Effective Positivity Problems for Simple Linear Recurrence Sequences*

Joël Ouaknine and James Worrell

Department of Computer Science, Oxford University, UK
{joel,jbw}@cs.ox.ac.uk

Abstract. We consider two computational problems for linear recurrence sequences (LRS) over the integers, namely the *Positivity Problem* (determine whether all terms of a given LRS are positive) and the *effective Ultimate Positivity Problem* (determine whether all but finitely many terms of a given LRS are positive, and if so, compute an index threshold beyond which all terms are positive). We show that, for simple LRS (those whose characteristic polynomial has no repeated roots) of order 9 or less, Positivity is decidable, with complexity in the Counting Hierarchy, and effective Ultimate Positivity is solvable in polynomial time.

1 Introduction

A (real) ***linear recurrence sequence (LRS)*** is an infinite sequence $\mathbf{u} = \langle u_0, u_1, u_2, \ldots \rangle$ of real numbers having the following property: there exist constants $a_1, a_2, \ldots, a_k$ (with $a_k \neq 0$) such that, for all $n \geq 0$,

$$u_{n+k} = a_1 u_{n+k-1} + a_2 u_{n+k-2} + \ldots + a_k u_n \,. \tag{1}$$

If the initial values $u_0, \ldots, u_{k-1}$ of the sequence are provided, the recurrence relation defines the rest of the sequence uniquely. Such a sequence is said to have ***order*** k.[1]

The best-known example of an LRS was given by Leonardo of Pisa in the 12th century: the Fibonacci sequence $\langle 0, 1, 1, 2, 3, 5, 8, 13, \ldots \rangle$, which satisfies the recurrence relation $u_{n+2} = u_{n+1} + u_n$. Leonardo of Pisa introduced this sequence as a means to model the growth of an idealised population of rabbits. Not only has the Fibonacci sequence been extensively studied since, but LRS now form a vast subject in their own right, with

* This research was partially supported by EPSRC. We are also grateful to Matt Daws for considerable assistance in the initial stages of this work.

[1] Some authors define the order of an LRS as the *least* k such that the LRS obeys such a recurrence relation. The definition we have chosen allows for a simpler presentation of our results and is algorithmically more convenient.

numerous applications in mathematics and other sciences. A deep and extensive treatise on the mathematical aspects of recurrence sequences is the recent monograph of Everest *et al.* [13].

Given an LRS $\mathbf{u}$ satisfying the recurrence relation (1), the ***characteristic polynomial*** of $\mathbf{u}$ is

$$p(x) = x^n - a_1 x^{n-1} - \ldots - a_{k-1} x - a_k \,.$$

An LRS is said to be ***simple*** if its characteristic polynomial has no repeated roots. Simple LRS, such as the Fibonacci sequence, possess a number of desirable properties which considerably facilitates their analysis—see, e.g., [13, 14]. They constitute a large and well-studied class of sequences, and correspond to *diagonalisable* matrices in the matricial formulation of LRS—see Section 2.

In this paper, we focus on two key computational problems for simple LRS over the integers (or equivalently, for our purposes, the rationals):

– The ***Positivity Problem***: given an LRS $\mathbf{u}$, are all terms of $\mathbf{u}$ positive?
– The ***(effective) Ultimate Positivity Problem***: given an LRS $\mathbf{u}$, are all but finitely many terms of $\mathbf{u}$ positive?[2] *Effectiveness* requires in addition that, when $\mathbf{u}$ is ultimately positive, a threshold beyond which all terms are positive be explicitly produced.

As detailed in [28], these problems (and assorted variants) have applications in a wide array of scientific areas, including theoretical biology, economics, software verification, probabilistic model checking, quantum computing, discrete linear dynamical systems, combinatorics, formal languages, statistical physics, generating functions, etc.

Both Positivity and effective Ultimate Positivity bear an important relationship to the well-known *Skolem Problem*: does a given LRS have a zero? The decidability of the Skolem Problem is generally considered to have been open since the 1930s (notwithstanding the fact that algorithmic decision issues had not at the time acquired the importance that they have today—see [18] for a discussion on this subject; see also [36] and [21], in which this state of affairs—the enduring openness of decidability for the Skolem Problem—is described as "faintly outrageous" by Tao and a "mathematical embarrassment" by Lipton). A breakthrough occurred in the mid-1980s, when Mignotte *et al.* [25] and Vereshchagin [39] independently showed decidability for real algebraic LRS of order 4 or less. These

[2] Note that both problems come in two natural flavours, according to whether strict or non-strict positivity is required. This paper focusses on the non-strict version, but alternatives and extensions (including strictness) are discussed in Section 5.

deep results make essential use of Baker's theorem on linear forms in logarithms (which earned Baker the Fields medal in 1970), as well as a p-adic analogue of Baker's theorem due to van der Poorten. Unfortunately, little progress on that front has since been recorded.[3]

It is considered folklore that the decidability of either Positivity or *effective* Ultimate Positivity (for arbitrary LRS) would entail that of the Skolem Problem [28], noting however that the reduction increases the order of LRS quadratically. Nevertheless, the earliest explicit references in the literature to the Positivity and Ultimate Positivity Problems that we have found are from the 1970s (see, e.g., [34, 33, 5]). In [34], the Skolem and Positivity Problems are described as "very difficult", whereas in [32], the authors assert that the Skolem, Positivity, and Ultimate Positivity Problems are "generally conjectured [to be] decidable". Positivity and/or Ultimate Positivity are again stated as open in [17, 4, 20, 22, 37, 28], among others.

Unsurprisingly, progress on the Positivity and Ultimate Positivity Problems has been fairly slow. In the early 1980s, Burke and Webb showed that effective Ultimate Positivity is decidable for LRS of order 2 [9], and nine years later Nagasaka and Shiue [26] showed the same for LRS of order 3 that have repeated characteristic roots. Much more recently, Halava *et al.* showed that Positivity is decidable for integer LRS of order 2 [17], and three years later Laohakosol and Tangsupphathawat proved that both Positivity and effective Ultimate Positivity are decidable for integer LRS of order 3 [20]. In 2012, an article claiming to show decidability of Positivity for LRS of order 4 was published [35], with the authors noting being unable to tackle the case of order 5. Unfortunately, as pointed out in [28] and acknowledged by the authors themselves [19], that paper contains a major error. Very recently, Positivity and effective Ultimate Positivity were shown decidable for arbitrary integer LRS of order 5 or less [28], with complexity in the Counting Hierarchy for the former and in polynomial time for the latter; moreover, the same paper shows by way of hardness that the decidability of either Positivity or Ultimate Positivity for integer LRS of order 6 would entail major breakthroughs in analytic number theory (certain Diophantine approximation problems long considered to be hard would become solvable). In [29], the authors show that *non-effective* Ultimate Positivity for simple integer LRS of unrestricted order is decidable within PSPACE, and in polynomial time if the order is fixed.

[3] A proof of decidability of Skolem's Problem for LRS of order 5 was announced in [18]. However, as pointed out in [27], the proof seems to have a serious gap.

The main results of this paper are as follows:[4]

- The Positivity Problem for simple integer LRS of order 9 or less is decidable in $\mathrm{coNP}^{\mathrm{PP}^{\mathrm{PP}^{\mathrm{PP}}}}$, i.e., within the fourth level of the Counting Hierarchy.
- The effective Ultimate Positivity Problem for simple integer LRS of order 9 or less is decidable in polynomial time. When the LRS is ultimately positive, an index threshold of at most exponential magnitude can be computed in polynomial time as well, beyond which the remaining terms of the sequence are positive.

It is important to note the fundamental difference between the results presented here—and in particular the *effectiveness* component of the decidability of Ultimate Positivity for simple LRS of order 9 or less—with those of [29], in which among others Ultimate Positivity is shown to be decidable for simple LRS of all orders, but in a *non-effective* sense: no threshold whatsoever is provided. In fact, as noted in [29], merely obtaining an effective threshold for Ultimate Positivity for simple LRS of order at most 14 would immediately entail the decidability of the Skolem problem for arbitrary LRS of order 5, a longstanding and major open problem.

Comparison with Related Work. As this paper deals with linear recurrence sequences, it naturally includes and summarises a certain amount of standard and routine material on the subject. We also recall the statements of various mathematical tools needed in our development, notably Baker's theorem on linear forms in logarithms, Masser's results on multiplicative relationships among algebraic numbers, Kronecker's theorem on simultaneous Diophantine approximation, and Renegar's work on the fine-grained complexity of quantifier elimination in the first-order theory of the reals.

Our overall approach is similar to that followed in [28], attacking the problem using sophisticated tools from analytic and algebraic number theory, Diophantine geometry and approximation, and real algebraic geometry. However the present paper makes vastly greater and deeper use of real algebraic geometry, particularly in the form of Lemmas 11

[4] The complexities are given as a function of the bit length of standard representations of integer LRS of order k; for an LRS as defined by Equation (1), this representation consists of the $2k$-tuple $(a_1, \ldots, a_k, u_0, \ldots, u_{k-1})$ of integers.

Note also that the Counting-Hierarchy complexity class does not require parenthesising since $\mathrm{co}(\mathrm{NP}^{\mathrm{PP}^{\mathrm{PP}^{\mathrm{PP}}}}) = (\mathrm{coNP})^{\mathrm{PP}^{\mathrm{PP}^{\mathrm{PP}}}}$.

and 12 (which serve to establish the key fact that certain varieties are zero-dimensional, enabling our application of Baker's theorem in higher dimensions), and throughout the whole of Subsection 4.2, which handles what is by far the most difficult and complex critical case in our analysis. Both Lemmas 11 and 12, as well as the development of Subsection 4.2, are entirely new.

The present paper also markedly differs from [29]. In fact, aside from sharing standard material on LRS, the non-effective approach of [29] eschews most of the real-algebraic treatment of the present paper, as well as Baker's theorem, and is underpinned instead by non-constructive lower bounds on sums of S-units, which in turn follow from deep results in Diophantine approximation (Schlickewei's p-adic generalisation of Schmidt's Subspace theorem). It is also worth noting that, due to its intrinsically non-effective nature, it does not seem possible to use the approach of [29] to decide the *Positivity* Problem for simple LRS of any order.

Section 2 summarises standard facts on LRS, and Section 3 presents the main mathematical tools used in our development. A high-level overview of our proof strategy—split in two parts—can be found at the beginning of Section 4; the remainder of the section is then devoted to the proof of our main results. Various extensions and generalisations, along with avenues for future work, are finally discussed in Section 5.

2 Linear Recurrence Sequences

We recall some fundamental properties of (simple) linear recurrence sequences. Results are stated without proof, and we refer the reader to [13, 18] for details.

Let $\mathbf{u} = \langle u_n \rangle_{n=0}^{\infty}$ be an LRS of order k over the reals satisfying the recurrence relation $u_{n+k} = a_1 u_{n+k-1} + \ldots + a_k u_n$, where without loss of generality we may assume that $a_k \neq 0$. We denote by $||\mathbf{u}||$ the bit length of its representation as a $2k$-tuple of integers, as discussed in the previous section. The ***characteristic roots*** of $\mathbf{u}$ are the roots of its characteristic polynomial (as defined earlier), and the ***dominant roots*** are the roots of maximum modulus.

The characteristic roots divide naturally into ones that are real and ones that are not. As we exclusively deal with LRS over the reals, the characteristic polynomial has real coefficients and non-real roots therefore always come in conjugate pairs. Thus we may write $\{\rho_1, \ldots, \rho_\ell, \gamma_1, \overline{\gamma_1}, \ldots, \gamma_m, \overline{\gamma_m}\}$ to represent the set of characteristic roots of $\mathbf{u}$, where each $\rho_i \in \mathbb{R}$ and each $\gamma_j \in \mathbb{C} \setminus \mathbb{R}$. If $\mathbf{u}$ is a simple LRS, there are algebraic constants $a_1, \ldots, a_\ell \in \mathbb{R}$ and $c_1, \ldots, c_m$ such that, for all $n \geq 0$,

$$u_n = \sum_{i=1}^{\ell} a_i \rho_i^n + \sum_{j=1}^{m} \left(c_j \gamma_j^n + \overline{c_j \gamma_j}^n \right) . \tag{2}$$

This expression is referred to as the ***exponential polynomial*** solution of $\mathbf{u}$. For fixed k, all constants a_i and c_j can be computed in time polynomial in $||\mathbf{u}||$, since they can be obtained by solving a system of linear equations involving the first k instances of Equation (2). See Section 3 for further details on algebraic-number manipulations.

An LRS is said to be ***non-degenerate*** if it does not have two distinct characteristic roots whose quotient is a root of unity. As pointed out in [13], the study of arbitrary LRS can effectively be reduced to that of non-degenerate LRS, by partitioning the original LRS into finitely many subsequences, each of which is non-degenerate. In general, such a reduction will require exponential time. However, when restricting ourselves to LRS of bounded order (in our case, of order at most 9), the reduction can be carried out in polynomial time. In particular, any LRS of order 9 or less can be partitioned in polynomial time into at most $3.9 \cdot 10^7$ non-degenerate LRS of the same order or less.[5] Note that if the original LRS is simple, this process will yield a collection of simple non-degenerate subsequences. In the rest of this paper, we shall therefore assume that all LRS we are given are non-degenerate.

Any LRS $\mathbf{u}$ of order k can alternately be given in matrix form, in the sense that there is a square matrix M of dimension $k \times k$, together with k-dimensional column vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, such that, for all $n \geq 0$, $u_n = \boldsymbol{v}^T M^n \boldsymbol{w}$. It suffices to take M to be the transpose of the companion matrix of the characteristic polynomial of $\mathbf{u}$, let $\boldsymbol{v}$ be the vector $(u_{k-1}, \ldots, u_0)$ of initial terms of $\mathbf{u}$ in reverse order, and take $\boldsymbol{w}$ to be the vector whose first $k-1$ entries are 0 and whose kth entry is 1. It is worth noting that the characteristic roots of $\mathbf{u}$ correspond precisely to the eigenvalues of M, and that if $\mathbf{u}$ is simple then M is diagonalisable. This translation is instrumental in Section 4 to place the Positivity Problem for simple LRS of order at most 9 within the counting hierarchy.

Conversely, given any square matrix M of dimension $k \times k$, and any k-dimensional vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, let $u_n = \boldsymbol{v}^T M^n \boldsymbol{w}$. Then $\langle \boldsymbol{v}^T M^n \boldsymbol{w} \rangle_{n=k}^{\infty}$ is an LRS of order at most k whose characteristic polynomial divides that

[5] We obtained this value using a bespoke enumeration procedure for order 9. A bound of $e^{2\sqrt{6 \cdot 9 \log 9}} \leq 2.9 \cdot 10^9$ can be obtained from Corollary 3.3 of [40].

of M, as can be seen by applying the Cayley-Hamilton Theorem.[6] When M is diagonalisable, the resulting LRS is simple.

3 Mathematical Tools

In this section we introduce the key technical tools used in this paper.

For $p \in \mathbb{Z}[x_1, \ldots, x_m]$ a polynomial with integer coefficients, let us denote by $||p||$ the bit length of its representation as a list of coefficients encoded in binary. Note that the degree of p is at most $||p||$, and the height of p—i.e., the maximum of the absolute values of its coefficients—is at most $2^{||p||}$.

We begin by summarising some basic facts about algebraic numbers and their (efficient) manipulation. The main references include [11, 3, 31].

A complex number α is ***algebraic*** if it is a root of a single-variable polynomial with integer coefficients. The ***defining polynomial*** of α, denoted p_α, is the unique polynomial of least degree, and whose coefficients do not have common factors, which vanishes at α. The ***degree*** and ***height*** of α are respectively those of p_α.

A standard representation[7] for algebraic numbers is to encode α as a tuple comprising its defining polynomial together with rational approximations of its real and imaginary parts of sufficient precision to distinguish α from the other roots of p_α. More precisely, α can be represented by $(p_\alpha, a, b, r) \in \mathbb{Z}[x] \times \mathbb{Q}^3$ provided that α is the unique root of p_α inside the circle in $\mathbb{C}$ of radius r centred at $a + bi$. A separation bound due to Mignotte [24] asserts that for roots $\alpha \neq \beta$ of a polynomial $p \in \mathbb{Z}[x]$, we have

$$|\alpha - \beta| > \frac{\sqrt{6}}{d^{(d+1)/2} H^{d-1}}, \tag{3}$$

where d and H are respectively the degree and height of p. Thus if r is required to be less than a quarter of the root-separation bound, the representation is well-defined and allows for equality checking. Given a polynomial $p \in \mathbb{Z}[x]$, it is well-known how to compute standard representations of each of its roots in time polynomial in $||p||$ [30, 11, 3]. Thus given α an algebraic number for which we have (or wish to compute) a standard representation, we write $||\alpha||$ to denote the bit length of this representation. From now on, when referring to computations on algebraic numbers, we always implicitly refer to their standard representations.

[6] In fact, if none of the eigenvalues of M are zero, it is easy to see that the full sequence $\langle \boldsymbol{v}^T M^n \boldsymbol{w} \rangle_{n=0}^{\infty}$ is an LRS (of order at most k).

[7] Note that this representation is not unique.

Note that Equation (3) can be used more generally to separate arbitrary algebraic numbers: indeed, two algebraic numbers α and β are always roots of the polynomial $p_\alpha p_\beta$ of degree at most the sum of the degrees of α and β, and of height at most the product of the heights of α and β.

Given algebraic numbers α and β, one can compute $\alpha+\beta$, $\alpha\beta$, $1/\alpha$ (for non-zero α), $\overline{\alpha}$, and $|\alpha|$, all of which are algebraic, in time polynomial in $||\alpha|| + ||\beta||$. Likewise, it is straightforward to check whether $\alpha = \beta$. Moreover, if $\alpha \in \mathbb{R}$, deciding whether $\alpha > 0$ can be done in time polynomial in $||\alpha||$. Efficient algorithms for all these tasks can be found in [11, 3].

Remarkably, integer multiplicative relationships among a fixed number of algebraic numbers can be elicited systematically in polynomial time:

Theorem 1. *Let m be fixed, and let $\lambda_1, \ldots, \lambda_m$ be complex algebraic numbers of modulus 1. Consider the abelian group L under addition given by*

$$L = \{(v_1, \ldots, v_m) \in \mathbb{Z}^m : \lambda_1^{v_1} \ldots \lambda_m^{v_m} = 1\}\,.$$

L has a basis $\{\boldsymbol{\ell_1}, \ldots, \boldsymbol{\ell_p}\} \subseteq \mathbb{Z}^m$ (with $p \leq m$), where the entries of each of the $\boldsymbol{\ell_j}$ are all polynomially bounded in $||\lambda_1|| + \ldots + ||\lambda_m||$. Moreover, such a basis can be computed in time polynomial in $||\lambda_1|| + \ldots + ||\lambda_m||$.

Note in the above that the bound is on the *magnitude* of the entries of the $\boldsymbol{\ell_j}$ (rather than their bit length), which follows from a deep result of Masser [23]. For a proof of Theorem 1, see also [15, 10].

We now turn to the first-order theory of the reals. Let $\boldsymbol{x} = (x_1, \ldots, x_m)$ and $\boldsymbol{y} = (y_1, \ldots, y_r)$ be tuples of real-valued variables, and let $\sigma(\boldsymbol{x}, \boldsymbol{y})$ be a Boolean combination of atomic predicates of the form $g(\boldsymbol{x}, \boldsymbol{y}) \sim 0$, where each $g(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{Z}[\boldsymbol{x}, \boldsymbol{y}]$ is a polynomial with integer coefficients over these variables, and $\sim$ is either $>$ or $=$. A ***formula of the first-order theory of the reals*** is of the form

$$Q_1 \boldsymbol{x_1} \ldots Q_m \boldsymbol{x_m}\, \sigma(\boldsymbol{x}, \boldsymbol{y})\,, \tag{4}$$

where each Q_i is one of the quantifiers $\exists$ or $\forall$. Let us denote the above formula by $\tau(\boldsymbol{y})$, whose free variables are contained in $\boldsymbol{y}$. When τ has no free variables, we refer to it as a ***sentence***. Naturally, $||\tau(\boldsymbol{y})||$ denotes the bit length of the syntactic representation of the formula, and the ***degree*** and ***height*** of $\tau(\boldsymbol{y})$ refer to the maximum degree and height of the polynomials $g(\boldsymbol{x}, \boldsymbol{y})$ appearing in $\tau(\boldsymbol{y})$.

Tarski [38] famously showed that the first-order theory of the reals admits *quantifier elimination*: that is, given $\tau(\boldsymbol{y})$ as above, there is a quantifier-free formula $\chi(\boldsymbol{y})$ that is equivalent to τ: for any tuple $\hat{\boldsymbol{y}} = (\hat{y}_1, \ldots, \hat{y}_r) \in \mathbb{R}^r$ of real numbers, $\tau(\hat{\boldsymbol{y}})$ holds iff $\chi(\hat{\boldsymbol{y}})$ holds. An immediate corollary is that the first-order theory of the reals is decidable.

Tarski's procedure, however, has non-elementary complexity. Many substantial improvements followed over the years, starting with Collins's technique of cylindrical algebraic decomposition [12]. For our purposes, we require bounds not only on the computation time, but also on the degree and height of the resulting equivalent quantifier-free formula, as well as on the number of atomic predicates it comprises. Such bounds are available thanks to the work of Renegar [31]. In this paper, we focus exclusively on the situation in which the number of variables is uniformly bounded.

Theorem 2 (Renegar). *Let $M \in \mathbb{N}$ be fixed. Let $\tau(\boldsymbol{y})$ be of the form (4) above. Assume that the number of (free and bound) variables in $\tau(\boldsymbol{y})$ is bounded by M (i.e., $m + r \leq M$). Denote the degree of $\tau(\boldsymbol{y})$ by d and the number of atomic predicates in $\tau(\boldsymbol{y})$ by n.*

Then there is a procedure which computes an equivalent quantifier-free formula

$$\chi(\boldsymbol{y}) = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} h_{i,j}(\boldsymbol{y}) \sim_{i,j} 0$$

in disjunctive normal form, where each $\sim_{i,j}$ is either $>$ or $=$, with the following properties:

1. *Each of I and J_i (for $1 \leq i \leq I$) is bounded by $(nd)^{\mathcal{O}(1)}$;*
2. *The degree of $\chi(\boldsymbol{y})$ is bounded by $(nd)^{\mathcal{O}(1)}$;*
3. *The height of $\chi(\boldsymbol{y})$ is bounded by $2^{||\tau(\boldsymbol{y})||(nd)^{\mathcal{O}(1)}}$.*

Moreover, this procedure runs in time polynomial in $||\tau(\boldsymbol{y})||$.

Note in particular that, when τ is a sentence, its truth value can be determined in polynomial time.

Theorem 2 follows immediately from Theorem 1.1 (for the case in which τ is a sentence) and Theorem 1.2 of [31].

Our next result is a special case of Kronecker's famous theorem on simultaneous Diophantine approximation, a statement and proof of which can be found in [7, Chap. 7, Sec. 1.3, Prop. 7].

For $x \in \mathbb{R}$, write $[x]_{2\pi}$ to denote the distance from x to the closest integer multiple of 2π: $[x]_{2\pi} = \min\{|x - 2\pi j| : j \in \mathbb{Z}\}$.

Theorem 3 (Kronecker). *Let $t_1, \ldots, t_m, x_1, \ldots, x_m \in [0, 2\pi)$. The following are equivalent:*

1. *For any $\varepsilon > 0$, there exists $n \in \mathbb{Z}$ such that, for $1 \le j \le m$, we have $[nt_j - x_j]_{2\pi} \le \varepsilon$.*
2. *For every tuple $(v_1, \ldots, v_m)$ of integers such that $[v_1 t_1 + \ldots + v_m t_m]_{2\pi} = 0$, we have $[v_1 x_1 + \ldots + v_m x_m]_{2\pi} = 0$.*

We can strengthen Theorem 3 by requiring that $n \in \mathbb{N}$ in the first assertion. Indeed, suppose that in a given instance, we find that $n < 0$. A straightforward pigeonhole argument shows that there exist arbitrarily large positive integers g such that $[gt_j]_{2\pi} \le \varepsilon$ for $1 \le j \le m$. It follows that $[(g + n)t_j - x_j]_{2\pi} \le 2\varepsilon$, which establishes the claim for sufficiently large g (noting that ε is arbitrary).

Let $\lambda_1, \ldots, \lambda_m$ be complex algebraic numbers of modulus 1. For each $j \in \{1, \ldots, m\}$, write $\lambda_j = e^{i\theta_j}$ for some $\theta_j \in [0, 2\pi)$. Let

$$\begin{aligned} L &= \{(v_1, \ldots, v_m) \in \mathbb{Z}^m : \lambda_1^{v_1} \ldots \lambda_m^{v_m} = 1\} \\ &= \{(v_1, \ldots, v_m) \in \mathbb{Z}^m : [v_1\theta_1 + \ldots + v_m\theta_m]_{2\pi} = 0\}. \end{aligned}$$

Recall from Theorem 1 that L is an abelian group under addition with basis $\{\boldsymbol{\ell_1}, \ldots, \boldsymbol{\ell_p}\} \subseteq \mathbb{Z}^m$, where $p \le m$.

For each $j \in \{1, \ldots, p\}$, let $\boldsymbol{\ell_j} = (\ell_{j,1}, \ldots, \ell_{j,m})$. Write

$$R = \{\boldsymbol{x} = (x_1, \ldots, x_m) \in [0, 2\pi)^m : [\boldsymbol{\ell_j} \cdot \boldsymbol{x}]_{2\pi} = 0 \text{ for } 1 \le j \le p\}.$$

By Theorem 3, for an arbitrary tuple $(x_1, \ldots, x_m) \in [0, 2\pi)^m$, it is the case that, for all $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that, for $j \in \{1, \ldots, m\}$, $[n\theta_j - x_j]_{2\pi} \le \varepsilon$ iff $(x_1, \ldots, x_m) \in R$.

Now observe that $(x_1, \ldots, x_m) \in R$ iff $(e^{ix_1}, \ldots, e^{ix_m}) \in T$, where

$$\begin{aligned} T = \{(z_1, \ldots, z_m) \in \mathbb{C}^m : |z_1| = \ldots = |z_m| = 1 \text{ and}, \\ \text{for each } j \in \{1, \ldots, p\},\ z_1^{\ell_{j,1}} \ldots z_m^{\ell_{j,m}} = 1\}. \end{aligned}$$

Since $e^{in\theta_j} = \lambda_j^n$, we immediately have the following:

Corollary 4. *Let $\lambda_1, \ldots, \lambda_m$ and T be as above. Then $\{(\lambda_1^n, \ldots, \lambda_m^n) : n \in \mathbb{N}\}$ is a dense subset of T.*

Finally, we give a version of Baker's deep theorem on linear forms in logarithms. The particular statement we have chosen is a sharp formulation due to Baker and Wüstholz [2].

In what follows, log refers to the principal value of the complex logarithm function given by $\log z = \log |z| + i \arg z$, where $-\pi < \arg z \le \pi$.

Theorem 5 (Baker and Wüstholz). *Let $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$ be algebraic numbers different from 0 or 1, and let $b_1, \ldots, b_m \in \mathbb{Z}$ be integers. Write*

$$\Lambda = b_1 \log \alpha_1 + \ldots + b_m \log \alpha_m \,.$$

Let $A_1, \ldots, A_m, B \geq e$ be real numbers such that, for each $j \in \{1, \ldots, m\}$, A_j is an upper bound for the height of α_j, and B is an upper bound for $|b_j|$. Let d be the degree of the extension field $\mathbb{Q}(\alpha_1, \ldots, \alpha_m)$ over $\mathbb{Q}$.

If $\Lambda \neq 0$, then $\log |\Lambda| > -(16md)^{2(m+2)} \log A_1 \ldots \log A_m \log B$.

Corollary 6. *There exists $D \in \mathbb{N}$ such that, for any algebraic numbers $\lambda, \zeta \in \mathbb{C}$ of modulus 1, and for all $n \geq 2$, whenever $\lambda^n \neq \zeta$, then*

$$|\lambda^n - \zeta| > \frac{1}{n^{(||\lambda||+||\zeta||)^D}} \,.$$

Proof. We can clearly assume that $\lambda \neq 1$, otherwise the result follows immediately from Equation (3). Likewise, the case $\zeta = 1$ is easily handled along the same lines as the proof below, so we assume $\zeta \neq 1$.

Let $\theta = \arg \lambda$ and $\varphi = \arg \zeta$. Then for all $n \in \mathbb{N}$, there is $j \in \mathbb{Z}$ with $|j| \leq n$ such that

$$|\lambda^n - \zeta| > \frac{1}{2}|n\theta - \varphi - 2j\pi| = \frac{1}{2}|n \log \lambda - \log \zeta - 2j \log(-1)| \,.$$

Let $H \geq e$ be an upper bound for the heights of λ and ζ, and let d be the largest of the degrees of λ and ζ. Notice that the degree of $\mathbb{Q}(\lambda, \zeta)$ over $\mathbb{Q}$ is at most d^2. Applying Theorem 5 to the right-hand side of the above equation, we get

$$|\lambda^n - \zeta| > \frac{1}{2} \exp\left(-(48d^2)^{10} \log^2 H \log(2n+1)\right) = \frac{1}{2(2n+1)^{(\log^2 H)(48d^2)^{10}}} \,.$$

for $n \geq 1$, provided $\lambda^n \neq \zeta$.

The required result follows by noting that $\log H \leq ||\lambda|| + ||\zeta||$ and $d \leq ||\lambda|| + ||\zeta||$. □

Finally, we record the following fact, whose straightforward proof is left to the reader.

Proposition 7. *Let $a \geq 2$ and $\varepsilon \in (0, 1)$ be real numbers. Let $B \in \mathbb{Z}[x]$ have degree at most a^{D_1} and height at most $2^{a^{D_2}}$, and assume that $1/\varepsilon \leq 2^{a^{D_3}}$, for some $D_1, D_2, D_3 \in \mathbb{N}$. Then there is $D_4 \in \mathbb{N}$ depending only on D_1, D_2, D_3 such that, for all $n \geq 2^{a^{D_4}}$, $\dfrac{1}{B(n)} > (1 - \varepsilon)^n$.*

4 Decidability and Complexity

Let $\mathbf{u} = \langle u_n \rangle_{n=0}^{\infty}$ be an integer LRS of order k. As discussed in the Introduction, we assume that u is presented as a $2k$-tuple of integers $(a_1, \ldots, a_k, u_0, \ldots, u_{k-1}) \in \mathbb{Z}^{2k}$, such that for all $n \geq 0$,

$$u_{n+k} = a_1 u_{n+k-1} + \ldots + a_k u_n \,. \tag{5}$$

The ***Positivity Problem*** asks, given such an LRS $\mathbf{u}$, whether for all $n \geq 0$, it is the case that $u_n \geq 0$. When this holds, we say that $\mathbf{u}$ is ***positive***.

The ***(effective) Ultimate Positivity Problem*** asks, given such an LRS $\mathbf{u}$, whether there exists a threshold $N \geq 0$ such that, for all $n \geq N$, it is the case that $u_n \geq 0$. When this holds, we say that $\mathbf{u}$ is ***ultimately positive***. *Effectiveness* requires in addition that the threshold N be explicitly produced.

In this section, we establish the following main results:

Theorem 8. *The Positivity Problem for simple integer LRS of order* 9 *or less is decidable in* $\mathrm{coNP}^{\mathrm{PP}^{\mathrm{PP}^{\mathrm{PP}}}}$.

Theorem 9. *The effective Ultimate Positivity Problem for simple integer LRS of order* 9 *or less is decidable in polynomial time. When the LRS is ultimately positive, an index threshold of at most exponential magnitude can be computed in polynomial time as well, beyond which the remaining terms of the sequence are positive.*

Note that deciding whether the characteristic roots are simple can easily be done in polynomial time.

Observe also that the above results immediately carry over to rational LRS. To see this, consider a rational LRS $\mathbf{u}$ obeying the recurrence relation (5). Let ℓ be the least common multiple of the denominators of the rational numbers $a_1, \ldots, a_k, u_0, \ldots, u_{k-1}$, and define an integer sequence $\mathbf{v} = \langle v_n \rangle_{n=0}^{\infty}$ by setting $v_n = \ell^{n+1} u_n$ for all $n \geq 0$. It is easily seen that $\mathbf{v}$ is an integer LRS of the same order as $\mathbf{u}$, and that for all n, $v_n \geq 0$ iff $u_n \geq 0$. Moreover, $\mathbf{v}$ is simple iff $\mathbf{u}$ is simple.

Positivity—High-Level Synopsis. At a high level, the algorithm upon which Theorem 8 rests proceeds as follow. Given an LRS $\mathbf{u}$, we first decide whether or not $\mathbf{u}$ is ultimately positive by studying its exponential polynomial solution—further details on this task are provided below. We show that whenever $\mathbf{u}$ is an ultimately positive LRS of order 9 or less, there is an effective bound N of at most exponential magnitude such that all terms of $\mathbf{u}$ beyond N are positive. Next, observe that $\mathbf{u}$ cannot be positive unless it is ultimately positive. Now in order to assert that an ultimately positive LRS $\mathbf{u}$ is *not* positive, we use a *guess-and-check* procedure: find $n \leq N$ such that $u_n < 0$. By writing $u_n = \boldsymbol{v}^T M^n \boldsymbol{w}$, for some square integer matrix M and vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ (cf. Section 2), we can decide whether $u_n < 0$ in PosSLP[8] via iterative squaring, which yields an $\mathrm{NP}^{\mathrm{PosSLP}}$ procedure for non-Positivity. Thanks to the work of Allender *et al.* [1], which asserts that PosSLP $\subseteq \mathrm{P}^{\mathrm{PP}^{\mathrm{PP}^{\mathrm{PP}}}}$, we obtain the required $\mathrm{coNP}^{\mathrm{PP}^{\mathrm{PP}^{\mathrm{PP}}}}$ algorithm for deciding Positivity.

The following is an old result concerning LRS; proofs can be found in [16, Thm. 7.1.1] and [4, Thm. 2]. It also follows easily and directly from either Pringsheim's theorem or from [8, Lem. 4]. It plays an important role in our approach by enabling us to significantly cut down on the number of subcases that must be considered, avoiding the sort of quagmire alluded to in [26].

Proposition 10. *Let $\langle u_n \rangle_{n=0}^{\infty}$ be an LRS with no real positive dominant characteristic root. Then there are infinitely many n such that $u_n < 0$ and infinitely many n such that $u_n > 0$.*

By Proposition 10, it suffices to restrict our attention to LRS whose dominant characteristic roots include one real positive value. Given an integer LRS $\mathbf{u}$, note that determining whether the latter holds is easily done in time polynomial in $||\mathbf{u}||$.

Thus let $\mathbf{u}$ be a non-degenerate simple integer LRS of order $k \leq 9$ having a real positive dominant characteristic root $\rho > 0$. Note that $\mathbf{u}$ cannot have a real negative dominant characteristic root (which would be $-\rho$), since otherwise the quotient $-\rho/\rho = -1$ would be a root of unity, contradicting non-degeneracy. Let us therefore write the characteristic roots as $\{\rho, \gamma_1, \overline{\gamma_1}, \dots, \gamma_m, \overline{\gamma_m}\} \cup \{\gamma_{m+1}, \gamma_{m+2}, \dots, \gamma_\ell\}$, where we assume that the roots in the first set all have common modulus ρ, whereas the roots in the second set all have modulus strictly smaller than ρ.

Let $\lambda_i = \gamma_i/\rho$ for $1 \leq i \leq \ell$. We can then write

$$\frac{u_n}{\rho^n} = a + \sum_{j=1}^{m} \left(c_j \lambda_j^n + \overline{c_j}\,\overline{\lambda_j}^n \right) + r(n)\,, \tag{6}$$

[8] Recall that PosSLP is the problem of determining whether an arithmetic circuit, with addition, multiplication, and subtraction gates, evaluates to a positive integer.

for some real algebraic constant a and complex algebraic constants $c_1, \ldots, c_m$, where $r(n)$ is a term tending to zero exponentially fast.

Note that none of $\lambda_1, \ldots, \lambda_m$, all of which have modulus 1, can be a root of unity, as each λ_i is a quotient of characteristic roots and $\mathbf{u}$ is assumed to be non-degenerate. Likewise, for $i \neq j$, λ_i/λ_j and $\overline{\lambda_i}/\lambda_j$ cannot be roots of unity.

For $i \in \{1, \ldots, \ell\}$, observe also that as each λ_i is a quotient of two roots of the same polynomial of degree k, it has degree at most $k(k-1)$. In fact, it is easily seen that $||\lambda_i|| = ||\mathbf{u}||^{\mathcal{O}(1)}$, $||a|| = ||\mathbf{u}||^{\mathcal{O}(1)}$, and $||c_i|| = ||\mathbf{u}||^{\mathcal{O}(1)}$.

Finally, we place bounds on the rate of convergence of $r(n)$. We have $r(n) = c_{m+1}\lambda_{m+1}^n + \ldots + c_\ell \lambda_\ell^n$. Combining our estimates on the height and degree of each λ_i together with the root-separation bound given by Equation (3), we get $\left|\frac{1}{1-\lambda_i}\right| = 2^{||\mathbf{u}||^{\mathcal{O}(1)}}$, for $m+1 \leq i \leq \ell$. Thanks also to the bounds on the height and degree of the constants c_i, it follows that we can find $\varepsilon \in (0,1)$ and $N \in \mathbb{N}$ such that:

$$1/\varepsilon = 2^{||\mathbf{u}||^{\mathcal{O}(1)}} \tag{7}$$

$$N = 2^{||\mathbf{u}||^{\mathcal{O}(1)}} \tag{8}$$

$$\text{For all } n > N,\ |r(n)| < (1-\varepsilon)^n\,. \tag{9}$$

In addition, we can compute such ε and N in time polynomial in $||\mathbf{u}||$. Naturally, given k, we can also assume that we have calculated explicitly once and for all the constants implicit in the various instances of the $\mathcal{O}(1)$ notation.

We now seek to answer Positivity and Ultimate Positivity questions for the LRS $\mathbf{u} = \langle u_n \rangle_{n=0}^\infty$ by studying the same for $\langle u_n/\rho^n \rangle_{n=0}^\infty$.

In what follows, we assume that $\mathbf{u}$ is as given above; in particular, $\mathbf{u}$ is a non-degenerate simple integer LRS having a real positive dominant characteristic root $\rho > 0$.

Effective Ultimate Positivity—High-Level Synopsis. Before launching into technical details, let us provide a high-level overview of our proof strategy for deciding Ultimate Positivity. Let us rewrite Equation (6) as

$$\frac{u_n}{\rho^n} = a + h(\lambda_1^n, \ldots, \lambda_m^n) + r(n)\,, \tag{10}$$

where $h : \mathbb{C}^m \to \mathbb{R}$ is a continuous function. In general, there will be integer multiplicative relationships among the $\lambda_1, \ldots, \lambda_m$, for which we can compute a basis B thanks to Theorem 1. These multiplicative relationships define a torus $T \subseteq \mathbb{C}^m$ on which the joint iterates $(\lambda_1^n, \ldots, \lambda_m^n)$ are dense, as per Kronecker's theorem (in the form of Corollary 4).

Now the critical case arises when

$$a + \min h\restriction_T = 0\,,$$

where $h\restriction_T$ denotes the function h restricted to the torus T. Provided that $h\restriction_T$ achieves its minimum $-a$ at only finitely many points, we can use Baker's theorem (in the form of Corollary 6) to bound the iterates $(\lambda_1^n, \ldots, \lambda_m^n)$ away from these points by an inverse polynomial in n. By combining Renegar's results (Theorem 2) with techniques from real algebraic geometry, we then argue that $h(\lambda_1^n, \ldots, \lambda_m^n)$ is itself eventually bounded away from the minimum $-a$ by a (different) inverse polynomial in n, and since $r(n)$ decays to zero exponentially fast, we are able to conclude that u_n/ρ^n is ultimately positive, and can compute a bound N after which all terms u_n (for $n > N$) are positive.

Note in the above that a key component is the requirement that $h\restriction_T$ achieve its minimum at finitely many points. Lemmas 11 and 12 show that this is the case provided that B, the basis of the integer multiplicative relationships among the $\lambda_1, \ldots, \lambda_m$, has cardinality 0, 1, $m-1$, or m. In fact, simple counterexamples can be manufactured in the other instances, which seems to preclude the use of Baker's theorem. Since non-real characteristic roots always come in conjugate pairs, the earliest appearance of this vexing state of affairs is at order 10: one real dominant root, $m = 4$ pairs of complex dominant roots, one non-dominant root ensuring that the term $r(n)$ is not identically 0, and a basis B of cardinality 2. The difficulty encountered there is highly reminiscent of (if technically different from) that of the critical unresolved case for the Skolem Problem at order 5, as described in [27].

Let us begin by stating and proving Lemmas 11 and 12, which constitute the cornerstone of our approach involving Baker's theorem.

Lemma 11. *Let $a_1, \ldots, a_m \in \mathbb{R}$ and $\varphi_1, \ldots, \varphi_m \in \mathbb{R}$ be two collections of real numbers, with each of the a_i non-zero, and let $\ell_1, \ldots, \ell_m \in \mathbb{Z}$ be m integers. Define $f, g : \mathbb{R}^m \to \mathbb{R}$ by setting*

$$f(x_1, \ldots, x_m) = \sum_{i=1}^{m} a_i \cos(x_i + \varphi_i) \quad \text{and} \quad g(x_1, \ldots, x_m) = \sum_{i=1}^{m} \ell_i x_i\,.$$

Assume that $g(x_1, \ldots, x_m)$ is not of the form $\ell(x_i \pm x_j)$, for some non-zero $\ell \in \mathbb{Z}$ and indices $i \neq j$. Let $\psi \in \mathbb{R}$.

Then the function f, subject to the constraint $g(x_1, \ldots, x_m) = \psi$, achieves its minimum at only finitely many points over the domain $[0, 2\pi)^m$.

Proof. We will establish the slightly stronger statement that f, subject to the constraint $g = \psi$, achieves its minimum over $\mathbb{R}^m$ only finitely often modulo 2π.

Note that by performing the substitutions $x_i' = x_i + \varphi_i$ (for $1 \leq i \leq m$) and $\psi' = \psi + \sum_{i=1}^m \varphi_i$, and rephrasing the statement in terms of the primed variables and constant ψ', we see that we may assume without loss of generality that each $\varphi_i = 0$.

Observe that if each $\ell_i = 0$ (corresponding to there being no constraint), the result is immediate: f is minimised when each x_i is either an odd or even multiple of π, depending on the sign of a_i. Without loss of generality, let us therefore assume that ℓ_1 is non-zero. The case of $m = 1$ is also immediate, since the constraint then reduces the domain of the unique variable x_1 to a singleton. Let us therefore assume that $m \geq 2$.

We use the method of Lagrange multipliers. Minima of f subject to the constraint $g = \psi$ must satisfy $\nabla f = \lambda \nabla g$ for some $\lambda \in \mathbb{R}$, i.e., $-a_i \sin x_i = \lambda \ell_i$, for $1 \leq i \leq m$. Note that λ must satisfy $|\lambda| \leq \frac{|a_i|}{|\ell_i|}$ for all $1 \leq i \leq m$. Observe also that each choice of λ gives rise to only finitely many choices of $x_1, \ldots, x_m$ (modulo 2π) which satisfy these equations.

From $-a_i \sin x_i = \lambda \ell_i$, it follows that $\cos^2 x_i = 1 - \frac{\lambda^2 \ell_i^2}{a_i^2}$. Taking square roots gives us 2^m choices of signs, and for each choice let us write

$$\tilde{f}(\lambda) = \sum_{i=1}^{m} \pm a_i \sqrt{1 - \lambda^2 \frac{\ell_i^2}{a_i^2}} \,.$$

Let us denote by μ the global minimum of f subject to $g = \psi$. Suppose that there are infinitely many values of $(x_1, \ldots, x_m)$ (modulo 2π) such that $g(x_1, \ldots, x_m) = \psi$ and $f(x_1, \ldots, x_m) = \mu$. It then follows that, for some fixed choice of signs, there must be infinitely many values of λ such that $\tilde{f}(\lambda) = \mu$.

Assume without loss of generality that $\frac{|a_1|}{|\ell_1|} \leq \frac{|a_i|}{|\ell_i|}$ for $2 \leq i \leq m$. Notice that $\tilde{f}(\lambda)$ is analytic (equal to its Taylor power series) on $(\frac{-|a_1|}{|\ell_1|}, \frac{|a_1|}{|\ell_1|})$. Now if the set of λ such that $\tilde{f}(\lambda) = \mu$ has an accumulation point in $(\frac{-|a_1|}{|\ell_1|}, \frac{|a_1|}{|\ell_1|})$, then $\tilde{f}$ is identically equal to μ on $[\frac{-|a_1|}{|\ell_1|}, \frac{|a_1|}{|\ell_1|}]$. Thus in any case the set of λ such that $\tilde{f}(\lambda) = \mu$ must have an accumulation point at $\frac{|a_1|}{|\ell_1|}$.

Observe that if $\frac{|a_1|}{|\ell_1|} < \frac{|a_i|}{|\ell_i|}$ for all $2 \leq i \leq m$, then a contradiction is reached as $\tilde{f}$ cannot infinitely often take on the constant value μ as λ approaches $\frac{|a_1|}{|\ell_1|}$. To see this, examine the derivative of each term of the form $\sqrt{1-\lambda^2\frac{\ell_i^2}{a_i^2}}$: they remain bounded for $i \neq 1$, but tends to $-\infty$ for $i = 1$.

Let I be the set of indices $i \in \{1, \dots, m\}$ such that $\frac{|a_i|}{|\ell_i|} = \frac{|a_1|}{|\ell_1|}$. By the same argument as above, for the given choice of signs in $\tilde{f}$, we must have $\sum_{i \in I} \pm a_i = 0$, and therefore for all $\lambda \in [\frac{-|a_1|}{|\ell_1|}, \frac{|a_1|}{|\ell_1|}]$,

$$\tilde{f}(\lambda) = \sum_{i \notin I} \pm a_i \sqrt{1 - \lambda^2 \frac{\ell_i^2}{a_i^2}} \,. \tag{11}$$

Observe that $|I| \geq 2$. Two cases now arise, according to whether (i) $|I| \geq 3$ or (ii) $|I| = 2$. In both cases, we derive a contradiction by showing that f subject to $g = \psi$ can achieve a value strictly lower than μ.

(i) Suppose without loss of generality that $I = \{1, 2, \dots, p\}$, where $p \geq 3$, and that $|a_p| \leq |a_i|$ for $1 \leq i \leq p-1$. Pick $\hat{x}_1, \dots, \hat{x}_m \in \mathbb{R}$ such that $f(\hat{x}_1, \dots, \hat{x}_m) = \mu$ and $g(\hat{x}_1, \dots, \hat{x}_m) = \psi$. There is some value $\hat{\lambda} \in [\frac{-|a_1|}{|\ell_1|}, \frac{|a_1|}{|\ell_1|}]$ such that $-a_i \sin \hat{x}_i = \hat{\lambda} \ell_i$, for $1 \leq i \leq m$. Now for the given choice of signs in $\tilde{f}$,

$$\sum_{i=1}^{p} \pm a_i \sqrt{1 - \hat{\lambda}^2 \frac{\ell_i^2}{a_i^2}} = 0 \quad \text{and} \quad \sum_{i=p+1}^{m} \pm a_i \sqrt{1 - \hat{\lambda}^2 \frac{\ell_i^2}{a_i^2}} = \mu \,,$$

or equivalently,

$$\sum_{i=1}^{p} a_i \cos \hat{x}_i = 0 \quad \text{and} \quad \sum_{i=p+1}^{m} a_i \cos \hat{x}_i = \mu \,. \tag{12}$$

In order to make f assume a value strictly smaller than μ, pick $\check{x}_1, \dots, \check{x}_{p-1}$ to be π or 0 depending respectively on the signs of $a_1, \dots, a_{p-1}$, and pick $\check{x}_p$ so that $g(\check{x}_1, \dots, \check{x}_p, \hat{x}_{p+1}, \dots, \hat{x}_m) = \psi$ (noting that $\ell_p \neq 0$ since $p \in I$). Then

$$\sum_{i=1}^{p} a_i \cos \check{x}_i \leq -\left(\sum_{i=1}^{p-1} |a_i| \right) + |a_p| < 0 \,,$$

where the strict inequality follows from the fact that $p \geq 3$ and $|a_p| \leq |a_i|$ for $1 \leq i \leq p-1$.

It then follows by the right-hand side of (12) that

$$f(\check{x}_1, \ldots, \check{x}_p, \hat{x}_{p+1}, \ldots, \hat{x}_m) < \mu \,,$$

concluding Case (i).

(ii) Without loss of generality, let us have $I = \{1, 2\}$, so that $|a_1| = |a_2|$ and $|\ell_1| = |\ell_2|$. Note that we then cannot have $\ell_3, \ldots, \ell_m$ all zero, otherwise g would be of the form $\ell_1(x_1 \pm x_2)$, violating one of our hypotheses. It therefore also follows that $m \geq 3$.

We can thus assume without loss of generality that ℓ_3 is non-zero, and furthermore that $\frac{|a_3|}{|\ell_3|} \leq \frac{|a_i|}{|\ell_i|}$ for $4 \leq i \leq m$. From Equation (11), we see that $\tilde{f}$ can be analytically extended to the larger domain $(\frac{-|a_3|}{|\ell_3|}, \frac{|a_3|}{|\ell_3|})$, and by a similar line of reasoning as earlier, we can then conclude that there must be a non-empty set $J \subseteq \{3, \ldots, m\}$ such that, for all $j \in J$, $\frac{|a_j|}{|\ell_j|} = \frac{|a_3|}{|\ell_3|}$ and moreover $\sum_{j \in J} \pm\, a_j = 0$ for the given choice of signs in $\tilde{f}$. We can therefore write

$$\tilde{f}(\lambda) = \sum_{i \notin I \cup J} \pm a_i \sqrt{1 - \lambda^2 \frac{\ell_i^2}{a_i^2}} \,.$$

But this situation is entirely similar to Case (i) since $|I \cup J| \geq 3$, which concludes Case (ii) and the proof of Lemma 11. ☐

Lemma 12. *Let* $\mathbf{u}$ *be a non-degenerate simple LRS, with dominant characteristic roots* $\rho \in \mathbb{R}$ *and* $\gamma_1, \overline{\gamma_1}, \ldots, \gamma_m, \overline{\gamma_m} \in \mathbb{C} \setminus \mathbb{R}$. *Write* $\lambda_i = \gamma_i/\rho$ *for* $1 \leq i \leq m$, *and let* $L = \{(v_1, \ldots, v_m) \in \mathbb{Z}^m : \lambda_1^{v_1} \ldots \lambda_m^{v_m} = 1\}$. *Let* $\{\boldsymbol{\ell_1}, \ldots, \boldsymbol{\ell_{m-1}}\}$ *be a basis for* L *of cardinality* $m-1$, *and write* $\boldsymbol{\ell_j} = (\ell_{j,1}, \ldots, \ell_{j,m})$ *for* $1 \leq j \leq m-1$. *Let*

$$M = \begin{pmatrix} \ell_{1,1} & \ell_{1,2} & \ldots & \ell_{1,m-1} & \ell_{1,m} \\ \ell_{2,1} & \ell_{2,2} & \ldots & \ell_{2,m-1} & \ell_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \ell_{m-1,1} & \ell_{m-1,2} & \ldots & \ell_{m-1,m-1} & \ell_{m-1,m} \end{pmatrix} .$$

Let $a_1, \ldots, a_m \in \mathbb{R}$ *and* $\varphi_1, \ldots, \varphi_m \in \mathbb{R}$ *be two collections of* m *real numbers, with each of the* a_i *non-zero, and let* $\boldsymbol{q} = (q_1, \ldots, q_{m-1}) \in \mathbb{Z}^{m-1}$ *be a column vector of* $m-1$ *integers. Let us further write* $\boldsymbol{x} = (x_1, \ldots, x_m)$ *to denote a column vector of* m *real-valued variables.*

Then the function

$$f(x_1, \ldots, x_m) = \sum_{i=1}^{m} a_i \cos(x_i + \varphi_i) \,,$$

subject to the constraint $M\boldsymbol{x} = 2\pi\boldsymbol{q}$, achieves its minimum at only finitely many points over the domain $[0, 2\pi)^m$.

Proof. By repeatedly making use of the following row operations:

1. Swapping two rows,
2. Multiplying any row by a non-zero integer, and
3. Adding to any row any integer linear combination of any of the other rows,

we can transform the augmented matrix $(M|\boldsymbol{q})$ into the matrix

$$(N|\boldsymbol{p}) = \left(\begin{array}{cccccc|c} n_{1,1} & 0 & \dots\dots\dots\dots & & 0 & b_1 & p_1 \\ 0 & n_{2,2} & 0 & \dots\dots\dots & 0 & b_2 & p_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & n_{m-2,m-2} & 0 & b_{m-2} & p_{m-2} \\ 0 & \dots\dots & & 0 & n_{m-1,m-1} & b_{m-1} & p_{m-1} \end{array}\right).$$

In the above, all entries are integer. Without loss of generality (relabelling variables and constants if necessary), we can assume that this was achieved without the need for any row-swapping operations.

Note that the rows of N remain in L (though need no longer form a basis). Hence for each $i \in \{1, \dots, m-1\}$, the $\lambda_1, \dots, \lambda_m$ satisfy the equation $\lambda_i^{n_{i,i}} \lambda_m^{b_i} = 1$. Since N has rank $m-1$, no row can be $\mathbf{0}$. From this and the fact that the LRS $\mathbf{u}$ is non-degenerate we may conclude that no $n_{i,i}$ can be zero (otherwise λ_m would be a root of unity), and likewise no b_i can be zero (otherwise λ_i would be a root of unity). Furthermore, we can never have $n_{i,i} = -b_i$ (otherwise λ_i/λ_m would be a root of unity) nor can we have $n_{i,i} = b_i$ (otherwise $\overline{\lambda_i}/\lambda_m$ would be a root of unity). In other words, we always have $n_{i,i}^2 \neq b_i^2$. Finally, for $i \neq j$, $\frac{b_i}{n_{i,i}} \neq \frac{b_j}{n_{j,j}}$: indeed, since $\lambda_i^{n_{i,i}} \lambda_m^{b_i} = 1$, we have $\lambda_i^{n_{i,i} b_j} \lambda_m^{b_i b_j} = 1$, and likewise $\lambda_j^{n_{j,j} b_i} \lambda_m^{b_i b_j} = 1$, from which we deduce that $\lambda_i^{n_{i,i} b_j} = \lambda_j^{n_{j,j} b_i}$. But if we had $\frac{b_i}{n_{i,i}} = \frac{b_j}{n_{j,j}}$, it would follow that λ_i/λ_j is a root of unity. Similarly, by noting that $\overline{\lambda_j}^{-n_{j,j}} = \lambda_j^{n_{j,j}}$ and repeating the calculation, we deduce that $\frac{b_i}{n_{i,i}} \neq -\frac{b_j}{n_{j,j}}$ for $i \neq j$. Combining the last two disequalities, we have that $\frac{b_i^2}{n_{i,i}^2} \neq \frac{b_j^2}{n_{j,j}^2}$ for $i \neq j$.

It is clear that the constraints $M\boldsymbol{x} = 2\pi\boldsymbol{q}$ and $N\boldsymbol{x} = 2\pi\boldsymbol{p}$ are equivalent. From the latter, we may write $x_i = \frac{p_i}{n_{i,i}} - \frac{b_i}{n_{i,i}} x_m$ for $1 \leq i \leq m-1$.

For ease of notation, let us set

$$d_i = -\frac{b_i}{n_{i,i}} \text{ for } 1 \le i \le m-1, \text{ and } d_m = 1\,;$$
$$\nu_i = \frac{p_i}{n_{i,i}} + \varphi_i \text{ for } 1 \le i \le m-1, \text{ and } \nu_m = \varphi_m\,.$$

From our earlier observations, let us record that:

1. Each d_i is non-zero, and
2. For $1 \le i < j \le m$, we have $d_i^2 \ne d_j^2$.

Indeed, we have already seen that the second assertion holds when $j \le m-1$. But since $n_{i,i}^2 \ne b_i^2$, for $1 \le i \le m-1$ we have that $d_i^2 \ne 1 = d_m^2$.

Substituting into f yields

$$\tilde{f}(x_m) = \sum_{i=1}^{m} a_i \cos(d_i x_m + \nu_i)\,,$$

where $\tilde{f}$ is now unconstrained. Since any value of x_m in $[0, 2\pi)$ minimising $\tilde{f}$ yields at most one point in $[0, 2\pi)^m$ at which f is minimal, it remains to show that $\tilde{f}$ can achieve its minimum only finitely often over $[0, 2\pi)$.

Thus suppose, to the contrary, that $\tilde{f}$ achieves its global minimum at infinitely many points in $[0, 2\pi)$. These points must accumulate, and since $\tilde{f}$ is analytic over $\mathbb{R}$, $\tilde{f}$ must be identically equal to its minimum all over the reals. It follows that derivatives of all orders must vanish everywhere. Now for $j \ge 1$, the $(2j-1)$th derivative of $\tilde{f}$ is given by

$$f^{(2j-1)}(x_m) = \sum_{i=1}^{m} (-1)^j d_i^{2j-1} a_i \sin(d_i x_m + \nu_i)\,.$$

Writing

$$D = \begin{pmatrix} 1 & 1 & \cdots\cdots & 1 \\ -d_1^2 & -d_2^2 & \cdots\cdots & -d_m^2 \\ d_1^4 & d_2^4 & \cdots\cdots & d_m^4 \\ \vdots & \vdots & \vdots & \vdots \\ (-1)^{m-1} d_1^{2(m-1)} & (-1)^{m-1} d_2^{2(m-1)} & \dots & (-1)^{m-1} d_m^{2(m-1)} \end{pmatrix},$$

we therefore have that

$$\begin{pmatrix} f^{(1)}(x_m) \\ f^{(3)}(x_m) \\ \vdots \\ f^{(2m-1)}(x_m) \end{pmatrix} = D \begin{pmatrix} -d_1 a_1 \sin(d_1 x_m + \nu_1) \\ -d_2 a_2 \sin(d_2 x_m + \nu_2) \\ \vdots \\ -d_m a_m \sin(d_m x_m + \nu_m) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

must hold for all $x_m \in \mathbb{R}$.

But this is a contradiction since D is a Vandermonde matrix which is invertible (given that for $i \neq j$, we have $-d_i^2 \neq -d_j^2$) and the vector

$$\begin{pmatrix} -d_1 a_1 \sin(d_1 x_m + \nu_1) \\ -d_2 a_2 \sin(d_2 x_m + \nu_2) \\ \vdots \\ -d_m a_m \sin(d_m x_m + \nu_m) \end{pmatrix}$$

clearly cannot be identically $\mathbf{0}$. □

We now proceed with the algorithm for deciding Positivity and Ultimate Positivity. Recall that we are given a non-degenerate simple LRS $\mathbf{u}$ of order most 9, with a real positive dominant characteristic root $\rho > 0$ and complex dominant roots $\gamma_1, \overline{\gamma_1}, \ldots, \gamma_m, \overline{\gamma_m} \in \mathbb{C} \setminus \mathbb{R}$. We write $\lambda_j = \gamma_j/\rho$ for $1 \leq j \leq m$.

Since the number of dominant roots is odd and at most 9, we split our analysis into two cases, there being exactly 9 dominant roots (Subsection 4.1), and there being 7 or fewer dominant roots (Subsection 4.2). Our starting point is Equation (6).

Let $L = \{(v_1, \ldots, v_m) \in \mathbb{Z}^m : \lambda_1^{v_1} \ldots \lambda_m^{v_m} = 1\}$, and let $\{\boldsymbol{\ell_1}, \ldots, \boldsymbol{\ell_p}\}$ be a basis for L of cardinality p. Write $\boldsymbol{\ell_q} = (\ell_{q,1}, \ldots, \ell_{q,m})$ for $1 \leq q \leq p$. Recall from Theorem 1 that such a basis may be computed in polynomial time, and moreover that each $\ell_{q,j}$ may be assumed to have magnitude polynomial in $||\mathbf{u}||$.

4.1 Nine Dominant Roots.

If $\mathbf{u}$ has 9 dominant roots, then $m = 4$ and $r(n)$ is identically 0 in Equation (6).

Write

$$T = \{(z_1, \ldots, z_4) \in \mathbb{C}^4 : |z_1| = \ldots = |z_4| = 1 \text{ and, } \text{for each } q \in \{1, \ldots, p\},\ z_1^{\ell_{q,1}} \ldots z_4^{\ell_{q,4}} = 1\}\,.$$

Define $h : T \to \mathbb{R}$ by setting $h(z_1, \ldots, z_4) = \sum_{j=1}^4 (c_j z_j + \overline{c_j z_j})$, so that for all n, $u_n/\rho^n = a + h(\lambda_1^n, \ldots, \lambda_4^n)$. By Corollary 4, the set $\{(\lambda_1^n, \ldots, \lambda_4^n) : n \in \mathbb{N}\}$ is a dense subset of T. Since h is continuous, we immediately have that $\inf\{u_n/\rho^n : n \in \mathbb{N}\} = a + \min h\restriction_T$. It follows that $\mathbf{u}$ is ultimately positive iff $\mathbf{u}$ is positive iff $\min h\restriction_T \geq -a$ iff

$$\forall (z_1, z_2, z_3, z_4) \in T,\ h(z_1, z_2, z_3, z_4) \geq -a\,. \tag{13}$$

We now show how to rewrite Assertion (13) as a sentence in the first-order theory of the reals, i.e., involving only real-valued variables and first-order quantifiers, Boolean connectives, and integer constants together with the arithmetic operations of addition, subtraction, multiplication, and division.[9] The idea is to separately represent the real and imaginary parts of each complex quantity appearing in Assertion (13), and combine them using real arithmetic so as to mimic the effect of complex arithmetic operations.

To this end, we use two real variables x_j and y_j to represent each of the z_j: intuitively, $z_j = x_j + iy_j$. Since the real constant a is algebraic, there is a formula $\sigma_a(x)$ which is true over the reals precisely for $x = a$. Likewise, the real and imaginary parts $\mathrm{Re}(c_j)$ and $\mathrm{Im}(c_j)$ of the complex algebraic constants c_j are themselves real algebraic, and can be represented as formulas in the first-order theory of the reals. All such formulas can readily be shown to have size polynomial in $||u||$.

Terms of the form $z_j^{\ell_{q,j}}$ are simply expanded: for example, if $\ell_{q,j}$ is positive, then $z_j^{\ell_{q,j}} = (x_j + iy_j)^{\ell_{q,j}} = A_{q,j}(x_j) + iB_{q,j}(y_j)$, where A and B are polynomials with integer coefficients. Note that since the magnitude of $\ell_{q,j}$ is polynomial in $||\mathbf{u}||$, so are $||A||$ and $||B||$. The case in which $\ell_{q,j}$ is negative is handled similarly, with the additional use of a division operation.

Combining everything, we obtain a sentence τ of the first-order theory of the reals with division which is true iff Assertion (13) holds. τ makes use of at most 17 real variables: two for each of $z_1, \ldots, z_4$, one for a, and one for each of $\mathrm{Re}(c_1), \mathrm{Im}(c_1), \ldots, \mathrm{Re}(c_4), \mathrm{Im}(c_4)$. In removing divisions from τ, the number of variables potentially swells to 29. Finally, the size of τ is polynomial in $||\mathbf{u}||$. We can therefore invoke Theorem 2 to conclude that Assertion (13) can be decided in time polynomial in $||\mathbf{u}||$.

4.2 Seven or Fewer Dominant Roots.

We now turn to the situation in which $\mathbf{u}$ has 7 dominant roots, so that $m = 3$ in Equation (6). The cases of 1, 3, and 5 dominant roots are very similar—if slightly simpler—and are therefore omitted.

[9] In Section 3, we did not have division as an allowable operation when we introduced the first-order theory of the reals; however instances of division can always be removed in linear time at the cost of introducing a linear number of existentially quantified fresh variables.

As before, let us write

$$T = \{(z_1, z_2, z_3) \in \mathbb{C}^3 : |z_1| = |z_2| = |z_3| = 1 \text{ and,}$$
$$\text{for each } q \in \{1, \ldots, p\},\ z_1^{\ell_{q,1}} z_2^{\ell_{q,2}} z_3^{\ell_{q,3}} = 1\}\,.$$

Define $h : T \to \mathbb{R}$ by setting $h(z_1, z_2, z_3) = \sum_{j=1}^{3}(c_j z_j + \overline{c_j z_j})$, so that for all n,

$$\frac{u^n}{\rho^n} = a + h(\lambda_1^n, \lambda_2^n, \lambda_3^n) + r(n)\,. \tag{14}$$

By Corollary 4, the set $\{(\lambda_1^n, \lambda_2^n, \lambda_3^n) : n \in \mathbb{N}\}$ is a dense subset of T. Since h is continuous, we have $\inf\{h(\lambda_1^n, \lambda_2^n, \lambda_3^n) : n \in \mathbb{N}\} = \min h\restriction_T = \mu$, for some $\mu \in \mathbb{R}$.

We can represent μ via the following formula $\tau(y)$:

$$\exists(\zeta_1, \zeta_2, \zeta_3) \in T\ (h(\zeta_1, \zeta_2, \zeta_3) = y \wedge \forall(z_1, z_2, z_3) \in T,\ y \le h(z_1, z_2, z_3))\,.$$

Similarly to the translation carried out in Section 4.1, we can construct an equivalent formula $\tau'(y)$ in the first-order theory of the reals, over a bounded number of real variables, with $||\tau'(y)|| = ||\mathbf{u}||^{\mathcal{O}(1)}$. According to Theorem 2, we can then compute in polynomial time an equivalent quantifier-free formula

$$\chi(y) = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} h_{i,j}(y) \sim_{i,j} 0\,.$$

Recall that each $\sim_{i,j}$ is either $>$ or $=$. Now $\chi(y)$ must have a satisfiable disjunct, and since the satisfying assignment to y is unique (namely $y = \mu$), this disjunct must comprise at least one equality predicate. Since Theorem 2 guarantees that the degree and height of each $h_{i,j}$ are bounded by $||\mathbf{u}||^{\mathcal{O}(1)}$ and $2^{||\mathbf{u}||^{\mathcal{O}(1)}}$ respectively, we immediately conclude that μ is an algebraic number and moreover that $||\mu|| = ||\mathbf{u}||^{\mathcal{O}(1)}$.

Returning to Equation (14), we see that if $a+\mu < 0$, then $\mathbf{u}$ is neither positive nor ultimately positive, whereas if $a+\mu > 0$ then $\mathbf{u}$ is ultimately positive. In the latter case, thanks to our bounds on $||\mu||$, together with the root-separation bound given by Equation (3), we have $\frac{1}{a+\mu} = 2^{||\mathbf{u}||^{\mathcal{O}(1)}}$. The latter, together with Equations (7)–(9), implies an exponential upper bound on the index of possible violations of positivity, as required.

It remains to analyse the case in which $\mu = -a$. To this end, let $\lambda_j = e^{i\theta_j}$ for $1 \le j \le 3$. From Equation (6), we have:

$$\frac{u_n}{\rho^n} = a + \sum_{j=1}^{3} 2|c_j|\cos(n\theta_j + \varphi_j) + r(n)\,.$$

In the above, $c_j = |c_j|e^{i\varphi_j}$ for $1 \leq j \leq 3$. We make the further assumption that each c_j is non-zero; note that if this did not hold, we could simply recast our analysis in a lower dimension.

We first claim that the function h achieves its minimum μ only finitely many times over T. To establish the claim, we proceed according to the cardinality p of the basis $\{\boldsymbol{\ell_1}, \ldots, \boldsymbol{\ell_p}\}$ of L:

(i) We first consider the case in which $p = 1$, and handle the case $p = 0$ immediately afterwards. Let $\boldsymbol{\ell_1} = (\ell_{1,1}, \ell_{1,2}, \ell_{1,3}) \in \mathbb{Z}^3$ be the sole vector spanning L. For $x \in \mathbb{R}$, recall that we denote by $[x]_{2\pi}$ the distance from x to the closest integer multiple of 2π. Write

$$R = \{(x_1, x_2, x_3) \in [0, 2\pi)^3 : [\ell_{1,1}x_1 + \ell_{1,2}x_2 + \ell_{1,3}x_3]_{2\pi} = 0\} .$$

Clearly, for any $(x_1, x_2, x_3) \in [0, 2\pi)^3$, we have $(x_1, x_2, x_3) \in R$ iff $(e^{ix_1}, e^{ix_2}, e^{ix_3}) \in T$. Define $f : \mathbb{R}^3 \to \mathbb{R}$ by setting

$$f(x_1, x_2, x_3) = \sum_{j=1}^{3} 2|c_j| \cos(x_j + \varphi_j) .$$

Clearly, for all $(x_1, x_2, x_3) \in [0, 2\pi)^3$, we have $f(x_1, x_2, x_3) = h(e^{ix_1}, e^{ix_2}, e^{ix_3})$, and therefore the minima of f over R are in one-to-one correspondence with those of h over T.

Define $g : \mathbb{R}^3 \to \mathbb{R}$ by setting

$$g(x_1, x_2, x_3) = \ell_{1,1}x_1 + \ell_{1,2}x_2 + \ell_{1,3}x_3 .$$

Note that $g(x_1, x_2, x_3)$ cannot be of the form $\ell(x_i - x_j)$, for non-zero $\ell \in \mathbb{Z}$ and $i \neq j$, otherwise $\lambda_i^{\ell}\lambda_j^{-\ell} = 1$, i.e., λ_i/λ_j would be a root of unity, contradicting the non-degeneracy of $\mathbf{u}$. Likewise, g cannot be of the form $\ell(x_i + x_j)$, otherwise $\lambda_i/\overline{\lambda_j}$ would be a root of unity.

Finally, observe that for $(x_1, x_2, x_3) \in [0, 2\pi)^3$, we have $(x_1, x_2, x_3) \in R$ iff $\ell_{1,1}x_1 + \ell_{1,2}x_2 + \ell_{1,3}x_3 = 2\pi q$, for some $q \in \mathbb{Z}$ with $|q| \leq |\ell_{1,1}| + |\ell_{1,2}| + |\ell_{1,3}|$. For each of these finitely many q, we can invoke Lemma 11 with f, g, and $\psi = 2\pi q$, to conclude that f achieves its minimum μ finitely many times over R, and therefore that h achieves the same minimum finitely many times over T.

The case $p = 0$, i.e., in which there are no non-trivial integer multiplicative relationships among $\lambda_1, \lambda_2, \lambda_3$, is now a special case of the above, where we have $\ell_{1,1} = \ell_{1,2} = \ell_{1,3} = 0$.

(ii) We now turn to the case $p = 2$. We have $\boldsymbol{\ell_1} = (\ell_{1,1}, \ell_{1,2}, \ell_{1,3}) \in \mathbb{Z}^3$ and $\boldsymbol{\ell_2} = (\ell_{2,1}, \ell_{2,2}, \ell_{2,3}) \in \mathbb{Z}^3$ spanning L. Let $\boldsymbol{x}$ denote the column vector

(x_1, x_2, x_3), and write

$$R = \{(x_1, x_2, x_3) \in [0, 2\pi)^3 : [\boldsymbol{\ell_1} \cdot \boldsymbol{x}]_{2\pi} = 0 \text{ and } [\boldsymbol{\ell_2} \cdot \boldsymbol{x}]_{2\pi} = 0\}\,.$$

Define $f : \mathbb{R}^3 \to \mathbb{R}$ by setting $f(x_1, x_2, x_3) = \sum_{j=1}^{3} 2|c_j| \cos(x_j + \varphi_j)$. As before, the minima of f over R are in one-to-one correspondence with those of h over T.

For $(x_1, x_2, x_3) \in [0, 2\pi)^3$, we have $[\boldsymbol{\ell_1} \cdot \boldsymbol{x}]_{2\pi} = 0$ and $[\boldsymbol{\ell_2} \cdot \boldsymbol{x}]_{2\pi} = 0$ iff there exist $q_1, q_2 \in \mathbb{Z}$, with $|q_1| \leq |\ell_{1,1}| + |\ell_{1,2}| + |\ell_{1,3}|$ and $|q_2| \leq |\ell_{2,1}| + |\ell_{2,2}| + |\ell_{2,3}|$, such that $\boldsymbol{\ell_1} \cdot \boldsymbol{x} = 2\pi q_1$ and $\boldsymbol{\ell_2} \cdot \boldsymbol{x} = 2\pi q_2$. For each of these finitely many $\boldsymbol{q} = (q_1, q_2)$, we can invoke Lemma 12 with f, $M = \begin{pmatrix} \ell_{1,1} & \ell_{1,2} & \ell_{1,3} \\ \ell_{2,1} & \ell_{2,2} & \ell_{2,3} \end{pmatrix}$, and $\boldsymbol{q}$, to conclude that f achieves its minimum μ finitely many times over R, and therefore that h achieves the same minimum finitely many times over T.

(iii) Finally, we observe that the case $p = 3$ cannot occur: indeed, a basis for L of dimension 3 would immediately entail that every λ_j is a root of unity.

This concludes the proof of the claim that h achieves its minimum at a finite number of points $Z = \{(\zeta_1, \zeta_2, \zeta_3) \in T : h(\zeta_1, \zeta_2, \zeta_3) = \mu\}$. We concentrate on the set Z_1 of first coordinates of Z. Write

$$\begin{aligned} \tau_1(x) &= \exists z_1\, (\mathrm{Re}(z_1) = x \wedge z_1 \in Z_1) \\ \tau_2(y) &= \exists z_1\, (\mathrm{Im}(z_1) = y \wedge z_1 \in Z_1)\,. \end{aligned}$$

Similarly to our earlier constructions, $\tau_1(x)$ is equivalent to a formula $\tau_1'(x)$ in the first-order theory of the reals, over a bounded number of real variables, with $||\tau_1'(x)|| = ||\mathbf{u}||^{\mathcal{O}(1)}$. Thanks to Theorem 2, we then obtain an equivalent quantifier-free formula

$$\chi_1(x) = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} h_{i,j}(x) \sim_{i,j} 0\,.$$

Note that since there can only be finitely many $\hat{x} \in \mathbb{R}$ such that $\chi_1(\hat{x})$ holds, each disjunct of $\chi_1(x)$ must comprise at least one equality predicate, or can otherwise be entirely discarded as having no solution.

A similar exercise can be carried out with $\tau_2(y)$. The bounds on the degree and height of each $h_{i,j}$ in $\chi_1(x)$ and $\chi_2(y)$ then enable us to conclude that any $\zeta_1 = \hat{x} + i\hat{y} \in Z_1$ is algebraic, and moreover satisfies

$||\zeta_1|| = ||\mathbf{u}||^{\mathcal{O}(1)}$. In addition, bounds on I and J_i guarantee that the cardinality of Z_1 is at most polynomial in $||\mathbf{u}||$.

Since λ_1 is not a root of unity, for each $\zeta_1 \in Z_1$ there is at most one value of n such that $\lambda_1^n = \zeta_1$. Theorem 1 then entails that this value (if it exists) is at most $M = ||\mathbf{u}||^{\mathcal{O}(1)}$, which we can take to be uniform across all $\zeta_1 \in Z_1$. We can now invoke Corollary 6 to conclude that, for $n > M$, and for all $\zeta_1 \in Z_1$, we have

$$|\lambda_1^n - \zeta_1| > \frac{1}{n^{||\mathbf{u}||^D}}\,, \tag{15}$$

where $D \in \mathbb{N}$ is some absolute constant.

Let $b > 0$ be minimal such that the set

$$\{z_1 \in \mathbb{C} : |z_1| = 1 \text{ and, for all } \zeta_1 \in Z_1,\ |z_1 - \zeta_1| \geq \frac{1}{b}\}$$

is non-empty. Thanks to our bounds on the cardinality of Z_1, we can use the first-order theory of the reals, together with Theorem 2, to conclude that b is algebraic and $||b|| = ||\mathbf{u}||^{\mathcal{O}(1)}$.

Define the function $g : [b, \infty) \to \mathbb{R}$ as follows:

$$g(x) = \min\{h(z_1, z_2, z_3) - \mu : (z_1, z_2, z_3) \in T \text{ and, } \\ \text{for all } \zeta_1 \in Z_1,\ |z_1 - \zeta_1| \geq \frac{1}{x}\}\,.$$

It is clear that g is continuous and $g(x) > 0$ for all $x \in [b, \infty)$. Moreover, g can be translated in polynomial time into a function in the first-order theory of the reals over a bounded number of variables. It follows from Proposition 2.6.2 of [6] (invoked with the function $1/g$) that there is a polynomial $P \in \mathbb{Z}[x]$ such that, for all $x \in [b, \infty)$,

$$g(x) \geq \frac{1}{P(x)}\,. \tag{16}$$

Moreover, an examination of the proof of [6, Prop. 2.6.2] reveals that P is obtained through a process which hinges on quantifier elimination. By Theorem 2, we are therefore able to conclude that $||P|| = ||\mathbf{u}||^{\mathcal{O}(1)}$, a fact which relies among others on our upper bounds for $||b||$.

By Equations (7)–(9), we can find $\varepsilon \in (0, 1)$ and $N = 2^{||\mathbf{u}||^{\mathcal{O}(1)}}$ such that for all $n > N$, we have $|r(n)| < (1-\varepsilon)^n$, and moreover $1/\varepsilon = 2^{||\mathbf{u}||^{\mathcal{O}(1)}}$. Moreover, by Proposition 7, there is $N' = 2^{||\mathbf{u}||^{\mathcal{O}(1)}}$ such that

$$\frac{1}{P(n^{||\mathbf{u}||^D})} > (1 - \varepsilon)^n \tag{17}$$

for all $n \geq N'$.

Combining Equations (14)–(17), we get

$$\begin{aligned}
\frac{u^n}{\rho^n} &= a + h(\lambda_1^n, \lambda_2^n, \lambda_3^n) + r(n) \\
&\geq -\mu + h(\lambda_1^n, \lambda_2^n, \lambda_3^n) - (1-\varepsilon)^n \\
&\geq g(n^{||\mathbf{u}||^D}) - (1-\varepsilon)^n \\
&\geq \frac{1}{P(n^{||\mathbf{u}||^D})} - (1-\varepsilon)^n \\
&\geq 0\,,
\end{aligned}$$

provided $n > \max\{M, N, N'\}$, which establishes ultimate positivity of $\mathbf{u}$ and provides an exponential upper bound on the index of possible violations of positivity, as required.

This completes the proof of Theorems 8 and 9.

5 Extensions and Future Work

Several of the results presented in this paper have natural extensions or generalisations, some of which we briefly mention here.

Define an LRS $\mathbf{u} = \langle u_n \rangle_{n=0}^{\infty}$ to be ***strictly positive*** (respectively ***ultimately strictly positive***) if $u_n > 0$ for all n (respectively for all sufficiently large n). An examination of our proofs readily shows that all our effective decidability and complexity results carry over without difficulty to the analogous strict formulation for simple integer LRS of order 8 or less.[10] This can also be seen by observing that for $\mathbf{u}$ a given simple integer LRS of order k, the sequence $\langle u_n - 1 \rangle_{n=0}^{\infty}$ is a simple LRS of order at most $k+1$.

All our decidability results also carry over to LRS over real algebraic numbers, as can readily be seen by examining the relevant proofs. Our complexity upper bounds, however, are more delicate, and it is an open question whether they continue to hold in the algebraic setting.

It seems likely that the techniques developed in this paper could be usefully deployed to tackle other natural problems for simple LRS or sequences of powers of diagonalisable matrices, which in turn may find applications in some of the areas mentioned in the Introduction, such as the analysis of termination of linear programs or the behaviour of discrete

[10] The decidability of Strict Positivity and *effective* Ultimate Strict Positivity for LRS of order 9 appear to be difficult problems and are presently open.

linear dynamical systems. More ambitiously, one could seek to explore computational problems for *parametric* simple LRS, where the aim is to characterise ranges for the parameters guaranteeing certain behaviours, etc.

References

[1] E. Allender, P. Bürgisser, J. Kjeldgaard-Pedersen, and P. B. Miltersen. On the complexity of numerical analysis. *SIAM J. Comput.*, 38(5), 2009.
[2] A. Baker and G. Wüstholz. Logarithmic forms and group varieties. *Jour. Reine Angew. Math.*, 442, 1993.
[3] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry.* Springer, 2nd edition, 2006.
[4] J. P. Bell and S. Gerhold. On the positivity set of a linear recurrence. *Israel Jour. Math.*, 57, 2007.
[5] J. Berstel and M. Mignotte. Deux propriétés décidables des suites récurrentes linéaires. *Bull. Soc. Math. France*, 104, 1976.
[6] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry.* Springer, 1998.
[7] N. Bourbaki. *Elements of Mathematics: General Topology (Part 2).* Addison-Wesley, 1966.
[8] M. Braverman. Termination of integer linear programs. In *Proc. Intern. Conf. on Computer Aided Verification (CAV)*, volume 4144 of *LNCS*. Springer, 2006.
[9] J. R. Burke and W. A. Webb. Asymptotic behavior of linear recurrences. *Fib. Quart.*, 19(4), 1981.
[10] J.-Y. Cai, R. J. Lipton, and Y. Zalcstein. The complexity of the A B C problem. *SIAM J. Comput.*, 29(6), 2000.
[11] H. Cohen. *A Course in Computational Algebraic Number Theory.* Springer-Verlag, 1993.
[12] G. E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. In *Proc. 2nd GI Conf. Automata Theory and Formal Languages.* Springer-Verlag, 1975.
[13] G. Everest, A. van der Poorten, I. Shparlinski, and T. Ward. *Recurrence Sequences.* American Mathematical Society, 2003.
[14] J.-H. Evertse, H. P. Schlickewei, and W. M. Schmidt. Linear equations in variables which lie in a multiplicative group. *Ann. Math.*, 155(3), 2002.
[15] G. Ge. *Algorithms Related to Multiplicative Representations of Algebraic Numbers.* PhD thesis, U.C. Berkeley, 1993.
[16] I. Gyori and G. Ladas. *Oscillation Theory of Delay Differential Equations.* Oxford Mathematical Monographs. Oxford University Press, 1991.
[17] V. Halava, T. Harju, and M. Hirvensalo. Positivity of second order linear recurrent sequences. *Discrete Appl. Math.*, 154(3), 2006.
[18] V. Halava, T. Harju, M. Hirvensalo, and J. Karhumäki. Skolem's problem — on the border between decidability and undecidability. Technical Report 683, Turku Centre for Computer Science, 2005.
[19] V. Laohakosol. Personal communication, July 2013.
[20] V. Laohakosol and P. Tangsupphathawat. Positivity of third order linear recurrence sequences. *Discrete Appl. Math.*, 157(15), 2009.
[21] R. J. Lipton. Mathematical embarrassments. *Blog entry*, December 2009. `http://rjlipton.wordpress.com/2009/12/26/mathematical-embarrassments/`.

[22] L. L. Liu. Positivity of three-term recurrence sequences. *Electr. J. Comb.*, 17(1), 2010.

[23] D. W. Masser. Linear relations on algebraic groups. In *New Advances in Transcendence Theory.* Cambridge University Press, 1988.

[24] M. Mignotte. Some useful bounds. In *Computer Algebra*, 1982.

[25] M. Mignotte, T. N. Shorey, and R. Tijdeman. The distance between terms of an algebraic recurrence sequence. *Journal für die reine und angewandte Mathematik*, 349, 1984.

[26] K. Nagasaka and J.-S. Shiue. Asymptotic positiveness of linear recurrence sequences. *Fib. Quart.*, 28(4), 1990.

[27] J. Ouaknine and J. Worrell. Decision problems for linear recurrence sequences. In *Proc. Intern. Workshop on Reachability Problems (RP)*, volume 7550 of *LNCS.* Springer, 2012.

[28] J. Ouaknine and J. Worrell. Positivity problems for low-order linear recurrence sequences. *CoRR*, abs/1307.2779, 2013.

[29] J. Ouaknine and J. Worrell. Ultimate Positivity is decidable for simple linear recurrence sequences. In preparation, 2013.

[30] V. Pan. Optimal and nearly optimal algorithms for approximating polynomial zeros. *Computers & Mathematics with Applications*, 31(12), 1996.

[31] J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. Part I: Introduction. Preliminaries. The geometry of semi-algebraic sets. The decision problem for the existential theory of the reals. *J. Symb. Comp.*, 1992.

[32] G. Rozenberg and A. Salomaa. *Cornerstones of Undecidability.* Prentice Hall, 1994.

[33] A. Salomaa. Growth functions of Lindenmayer systems: Some new approaches. In A. Lindenmayer and G. Rozenberg, editors, *Automata, Languages, Development.* North-Holland, 1976.

[34] M. Soittola. On D0L synthesis problem. In A. Lindenmayer and G. Rozenberg, editors, *Automata, Languages, Development.* North-Holland, 1976.

[35] P. Tangsupphathawat, N. Punnim, and V. Laohakosol. The positivity problem for fourth order linear recurrence sequences is decidable. *Colloq. Math.*, 128(1), 2012.

[36] T. Tao. Open question: effective Skolem-Mahler-Lech theorem. *Blog entry*, May 2007. `http://terrytao.wordpress.com/2007/05/25/open-question-effective-skolem-mahler-lech-theorem/`.

[37] S. P. Tarasov and M. N. Vyalyi. Orbits of linear maps and regular languages. In *Proc. Intern. Comp. Sci. Symp. in Russia (CSR)*, volume 6651 of *LNCS.* Springer, 2011.

[38] A. Tarski. *A Decision Method for Elementary Algebra and Geometry.* University of California Press, 1951.

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

[40] K. Yokoyama, Z. Li, and I. Nemes. Finding roots of unity among quotients of the roots of an integral polynomial. In *Proc. Intern. Symp. on Symb. and Algebraic Comp.*, 1995.