F)}}{j(F)!}\right) \geq V'(F)$$

and

$$V'(F) = v_p\left(\frac{\beta_{k(F)}}{k(F)!}\right) = v_p(b_{k(F)}) \geq V(F)\,.$$

The combination of these inequalities give $V(F) = V'(F)$ and $j(F) = k(F)$. □

2.5 Interpolation of Linear Recurrence Sequences

It is well known that for a suitable prime p, a linear recurrence sequence admits a decomposition into finitely many subsequences such that each subsequence can be interpolated to a p-adic analytic function. This fact is the basis of the Skolem's original proof of the Skolem-Mahler-Lech theorem. In this section, we give a self-contained account of this construction. Furthermore, for each of the p-adic analytic functions F that arise from the interpolation of an LRS we give a characterisation of $j(F)$ in terms of the first $d-1$ discrete derivatives of the LRS at zero, where d is the order of the recurrence. In the process, we recover a result of [21, Lemma 2] that $j(F) < d-1$ for p sufficiently large relative to the order d of the recurrence:

Theorem 2.5. Let $\boldsymbol{u}$ satisfy the order-d linear recurrence (1.1) and let $p > d+1$ be a prime not dividing the constant term a_0 of the recurrence. Then there exists $M < p^{d^2}$ and analytic functions $F_\ell : \mathbb{Z}_p \to \mathbb{Z}_p$, $\ell \in \{0, \ldots, M-1\}$, such that for all ℓ we have

1. $F_\ell(n) = u_{Mn+\ell}$ for all $n \in \mathbb{N}$;
2. $V(F_\ell) = \min\{v_p((\Delta^j F_\ell)(0)) : 0 \leq j \leq d-1\}$;
3. $j(F_\ell) = \max\{j \in \{0,\ldots,d-1\} : v_p((\Delta^j F_\ell)(0)) = V(F_\ell)\}$.

Proof. Write $u_n = x^\top A^n y$, where A is the companion matrix of the recurrence and

$$x = \begin{pmatrix} 0 & 0 & \ldots & 1 \end{pmatrix}, \qquad\qquad y = \begin{pmatrix} u_{d-1} & u_{d-2} & \ldots & u_0 \end{pmatrix}^T .$$

Since $p \nmid a_0$, we have $p \nmid \det(A)$ and hence A is invertible in $\mathrm{GL}_d(\mathbb{Z}/p\mathbb{Z})$. It follows that $A^M \equiv I \pmod p$, for some $0 < M < p^{d^2}$. Write $A^M = I + pB$ for some $d \times d$ integer matrix B. Then for $\ell \in \{0,\ldots,M\}$ we have

$$\begin{aligned} u_{Mn+\ell} &= x^T A^{Mn+\ell} y \\ &= x^T (I+pB)^n A^\ell y \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k \, x^T B^k A^\ell y \\ &= \sum_{k=0}^{\infty} \beta_{\ell,k} \binom{n}{k}, \end{aligned}$$

where $\beta_{\ell,k} := p^k \, x^T B^k A^\ell y$. Since $v_p(\beta_{\ell,k}) \geq k$, using the standard bound $v_p(k!) \leq \frac{k}{p-1}$, we have $v_p(\beta_{\ell,k}/k!) \geq \frac{k(p-2)}{p-1}$. Thus $v_p(\beta_{\ell,k}/k!) \to \infty$ as $k \to \infty$ and so the Mahler series $F_\ell(x) := \sum_{k=0}^\infty \beta_{\ell,k}\binom{x}{k}$ defines an analytic function such that $F_\ell(n) = u_{Mn+\ell}$ for all $n \in \mathbb{Z}$.

Fix $\ell \in \{0,\ldots,M-1\}$ and write $\gamma_{\ell,k} := x^\top B^k A^\ell y$, so that $\beta_{\ell,k} = p^k \gamma_{\ell,k}$. By the Cayley-Hamilton theorem, there are integers $b_0,\ldots,b_{d_1}$ such that $B^d = b_0 B^{d-1} + \cdots b_{d-1} I$. Hence

$$\gamma_{\ell,k} = x^\top B^k A^\ell y = \sum_{i=0}^{d-1} b_i \, x^\top B^{k-1-i} A^\ell y = \sum_{i=0}^{d-1} b_i \gamma_{\ell,k-1-i} \tag{2.6}$$

for all $k \geq d$. Let $m := \min\{v_p(\gamma_{\ell,k}) : 0 \leq k \leq d-1\}$. Then, by the Equation 2.6, it follows by strong induction on k that $v_p(\gamma_{\ell,k}) \geq m$ for all $k \in \mathbb{N}$.

For all $k \geq d$, using the standard bound $v_p(k!) \leq \frac{k-1}{p}$, we have:

$$\begin{aligned} v_p(\beta_{\ell,k}/k!) &= v_p(\gamma_{\ell,k}) + v_p(p^k/k!) \\ &\geq m + \tfrac{k(p-2)}{p-1} \\ &\geq m + \tfrac{d(p-2)}{p-1} \\ &> m + d - 1 \quad \text{(since } p > d+1\text{)}. \end{aligned}$$

On the other hand, for $0 \leq k \leq d-1$ we have $v_p(\beta_{\ell,k}/k!) \leq v_p(\gamma_{\ell,k}) + d - 1$ and hence

$$\min_{0 \leq k \leq d-1} v_p(\beta_{\ell,k}/k!) \leq m + d - 1 \, .$$

We conclude that $k(F_\ell) \leq d-1$.

Applying Proposition 2.4, since $d-1 < p$, we have

$$V(F_\ell) = V'(F_\ell) = \min_{0 \leq k \leq d-1} v_p(\beta_{\ell,k}/k!) = \min_{0 \leq k \leq d-1} v_p(\beta_{\ell,k}) = \min_{0 \leq k \leq d-1} v_p((\Delta^k F_\ell)(0))$$

and $j(F) = k(F)$ is the maximum value of $k \in \{0,\ldots,d-1\}$ such that $v_p(\beta_{\ell,k}) = V'(F)$. □

3 Algorithm and Complexity Analysis

Let LRS(d) denote the class of LRS of order d. Throughout this section, we assume we are working inside the class LRS(d) for some $d \in \mathbb{N}$. Our complexity estimates are in terms of $\|\boldsymbol{u}\|$, the length of the description of initial values and recurrence that define $\boldsymbol{u}$, and the bit length of the threshold N describing the range $\{0, \ldots, N\}$ in which we seek zeros of LRS.

The goal of this section is to prove the following result:

Theorem 3.1. For every $d \in \mathbb{N}$, the Bounded Skolem Problem on LRS(d) is disjunctively Turing reducible to EqSLP and hence lies in coRP.

Note above that one can test whether at least one among a finite collection of arithmetic circuits is zero by testing the product for zeroness, which can be done in coRP [17].

3.1 Informal outline of the algorithm

We will first outline the general ideas behind the algorithm and the ingredients required to prove the required complexity bounds on the Bounded Skolem Problem. Given an LRS $\boldsymbol{u}$ and integer upper bound N written in binary, let $\|\boldsymbol{u}\|$ and $\|N\|$ denote the lengths of their respective binary strings. Recall we are required to search for integers $0 \leq n \leq N$ with $u_n = 0$.

In short, the idea is to find p-adic approximations of any possible zeros of $\boldsymbol{u}$ by finding discs of decreasing size in which zeros may lie. We do this approximation until for each disc there is only one possible integer that lies below the upper bound N and lies inside the disc, and then we check all the candidate integer zeros. In more detail, recalling the definition (2.1) of the disc $\overline{D}(z, r)$, we do the following:

1. Choose a prime p that does not divide the last coefficient a_0 of the recurrence.

2. Split the LRS into subsequences $u_{Mn+\ell}$ for some integer M and $0 \leq \ell \leq M-1$ such that $u_{Mn+\ell}$ may be interpolated by a p-adic analytic function $F_\ell : \mathbb{Z}_p \to \mathbb{Z}_p$. We have F_ℓ is identically zero if and only if $u_{Mn+\ell} = 0$ for $n = 0, 1, \ldots, d-1$. This may be checked in polynomial-time as $Mn + \ell = \mathsf{poly}(\|\boldsymbol{u}\|)$ for these values of n. We analyse those functions F_ℓ for which the corresponding sequence $u_{Mn+\ell}$ is not identically zero.

3. For R such that $Mp^R + \ell \geq N$, find all discs $\overline{D}\left(z, R\right)$ for $0 \leq z \leq p^R - 1$ containing a zero $x_0 \in \mathcal{O}_p$ of F_ℓ. This identifies all residue classes mod p^R that could contain an integer zero of $u_{Mn+\ell}$.

   - To do this efficiently, we inductively find each disc $\overline{D}(z, r)$ containing a zero of F_ℓ for each $0 \leq r \leq R$ and $0 \leq z \leq p^r - 1$.
   - If we have found that the disc $\overline{D}(z, r)$ contains a zero of F_ℓ for integer $0 \leq z \leq p^r - 1$, then check each disc $\overline{D}(z + p^r a, r + 1)$ to determine whether it has any zeros of F_ℓ, for integers $0 \leq a \leq p - 1$.
   - Checking whether the disc $\overline{D}(z, r)$ contains any zeros of F_ℓ may be done by computing $j(h_{z,r})$, where $h_{z,r}(x) = F(p^r x + z)$, by Corollary 2.3. In practice we do

this by looking at the p-adic valuations of the Mahler-series coefficients $\beta_{\ell,k}$ for each k (as defined in Section 2.5), using Theorem 2.5.

4. Check whether $u_{Mz+\ell} = 0$ for each disc $\overline{D}(z,R)$ computed.

Correctness of the algorithm is straightforward to see as by definition of R, any disc $\overline{D}(z,R)$ contains at most one integer n such that $0 \leq Mn + \ell \leq N$, namely $n = z$ being the only possible candidate. The discs found by the end of the computation are the only discs $\overline{D}(z,R)$ containing a p-adic zero $x_0 \in \mathcal{O}_p$ of F_ℓ, so they obviously contain all integer zeros of $F_\ell(n) = u_{Mn+\ell}$. The bulk of the problem is to prove that steps 1-3 above may be done in $\mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ time. The strategy of the next subsection will be to show that:

1. p, M may be taken to be $\mathsf{poly}(\|\boldsymbol{u}\|)$ in magnitude.

2. Computing whether the disc $\overline{D}(z,r)$ contains any zeros of F_ℓ takes $\mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ time for any integer $0 \leq z \leq p^r - 1$ and $1 \leq r \leq R$ for R smallest integer such that $Mp^R + \ell \geq N$

3. It takes $\mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ many of the above computations to determine all the residue classes mod p^R possibly containing a zero of F_ℓ

3.2 Proofs of complexity bounds

In this subsection we prove that all the parts of the algorithm outlined have the required complexity bounds.

Lemma 3.2. We may find a prime $p > d + 1$ and integer M with $\mathsf{poly}(\|\boldsymbol{u}\|)$ magnitude in $\mathsf{poly}(\|\boldsymbol{u}\|)$ time that satisfy the conditions of Theorem 2.5.

Proof. Choose any prime $p \geq d + 2$ such that $p \nmid a_0$, the last coefficient of the recurrence relation (1.1) that $\boldsymbol{u}$ satisfies. By Bertrand's postulate, there exists a prime p satisfying $\max\{d + 2, |a_0|, 4\} < p < 2\max\{d + 2, |a_0|, 4\}$, so to find the required prime p it suffices to search this interval for a prime, which may be done in $\mathsf{poly}(\|\boldsymbol{u}\|)$ time.

Now, by Theorem 2.5 we can take $M < p^{d^2}$, in particular the smallest positive integer such that $A^M \equiv I \mod p$, where A is the companion matrix of $\boldsymbol{u}$. Such M is easily found in $\mathsf{poly}(\|\boldsymbol{u}\|)$ time by directly checking if $A^m \equiv I \mod p$ for each $0 < m < p^{d^2}$. □

We would now like to prove that the search for p-adic discs containing zeros of F_ℓ takes $\mathsf{poly}(\|\boldsymbol{u}\|)$ time. To do this, we need to be able to find $j(h_{z,r})$ quickly, where $h_{z,r}(x) = F(p^r x + z)$, so we need to show $V(h_{z,r})$ is not too large.

Lemma 3.3. Let $F_\ell : \mathbb{Z}_p \to \mathbb{Z}_p$ interpolate the subsequence $u_{Mn+\ell}$ for p, M as in Lemma 3.2, and suppose that the subsequence $u_{Mn+\ell}$ is not identically zero. Let $h_{z,r}(x) = F_\ell(p^r x + z)$ for integers r and $0 \leq z \leq p^r - 1$. Then

1. $j(h_{z,r}) \leq j(F_\ell) \leq d - 1$

2. $V(h_{z,r}) \leq V(F_\ell) + rj(F_\ell) = \mathsf{poly}(\|\boldsymbol{u}\|, r)$.

Proof. We have $j(h_{z,r}) \leq j(F_\ell)$ by Corollary 2.3 and we have $j(F_\ell) \leq d-1$ by Theorem 2.5. This proves Item 1.

For the bound on $V(h_{z,r})$ in Item 2., note that by Item 2 of Theorem 2.5 we have

$$V(F_\ell) \in \{v_p((\Delta^0 F_\ell)(0)), v_p((\Delta^1 F_\ell)(0)), \dots, v_p((\Delta^{d-1} F_\ell)(0))\}\,.$$

Then from the formula

$$(\Delta^k F_\ell)(0) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} F_\ell(j)$$

we have

$$V(F_\ell) \leq \log\left(\max_{0\leq k\leq d-1} |(\Delta^k F_\ell)(0)|\right) \leq \log\left(\sum_{j=0}^{d-1} \binom{d-1}{k} |F_\ell(j)|\right) \leq \log\left(\sum_{j=0}^{d-1} 2^d |u_{Mj+\ell}|\right)$$

and using that $|u_{Mj+\ell}| \leq 2^{(Mj+\ell)\|\boldsymbol{u}\|} d^{Mj+\ell}$, which is easily shown by induction with the recurrence relation (1.1), we get $V(F_\ell) = \mathsf{poly}(\|\boldsymbol{u}\|)$.

For the bound on $V(h_{z,r})$, note the following. For any analytic $F(x) = \sum_{j=0}^{\infty} b_j x^j$, for $0 \leq k \leq d-1$ we have $F^{(k)}(0) = b_k k!$ so since $p \geq d+2$ we have $v_p(F^{(k)}(0)) = v_p(b_k)$. Furthermore, we have $v_p(b_{j(F)}) < v_p(b_k)$ for all $k > j(F)$ so $v_p(F^{(j(F))}(z)) = v_p(F^{(j(F))}(0))$ for all $z \in \mathbb{Z}$. From this we can deduce

$$V(h_{z,r}) = v_p(h_{z,r}^{(j(h_{z,r}))}(0)) \leq v_p(h_{z,r}^{(j(F_\ell))}(0)) = v_p(p^{rj(F_\ell)} F_\ell^{(j(F_\ell))}(z)) = V(F_\ell) + rj(F_\ell)$$

and we have $V(F_\ell) + rj(F_\ell) \leq V(F_\ell) + r(d-1) = \mathsf{poly}(\|\boldsymbol{u}\|, r)$.

□

Having shown that $V(h_{z,r})$ cannot be too large, we now show we can compute $j(h_{z,r})$ sufficiently quickly.

Proposition 3.4. Let $F_\ell : \mathbb{Z}_p \to \mathbb{Z}_p$ interpolate a not-identically-zero subsequence $u_{Mn+\ell}$ with $p, M = \mathsf{poly}(\|\boldsymbol{u}\|)$. For any integers $r > 0$ and $0 \leq z \leq p^r - 1$, we may find the number of zeros (counted with multiplicity) of F_ℓ in the disc $\overline{D}(z, r)$ in $\mathsf{poly}(\|\boldsymbol{u}\|, r)$ time.

Proof. By Corollary 2.3 it suffices to find $j(h_{z,r})$, which by Lemma 3.3 satisfies $j(h_{z,r}) \leq d-1$. By Theorem 2.5, $j(F)$ is simply the largest integer $k \in \{0, \dots, d-1\}$ such that $v_p((\Delta^k h_{z,r})(0))$ is minimal (and equal to $V(h_{z,r})$). By Lemma 3.3 we have that $V(h_{z,r}) < \nu$ for some $\nu = \mathsf{poly}(\|\boldsymbol{u}\|, r)$.[2]

Then at least one of $(\Delta^0 h_{z,r})(0), (\Delta^1 h_{z,r})(0), \dots, (\Delta^{d-1} h_{z,r})(0)$ is non-zero mod p^ν, so to determine $j(h_{z,r})$ it suffices to compute these quantities mod p^ν. We have the formula

$$(\Delta^k h_{z,r})(0) = \sum_{n=0}^{k} (-1)^{k-n} \binom{k}{n} h_{z,r}(n) = \sum_{n=0}^{k} (-1)^{k-n} \binom{k}{n} u_{M(p^r n+z)+\ell}\,.$$

[2] By inspecting the proof of Lemma 3.3 we may take $\nu = \log(2^d 2^{(Md+\ell)\|\boldsymbol{u}\|} d^{Md+\ell+1}) + rd + 1$ for actual computations.

Thus $(\Delta^k h_{z,r})(0) \mod p^\nu$ may be computed in $\mathsf{poly}(\|\boldsymbol{u}\|)$ time if $u_{M(p^r n+z)+\ell} \mod p^\nu$ may be computed in $\mathsf{poly}(\|\boldsymbol{u}\|)$ time, for $0 \leq n \leq d-1$. Writing $\alpha = \begin{pmatrix} 0 & 0 \dots & 1 \end{pmatrix}$, $\beta = \begin{pmatrix} u_{d-1} & u_{d-2} \dots & u_0 \end{pmatrix}^T$ and

$$A = \begin{pmatrix} a_{d-1} & \dots & a_1 & a_0 \\ 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & 0 \\ 0 & \dots & 1 & 0 \end{pmatrix}$$

the companion matrix, we have

$$u_m = \alpha A^m \beta \,.$$

For $0 \leq n \leq d-1$, we have $M(p^r n+z)+\ell \leq q_1(\|\boldsymbol{u}\|)^{q_2(r)}$ for polynomials q_1, q_2. We may compute $A^{M(p^r n+z)} \mod p^\nu$ using $O(\log(M(p^r n+z)+\ell)) = O(q_2(r)\log(q_1(\|\boldsymbol{u}\|))) = \mathsf{poly}(\|\boldsymbol{u}\|, r)$ multiplications of matrices mod p^ν, using repeated squaring. Since $\nu = \mathsf{poly}(\|\boldsymbol{u}\|, r)$ and $p = \mathsf{poly}(\|\boldsymbol{u}\|)$, the entries of the matrices mod p^ν may be written with $\mathsf{poly}(\|\boldsymbol{u}\|, r)$ many binary digits, so each multiplication of matrices may be done in $\mathsf{poly}(\|\boldsymbol{u}\|, r)$ time.

In conclusion, $(\Delta^k h_{z,r})(0) \mod p^\nu$ may be computed in $\mathsf{poly}(\|\boldsymbol{u}\|)$ time for all $0 \leq k \leq d-1$, from which we can immediately find $j(h_{z,r})$ and conclude. □

Corollary 3.5. For each $0 \leq \ell \leq M-1$, for integer N written in binary, we can take $R = O(\|N\|)$ such that $Mp^R + \ell \geq N$ and we may find all discs $\overline{D}(z, R)$ containing a zero $x_0 \in \mathcal{O}_p$ of F_ℓ in $\mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ time.

Proof. Firstly, note that we can take $R = \left\lceil \frac{\log N}{\log p} \right\rceil$ so $R = O(\|N\|)$. By Proposition 3.4, for each integer $1 \leq r \leq R$ and each integer $0 \leq z \leq p^r - 1$ it takes $\mathsf{poly}(\|\boldsymbol{u}\|, r) = \mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ time to determine whether $\overline{D}(z, r)$ contains any zeros of F_ℓ. Consider the following routine:

1. Determine whether F_ℓ has any zeros in $\overline{D}(0, 0) = \mathcal{O}_p$.

2. Given a disc $\overline{D}(z, r)$ that has been determined to have zeros of F_ℓ, for each integer $a = 0, 1, \dots, p-1$, determine whether $\overline{D}(z + p^r a, r+1)$ has any zeros of F_ℓ.

3. Repeat step 2 until $r = R$, output all discs $\overline{D}(z, R)$ found and then stop.

This routine inductively finds all the discs $\overline{D}(z, R)$ containing zeros of F_ℓ for $0 \leq z \leq p^R - 1$. By Theorem 2.5, F_ℓ has at most $d-1$ zeros, so we end up with at most $d-1$ of these discs. Each such disc $\overline{D}(z, R)$ takes at most $pR = \mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ operations of checking particular discs $\overline{D}(z', r)$ for zeros, and each operation takes $\mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ time, so the entire routine takes $\mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$ time. □

We are now ready to prove our main result.

Theorem 3.1. For every $d \in \mathbb{N}$, the Bounded Skolem Problem on LRS(d) is disjunctively Turing reducible to EqSLP and hence lies in coRP.

Proof. We can find $p, M = \mathsf{poly}(\|\boldsymbol{u}\|)$ in $\mathsf{poly}(\|\boldsymbol{u}\|)$-time satisfying the conditions of Theorem 2.5 by Lemma 3.2. For each subsequence $u_{Mn+\ell}$ for $0 \le \ell \le M-1$ we check whether $u_{Mn+\ell}$ is identically zero by checking directly whether $u_{Mn+\ell} = 0$ for $n = 0, 1, \ldots, d-1$. This takes $\mathsf{poly}(\|\boldsymbol{u}\|)$ time as $Mn + \ell = \mathsf{poly}(\|\boldsymbol{u}\|)$ for such values of n. For each ℓ such that $u_{Mn+\ell}$ is not identically zero we analyse its p-adic interpolation $F_\ell : \mathbb{Z}_p \to \mathbb{Z}_p$.

By Corollary 3.5, for each $0 \le \ell \le M-1$ and for $R = O(\|N\|)$ such that $Mp^R + \ell \ge N$ we may find all discs $\overline{D}(z, R)$ for integers $0 \le z \le p^R - 1$ containing a zero of F_ℓ. By construction of R, the only positive integer inside each disc $\overline{D}(z, R)$ possibly corresponding to an integer zero $0 \le n \le N$ of $\boldsymbol{u}$ is z itself. Therefore, for each such disc we need only check whether $u_{Mz+\ell} = 0$.

Let the collection of all candidate integer zeros be $n_1, \ldots, n_m$. Note each $n_i = Mz_i + \ell$ for some $0 \le \ell \le M-1$ and $0 \le z_i \le p^R$. By Theorem 2.5, each F_ℓ has at most $d-1$ zeros, so $m \le (d-1)M = \mathsf{poly}(\|\boldsymbol{u}\|)$. However, we cannot check whether $u_{n_i} = 0$ directly since in general $n_i = 2^{\mathsf{poly}(\log \|\boldsymbol{u}\|, \|N\|)}$, so one more trick is required. Consider the product $U = u_{n_1} u_{n_2} \ldots u_{n_m}$. Since we have

$$u_{n_i} = \alpha A^{n_i} \beta$$

for certain vectors α, β and the companion matrix A, each u_{n_i} may be written as the output of a $\mathsf{poly}(\|\boldsymbol{u}\|, \|N\|)$-sized circuit, using repeated squaring. Therefore, since $m = \mathsf{poly}(\|\boldsymbol{u}\|)$, we may write U as the output of a $\mathsf{poly}(\|\boldsymbol{u}\|)$-sized circuit. The LRS $\boldsymbol{u}$ has a zero $0 \le n \le N$ if and only if $U = 0$, and the problem of checking whether $U = 0$ is in EqSLP. □

Since the described algorithm identifies all possible candidates for zeros $0 \le n \le N$ of $\boldsymbol{u}$, we can improve the complexity of the function problem version of the Bounded Skolem Problem.

Theorem 3.6. Let $d \in \mathbb{N}$. Given $\boldsymbol{u} \in \mathrm{LRS}(d)$ and $N \in \mathbb{N}$, the problem of computing the set of all $0 \le n \le N$ with $u_n = 0$ (represented as the union of a finite set and finitely many arithmetic progressions) is in $\mathsf{FP}^{\mathsf{EqSLP}} = \mathsf{FP}^{\mathsf{RP}}$.

Proof. We may follow the same exact proof as for Theorem 3.1, except using an EqSLP oracle to decide whether $u_n = 0$ for each candidate zero n. □

Corollary 3.7. The Skolem Problem for LRS of order at most 4 is disjunctively Turing reducible to EqSLP and hence lies in coRP. The function problem of determining all zeros of an LRS of order at most 4 is in $\mathsf{FP}^{\mathsf{EqSLP}} = \mathsf{FP}^{\mathsf{RP}}$.

Proof. Follows from Theorems 3.1 and 3.6 and the fact that there is an upper bound $N = 2^{\mathsf{poly}(\|\boldsymbol{u}\|)}$ on the size of the largest zero lying outside an arithmetic progression of zeros. We deduce this from [10, Appendix C] which states that a *non-degenerate* sequence $\boldsymbol{u}$ of order at most 4 has an upper bound $N = 2^{O(\|\boldsymbol{u}\|')}$ on the size of the largest zero, where $\|\boldsymbol{u}\|'$ is the size of a description of the exponential-polynomial representation of $\boldsymbol{u}$ in binary. That is, if

$$u_n = \sum_{i=1}^{s} P_i(n) \lambda_i^n$$

for polynomials P_i with algebraic coefficients and algebraic numbers λ_i, then $\|\boldsymbol{u}\|' = \sum_i \|P_i\| + \|\lambda_i\|$. The description of the coefficients of P_i and λ_i are the standard ways to represent an algebraic number by its minimal polynomial and an appropriate approximation, see [10] for more details. It is noted there also that there is an integer $L(d)$ depending only on the order d of the LRS such that the subsequences $u_{L(d)n+r}$ are identically zero or non-degenerate. These subsequences have exponential polynomial representation

$$u_{L(d)n+r} = \sum_{i=1}^{s} P_i(L(d)n+r)\lambda_i^r \lambda_i^{L(d)n}$$

and by [10, Lemma A.1] we have $\|Q_i\|, \|\lambda_i^{L(d)}\| = \mathsf{poly}(\|\boldsymbol{u}\|')$ where $Q_i(n) = P_i(L(d)n+r)$, so each subsequence $v_n^{(r)} = u_{L(d)n+r}$ has $\|v^{(r)}\|' = \mathsf{poly}(\|\boldsymbol{u}\|')$. Thus, since $\|\lambda_i\|, \|P_i\| = \mathsf{poly}(\|\boldsymbol{u}\|)$ (e.g. see [2, Lemma 6]) we have $\|\boldsymbol{u}\|' = \mathsf{poly}(\|\boldsymbol{u}\|)$. By [10] there's a $N_r = 2^{O(\|\boldsymbol{v}^{(r)}\|'}$ for each not-identically-zero $\boldsymbol{v}^{(r)}$ so we may take $N = \max_{0 \le r \le L(d)-1} N_r = 2^{\mathsf{poly}(\|\boldsymbol{u}\|)}$. □

In fact, the Skolem Problem is known to be decidable for a wider class of LRS. Named after the authors of the papers [20, 22] in which decidability is established, the *MSTV class* is the class of LRS with at most 3 dominant roots with respect to an Archimedean absolute value or at most 2 dominant roots with respect to a non-Archimedean absolute value. See also the exposition of [7] on these results. Let MSTV(d) denote the class of LRS of order d lying in the MSTV class. We claim that by carrying the constants through more precisely in the proof of decidability, there is a bound $N = 2^{\mathsf{poly}(\|\boldsymbol{u}\|)}$ on the size of the largest zero, and hence for any integer $d \ge 1$ the Skolem Problem for the class MSTV(d) is disjunctively Turing reducible to EqSLP and therefore in coRP.

Finally, one more class of LRS for which an exponential upper bound on the size of the largest zero was shown is the class of LRS whose characteristic roots are all roots of rational numbers [2]. For this class we likewise obtain a coRP bound to decide the Skolem Problem for LRS of every fixed order.

3.3 Pseudocode

For completeness, we give the full pseudocode for the algorithm to find candidate zeros of $\boldsymbol{u}$ previously described. The algorithm takes as input an LRS $\boldsymbol{u}$ and integer N and finds in polynomial-time a description of a set containing all possible integer zeros of $\boldsymbol{u}$ in the interval $[0, N]$. In particular this set is comprised of finitely many arithmetic progressions $(Mn+\ell)_{n\in\mathbb{N}}$ for which $u_{Mn+\ell} = 0$, and polynomially-many exceptional integers $0 \le n \le N$ lying outside those arithmetic progressions. To solve the Bounded Skolem Problem, one simply has to check whether $u_{n_1} \dots u_{n_s} = 0$ where $n_1, \dots, n_s$ are all the exceptional candidate zeros. To solve the function problem version, one checks whether $u_{n_j} = 0$ for each $j = 1, \dots, s$.

Algorithm 1: Polynomial-time algorithm for finding all arithmetic progressions of zeros and polynomially-many extra candidate integer zeros of an LRS

```
Input: LRS $\boldsymbol{u}$ satisfying (1.1) of order $d$, integer $N$
Output: List of pairs $(M, \ell)$ for which $u_{Mn+\ell}$ is identically zero, and a list of
        polynomially many exceptional candidate zeros $0 \leq n \leq N$

define zeroSearch $(F, d, N, \nu, M, p)$:
    discs $\leftarrow$ empty list;
    zeros $\leftarrow$ empty list;
    Compute $(\Delta^k F)(0) \mod p^\nu$ for each $0 \leq k \leq d-1$;
    Compute $v_p((\Delta^k F)(0))$ from this for each $0 \leq k \leq d-1$;
    $j(F) \leftarrow$ largest integer $k$ for which $v_p((\Delta^k F)(0)) = \min_{0 \leq j \leq d-1} v_p((\Delta^j F)(0))$;
    if $j(F) > 0$ then append $\overline{D}(0,0)$ to discs;
    while discs $\neq \emptyset$ do
        foreach $\overline{D}(z,r) \in$ discs do
            foreach $0 \leq a \leq p-1$ do
                $z' \leftarrow z + p^r a$;
                Compute $(\Delta^k h_{z',r+1})(0) \mod p^\nu$ for each $0 \leq k \leq d-1$;
                Compute $\min\{v_p((\Delta^k h_{z',r+1})(0)), \nu\}$ from this for each $0 \leq k \leq d-1$;
                $j(h_{z',r+1}) \leftarrow$ largest integer $k$ for which
                  $v_p((\Delta^k h_{z',r+1})(0)) = \min_{0 \leq j \leq d-1} v_p((\Delta^j h_{z',r+1})(0))$;
                if $j(h_{z',r+1}) > 1$ then
                    if $p^{r+1}M + \ell \geq N$ then append $z'$ to zeros;
                    else append $\overline{D}(z', r+1)$ to discs;
            Remove $\overline{D}(z,r)$ from discs;
    return zeros
candidates $\leftarrow$ empty list;
APs $\leftarrow$ empty list;
Compute smallest prime $p$ with $p \geq d+2$ and $p \nmid a_0$;
$A \leftarrow$ companion matrix of $\boldsymbol{u}$;
$M \leftarrow$ the smallest positive integer such that $A^M \equiv I \mod p$;
$\nu \leftarrow \log(2^d 2^{(Md+k)\|\boldsymbol{u}\|} d^{Mn+k+1}) + rd + 1$;
foreach $0 \leq \ell \leq M-1$ do
    if $u_{Mn+\ell} = 0$ for every $n = 0, 1, \ldots, d-1$ then
        append $(M, \ell)$ to APs;
        next $\ell$
    else
        $F_\ell \leftarrow$ the function $n \mapsto u_{Mn+\ell}$ on $\mathbb{N}$;
        foreach $z$ in zeroSearch $(F_\ell, d, N, \nu, M, p)$ do
            append $Mz + \ell$ to candidates
        next $\ell$
output candidates
```

3.4 An example

Here we show an example of how the algorithm works for a particular LRS. Note this is only supposed to be illustrative so ν is chosen smaller than in Algorithm 1. Let $\boldsymbol{u}$ be defined by

$$u_{n+3} = 2u_{n+2} - 3u_{n+1} + u_n$$

and initial values $u_0 = -1, u_1 = 1, u_2 = 7$. We may take prime $p = 5$, and we can see that the companion matrix A satisfies $A^8 \equiv I \mod p$, thus we can take $M = 8$. Suppose that we are also given upper bound $N = 200$. For the benefit of presentation, we take $\nu = 5$. Following the algorithm, consider the subsequences $u_{8n}, u_{8n+1}, \ldots, u_{8n+7}$ and their corresponding p-adic analytic interpolations $F_0, F_1, \ldots, F_7$. It is easy to check that none of the subsequences are identically zero. To compute the number of zeros of F_ℓ in $\overline{D}(0,0)$, we compute each of $(\Delta^0 F_\ell)(0), (\Delta^1 F_\ell)(0), (\Delta^2 F_\ell)(0) \mod 5^5$.

ℓ	$(\Delta^0 F_\ell)(0)$	$(\Delta^1 F_\ell)(0)$	$(\Delta^2 F_\ell)(0)$
0	3124	80	400
1	1	130	2875
2	7	15	25
3	10	2845	1190
4	0	2650	2075
5	3102	3030	575
6	3089	955	2375
7	3122	1720	1975

Table 1: The values of $(\Delta^k F)(0) \mod 5^5$ for $0 \leq \ell \leq 7$

z	$(\Delta^0 h_{1,z})(0)$	$(\Delta^1 h_{1,z})(0)$	$(\Delta^0 h_{2,2+5z})(0)$	$(\Delta^1 h_{2,2+5z})(0)$
0	10	2475	650	500
1	2855	350	2625	500
2	650	1975	225	500
3	395	1100	2825	500
4	2215	850	1050	500

Table 2: The values of $(\Delta^k h_{1,z})(0) \mod 5^5$ for $0 \leq z \leq 4$ where $h_{1,z}(x) = F_3(px + z)$

z	$(\Delta^0 h_{1,z})(0)$	$(\Delta^1 h_{1,z})(0)$	$(\Delta^2 h_{1,z})(0)$	$(\Delta^0 h_{2,5z})(0)$	$(\Delta^1 h_{2,5z})(0)$	$(\Delta^0 h_{2,2+5z})(0)$	$(\Delta^1 h_{2,2+5z})(0)$
0	0	2750	1875	0	1250	1125	1875
1	2650	625	1875	2750	1250	2750	1875
2	1125	1625	1875	1125	1250	0	1875
3	2300	2625	1875	1375	1250	2250	1875
4	1800	500	1875	375	1250	125	1875

Table 3: The values of $(\Delta^k h_{1,z})(0) \mod 5^5$ for $0 \leq z \leq 4$ where $h_{1,z}(x) = F_4(px + z)$

For $\ell = 0, 1, 2, 5, 6, 7$ we have $v_5(\Delta^0 F_\ell)(0) < v_5(\Delta^1 F_\ell)(0), v_5(\Delta^2 F_\ell)(0)$ so $j(F_\ell) = 0$, therefore F_ℓ has no zeros in $\overline{D}(0, 0)$ for these values of ℓ. For $\ell = 3, 4$ we see that $j(F_3) = 1$ and $j(F_4) = 2$ so F_3, F_4 have 1 and 2 zeros in $\overline{D}(0, 0)$ respectively.

Next, we find the zeros of F_3. Since $j(F_3) = 1$, only $(\Delta^0 h_{z,r})(0), (\Delta^1 h_{z,r})(0)$ need to be computed to determine where the zero of F_3 lies, so for presentation $(\Delta^2 h_{z,r})(0)$ has been suppressed from the tables. Table 2 shows the only z for which $v_p((\Delta^0 h_{1,z})(0)) \geq v_p((\Delta^1 h_{1,z})(0))$ is $z = 2$ so the zero of F_3 lies in $\overline{D}(1, 2)$. Similarly, $v_p((\Delta^0 h_{2,7})(0)) \geq v_p((\Delta^1 h_{2,7})(0))$ so the zero of F_3 lies in $\overline{D}(2, 7)$. Since $N = 200$ and $Mp^2 + \ell = 203 \geq 200$, there is only one possible candidate integer zero of $\boldsymbol{u}$ corresponding to this; we need only check $u_{8\cdot 7+3} = u_{59}$.

We repeat the same process with F_4. Analysing the values of $(\Delta^k h_{z,r})(0)$ given in Table 3 shows $\overline{D}(1, 0) \supset \overline{D}(2, 0)$ and $\overline{D}(1, 2) \supset \overline{D}(2, 12)$ contain one zero of F_4 each. Again, since $Mp^2 + \ell = 204 \geq 200$, we need only check $u_{8\cdot 0+4} = u_4$ and $u_{8\cdot 12+4} = u_{100}$.

At this point, to decide the Skolem Problem we use an algorithm for EqSLP to decide whether $u_4 u_{59} u_{100} = 0$. To solve the function problem of determining all zeros, call an EqSLP oracle to check whether each u_4, u_{59}, $u_{100} = 0$. It turns out in this case that the only zero is $u_4 = 0$.

4 Conclusion

We have exhibited an algorithm that shows that for every $d \in \mathbb{N}$ the Bounded Skolem Problem lies in coRP for LRS of order at most d. As a corollary we showed that the Skolem Problem for LRS of order 4 lies also lies in coRP. The most natural route to prove decidability of Skolem's Problem is to exhibit an effective upper bound for the largest integer zero of a non-degenerate LRS (as is done in the order-4 case). It is plausible that such a bound has magnitude at most exponential in the description of the LRS—there is no known family of LRS for which the largest zero is super-exponential. The existence of such an exponential bound would yield a polynomial-time reduction of Skolem's Problem to the Bounded Skolem Problem.

References

[1] Manindra Agrawal, S. Akshay, Blaise Genest, and P. S. Thiagarajan. Approximate Verification of the Symbolic Dynamics of Markov Chains. *J. ACM*, 62(1):2:1–2:34, March 2015. doi:10.1145/2629417.

[2] S. Akshay, N. Balaji, A. Murhekar, R. Varma, and N. Vyas. Near-Optimal Complexity Bounds for Fragments of the Skolem Problem. In *37th International Symposium on Theoretical Aspects of Computer Science (STACS 2020)*, volume 154 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 37:1–37:18, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.STACS.2020.37.

[3] Shaull Almagor, Toghrul Karimov, Edon Kelmendi, Joël Ouaknine, and James Worrell. Deciding ω-regular properties on linear recurrence sequences. *Proc. ACM Program. Lang.*, 5(POPL):48:1–48:24, January 2021. doi:10.1145/3434329.

[4] Gilles Barthe, Charlie Jacomme, and Steve Kremer. Universal equivalence and majority of probabilistic programs over finite fields. In *Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '20, pages 155–166, New York, NY, USA, July 2020. Association for Computing Machinery. doi:10.1145/3373718.3394746.

[5] Paul C. Bell, Igor Potapov, and Pavel Semukhin. On the mortality problem: From multiplicative matrix equations to linear recurrence sequences and beyond. *Information and Computation*, 281:104736, December 2021. doi:10.1016/j.ic.2021.104736.

[6] Jean Berstel and Christophe Reutenauer. *Noncommutative Rational Series with Applications.* Cambridge University Press, 1 edition, 10 2010. doi:10.1017/CBO9780511760860.

[7] Y. Bilu. Skolem problem for linear recurrence sequences with 4 dominant roots (after mignotte, shorey, tijdeman, vereshchagin and bacik), 2025. arXiv:2501.16290.

[8] V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, September 2000. doi:10.1016/S0005-1098(00)00050-9.

[9] Vincent D. Blondel and Natacha Portier. The presence of a zero in an integer linear recurrent sequence is np-hard to decide. *Linear Algebra and its Applications*, 351-352:91–98, 2002. Fourth Special Issue on Linear Systems and Control. doi:10.1016/S0024-3795(01)00466-9.

[10] V. Chonev, J. Ouaknine, and J. Worrell. On the complexity of the orbit problem. *J. ACM*, 63(3), June 2016. doi:10.1145/2857050.

[11] Graham Everest, Alf Van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence sequences.* Number v. 104 in Mathematical surveys and monographs. American Mathematical Society, Providence, RI, 2003.

[12] Christer Lech. A note on recurring series. *Arkiv för Matematik*, 2(5):417–421, 8 1953. doi:10.1007/BF02590997.

[13] K. Mahler. Eine arithmetische Eigenschaft der Taylor-Koeffizienten rationaler Funktionen. *Proc. Akad. Wet. Amsterdam*, 38:50–60, 1935.

[14] Joël Ouaknine and James Worrell. On linear recurrence sequences and loop termination. *ACM SIGLOG News*, 2(2):4–13, April 2015. doi:10.1145/2766189.2766191.

[15] A. Robert. *A course in p-adic analysis*. Number 198 in Graduate texts in mathematics. Springer, New York, NY, 2000. doi:10.1007/978-1-4757-3254-2.

[16] A. Salomaa and M. Soittola. *Automata-theoretic aspects of formal power series*. Texts and monographs in computer science. Springer, 1978.

[17] Arnold Schönhage. On the power of random access machines. In Hermann A. Maurer, editor, *Automata, Languages and Programming*, pages 520–529, Berlin, Heidelberg, 1979. Springer Berlin Heidelberg.

[18] Th. Skolem. Ein Verfahren zur Behandlung gewisser exponentialer Gleichungen und diophantischer Gleichungen. 8. Skand. Mat.-Kongr., 163-188 (1935)., 1935.

[19] T. Tao. *Structure and Randomness*. American Mathematical Society, 2008.

[20] R. Tijdeman, M. Mignotte, and T.N. Shorey. The distance between terms of an algebraic recurrence sequence. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1984(349):63–76, May 1984. doi:10.1515/crll.1984.349.63.

[21] A.J. van der Poorten and H.P. Schlickewei. Zeros of recurrence sequences. *Bulletin of the Australian Mathematical Society*, 44(2):215–223, 1991. doi:10.1017/S0004972700029646.

[22] N. K. Vereshchagin. Occurrence of zero in a linear recursive sequence. *Mathematical Notes of the Academy of Sciences of the USSR*, 38(2):609–615, 1985. doi:10.1007/BF01156238.