On Large Zeros of Linear Recurrence Sequences

Florian Luca
Mathematics Division, Stellenbosch University, South Africa

Joël Ouaknine
Max Planck Institute for Software Systems, Germany

James Worrell
Department of Computer Science, Oxford University, UK

Abstract

The Skolem Problem asks to determine whether a given integer linear recurrence sequence (LRS) has a zero term. This problem, whose decidability has been open for many decades, arises across a wide range of topics in computer science, including loop termination, formal languages, automata theory, and probabilistic model checking, amongst many others.

In the present paper, we introduce a notion of "large" zeros of linear recurrence sequences, i.e., zeros occurring at an index larger than a sixth-fold exponential of the size of the data defining the given LRS. We establish two main results. First, we show that large zeros are very sparse: the set of positive integers that can possibly arise as large zeros of some LRS has null density. This in turn immediately yields a Universal Skolem Set of density one, answering a question left open in the literature. Second, we define an infinite set of prime numbers, termed "good", having density one amongst all prime numbers, with the following property: for any large zero of a given LRS, there is an interval around the large zero together with an upper bound on the number of good primes possibly present in that interval. The bound in question is much lower than one would expect if good primes were similarly distributed as ordinary prime numbers, as per the Cramér model in number theory. We therefore conjecture that large zeros do not exist, which would entail decidability of the Skolem Problem.

2012 ACM Subject Classification Mathematics of computing → Discrete mathematics

Keywords and phrases Skolem Problem, linear recurrence sequences, decidability, Cramér conjecture

1 Introduction

An (integer) linear recurrence sequence (LRS) $\langle u_n \rangle_{n=0}^{\infty}$ is a sequence of integers satisfying a recurrence of the form

$$u_{n+k} = a_1 u_{n+k-1} + \cdots + a_k u_n \tag{1}$$

where the coefficients $a_1, \ldots, a_k$ are integers. The celebrated theorem of Skolem, Mahler, and Lech [27, 19, 15] describes the set of zero terms of such a recurrence:

▶ **Theorem 1.** *Given an integer linear recurrence sequence $\langle u_n \rangle_{n=0}^{\infty}$, the set $\{n \in \mathbb{N} : u_n = 0\}$ is a union of finitely many arithmetic progressions together with a finite set.*

The statement of Thm. 1 can be refined by considering the notion of *non-degeneracy* of an LRS. An LRS is non-degenerate if in its minimal recurrence no quotient of two distinct roots of the characteristic polynomial is a root of unity.[1] A given LRS can be effectively decomposed as the merge 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 zero terms. Unfortunately, all

[1] For basic definitions, facts, and properties concerning linear recurrence sequences, we refer the reader to standard texts such as [9, Chaps. 1 and 2], [14, Chap. 4], or [28, Chap. 4].

known proofs of this last result are ineffective: it is not known how to compute the finite set of zeros of a given non-degenerate linear recurrence sequence. It is readily seen that existence of a procedure to do so is equivalent to the existence of a procedure to determine whether an arbitrary given LRS has a zero term; the latter is known as the Skolem Problem. We refer to [4, Chap. 6] and [29, Chap. X] for expository accounts of the Skolem-Mahler-Lech theorem and discussion of the ineffectiveness of known proofs.

Decidability of the Skolem Problem is known only for certain special cases, based on the relative order of the absolute values of the characteristic roots. Say that a characteristic root λ is *dominant* if its absolute value is maximal among all the characteristic roots. Decidability is known in case there are at most 3 dominant characteristic roots, and also for recurrences of order at most 4 [20, 30]. However for LRS of order 5 it is not currently known how to decide the Skolem Problem. For a (highly restricted) subclass of LRS, the paper [1] obtains nearly matching complexity lower and upper bounds for the problem.

Some recent lines of research have succeeded in establishing conditional decidability of the Skolem Problem for simple LRS (i.e., LRS none of whose characteristic roots are repeated), assuming certain classical number-theoretic conjectures [16, 5]. Nevertheless, to the best of our knowledge, no putative algorithm has to date been proposed to solve the Skolem Problem in full generality.

A different approach was initiated in [17, 18] via the notion of *Universal Skolem Sets*. An infinite, recursive set $\mathcal{S} \subseteq \mathbb{N}$ is a Universal Skolem Set if there is some algorithm which, given any LRS, determines whether or not the LRS has a zero in $\mathcal{S}$. Decidability of the Skolem Problem is then of course equivalent to the assertion that $\mathbb{N}$ is itself a Universal Skolem Set. The authors of [17] succeded in exhibiting a *sparse* Universal Skolem Set, i.e., a set having null density, and left open the question of whether Universal Skolem Sets of strictly positive density, or even density one, could be constructed (the interest in such sets being that they approximate $\mathbb{N}$ more and more closely). The question was partially answered in [18], which presented a positive-density Universal Skolem Set albeit restricted to simple LRS.

In computer science, the Skolem Problem lies at the heart of key decision problems in formal power series [25, 3], stochastic model checking [24], control theory [6, 10], and loop termination [23]. The problem is also closely related to membership problems on commutative matrix groups and semigroups, as considered in [7, 13]. We note that in several of the above-mentioned citations, the Skolem Problem is used as a reference benchmark to establish hardness of other open decision problems.

In this paper we propose an explicit bound for the largest zero of a non-degenerate LRS in terms of the data describing the LRS. We call zeros that exceed this bound *large zeros* of the LRS. Evidently, decidability of the Skolem Problem would follow from a proof that large zeros do not exist. Using known upper bounds on the cardinality of the set of zeros of non-degenerate LRS, it is relatively straightforward to show that the set of integers arising as large zeros of some non-degenerate LRS has null density, which in turn yields a Universal Skolem Set of density one. While a proof that large zeros do not exist currently seems well out of reach, we give a heuristic argument as to why this should nevertheless be expected. This argument is based on an analogue of the well-known Cramér conjecture on gaps between consecutive primes. This conjecture, originally formulated by Cramér in 1936 [8] and subsequently refined by various number theorists into its present form, asserts that, for some constant $\kappa > 1$, for every prime p the distance to the next largest prime is at most $\kappa(\log p)^2$. The conjecture is based on the heuristic that the sequence of prime numbers behaves similarly to a Poisson-like random process in which the probability of a number x being prime is $1/\log x$. The largest observed prime gaps are of the order of $0.5(\log p)^2$ [22],

however the best known upper bound on prime gaps is $O(p^{0.525})$, due to Baker, Harman, and Pintz [2], which is far from Cramér's conjectured bound. Cramér himself proved that, under the Riemann hypothesis, prime gaps are bounded above by $O(p^{0.5} \log p)$ [8]. On the other hand, the best known lower bound is $\Omega(\log p \log\log p)$, which is some way from the conjectured upper bound. We refer to [12] for a discussion of Cramér's conjecture and its refinements.

Here we define a subset of so-called *good* primes based on divisibility properties of LRS. We show that the set of good primes has density one in the set of all primes. We further show that if the Cramér conjecture applies also to gaps between consecutive good primes, then large zeros of LRS cannot exist. The proof of the latter result proceeds by establishing an upper bound on the number of good primes in the neighbourhood of a large zero that violates the conjectured upper bound on gaps between good primes. In other words, if good primes are distributed according to Cramér's heuristic then large zeros cannot exist and the Skolem Problem is decidable.

2 Background

We will need some basic notions concerning algebraic numbers. All material can be found in [11]. Recall that a *number field* $\mathbb{K}$ is a subfield of $\mathbb{C}$ that is finite dimensional as a vector space over $\mathbb{Q}$. We assume that $\mathbb{K}$ is a Galois extension of $\mathbb{Q}$, that is, it arises as the splitting field of a polynomial with integer coefficients. All elements of $\mathbb{K}$ are algebraic over $\mathbb{Q}$, that is, they arise as roots of polynomials with integer coefficients. Those elements that arise more specifically as roots of monic polynomials with integer coefficients are called *algebraic integers*. The algebraic integers in $\mathbb{K}$ form a subring, denoted $\mathcal{O}_{\mathbb{K}}$.

For a number field $\mathbb{K}$, we denote by $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ the group of field automorphisms of $\mathbb{K}$. Given $\alpha \in \mathbb{K}$, the *norm* of α is defined by

$$N_{\mathbb{K}/\mathbb{Q}}(\alpha) = \prod_{\sigma \in \mathrm{Gal}(\mathbb{K}/\mathbb{Q})} \sigma(\alpha)\,.$$

The norm $N_{\mathbb{K}/\mathbb{Q}}(\alpha)$ is rational for all $\alpha \in \mathbb{K}$; moreover $N_{\mathbb{K}/\mathbb{Q}}(\alpha) = 0$ iff $\alpha = 0$, and $N_{\mathbb{K}/\mathbb{Q}}(\alpha)$ is an integer if $\alpha \in \mathcal{O}_{\mathbb{K}}$. Clearly we have $|N_{\mathbb{K}/\mathbb{Q}}(\alpha)| \leq M^{d_{\mathbb{K}}}$, where $d_{\mathbb{K}}$ is the degree of $\mathbb{K}$ and

$$M := \max_{\sigma \in \mathrm{Gal}(\mathbb{K}/\mathbb{Q})} |\sigma(\alpha)|$$

is the *house* of α.

We recall that every ideal in $\mathcal{O}_{\mathbb{K}}$ can be written uniquely up to the order of its factors as the product of prime ideals. Given a rational prime $P \in \mathbb{Z}$, we say that a prime ideal $\mathfrak{p}$ *lies above* P if $\mathfrak{p}$ is a factor of $P\mathcal{O}_{\mathbb{K}}$. In this case we have that $P \mid N_{\mathbb{K}/\mathbb{Q}}(\alpha)$ for all $\alpha \in \mathfrak{p}$.

Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_{\mathbb{K}}$ lying above $P \in \mathbb{Z}$. Recall that the Frobenius automorphism $\sigma \in \mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ corresponding to $\mathfrak{p}$ is such that $\sigma(\alpha) \equiv \alpha^P \bmod \mathfrak{p}$ for all $\alpha \in \mathcal{O}_{\mathbb{K}}$.

3 Large Zeros

For an LRS $\mathbf{u} = \langle u_n \rangle_{n=0}^{\infty}$ as in (1), define its size to be

$$C_{\mathbf{u}} := \max\{k, |a_1|, \ldots, |a_k|, |u_0|, \ldots, |u_{k-1}|, 10\}\,.$$

Given a (partial) function $f : \mathbb{R} \to \mathbb{R}$ and a positive integer ℓ, let $f_\ell(x) = f \circ f \circ \cdots \circ f(x)$, where the iteration is ℓ-fold (thus $f_1 = f$). We say that n is a zero of $\mathbf{u}$ if $u_n = 0$, and we

say that it is a *large zero* if the inequality

$$n < 2\exp_6(C_{\mathbf{u}}) \tag{2}$$

fails.

As we argue later on, there are good reasons to expect that (2) holds for all zeros of all non-degenerate LRS, which in turn would establish decidability of the Skolem Problem. Unfortunately, we are unable to prove this assertion. Nevertheless, we will show that large zeros are very sparse, i.e., have null density amongst the positive integers.

Let us first define what it means for a prime to be *bad*. Let $X > 2\exp_7(1)$ be an integer. We say that the LRS $\mathbf{u}$ is *small at level* X if it has size $C_{\mathbf{u}} \leq \log_6(X/2)$. Let us write $C := C_{\mathbf{u}}$ (that is, we omit the dependence on $\mathbf{u}$). We can express the general term u_t of $\mathbf{u}$ as

$$u_t = \sum_{i=1}^{s} Q_i(t)\alpha_i^t\,,$$

where $s \leq C$ and $\alpha_1, \ldots, \alpha_s$ are the roots of the characteristic polynomial

$$x^k - a_1x^{k-1} - \cdots - a_k$$

of $\mathbf{u}$ and $Q_1, \ldots, Q_s$ are univariate polynomials. Recall that if α_i has multiplicity μ_i as a characteristic root then Q_i has degree at most $\mu_i - 1$. Let $\mathbb{K} := \mathbb{Q}(\alpha_1, \ldots, \alpha_s)$. The coefficients of each Q_i are in $\mathbb{K}$ and can be computed from the initial values $u_0, \ldots, u_{k-1}$ of the sequence by solving a system of linear equations.

By Cramer's determinant rule,[2] each of the coefficients of Q_i is the quotient of an algebraic integer by the determinant[3]

$$\Delta := \begin{vmatrix} 1 & \ldots & 0 & 1 & \ldots & 0 & 1 & \cdots \\ \alpha_1 & \ldots & \alpha_1 & \alpha_2 & \ldots & \alpha_{s-1} & \alpha_s & \ldots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \alpha_1^{k-1} & \ldots & (k-1)^{\mu_1-1}\alpha_1^{k-1} & \alpha_2^{k-1} & \ldots & (k-1)^{\mu_s-1}\alpha_{s-1}^{k-1} & \alpha_s^{k-1} & \ldots \end{vmatrix}.$$

By the Cauchy root bound we have $|\alpha_i| \leq 1 + C$ for $i \in \{1, \ldots, s\}$. It follows that the squared Euclidean norm of each column vector above is at most

$$k(k-1)^{2(k-1)}(1+C)^{2k} < k^{2k}(1+C)^{2k}\,.$$

Thus, by the Hadamard inequality,

$$\Delta^2 < (k^{2k}(1+C)^{2k})^k = (k(1+C))^{2k^2}\,.$$

Let us now replace $\mathbf{u}$ by $\mathbf{v} := \Delta^2\mathbf{u}$, noting that $\mathbf{u}$ and $\mathbf{v}$ have the same zeros. We then have that

$$P_i := \Delta^2 Q_i(t) = \sum_{j=0}^{\mu_i-1} c_{i,j}t^j\,,$$

---

[2] This rule is named after the 18th-century Genevan mathematician Gabriel Cramer, who is presumably unrelated to the 20th-century Swedish mathematician Harald Cramér, whose work plays an important role in motivating the present article.

[3] The underlying matrix has s blocks, one for each characteristic root. For $\ell \in \{1, \ldots, s\}$ the ℓ-th block has dimension $k \times \mu_\ell$ and has (i,j)-th element $(i-1)^{(j-1)}\alpha_\ell^{(i-1)}$ for $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, \mu_\ell\}$.

where $c_{i,j} \in \mathcal{O}_{\mathbb{K}}$ are algebraic integers whose house is at most

$$(1+C)\Delta^2 < (1+C)^{4C^2+1} < C^{C^3}\,.$$

Let $\sigma \in \Sigma_s$ be any permutation of the first s integers and let

$$\beta_i := \alpha_{\sigma(i)} \quad \text{for} \quad i = 1, \ldots, s\,.$$

For some nonnegative integer m consider the algebraic integer

$$v_{m,\sigma} = \sum_{i=1}^{s} P_i(m)\beta_i\alpha_i^m\,. \tag{3}$$

▶ **Definition 2.** *We say that $P \in [X, 2X]$ is bad, if there exists an LRS $\mathbf{u}$ which is small at level X, a permutation $\sigma \in \Sigma_s$, and an integer $m \in [0, X^{1/4}]$, such that*

- *The algebraic integer $v_{m,\sigma}$ defined in (3) above is non-zero, and*
- *P is a prime factor of $\mathcal{N}_{\mathbb{K}/\mathbb{Q}}(v_{m,\sigma})$.*

Let $\mathcal{P}_{\rm bad}(X)$ be the set of bad primes in $[X, 2X]$.

▶ **Proposition 3.** *We have*

$$\#\mathcal{P}_{\rm bad}(X) < X^{2/3}$$

for all $X > X_0$, where X_0 is some effective absolute constant.

Proof. In order to estimate the size of $\mathcal{P}_{\rm bad}(X)$, we first need to find out:

1. How many such expressions (3) are there?
2. How large are they?

For (**1**), the coefficients $a_1, \ldots, a_k$ and initial values $u_0, \ldots, u_{k-1}$ are all in $[-C, C]$, an interval containing at most $2C + 1 < 3C$ integers. Altogether for fixed k there are at most $(3C)^{2k} \le (3C)^{2C}$ $2k$-tuples, and summing up over k we derive an upper bound of $C(3C)^{2C} < C^{3C}$ distinct possible LRS of size at most C. This in turn is an upper bound on the number of s-tuples $((Q_i, \alpha_i))_{i=1}^{s}$. We must then multiply this quantity with the number of possible permutations of the characteristic roots, which is at most $C! < C^C$. There are therefore at most C^{4C} linear recurrence sequences $\mathbf{w} = \langle w_m \rangle_{m=0}^{\infty}$ whose m-th term is given by

$$w_m = \sum_{i=1}^{s} P_i(m)\beta_i\alpha_i^m \quad \text{for all} \quad m \ge 0\,.$$

This answers (**1**). As for (**2**), recall that the coefficients of P_i are of size at most C^{C^3}. There are at most C terms, the largest monomial involved in $P_i(m)$ is at most $m^C < X^C$ and the largest root has magnitude at most $1 + C < 2C$. Thus each individual term is of absolute value at most

$$\begin{aligned} C^{C^3+1}(2C)X^C(2C)^{X^{1/4}} &= \exp\Big((C^3+1)\log C + \log(2C) + C\log X + X^{1/4}\log(2C)\Big) \\ &< \exp(X^{0.26}) \end{aligned}$$

for $X > X_0$, since C is tiny in comparison to X. Hence the norm of the number shown in (3) is of size at most

$$\exp(C!X^{0.26}) < \exp(X^{0.27}) \quad \text{for} \quad X > X_0\,,$$

since the degree of $\mathbb{K}$ is at most $C!$ (as $\mathbb{K}$ is the splitting field of a polynomial of degree at most C). Moreover, as noted earlier, there are at most C^{4C} such expressions. Thus a bad prime P divides an integer which is a product of such numbers and is of size at most

$$\exp(C^{4C}X^{0.27}) < \exp(X^{0.28}) \quad \text{for} \quad X > X_0\,.$$

Therefore the number of possible choices for P is at most $X^{0.28}$. Since the number of choices for m is at most $X^{0.25}$, we conclude that, for $X > X_0$, the cardinality of $\mathcal{P}_{\text{bad}}(X)$ is at most

$$X^{0.25+0.28} < X^{2/3}\,,$$

as required. ◀

We will shortly present heuristic arguments, in the spirit of Cramér, to the effect that large zeros cannot exist. Let us however first prove that the set of large zeros is, at the very least, sparse. To this end, let

$$\mathcal{L} := \{n \in \mathbb{N} : \text{there exists a non-degenerate LRS } \mathbf{u} \text{ such that } u_n = 0 \text{ and (2) fails}\}\,.$$

Thus $\mathcal{L}$ is the set of large zeros of *some* non-degenerate LRS.

▶ **Theorem 4.** *The set $\mathcal{L}$ has null density. In fact, writing $\mathcal{L}(X) = \mathcal{L} \cap [0, X]$, the inequality*

$$\#\mathcal{L}(X) = O\left(\frac{X}{(\log X)^B}\right)$$

holds with any constant $B > 0$.

Proof. Let $X > 2\exp_7(1)$, and put $C := \lceil \log_6(X/2) \rceil$. We wish to count the number of large zeros in the interval $[0, X]$. By definition, any such zero originates from an LRS $\mathbf{u}$ which is small at level X, i.e., having size $C_{\mathbf{u}} \leq C$. As shown in the proof of Prop. 3, there are at most C^{3C} such LRS. On the other hand, Schmidt [26] proved that any LRS of order k or less has at most $\exp_3(3k \log k) < \exp_4(C)$ zeros, assuming k is taken to be at most C. Hence the total number of zeros emanating from such LRS is at most $\exp_4(C)C^{3C}$, whence the inequality

$$\#\mathcal{L}(X) \ll \frac{X}{(\log X)^B}$$

easily follows for any $B > 0$. ◀

▶ **Corollary 5.** *The set $\mathcal{S} := \mathbb{N} \setminus \mathcal{L}$ is a Universal Skolem Set of density one.*

Proof. It is clear that the set $\mathcal{L}$ is recursive, and hence that $\mathcal{S}$ is recursive as well.

Density one follows from Thm. 4, and universality follows from the fact that $\mathcal{S}$, by definition, doesn't contain any large zeros. Thus given any nondegenerate LRS $\mathbf{u}$ of size $C_{\mathbf{u}}$, its only possible zeros in $\mathcal{S}$ can only lie in the interval $[0, 2\exp_6(C_{\mathbf{u}})]$, which can readily be checked. ◀

In the remainder of the paper we present a heuristic argument supporting the much stronger assertion that large zeros do not exist, or in other words that the set $\mathcal{L}$ is empty. The strategy is as follows. Recall that, according to Prop. 3 along with the prime number theorem, the set of bad primes has null density amongst the prime numbers.

We define an infinite set of prime numbers, termed "good", that has density one amongst all prime numbers. Assuming that good primes are distributed similarly as ordinary primes, according to the Cramér model in number theory, we would expect that Cramér's conjecture on gaps between primes applies also to good primes. More precisely, this conjecture postulates the existence of precise upper bounds on the largest possible gap between consecutive primes, and is predicated on the heuristic that the primes behave as a set of randomly distributed integers with asymptotic density conforming to the prime number theorem. However we show that around any large zero of an LRS there is an interval and an upper bound on the number of good primes in the interval that together contradict the above Cramér-type conjecture on gaps between good primes. We therefore conjecture that large zeros do not exist.

Let us write $\mathcal{P} = \{p_1, p_2, \ldots\}$ to denote the set of prime numbers, enumerated in increasing order, and let $\mathcal{P}_{\text{good}} := \mathcal{P} \backslash \mathcal{P}_{\text{bad}} = \{g_1, g_2, \ldots\}$ denote the set of *good* primes, again enumerated in increasing order.

▶ **Conjecture 6** (Cramér-Granville). *For some $\kappa > 1$,*

$$\limsup_{j\to\infty} \frac{p_{j+1} - p_j}{\log^2 p_j} = \kappa\,.$$

Cramér initially suggested that the constant κ in Conjecture 6 might be 1 [8], but several decades later, building on substantial developments in the field, Granville produced evidence that $\kappa \geq 2e^{-\gamma} \approx 1.1229\ldots$, where γ is the Euler–Mascheroni constant [12]. There is in any event considerable computational evidence in support of the Cramér-Granville conjecture [22, 21].

As noted earlier, thanks to Prop. 3 and the prime number theorem, good primes have density one amongst all prime numbers:

$$\lim_{X\to\infty} \frac{\#\left(\mathcal{P}_{\text{good}} \cap [0, X]\right)}{\#\left(\mathcal{P} \cap [0, X]\right)} = 1\,.$$

Accordingly, it seems reasonable to suppose that, asymptotically speaking, good primes should behave similarly to ordinary primes, or at least should exhibit similar "statistical" properties. We therefore formulate:

▶ **Conjecture 7.** *For some $\eta > 1$,*

$$\limsup_{j\to\infty} \frac{g_{j+1} - g_j}{\log^2 g_j} = \eta\,.$$

We now have the following result.

▶ **Theorem 8.** *Conjecture 7 implies that large zeros of LRS do not exist; or more precisely, that $\mathcal{L}$ is a finite set.*

Proof. Conjecture 7 can be reformulated as follows: there exist $\eta > 1$ and $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, the interval

$$[n - \eta(\log n)^2, n]$$

always contains some good prime. In turn, this implies that the interval $[n - \eta(\log n)^3, n]$ must contain at least $\log n$ distinct good primes for n sufficiently large (say $n \geq n_1 \geq \max\{n_0, 2\exp_7(1)\}$).

Thus let $n \geq n_1$, put $C := \log_6(n/2)$, and suppose that there is some LRS $\mathbf{u}$ with $C_{\mathbf{u}} \leq C$ such that $u_n = 0$—in other words, n is a large zero of $\mathbf{u}$. Write $n = P + m$,

where $P \in [n - \eta(\log n)^3, n]$ is a good prime and $0 \leq m < \eta(\log n)^3 < n^{1/4}$.[4] As in the proof of Prop. 3, let $\alpha_1, \ldots, \alpha_s$ be the characteristic roots of $\mathbf{u}$, and let Δ^2 be the smallest positive integer such that, writing $\mathbf{v} := \Delta^2\mathbf{u}$, every term of v_t of $\mathbf{v}$ has a representation as an exponential polynomial

$$v_t = \sum_{i=1}^{s} P_i(t)\alpha_i^t$$

in which all polynomials P_i have algebraic-integer coefficients.

Since $u_n = v_n = 0$, we get

$$0 = \sum_{i=1}^{s} P_i(P+m)\alpha_i^{P+m}\,.$$

We now reduce the above equation modulo $\mathfrak{p}$, where $\mathfrak{p}$ is some prime ideal of $\mathcal{O}_{\mathbb{K}}$ dividing P, from which we deduce that P divides

$$\mathcal{N}_{\mathbb{K}/\mathbb{Q}}\left(\sum_{i=1}^{s} P_i(m)\beta_i\alpha_i^m\right)\,, \tag{4}$$

where each $\beta_i = \sigma(\alpha_i)$ is obtained from applying the Frobenius automorphism induced by $\mathfrak{p}$ in $\mathbb{K}$ to α_i. If the above expression (4) were non-zero, we would have to conclude that $P \in \mathcal{P}_{\text{bad}}(n/2)$, contradicting our choice of P. Thus the expression (4) is zero.

Let us count how many expressions of the form (4) can vanish. More precisely, consider the (complex-valued) LRS $\mathbf{w} = \langle w_j\rangle_{j=0}^{\infty}$ whose j-th term is given by

$$w_j = \sum_{i=1}^{s} P_i(j)\beta_i\alpha_i^j \quad \text{for all} \quad j \geq 0\,,$$

and whose order is at most C. Schmidt [26] proves that the number of distinct positive integers m such that $w_m = 0$ is at most

$$\exp_3(3C\log C) < \exp_4(C)\,.$$

Of course, given $\mathbf{u}$, the s-tuple $(\beta_1, \ldots, \beta_s)$ can be chosen in at most $s! < C^C$ ways. Thus the total number of possible zeros for expression (4) is at most $C^C \exp_4(C) < \exp_5(C)$. Since distinct choices of P give rise to distinct such zeros, and (as noted earlier) there are at least $\log n$ possible choices for P, we conclude that

$$\log n < \exp_5(C)\,,$$

or equivalently $n < \exp_6(C) = n/2$, a contradiction. It thus follows, as claimed, that Conjecture 7 prohibits the existence of large zeros of LRS that are greater than the absolute constant n_1. ◀

Thanks to Thm. 8, Conjecture 7 implies the existence of an algorithm to solve the Skolem Problem. Given an LRS $\mathbf{u}$, first decompose $\mathbf{u}$ into finitely many non-degenerate LRS, and check that none of these is identically zero. Next, for each sub-LRS $\mathbf{v}$ of size $C_{\mathbf{v}}$, simply search for a zero up to index $2\exp_6(C_{\mathbf{v}})$.[5] If at the end of this process no zero has been found for any of the LRS, return that $\mathbf{u}$ has no zeros.

---

[4] Assume without loss of generality that n_1 has been chosen sufficiently large also to ensure the validity of the last of these inequalities.

[5] Technically speaking, the algorithm should examine all terms up to index $\max\{2\exp_6(C_{\mathbf{v}}), n_1\}$, where n_1 is the absolute constant appearing in the proof of Thm. 8. The *existence* of n_1 is implied by Conjecture 7, but its effectivity would depend on the effectivity of Conjecture 7.

References

1 S. Akshay, N. Balaji, A. Murhekar, R. Varma, and N. Vyas. Near-optimal complexity bounds for fragments of the Skolem problem. In *STACS*, volume 154 of *LIPIcs*, pages 37:1–37:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

2 R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes, ii. *Proceedings of the London Mathematical Society*, 83(3):532–562, 2001.

3 D. Beauquier, A. M. Rabinovich, and A. Slissenko. A logic of probability with decidable model checking. *J. Log. Comput.*, 16(4), 2006.

4 J. Berstel and C. Reutenauer. *Noncommutative Rational Series with Applications*. Cambridge University Press, 2010.

5 Y. Bilu, F. Luca, J. Nieuwveld, J. Ouaknine, D. Purser, and J. Worrell. Skolem meets Schanuel. In *MFCS*, volume 241 of *LIPIcs*, pages 20:1–20:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

6 V. Blondel and J. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, 2000.

7 J.-Y. Cai, R. J. Lipton, and Y. Zalcstein. The complexity of the A B C problem. *SIAM J. Comput.*, 29(6), 2000.

8 H. Cramér. On the order of magnitude of the difference between consecutive prime numbers. *Acta arithmetica*, 2:23–46, 1936.

9 G. Everest, A. van der Poorten, I. Shparlinski, and T. Ward. *Recurrence Sequences*. American Mathematical Society, 2003.

10 N. Fijalkow, J. Ouaknine, A. Pouly, J. Sousa Pinto, and J. Worrell. On the decidability of reachability in linear time-invariant systems. In *HSCC*, pages 77–86. ACM, 2019.

11 A. Fröhlich and M. J. Taylor. *Algebraic Number Theory*, volume 27 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1993.

12 A. Granville. Harald Cramér and the distribution of prime numbers. *Scandinavian Actuarial Journal*, 1995(1):12–28, 1995.

13 R. Kannan and R. J. Lipton. Polynomial-time algorithm for the orbit problem. *JACM*, 33(4), 1986.

14 M. Kauers and P. Paule. *The Concrete Tetrahedron - Symbolic Sums, Recurrence Equations, Generating Functions, Asymptotic Estimates*. Texts & Monographs in Symbolic Computation. Springer, 2011.

15 C. Lech. A note on recurring series. *Ark. Mat.*, 2, 1953.

16 R. Lipton, F. Luca, J. Nieuwveld, J. Ouaknine, D. Purser, and J. Worrell. On the Skolem problem and the Skolem conjecture. In *LICS*, pages 5:1–5:9. ACM, 2022.

17 F. Luca, J. Ouaknine, and J. Worrell. Universal Skolem sets. In *LICS*, pages 1–6. IEEE, 2021.

18 F. Luca, J. Ouaknine, and J. Worrell. A universal Skolem set of positive lower density. In *MFCS*, volume 241 of *LIPIcs*, pages 73:1–73:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

19 K. Mahler. Eine arithmetische Eigenschaft der Taylor Koeffizienten rationaler Funktionen. *Proc. Akad. Wet. Amsterdam*, 38, 1935.

20 M. Mignotte, T. Shorey, and R. Tijdeman. The distance between terms of an algebraic recurrence sequence. *J. für die reine und angewandte Math.*, 349, 1984.

21 T. R. Nicely. New maximal prime gaps and first occurrences. *Math. Comput.*, 68(227):1311–1315, 1999.

22 A. Odlyzko, M. Rubinstein, and M. Wolf. Jumping champions. *Experimental Mathematics*, 8(2):107–118, 1999.

23 J. Ouaknine and J. Worrell. On linear recurrence sequences and loop termination. *ACM SIGLOG News*, 2(2):4–13, 2015.

24 J. Piribauer and C. Baier. On Skolem-hardness and saturation points in Markov decision processes. In *ICALP*, volume 168 of *LIPIcs*, pages 138:1–138:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

25 G. Rozenberg and A. Salomaa. *Cornerstones of Undecidability.* Prentice Hall, 1994.

26 W. M. Schmidt. The zero multiplicity of linear recurrence sequences. *Acta Math.*, 182:243–282, 1999.

27 T. Skolem. Ein Verfahren zur Behandlung gewisser exponentialer Gleichungen. In *Comptes rendus du congrès des mathématiciens scandinaves*, 1934.

28 R. P. Stanley. Enumerative combinatorics. *Cambridge studies in advanced mathematics*, Volume 1, 2nd Edition, 2011.

29 T. Tao. *Structure and randomness: pages from year one of a mathematical blog.* American Mathematical Society, 2008.

30 N. K. Vereshchagin. The problem of appearance of a zero in a linear recurrence sequence (in Russian). *Mat. Zametki*, 38(2), 1985.