Vesa Halava | Tero Harju | Mika Hirvensalo | Juhani Karhumäki

Skolem's Problem – On the Border Between Decidability and Undecidability

Turku Centre *for* Computer Science

TUCS Technical Report
No 683, April 2005

Skolem's Problem – On the Border Between Decidability and Undecidability

Vesa Halava
Department of Mathematics, University of Turku
FIN-20014 Turku, Finland, and
Turku Centre for Computer Science
vehalava@utu.fi

Tero Harju
Department of Mathematics, University of Turku
FIN-20014 Turku, Finland, and
Turku Centre for Computer Science
harju@utu.fi

Mika Hirvensalo
Department of Mathematics, University of Turku
FIN-20014 Turku, Finland, and
Turku Centre for Computer Science
Supported by the Academy of Finland under grant 208797
mikhirve@cs.utu.fi

Juhani Karhumäki
Department of Mathematics, University of Turku
FIN-20014 Turku, Finland, and
Turku Centre for Computer Science
Supported by the Academy of Finland under grant 206039
karhumak@cs.utu.fi

TUCS Technical Report
No 683, April 2005

Abstract

We give a survey of Skolem's problem for linear recurrence sequences. We cover the known decidable cases for recurrence depths of at most 4, and give detailed proofs for these cases. Moreover, we shall prove that the problem is decidable for linear recurrences of depth 5.

Keywords: Skolem's Problem, Linear Recurrence, Undecidability, Matrix Problems

TUCS Laboratory
Discrete Mathematics for Information Technology

1 Introduction

1.1 Historical Background

Solutions to combinatorial problems often appear as sequences of numbers. These sequences are sometimes defined by some relations, where the nth element of the sequence can be counted by using the precedent elements of the sequence. Here we shall consider such sequences arising from linear recurrence relations. A sequence $(u_n)_{n=0}^{\infty}$ (or just (u_n) or u_n, for short) is a *linear recurrent sequence*, if it satisfies

$$u_n = a_{k-1}u_{n-1} + \ldots + a_1 u_{n-k+1} + a_0 u_{n-k} \tag{1}$$

for all $n \geq k$ with fixed $a_j \in \mathbb{Z}$ coefficients for $j = 0, \ldots, k-1$. We say that the relation (1) is a *linear recurrence relation* of *depth* (or *degree*) k. Note that we may assume that in (1), we have $a_0 \neq 0$, since $a_0 = 0$ would imply that the sequence satisfies a shorter linear recurrence.

The k first elements $u_0, u_1, \ldots, u_{k-1}$ of the linear recurrent sequence (u_n) in (1) are called the *initial conditions*. If the initial conditions are given, every element of the sequence is uniquely determined by the recurrence (1).

Example 1.1. *One of the most well known linear recurrence equations is that for the Fibonacci numbers: $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$ with the initial conditions are $F_0 = F_1 = 1$.*

In this article we are interested in the following *Skolem's problem*[1]:

> **Problem** SKOLEM: Given a linear recurrent sequence (u_i), that is, the linear recurrence relation and the initial conditions, determine whether or not there exists $i \geq 0$ such that $u_i = 0$.

We shall give a detailed survey of Skolem's problem covering the known (algorithmically) decidable cases, and, moreover, we shall prove that the problem is decidable for linear recurrences of depth 5. Also, we shall study the connections of Skolem's problem to some other problems in mathematics.

Let us use the following notation, for an integer sequence (u_n), let

$$Z(u_n) = \{i \in \mathbb{N} \mid u_i = 0\}$$

be the *set of zeroes of* (u_n). The starting point of the problem about zeros in recurrence sequences was in 1934 when Skolem [29] proved that $Z(u_n)$ is union of finitely many periodic sets and a finite set. In his paper Skolem used p-adic techniques to prove the result. In the days of Skolem's paper, algorithmic decidability issues were not yet as relevant as they are today. The result of Skolem

[1] In the literature Skolem's problem is sometimes referred to as *Pisot's problem*. The usage of the name of Norwegian mathematician Thoralf Skolem comes from the history as we shall see, but the usage of French mathematician Charles Pisot's name is not known by the authors.

was later proved also by Mahler [17] and Lech [23]. In Section 3 , we shall give the proof given by Hansel [13]. The result of Skolem is very strong, and, indeed, from Hansel's proof it follows that it is decidable whether or not a given linear recurrent sequence (u_n) has infinite $Z(u_n)$ or not. This was originally proved by Berstel and Mignotte in [5]. We give the decidability proof in Section 3. We also mention that it is decidable whether or not the set $Z(u_n)$ has a finite complement or $Z(u_n) = \mathbb{N}$, see [28].

Previously it has been shown that Skolem's problem is decidable for linear recurrent sequence of depth at most 4. The case of depth one is trivial, but for the depth two, the problem is already challenging. There is an old folklore proof for the case of depth 2, see [11]. In 1985 Vereshchagin [31] proved that Skolem's problem is decidable for sequences of depths 3 and 4. Vereshchagin's proof uses results of Baker [1] and of van der Poorten [25] for linear logarithms. Here we shall prove that Skolem's problem is decidable also when the depth of the linear recurrence relation is 5. On the way, we give new proofs for the cases where the depths are 2, 3, and 4. For these proofs, we use the same ideas as Vereshchagin: we study the *solutions* of the sequence, i.e., we study the problem when the nth element of the sequence is deduced from the sum

$$\sum_{i=1}^{r} p_i(n)\lambda_i, \tag{2}$$

where p_i's are polynomials, and λ_i's are complex numbers, called the *characteristic roots*. We prove that there is an effective upper bound M, deduced from the form (2) such that if $u_i = 0$ for some i, then there exists $j \leq M$ such that also $u_j = 0$. This upper bound is found using Baker's result [2] for logarithms. Mignotte, Shorey and Tijdeman [22] proved in 1984 that Skolem's problem is decidable for the depths 3 and 4, when the recurrence is nonsingular, i.e., in (2) λ_i/λ_j is not a root of unity for any $i \neq j$. Actually, Vereshchagin proved that any sequence can be reduced (or divided) to a finite set of nonsingular sequences, and we shall use this property in our proof also.

Algorithmic undecidability is close to Skolem's problem. As we shall see in Lemma 1.1, Skolem's problem can be equivalently stated in the following terms of matrices:

> Given a $k \times k$ integer matrix M, determine whether or not for some power $n \geq 0$, the element $(M^n)_{1k} = 0$.

The decidability results proved here yield that the matrix form is decidable for matrices with dimension $k \leq 5$. If we consider the problem where instead of one matrix we have a semigroup S generated by n integer matrices, then it is undecidable whether there exists $M \in S$ such that the right upper corner of M is zero. Indeed, this problem is undecidable for semigroups generated by seven 3×3 integer matrices. This follows by the proof of R.W. Floyd in [18], since the Post

Correspondence Problem is undecidable for instances of 7 letters, see [19] (see also [14]). On the other hand, this problem is undecidable for two matrices of dimension at least 24, see [8]. Therefore, when finding decidable cases of Skolem's problem, we are exploring the borderline between decidability and undecidability.

There are also other important undecidable problem in the theory of integer matrices, of which we like to mention a few. Let S be a finitely generated semigroups of $k \times k$ integer matrices. For example, for $k \geq 3$ it is undecidable whether or not the zero matrix is in S for $k \geq 3$ [24] (see also [12]), and whether or not the semigroup S is free [16] (see also [7]). Recently it was proved that it is undecidable for $k \geq 5$, whether or not a fixed diagonal matrix is in S [26].

1.2 Equivalent Formulations of the Problem

Recall from the introduction that Skolem's problem was stated as follows: Given a linear recurrence u_n (over $\mathbb{Z}$), does there exist an n such that $u_n = 0$?

The following lemma gives some equivalent representations for the problem.

Lemma 1.1. *For an integer sequence u_0, u_1, u_2, . . ., the following are equivalent:*

1. *Sequence u_n is a linear recurrent sequence.*

2. *For $n \geq 1$, $u_n = (M^n)_{1k}$, where $M \in \mathbb{Z}^{k\times k}$ for some k.*

3. *For $n \geq 1$, $u_n = \boldsymbol{v}M^n\boldsymbol{w}^T$, where $\boldsymbol{v}$, $\boldsymbol{w} \in \mathbb{Z}^k$ and $M \in \mathbb{Z}^{k\times k}$ for some k.*

Proof. Implication (1) $\implies$ (2): Assume that a sequence u_n is given by first fixing u_0, . . ., u_{k-1}, and for $n \geq k$ defined by recurrence

$$u_n = a_{k-1}u_{n-1} + \ldots + a_1u_{n-k+1} + a_0u_{n-k}.$$

Let

$$M_1 = \begin{pmatrix} a_{k-1} & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_2 & 0 & \ldots & 1 & 0 \\ a_1 & 0 & \ldots & 0 & 1 \\ a_0 & 0 & \ldots & 0 & 0 \end{pmatrix} \tag{3}$$

It is easy to see that for each $n \geq 0$, $u_n = \boldsymbol{v}M_1^n\boldsymbol{w}^T$, where $\boldsymbol{v} = (u_{k-1}, \ldots, u_1, u_0)$, and $\boldsymbol{w} = (0, \ldots, 0, 1)$. We denote $\mathbf{0} = (0, 0, \ldots, 0)$ and define a $(k+1) \times (k+1)$-matrix M by

$$M = \begin{pmatrix} 0 & \boldsymbol{v}M_1 \\ \mathbf{0}^T & M_1 \end{pmatrix}.$$

Inductively we see that

$$M^n = \begin{pmatrix} 0 & \boldsymbol{v}M_1^n \\ \mathbf{0}^T & M_1^n \end{pmatrix},$$

and, furthermore, that

$$(M^n)_{1,k+1} = (1\,\mathbf{0}) \begin{pmatrix} 0 & \boldsymbol{v}M_1^n \\ \mathbf{0}^T & M_1^n \end{pmatrix} \begin{pmatrix} 0 \\ \boldsymbol{w}^T \end{pmatrix} = \boldsymbol{v}M^n\boldsymbol{w}^T = u_n$$

whenever $n \geq 1$.

Implication (2) $\implies$ (3) follows directly from $(M^k)_{1k} = (1, \mathbf{0})M^n(\mathbf{0}, 1)^T$.

Implication (3) $\implies$ (1): Let $p(x) = x^k - a_{k-1}x^{k-1} - \ldots - a_1x - a_0$ be the characteristic polynomial of matrix M. According to the Cayley-Hamilton theorem [10]

$$M^k = a_{k-1}M^{k-1} + \ldots + a_1M + a_0I,$$

and consequently $M^n = a_{k-1}M^{n-1} + \ldots + a_1M^{n-k+1} + a_0M^{n-k}$ for any $n \geq k$. It follows that

$$\boldsymbol{v}M^n\boldsymbol{w}^T = a_{k-1}\boldsymbol{v}M^{n-1}\boldsymbol{w}^T + \ldots + a_1\boldsymbol{v}M^{n-k+1}\boldsymbol{w}^T + a_0\boldsymbol{v}M^{n-k}\boldsymbol{w}^T,$$

which is to say that

$$u_n = a_{k-1}u_{n-1} + \ldots + a_1u_{n-k+1} + a_0u_{n-k}, \tag{4}$$

so (4) is the desired recurrence. □

2 Mathematical Tools

In this section we give the mathematical tools needed for the construction of the algorithms presented in the later sections. We shall give an introduction to the theory of algebraic numbers, but first we would like to motivate the reader.

Let us consider a linear recurrent sequence u_n satisfying the relation,

$$u_n = a_{k-1}u_{n-1} + \cdots + a_0u_{n-k}.$$

The *characteristic polynomial* is the polynomial

$$p(x) = x^k - a_{k-1}x^{k-1} - a_{k-2}x^{k-2} - \cdots - a_1x - a_0. \tag{5}$$

The roots $\lambda_1, \lambda_2, \ldots, \lambda_k (\in \mathbb{C})$ of the characteristic polynomial are called the *characteristic roots*. Note that, since $a_0 \neq 0$, $\lambda_i \neq 0$ for all i.

We prove in Proposition 2.11, that the sequence u_n can be *solved* using the characteristic roots, that is,

$$u_n = p_1(n)\lambda_1^n + \ldots + p_r(n)\lambda_r^n,$$

where λ_i are distinct roots of $p(x)$ and $p_i(n)$'s are polynomials which can be effectively found. Here, the characteristic roots are algebraic numbers. Our algorithm uses this solved form of u_n, and the properties of the characteristic roots.

2.1 Algebraic Numbers

We now present the notions on algebraic numbers necessary in the continuation. For a more comprehensive representation on algebraic number theory, we refer to [10]. By an *algebraic number* we mean a (complex) number which is algebraic over the field of rational numbers. It is a well-known fact that all algebraic numbers form a field. The *minimal polynomial* of an algebraic number α is a monic polynomial $p(x) \in \mathbb{Q}[x]$ of the least possible degree having α as a zero. If c is the least common multiple of the nominators of the coefficients of $p(x)$, we call polynomial $cp(x) \in \mathbb{Z}[x]$ the *defining polynomial* of α. The defining polynomial is clearly unique. If the defining polynomial of α is monic, or equivalently, if the minimal polynomial of α is in $\mathbb{Z}[x]$, then α is an *algebraic integer*. Algebraic integers form an integral domain. The definition implies directly that for each algebraic number α there is a representation $\alpha = \frac{\beta}{m}$, where β is an algebraic integer and $m \in \mathbb{Z}$. Moreover, the minimal polynomial of β and m can be found algorithmically when the minimal polynomial of α is given. The degree of the minimal polynomial is called the *degree* of α and denoted by $\deg(\alpha)$. If $p(x) = p_d x^d + \ldots + p_1 x + p_0$ is the defining polynomial of α, we define the *height* of α as

$$H(\alpha) = \max\{|p_d|, \ldots, |p_1|, |p_0|\}.$$

Trivially any *root* (i.e, a solution of equation $x^n = \alpha$) of an algebraic number is again algebraic, and if $\beta^n = \alpha$, then $\deg(\beta) \le n \deg(\alpha)$. Thus, taking the root can increase the degree, but all the powers of α belong to the field generated by α, which implies that $\deg(\alpha^n) \le \deg(\alpha)$ for each $n \in \mathbb{N}$.

When introducing algebraic numbers we always assume that they are embedded in $\mathbb{C}$, and we fix an embedding in the following way: we say that a *description* of an algebraic number α is a quadruple $(p(x), \xi, \eta, \rho)$, where $p(x)$ is the defining polynomial of α, and ξ, η, and ρ are rational numbers which satisfy the following: the circle of radius ρ centered at $\xi + i\eta$ contains α but no other zeros of $p(x)$.

It may be useful to refer to the result in [21] to notice that if $p(x)$ is the defining polynomial of an algebraic number α with $d = \deg(\alpha)$ and $H = H(\alpha)$, then for any roots $\alpha_i \neq \alpha_j$ of $p(x)$, we have

$$|\alpha_i - \alpha_j| \ge \frac{\sqrt{6}}{d^{\frac{d+1}{2}} H^{d-1}}. \tag{6}$$

Therefore we will additionally require that in the description of an algebraic number, ρ should always be chosen smaller than a quarter of the above quantity (6). It is then possible to distinguish between all the roots of the minimal polynomial. It is clear that given a description, arbitrarily precise approximations of α with rational real and imaginary parts can be effectively computed by using for example Newton's method [9, p. 145]. It is also evident that, given the descriptions of α and β, the descriptions of $\alpha \pm \beta$, $\alpha\beta^{\pm 1}$, and nth roots ($n \in \{2, 3, \ldots\}$) of α can be found algorithmically. In fact, it is plain how to find good approximations of $a \pm \beta$, $\alpha\beta^{\pm 1}$, and the roots, and for the minimal polynomials of sums and products, one can use the resultant method, see [9, p. 157]. Moreover, it is easy to see that $\deg(\alpha+\beta)$ and $\deg(\alpha\beta)$ both are at most $\deg(\alpha)\deg(\beta)$.

If α is an algebraic number and $p(x)$ its minimal polynomial, then the roots $\alpha_1 = \alpha$, ..., α_d of $p(x)$ are called the *(Galois) conjugates* of α. Notice that since $p(x)$ is a minimal polynomial (and hence irreducible), numbers α_1, ..., α_d are necessarily distinct. If α is not a real number, then the *complex conjugate* $\overline{\alpha}$ of α is always among the (Galois) conjugates. It follows that $\alpha\overline{\alpha}$ is an algebraic number, and moreover that the *absolute value* (also called the *modulus*) $|\alpha|$ of α is an algebraic number, since $|\alpha|^2 = \alpha\overline{\alpha}$.

If α is an algebraic number of degree d and $\alpha_1 = \alpha$, ..., α_d the conjugates of α, then there are exactly d embeddings σ_1, ..., σ_d from field $\mathbb{Q}(\alpha)$ into $\mathbb{C}$, each defined as $\sigma_i(\alpha_1) = \alpha_i$ [10].

Remark 2.1. *It is also quite easy to see that each σ_i is a continuous mapping: $\sigma_i(\beta_1)$ and $\sigma_i(\beta_2)$ become arbitrarily close to each other, if $|\beta_1 - \beta_2|$ is chosen small enough. Notice also that if α is not a real number, then field $\mathbb{Q}(\alpha)$ is dense in $\mathbb{C}$, since 1 and α are then linearly independent over $\mathbb{R}$, and thus it is possible to have arbitrarily precise approximations z_n of any $z \in \mathbb{C}$ of form $z_n = a_n + b_n\alpha$, where a_n, $b_n \in \mathbb{Q}$. It follows also that we can extend the mapping σ_i to whole $\mathbb{C}$ uniquely. Moreover, if β is an algebraic number, it is possible to compute arbitrarily precise approximations of $\sigma_i(\beta)$ even though $\beta \notin \mathbb{Q}(\alpha)$.*

For each complex number α there is a representation $\alpha = |\alpha|\, e^{i\theta}$, where $\theta \in [0, 2\pi)$ is called the *phase* of α. If α is algebraic, then so is $|\alpha|$, and consequently $e^{i\theta}$ is an algebraic number as well.

Proposition 2.2. *Given descriptions $(p_1(x), \xi_1, \eta_1, \rho_1)$ and $(p_2(x), \xi_2, \eta_2, \rho_2)$ of algebraic numbers α and β respectively, the following questions are decidable:*

1. *$\alpha = \beta$?*

2. *$|\alpha| > |\beta|$?*

Proof. For deciding equality, if $p_1(x) \neq p_2(x)$, then certainly $\alpha \neq \beta$. If $p_1(x) = p_2(x)$ but $|\xi_1 + i\eta_1 - (\xi_2 + i\eta_2)| \geq \min\{\rho_1, \rho_2\}$, then $\alpha \neq \beta$, otherwise $\alpha = \beta$.

For the second question, it is possible to find descriptions for both $|\alpha|$ and $|\beta|$, and then to first decide whether $|\alpha| = |\beta|$. If $|\alpha| \neq |\beta|$, we can find arbitrarily precise approximations for both $|\alpha|$ and $|\beta|$ to decide whether $|\alpha| > |\beta|$. □

Let $r \in \mathbb{N}$. An rth *root of unity* is a (complex) number ζ satisfying $\zeta^r = 1$. A root of unity is clearly an algebraic number. The smallest positive r such that $\zeta^r = 1$ is called the *order* of ζ and denoted as $r = \operatorname{ord}(\zeta)$. Evidently $|\zeta| = 1$ for the roots of unity, and the number of rth roots of unity is exactly r. An rth root of unity is called *primitive* if $\zeta^k \neq 1$ for each $1 \leq k < r$, i.e., if its order is r. All rth roots of unity are obtained as powers of a primitive root of unity. It is possible always to choose $\zeta_r = e^{\frac{2\pi i}{r}}$ as a primitive rth root of unity. For any primitive rth root of unity ζ_r it is easy to see that ζ_r^k is primitive if and only if $\gcd(r, k) = 1$. The rth *cyclotomic polynomial* is defined as

$$\phi_r(x) = \prod_{\substack{k=0 \\ \gcd(k,r)=1}}^{r-1} (x - \zeta_r^k). \tag{7}$$

Clearly ϕ_r has all the primitive rth roots of unities as zeros. It can be shown also that $\phi_r(x) \in \mathbb{Z}[x]$, and that $\phi_r(x)$ is an irreducible polynomial. It follows that $\deg(\phi_r) = \varphi(r)$, where $\varphi(r)$ is the *Euler's function*. The cyclotomic polynomials $\phi_r(x)$ can be constructed algorithmically.

Proposition 2.3. *Given a description of an algebraic number α, it is decidable whether α is a root of unity. If α is a root of unity, then* $\mathrm{ord}(\alpha)$ *can be algorithmically found.*

Proof. We show that if α is a root of unity, then there is a computable number M such that $\mathrm{ord}(\alpha) \le M$. The claim follows, since one can construct representations for all numbers α, α^2, …, α^M and check if one of those numbers equals to 1.

Let $d = \deg(\alpha)$. If α is a primitive rth root of unity, then α is a zero of $\phi_r(x)$, which has degree $d = \varphi(r)$. On the other hand, it is known [27] that for $r \ge 3$,

$$\frac{r}{\varphi(r)} < e^{\gamma} \log\log r + \frac{2.50637}{\log\log r}, \tag{8}$$

where $\gamma = 0.5772156649\ldots$ is the *Euler's constant*. For $r \ge 3$, the right hand side of (8) can be estimated to get

$$\frac{r}{\varphi(r)} < 283 \log\log r, \tag{9}$$

which implies that $d = \varphi(r) > \frac{r}{283 \log\log r}$. It therefore suffices to take $M = \max(\{r \mid \frac{r}{283 \log\log r} < d\} \cup \{3\})$. □

As an *algebraic number field* we understand a finite extension of $\mathbb{Q}$. It is a well-known fact that each number field F is a *simple extension* of $\mathbb{Q}$, which means that F can be represented as $F = \mathbb{Q}(\alpha)$. Moreover, if representations of algebraic numbers α_1, …, α_r are given, it is possible to find algorithmically a representation of α such that $\mathbb{Q}(\alpha)$ contains all the numbers α_1, …, α_r [9]. If $p(x)$ is a polynomial, the *splitting field* of $p(x)$ is the smallest field where $p(x)$ can be decomposed into linear factors.

Remark 2.4. *Throughout the rest of Section 2, we assume, unless explicitly otherwise stated, that an algebraic number field F is chosen such that all the occurring algebraic numbers and their conjugates belong to F, and that the field F is* normal, *meaning that F is a splitting field of some polynomial in $\mathbb{Q}[x]$, or equivalently that every irreducible polynomial (over $\mathbb{Q}$) which has a zero in F, splits into linear factors over F. If $F_1 = \mathbb{Q}(\alpha_1)$ is a number field, then it is possible to find algorithmically an algebraic number α such that $F = \mathbb{Q}(\alpha)$ is a normal extension of $\mathbb{Q}$ containing F_1 [10].*

If α is an element of F and $\sigma_i : F \to \mathbb{C}$ are the embeddings of F into $\mathbb{C}$, we define the *norm* of α as

$$N(\alpha) = \prod \sigma_i(\alpha).$$

It is an easy consequence of the definition that $N(\alpha) \in \mathbb{Q}$, $N(\alpha\beta) = N(\alpha)N(\beta)$, and that $N(1) = 1$ always. It follows that $N : F^* \to \mathbb{Q}^*$ is a group morphism. By the definition, it is also clear that $N(\alpha)$ can be algorithmically computed for each $\alpha \in F$.

2.2 Ideals

One of the most interesting and fruitful part in the theory of algebraic numbers concerns the ideals of the ring of the algebraic integers of a given algebraic number field. We refer to [10] for the results and notions mentioned in this section.

The algebraic integers contained in an algebraic number field F form a ring, called the *ring of integers* of field F. Let us denote this ring by $\mathcal{O}$. It is known that all the ideals of $\mathcal{O}$ are finitely generated. If A and B are ideals of $\mathcal{O}$, then new ideals AB and $A + B$ can be defined as

$$AB = \{ab \mid a \in A, b \in B\}, \quad A + B = \{a + b \mid a \in A, b \in B\}.$$

It is rather straightforward to show that the ideals form a commutative ring, where $\mathcal{O}$ serves as a unit element and $\{0\}$ as a zero element. If we denote the principal ideal generated by $\alpha \in \mathcal{O}$ by $[\alpha]$, we can also write $\mathcal{O} = [1]$, $\{0\} = [0]$. It is easy to see that mapping $\alpha \to [\alpha]$ is a morphism from the multiplicative group of $\mathcal{O}$ into the multiplicative group of the ideals. The notion of divisibility among the ideals is easy to formulate: $B \mid A$ if and only if $A = BC$ for some ideal C. We say that $P \neq [0], [1]$ is a *prime* ideal if $A \mid P$ implies either that $A = [1]$ or $A = P$.

The ideal theory has been developed mainly because the factorization in ring $\mathcal{O}$ is not necessarily unique, whereas the *fundamental theorem of ideal theory* states the following:

Theorem 2.5. *Each ideal not equal to $\{0\}$ of ring $\mathcal{O}$ can be represented as a product of prime ideals. The representation is unique if the order of the prime ideals is ignored.*

Now we can fix a prime ideal P and define the *valuation* $v_P : \mathcal{O} \setminus \{0\} \to \mathbb{N} \cup \{0\}$ as follows: If $\alpha \neq 0$ and

$$[\alpha] = P_1^{k_1} \cdot \ldots \cdot P_r^{k_r} \tag{10}$$

is the representation of ideal $[\alpha]$ as a product of prime ideals, we define

$$v_P(\alpha) = \begin{cases} k_i, & \text{if } P = P_i \text{ in (10)} \\ 0, & \text{if } P \notin \{P_1, \ldots, P_r\}. \end{cases}$$

Moreover, we define symbolically $v_P(0) = \infty$. The definition of v_P can be extended to whole number field by noticing that always $\mathbb{Z} \subseteq \mathcal{O}$ and that if α

is not an algebraic integer, then there exists an $m \in \mathbb{Z}$ such that $\alpha = \alpha_1/m$, where α_1 is an algebraic integer. It is then possible to show that the definition $v_P(\alpha) = v_P(\alpha_1) - v_P(m)$ is independent of the choice of α_1 and m.

Unlike in $\mathbb{Z}$, the ideals of $\mathcal{O}$ need not to be principal, but anyway they are finitely generated, and the representation (10) can be found algorithmically [9, pp. 179–204]. It follows that given a number $\alpha \in F$, value $v_P(\alpha)$ can be found also algorithmically. The following important issues about v_P can be proven straightforwardly:

Proposition 2.6. *Let P be a prime ideal of the ring of integers $\mathcal{O}$ of an algebraic number field F. Then, for each $\alpha, \beta \in F$,*

1. $v_P(\alpha\beta) = v_P(\alpha) + v_P(\beta)$.
2. $v_P(\alpha + \beta) \geq \min\{v_P(\alpha), v_P(\beta)\}$.
3. *If $v_P(\alpha) < v_P(\beta)$, then $v_P(\alpha + \beta) = v_P(\alpha)$.*

A *unit of* ring $\mathcal{O}$ is an element γ that has also an inverse $\gamma^{-1} \in \mathcal{O}$. It is easy to see that $[\alpha_1] = [\alpha_2]$ if and only if $\alpha_1 = \gamma\alpha_2$ for some unit γ. It follows that if $\alpha = \frac{\alpha_1}{m}$ is not an algebraic integer (but α_1 is), the ideals $[\alpha_1]$ and $[m]$ are not equal. Therefore, the representations (10) of $[\alpha]$ and $[m]$ are not equal, and we get

Lemma 2.1. *If α is not an algebraic integer, then there is a prime ideal P such that $v_P(\alpha) \neq 0$.*

Such a prime ideal can also be found algorithmically [9].

A special case occurs if $F = \mathbb{Q}$. In ring $\mathbb{Z}$ all ideals are principal, and the prime ideals P are generated by the prime numbers. If p is a prime number and $P = [p]$ the ideal generated by it, we denote usually $v_P = v_p$. Given a prime number p, the valuation of an integer m can be then found as follows:

$$v_p(m) = \begin{cases} 0, & \text{if } p \nmid m \\ k, & \text{if } p^k \mid m \text{ but } p^{k+1} \nmid m. \end{cases}$$

For any rational number $r = \frac{m}{n}$, we have $v_p(r) = v_p(m) - v_p(n)$ according to the previous definition. Valuation v_p is called the *p-adic valuation*.

Finally we represent the notion of *norm* of an ideal. Let F be the number field in question and $\mathcal{O}$ its ring of integers. For any ideal $A \neq [0]$, we define its *norm* as the cardinality of the residue class ring: $N(A) = |\mathcal{O}/A|$. It can be shown that for each $A \neq [0]$, $N(A)$ is finite, and that $N(AB) = N(A)N(B)$ for all ideals A and B. Furthermore, $N([\alpha]) = |N(\alpha)|$ for each principal ideal $[\alpha]$, and for each prime ideal P there exist a unique prime number $p \in P$, and that $N(P) = p^f$ for some natural number f. By the definition, $N(P) \geq 2$ for any prime ideal, and as a consequence we see that

$$N(A) = N(P_1^{k_1} \dots P_k^{k_r}) \geq N(P_i)^{k_i} \geq 2^{k_i}.$$

It follows that $|N(\alpha)| = N([\alpha]) \geq 2^{v_P(\alpha)}$ for any $\alpha \neq 0$ and any prime ideal P.

Proposition 2.7. *Given descriptions of algebraic numbers α, $\beta \in F$, it is decidable whether $\alpha = \beta^n$ for some $n \in \mathbb{Z}$.*

Proof. Proposition 2.2 implies that it is possible to decide whether $|\beta| = 1$. Assume first that $|\beta| \neq 1$. Then $\log|\beta| \neq 0$ and equation $\alpha = \beta^n$ implies $|\alpha| = |\beta|^n$, which, in turn gives that

$$n = \frac{\log|\alpha|}{\log|\beta|}. \tag{11}$$

It is possible to compute arbitrarily precise approximations of the right hand side of (11), which allows to restrict the number of potential exponents n to at most one. Then it remains to compute a representation β^n for this potential n to check if $\alpha = \beta^n$.

Assume then that $|\beta| = 1$. According to Proposition 2.3 it is decidable whether β is a root of unity. If β is a root of unity of order r, then it remains to check if any of numbers β^0, β^1, . . ., β^{r-1} equals to α.

Next we assume that $|\beta| = 1$, and that β is not a root of unity. It follows that $\beta \notin \mathbb{R}$.

We split here into two subcases: 1) β is an algebraic integer, and 2) β is not an algebraic integer. It should be noted that to distinguish between 1) and 2) is an easy task when the minimal (or defining) polynomial of β is available.

In the first case, we show, following Kronecker's argumentation, that there exists a conjugate of β having absolute value greater than 1.

Assume the contrary: $|\sigma_i(\beta)| \leq 1$ for each conjugate $\sigma_i(\beta)$ and consider set $B = \{\beta, \beta^2, \beta^3, \ldots\}$. The minimal polynomial $p_n(x)$ of β^n is of form

$$p_n(x) = \prod_{i=1}^{d'} (x - \beta_i^n), \tag{12}$$

where β_1^n, . . ., $\beta_{d'}^n$ are conjugates of β^n, and $d' \leq \deg(\beta)$. Since β^n is always an algebraic integer, polynomials (12) belong to $\mathbb{Z}[x]$. The aforementioned conjugates are among numbers $\sigma_i(\beta^n) = \sigma_i(\beta)^n$, so the assumption implies that $|\beta_i^n| \leq 1$ for each conjugate in (12). It follows that the coefficients of (12) are bounded, and because polynomials in (12) are in $\mathbb{Z}[x]$, there are only finitely many polynomials $p_n(x)$. Consequently, there are only finitely many elements of B. It follows that $\beta^n = \beta^m$ for some $n < m$, and therefore $\beta^{m-n} = 1$, which contradicts the assumption that β is not a root of unity.

Now we can choose a conjugate $\beta_i = \sigma_i(\beta)$ with $|\beta_i| > 1$, and apply the homomorphism σ_i (or, to be precise, its extension to $\mathbb{C}$) to equation $\alpha = \beta^n$ to get

$$\sigma_i(\alpha) = \sigma_i(\beta)^n.$$

As mentioned in Remark 2.1, it is possible to compute arbitrarily precise approximations of $\sigma_i(\alpha)$, and because $|\sigma_i(\beta)| > 1$, we have $\log|\sigma_i(\beta)| \neq 0$, and we can restrict the number of potential exponents to at least one by computing approximations of the right hand side of

$$n = \frac{\log|\sigma_i(\alpha)|}{\log|\sigma_i(\beta)|}$$

precise enough.

Finally we assume that β is not an algebraic integer. Then it is possible to find efficiently a prime ideal P of the integer ring of F such that $v_P(\beta) \neq 0$. Thus

$$v_P(\alpha) = nv_P(\beta).$$

Because $2^{v_P(\alpha)} \leq |N(\alpha)|$, we have an estimation

$$\log_2|N(\alpha)| \geq |v_P(\alpha)| = |nv_P(\beta)| \geq |n|\,.$$

Moreover, we can compute arbitrarily precise approximations of $N(\alpha)$, so an upper bound for $|n|$ can be found algorithmically. □

Remark 2.8. *Proposition 2.7 occurred also in [31] in a less detailed form, and in [15] in form, where $\alpha = q(\beta)$ with $q \in \mathbb{Q}[x]$.*

Remark 2.9. *Notice also that when constructing the algorithm of Proposition 2.7, it is not necessary to compute any valuations $v_P(\alpha)$. The valuations were only introduced to prove the method correct. However, it is possible to compute the values $v_P(\alpha)$ algorithmically, see [9].*

Similarly to Proposition 2.7 we can prove the following, a little bit more complicated claim which is needed in the sequel.

Proposition 2.10. *Let α_1, α_2, β_1, . . ., β_4, and δ be algebraic numbers belonging to a number field F, which is a normal extension of $\mathbb{Q}$. If α_1/α_2 is not a root of unity, then it is decidable whether equation*

$$\beta_1\alpha_1^n + \beta_2\alpha_1^{-n} = \beta_3\alpha_2^n + \beta_4\alpha_2^{-n} + \delta \tag{13}$$

has a solution $n \in \mathbb{N}$.

Proof. By dividing by β_1 (or by β_3) and taking a common factor, we can assume that the equation (13) is in fact of form

$$\alpha_1^n + \beta_1\alpha_1^{-n} = \gamma(\alpha_2^n + \beta_2\alpha_2^{-n}) + \delta, \tag{14}$$

where all the occurring numbers are algebraic (if $\beta_1 = \beta_3 = 0$ in (13), then the situation is even simpler).

By changing the role of α_1 and α_1^{-1} we may assume that $|\alpha_1| \geq 1$ and similarly that $|\alpha_2| \geq 1$. Also we can assume that $|\alpha_1| \geq |\alpha_2|$.

If $|\alpha_1| > |\alpha_2|$, we have

$$\begin{aligned} |\alpha_1|^n - |\beta_1| \left|\alpha_1^{-1}\right|^n &\leq \left|\alpha_1^n + \beta_1\alpha_1^{-n}\right| = \left|\gamma(\alpha_2^n + \beta_2\alpha_2^{-n}) + \delta\right| \\ &\leq |\gamma| \, |\alpha_2|^n + |\gamma\beta_2| \left|\alpha_2^{-1}\right|^n + |\delta| \,. \end{aligned} \tag{15}$$

Now that $\left|\alpha_1^{-1}\right|$, $\left|\alpha_2^{-1}\right| \leq 1$ and $|\alpha_1| > |\alpha_2|$, the above equality ceases to be valid for n large enough. It is possible to compute approximations of the absolute values above, and find algorithmically a limit M such that (15) does not hold for $n \geq M$.

Assume then that $|\alpha_1| = |\alpha_2|$. If there is an embedding $\sigma_i : F \to \mathbb{C}$ such that $|\sigma_i(\alpha_1)| \neq |\sigma_i(\alpha_2)|$, we can apply the homomorphism σ_i to both sides of equation (13) and conclude as above.

Assume then that $|\sigma_i(\alpha_1)| = |\sigma_i(\alpha_2)|$ for each embedding $\sigma_i : F \to \mathbb{C}$. Now that $|\sigma_i(\alpha_1/\alpha_2)| = 1$ for each embedding σ_i, and α_1/α_2 is not a root of unity, we conclude, as in Proposition 2.7, that α_1/α_2 is not an algebraic integer. Therefore, there must be a prime ideal P of the integer ring of F such that $v_P(\alpha_1) \neq v_P(\alpha_2)$. By the symmetry of the roles of α_1 and α_2 we can assume that $v_P(\alpha_1) < v_P(\alpha_2)$. As above, we can also assume that $v_P(\alpha_1), v_P(\alpha_2) \leq 0$ (consequently, $v_P(\alpha_1)$, $v_P(\alpha_2) \geq 0$). Analogously to (15), we have

$$\begin{aligned} & \min\{nv_P(\alpha_1), v_P(\beta_1) + nv_P(\alpha_1^{-1})\} \\ = \;& v_P(\gamma) + \min\{nv_P(\alpha_2), v_P(\beta_2) + nv_P(\alpha_2^{-1}), v_P(\delta)\}. \end{aligned} \tag{16}$$

Now that $v_P(\alpha_1) < v_P(\alpha_2) \leq 0$ and $v_P(\alpha_1^{-1})$, $v_P(\alpha_2^{-1}) \geq 0$, It is clear that there is a number M such that (16) does not hold if $n \geq M$. Moreover, number M can be computed when values $v_P(\alpha_1)$, $v_P(\alpha_2)$, $v_P(\beta_1)$, $v_P(\beta_2)$, $v_P(\gamma)$ and $v_P(\delta)$ are known. The prime ideal P and the aforementioned values can be found algorithmically [9]. □

2.3 Solution of Linear Recurrent Sequence

The following well-known result is fundamental in the study of linear recurrences.

Proposition 2.11. *Let a_0, a_1, ..., a_{k-1}, and u_0, u_1, ..., u_{k-1} be fixed integers and recurrence*

$$u_n = a_{k-1}u_{n-1} + \ldots + a_1 u_{n-k+1} + a_0 u_{n-k} \tag{17}$$

define u_n for each $n \geq k$. Let also

$$p(x) = x^k - a_{k-1}x^{k-1} - \ldots - a_1 x - a_0$$

be the characteristic polynomial of the recurrence and F the splitting field of $p(x)$. If

$$p(x) = (x - \lambda_1)^{m_1} \cdot \ldots \cdot (x - \lambda_r)^{m_r},$$

where λ_1, ..., $\lambda_r \in F$ are the distinct zeros of $p(x)$, then there exist unique polynomials $p_1(x)$, ..., $p_r(x) \in F[x]$ such that $\deg(p_i) \leq m_i - 1$ and

$$u_n = p_1(n)\lambda_1^n + \ldots + p_r(n)\lambda_r^n \tag{18}$$

for each $n \geq 0$. *Conversely, any sequence* u_n *of form (18) (for polynomial* $p(x)$*) satisfies recurrence (17).*

Here, the form (18) is usually called the *solution* of u_n. Our algorithms for the cases of Skolem's problem uses this solved form of the sequence u_n. Note carefully, that the main issue in finding a solution of a sequence are the characteristic roots, and for the recurrent sequence of depth at least 5, this is impossible in most cases. This is also the reason for the need of the theory of the algebraic numbers.

Note also that Proposition 2.11 can be proved using the theory of *formal power series*. Our proof here is elementary using only linear algebra and algebraic properties polynomials.

Note also that the converse part can be proved by using the partial solution sequences of the form $v_n = n^j \lambda_i^n$, where λ_i is a characteristic roots and $j \leq m_i - 1$, which satisfy the relation (17). The claim follows by proving that for any sequences satisfying relation (17), any linear combination of then also satisfies (17). We shall give here a new proof for this converse, which may be of some reader's interest.

Before proceeding to the proof of Proposition 2.11, we need some auxiliary results.

Lemma 2.2. *Let*

$$p(x) = x^k - \sum_{s=0}^{k-1} a_s x^s$$

be a polynomial having λ *as a zero with multiplicity* $m \geq 1$. *Define* $p_0(x) = p(x)$ *and* $p_{i+1}(x) = xp_i'(x)$ *for each* $i \in \{0, 1, 2, \ldots\}$ *(*$p_i'(x)$ *stands for the derivative of polynomial* $p_i(x)$*). Then there exist polynomials* q_i *such that*

$$p_i(x) = (x - \lambda)^{m-i} q_i(x).$$

for each $i \in \{0, 1, \ldots, m-1\}$. *Conversely, if* $\lambda \neq 0$ *and* $p_i(\lambda) = 0$ *for each* $i \in \{0, 1, \ldots, m-1\}$*, then* $p_0(x) = (x - \lambda)^m q_0(x)$ *for some polynomial* $q_0(x)$.

Proof. We begin by showing that, for each $i \geq 1$, $p_i(x)$ can be represented as

$$p_i(x) = \sum_{j=1}^{i} c_j^{(i)} x^j p^{(j)}(x), \tag{19}$$

where each $c_j^{(k)}$ is an integer, and $c_1^{(i)} = c_i^{(i)} = 1$. The case $i = 1$ is clear: $p_1(x) = xp'(x)$. Assume then that (19) holds for some value of i. Then

$$\begin{aligned} p_i'(x) &= \sum_{j=1}^{i} c_j^{(i)} (jx^{j-1} p^{(j)}(x) + x^j p^{(j+1)}(x)) \\ &= \sum_{j=1}^{i} c_j^{(i)} j x^{j-1} p^{(j)}(x) + \sum_{j=2}^{i+1} c_{j-1}^{(i)} x^{j-1} p^{(j)}(x), \end{aligned}$$

and consequently

$$p_{i+1}(x) = xp'(x) + \sum_{j=2}^{i} (jc_j^{(i)} + c_{j-1}^{(i)})x^j p^{(j)}(x) + x^{i+1}p^{(i+1)}(x),$$

as claimed.

The rest of the proof is based on the following well-known fact: λ is a zero of $p(x)$ with multiplicity m, if and only if λ is a zero of $p^{(j)}(x)$ with multiplicity $m-j$. The first claim of the lemma follows directly from this and from representation (19), because then $p^{(j)}(x) = (x-\lambda)^{m-j}r_j(x)$ for some polynomial $r_j(x)$.

The converse statement follows from representation (19) by using induction. In fact, if the assumption holds for $m = 1$, then $0 = p_0(\lambda) = p(\lambda)$ and hence $(x-\lambda) \mid p(x)$. If claim the holds for all numbers less than m, and if $p_{m-1}(\lambda) = 0$ we have

$$0 = \sum_{j=0}^{m-1} c_j^{(m-1)} \lambda^j p^{(j)}(\lambda) = \lambda^{m-1} p^{(m-1)}(\lambda),$$

which implies that $p^{(m-1)}(\lambda) = 0$. The claim follows from this. □

Lemma 2.3. *Let the notations be as in the previous lemma. Then*

$$k^i \lambda^k = \sum_{s=0}^{k-1} s^i a_s \lambda^s$$

for each $i \in \{0, 1, \ldots, m-1\}$.

Proof. By induction we see that

$$p_i(x) = k^i x^k - \sum_{s=0}^{k-1} s^i a_s x^s$$

for each $i \in \{0, 1, 2, \ldots\}$. By the previous lemma, $p_i(\lambda) = 0$ for each $i \in \{0, 1, \ldots, m-1\}$, and the claim follows immediately. □

Lemma 2.4. *Let the notations be as in the previous lemmata. Then*

$$\sum_{s=1}^{k} a_{k-s} s^i \lambda^{k-s} = 0$$

for each $i \in \{1, 2, \ldots, m-1\}$.

Proof. A direct calculation gives

$$\sum_{s=1}^{k} a_{k-s} s^i \lambda^{k-s} = \sum_{s=0}^{k-1} a_s (k-s)^i \lambda^s = \sum_{t=0}^{i} (-1)^t \binom{i}{t} k^{i-t} \sum_{s=0}^{k-1} a_s s^t \lambda^s. \tag{20}$$

By the previous lemma, expression (20) can be written as

$$\sum_{t=0}^{i}(-1)^t\binom{i}{t}k^{i-t}\cdot k^t\lambda^k = k^i\lambda^k\sum_{t=0}^{i}(-1)^t\binom{i}{t}.$$

The latest expression equals to 0 whenever $i \in \{1, 2, \ldots, m-1\}$. □

Lemma 2.5. *Let the notations be as before. Then*

$$\sum_{s=1}^{k} a_{k-s}(n-s)^j\lambda^{k-s} = n^j\lambda^k$$

if $j \in \{0, 1, \ldots, m-1\}$.

Proof. A straightforward calculation shows that

$$\sum_{s=1}^{k} a_{k-s}(n-s)^j\lambda^{k-s} = \sum_{s=1}^{k} a_{k-s}\sum_{t=0}^{j}\binom{j}{t}n^{j-t}(-s)^t\lambda^{k-s}$$

$$= \sum_{t=0}^{j}\binom{j}{t}n^{j-t}(-1)^t\sum_{s=1}^{k} a_{k-s}s^t\lambda^{k-s} = n^j\lambda^k.$$

The last equality is due to Lemmata 2.3 and 2.4. □

The proof of Proposition 2.11. We will begin with the "converse part" of the statement. Let

$$u_n = \sum_{i=1}^{r} p_i(n)\lambda_i^n,$$

where λ_1, . . ., λ_r are distinct zeros of $p(x) = x^k - a_{k-1}x^{k-1} - \ldots - a_1x - a_0$ with multiplicities m_1, . . ., m_r, and $\deg(p_i) \leq m_i - 1$. Then, for each $n \geq k$,

$$\begin{aligned} & u_n - a_{k-1}u_{n-1} - \ldots - a_0u_{n-k} \\ = \ & \sum_{i=1}^{r} p_i(n)\lambda_i^n - \sum_{s=1}^{k} a_{k-s}\sum_{i=1}^{r} p_i(n-s)\lambda_i^{n-s}. \end{aligned} \tag{21}$$

By denoting $a_k = -1$ and

$$p_i(x) = \sum_{j=0}^{m_i-1} p_{ij}x^j$$

we can write (21) as

$$\begin{aligned} & -\sum_{s=0}^{k} a_{k-s}\sum_{i=1}^{r} p_i(n-s)\lambda_i^{n-s} \\ = \ & -\sum_{i=1}^{r}\sum_{s=0}^{k}\sum_{j=0}^{m_i-1} p_{ij}(n-s)^j a_{k-s}\lambda_i^{n-k}\lambda_i^{k-s} \\ = \ & -\sum_{i=1}^{r}\lambda_i^{n-k}\sum_{j=0}^{m_i-1} p_{ij}\sum_{s=0}^{k} a_{k-s}(n-s)^j\lambda_i^{k-s}. \end{aligned}$$

According to Lemma 2.5 we have (recall that $a_k = -1$)

$$\begin{aligned} \sum_{s=0}^{k} a_{k-s}(n-s)^j \lambda_i^{k-s} = -n^j\lambda_i^k + \sum_{s=1}^{k} a_{k-s}(n-s)^j \lambda_i^{k-s} \\ = \quad -n^j\lambda_i^k + n^j\lambda_i^k = 0. \end{aligned}$$

Hence

$$u_n - a_{k-1}u_{n-1} - \ldots - a_0 u_{n-k} = 0$$

for each $n \geq k$.

For the first part of the claim, we can assume without loss of generality that $a_0 \neq 0$. It follows that all the numbers λ_i are nonzero. Denote again

$$p_i(x) = \sum_{j=0}^{m_i-1} p_{ij}x^j$$

and regard p_{ij} as unknowns which are to be determined under conditions

$$u_n = \sum_{i=1}^{r} p_i(n)\lambda_i^n \tag{22}$$

for each $n \in \{0, 1, \ldots, k-1\}$. If numbers p_{ij} can be found, then the proof is complete, since according to the "converse part", sequence u_n in (22) satisfies the recurrence equation (17).

Equations (22) can be written more explicitly in form

$$u_n = \sum_{i=1}^{r} \sum_{j=0}^{m_i-1} p_{ij}n^j\lambda_i^n, \tag{23}$$

where $n \in \{0, 1, \ldots, k-1\}$. The determinant of system (23) is of form

$$\begin{vmatrix} 1 & 0 & \ldots & 0 & \ldots & 1 & 0 & \ldots & 0 \\ \lambda_1 & \lambda_1 & \ldots & \lambda_1 & \ldots & \lambda_r & \lambda_r & \ldots & \lambda_r \\ \lambda_1^2 & 2\lambda_1^2 & \ldots & 2^{m_1-1}\lambda_1^2 & \ldots & \lambda_r^2 & 2\lambda_r^2 & \ldots & 2^{m_r-1}\lambda_r^2 \\ \lambda_1^3 & 3\lambda_1^3 & \ldots & 3^{m_1-1}\lambda_1^3 & \ldots & \lambda_r^3 & 3\lambda_r^3 & \ldots & 3^{m_r-1}\lambda_r^3 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & (k-1)\lambda_1^{k-1} & & & & \lambda_r^{k-1} & & & \end{vmatrix}. \tag{24}$$

We will demonstrate that the determinant (24) is nonzero, which implies that the unknowns p_{ij} are determined uniquely. For that purpose, we show that the rows of (24) are linearly independent. To do this, assume that there are numbers c_0, c_1, . . ., c_{k-1} such that

$$\sum_{l=0}^{k-1} c_l l^i \lambda_j^l = 0 \tag{25}$$

for each $j \in \{1, 2, \ldots, r\}$ and $i \in \{0, 1, \ldots, m_j - 1\}$. By denoting

$$C(x) = \sum_{l=0}^{k-1} c_l x^l \tag{26}$$

and defining $C_0(x) = C(x)$, $C_{i+1}(x) = xC_i'(x)$ as in Lemma 2.2, relations (25) can be rewritten as

$$C_i(\lambda_j) = 0 \tag{27}$$

for each $j \in \{1, 2, \ldots, r\}$, $i \in \{0, 1, \ldots, m_j - 1\}$. Lemma 2.2 implies then that $C(x)$ is divisible by $(x - \lambda_1)^{m_1}$, …, $(x - \lambda_r)^{m_r}$, which shows that either $\deg(C(x)) \geq m_1 + \ldots + m_r = k$ or $C(x)$ is identically zero. By (26), the former option does not hold, so $C(x)$ is identically zero and hence $c_0 = c_1 = \ldots = c_{k-1} = 0$. □

Note that in the case of distinct roots, the determinant (24) is a Vandermonde's determinant, which is known to be nonzero.

Even more information can be extracted from the proof. The following proposition may be of reader's interest.

Proposition 2.12. *Determinant (24) is either in $\mathbb{R}$ or in $i\mathbb{R}$.*

Proof. We notice first that if λ is a root of a polynomial $p(x) \in \mathbb{R}$, then the multiplicity of λ and the one of its complex conjugate $\overline{\lambda}$ must coincide. In fact, if λ is a root of $p(x)$, then of course so is $\overline{\lambda}$, and $(x - \lambda)(x - \overline{\lambda}) \in \mathbb{R}[x]$ divides $p(x)$. Thus $p(x) = (x - \lambda)(x - \overline{\lambda})p_1(x)$, where $p_1(x) \in \mathbb{R}$, and we can apply the argumentation to $p_1(x)$ recursively.

Let us denote the determinant (24) by D. It is clear that the complex conjugate of D is obtained by taking the conjugates of all entries of (24). On the other hand, if there is a non-real root λ of $p(x)$ (recall that $p(x)$ is the characteristic polynomial of the recurrence in question) occurring exactly in m columns of (24), then there are also m columns of (24) which are identical to those m first ones, except that λ is replaced with $\overline{\lambda}$. This is because λ and $\overline{\lambda}$ have the same multiplicity. By swapping all the m columns where λ occurs with those ones where $\overline{\lambda}$ occurs, we can modify $\overline{D}$ into a form where the columns containing λ and $\overline{\lambda}$ are restored to the positions they were in D. Performing the same operation to each complex root of $p(x)$ we have again the original determinant D, since the rows containing occurrences of the real roots are not affected at all when taking the complex conjugates.

If there are $2r$ non-real roots λ_{i_1}, …, λ_{i_r} and their complex conjugates, (we assume that none of λ_{i_1}, …, λ_{i_r} are complex conjugates to each other) with multiplicities m_{i_1}, …, m_{i_r}, we learn that

$$\overline{D} = (-1)^{m_{i_1} + \ldots + m_{i_r}} D.$$

Now if $s = m_{i_1} + \ldots + m_{i_r}$ is even, we have $\overline{D} = D$, which implies $D \in \mathbb{R}$, whereas odd s means that $\overline{D} = -D$, and consequently $D \in i\mathbb{R}$. □

Proposition 2.13. *Let*

$$u_n = \sum_{i=1}^{r} p_i(n)\lambda_i^n$$

as in (22). If $\lambda_{i_2} = \overline{\lambda}_{i_1}$, then also $p_{i_2}(x) = \overline{p}_{i_1}(x)$, where $\overline{p}_{i_1}(x)$ stands for the polynomial which is obtained from $p_{i_1}(x)$ by replacing each coefficient by its complex conjugate.

Proof. Coefficients p_{ij} can be computed by Cramer's rule: $p_{ij} = D(i,j)/D$, where $D(i,j)$ stands for the determinant which is obtained from D by replacing the column $(0^j\lambda_i^0, 1^j\lambda_i^1, 2^j\lambda_i^2, \ldots, (k-1)^j\lambda_i^{k-1})^T$ in (24) by $(u_0, u_1, \ldots, u_{k-1})^T$. If then $\lambda_{i_2} = \overline{\lambda}_{i_1}$, we learn that, as in the previous proposition, that the determinant $D(i_2,j)$ can be obtained from $D(i_1,j)$ by first taking the complex conjugate of each entry of $D(i_1,j)$ and then swapping $s = m_{i_1} + \ldots + m_{i_r}$ columns (it is needed here that each u_n is real). It follows that $D(i_2,j) = (-1)^s\overline{D(i_1,j)}$, which together with $D = (-1)^s\overline{D}$ gives that

$$\overline{p}_{i_1 j} = \frac{\overline{D(i_1,j)}}{\overline{D}} = \frac{(-1)^s D(i_2,j)}{(-1)^s D} = \frac{D(i_2,j)}{D} = p_{i_2,j}.$$

□

2.4 Liner Forms of Logarithms

In the sequel we will need the following important results by Alan Baker [1].

Theorem 2.14. *Let α_1, ..., α_n be non-zero algebraic numbers with degrees at most d and heights at most A. Furthermore, let β_0, ..., β_n be algebraic numbers with degrees at most d and heights at most $B \geq 2$. Then, for*

$$\Lambda = \beta_0 + \beta_1 \log\alpha_1 + \ldots + \beta_n \log\alpha_n$$

we have either $\Lambda = 0$ or $|\Lambda| > B^{-C}$, where C is an algorithmically computable number depending only on n, d, A and the branch of the logarithms chosen.

The estimate for C is of form $C'(\log A)^\kappa$, where κ depends only on n, and C' depends only on n and d. In the case when $\beta_0 = 0$ and β_1, ..., β_n are rational integers, the theorem holds with $C = C'\Omega\log\Omega$, where

$$\Omega = \log A_1 \cdot \ldots \cdot \log A_n, \tag{28}$$

and the numbers $A_i \geq 4$ are chosen such that $H(\alpha_i) \leq A_i$ [1]. The following strengthened, quantitative version of the above theorem can be found in [2]:

Theorem 2.15. *Let the notations be as in the above (with the principal branch of the logarithms). If $\Lambda \neq 0$, then*

$$|\Lambda| > (B\Omega)^{-C\Omega\log\Omega'},$$

where $\Omega' = \Omega/\log A_n$ and $C = (16nd)^{200n}$.

Remark 2.16. *Recall that the height of an algebraic number α is defined as the maximum of the absolute values of the coefficients of the defining polynomial of α. However, the above theorem holds if one replaces the the height by the maximal absolute value of the coefficients of any polynomial in $\mathbb{Z}[x]$ having α as a zero [4]. See also the improved versions of the above inequalities expressed in the terms of logarithmic Weil height [3], as well as the p-adic analogues [25].*

3 Infinity of Zeros

In this section, we represent the proof of Skolem-Mahler-Lech theorem in a form given by G. Hansel [13]. The core of the proof by Hansel is the following theorem.

Theorem 3.1. *Let $p > 2$ be a prime number and d_i any sequence of integers, and define*

$$b_n = \sum_{i=0}^{n} \binom{n}{i} p^i d_i.$$

If $b_n = 0$ for infinitely many n, then $b_n = 0$ for each n.

Before giving the proof of Theorem 3.1, we need some lemmata and definitions. In the following lemmata, we assume the prime number $p > 2$ fixed, unless stated otherwise.

Lemma 3.1. *If p is any prime number and $n \in \mathbb{Z}$, then*

$$v_p(\frac{p^n}{n!}) \geq n\frac{p-2}{p-1}.$$

Proof. Anyway

$$v_p(\frac{p^n}{n!}) = v_p(p^n) - v_p(n!) = n - v_p(n!),$$

so it suffices to estimate $v_p(n!)$. It is plain that

$$\begin{aligned} v_p(n!) &= \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \ldots \\ &\leq \frac{n}{p} + \frac{n}{p^2} + \ldots = \frac{n}{p-1}, \end{aligned}$$

hence

$$v_p(\frac{p^n}{n!}) \geq n - \frac{n}{p-1} = n\frac{p-2}{p-1}.$$

□

Definition 3.2. *Given a polynomial*

$$P(x) = a_0 + a_1 x + \ldots + a_n x^n \in \mathbb{Q}[x],$$

let

$$\omega_k(P) = \begin{cases} \min\{v_p(a_j) \mid j \geq k\}, & \textit{if } k \leq n \\ \infty, & \textit{if } k > n. \end{cases}$$

Remark 3.3. *It is clear that* $\omega_0(P) \leq \omega_1(P) \leq \ldots$ *for each polynomial* $P(x)$.

Remark 3.4. *For a fixed value* $P(t)$ *(*$t \in \mathbb{Z}$*) we have of course*

$$\begin{array}{rcl} v_p(P(t)) & = & v_p(a_0 + a_1 t + \ldots a_n t^n) \\ & \geq & \min\{v_p(a_0), v_p(a_1 t), \ldots, v_p(a_n t^n)\} \\ & \geq & \min\{v_p(a_0), v_p(a_1), \ldots, v_p(a_n)\} = \omega_0(P). \end{array}$$

Lemma 3.2. *Let* $P(x)$, $Q(x) \in \mathbb{Q}[x]$ *and* $n_1, \ldots, n_k \in \mathbb{Z}$. *If*

$$P(x) = (x - n_1) \cdot \ldots \cdot (x - n_k) Q(x),$$

then $\omega_k(P) \leq \omega_0(Q)$.

Proof. We show that if $P(x) = (x - n_1)Q(x)$, then $\omega_{k+1}(P) \leq \omega_k(Q)$ for each k. The claim follows then by applying this result recursively to polynomial $Q(x)$. By writing

$$Q(x) = q_0 + q_1 x + \ldots + q_n x^n$$

and

$$P(x) = p_0 + p_1 x + \ldots + p_{n+1} x^{n+1}$$

we have that $p_{j+1} = q_j - n_1 q_{j+1}$, which implies that

$$q_j = p_{j+1} + n_1 p_{j+2} + n_1^2 p_{j+3} + \ldots + n_1^{n-j} p_{n+1},$$

which shows that

$$\begin{array}{rcl} v_p(q_j) & = & v_p(p_{j+1} + n_1 p_{j+2} + \ldots + n_1^{n-j} p_{n+1}) \\ & \geq & \min\{v_p(p_{j+1}), v_p(p_{j+2}), \ldots, v_p(p_{n+1})\} \\ & = & \omega_{j+1}(P). \end{array}$$

It follows that $\omega_k(Q) \geq \omega_{k+1}(P)$. □

Let $n \in \mathbb{N}$ be fixed and define $R(x) \in \mathbb{Q}[x]$ as

$$R(x) = \sum_{i=0}^{n} d_i p^i \frac{x(x-1) \cdot \ldots \cdot (x-i+1)}{i!}. \tag{29}$$

Lemma 3.3. *For each* k, *we have* $\omega_k(R) \geq k \frac{p-2}{p-1}$.

Proof. It is clear that $R(x)$ can be written as

$$\begin{array}{rcl} R(x) & = & \displaystyle\sum_{i=0}^{n} d_i \frac{p^i}{i!} x(x-1) \cdot \ldots (x-i+1) \\ & = & \displaystyle\sum_{i=0}^{n} d_i \frac{p^i}{i!} \sum_{j=0}^{i} s_{ij} x^j = \sum_{j=0}^{n} \sum_{i=j}^{n} d_i \frac{p^i}{i!} s_{ij} x^j, \end{array}$$

where s_{ij} are integers, so called *Stirling numbers (of the first kind)*. Therefore, the coefficient of x^j in polynomial $R(x)$ is given by

$$\sum_{i=j}^{n} d_i \frac{p^i}{i!} s_{ij} x^j,$$

and

$$\begin{aligned} v_p(\sum_{i=j}^{n} d_i \frac{p^i}{i!} s_{ij}) &\geq \min_{i \geq j}\{v_p(d_i \frac{p^i}{i!} s_{ij})\} \geq \min_{i \geq j}\{v_p(\frac{p^i}{i!})\} \\ &\geq \min_{i \geq j}\{i \cdot \frac{p-2}{p-1}\} \geq j \cdot \frac{p-2}{p-1}, \end{aligned}$$

so it follows that $\omega_j(R) \geq j \cdot \frac{p-2}{p-1}$. □

Proof of Theorem 3.1. We show that if $b_n = 0$ for $n \in \{n_1, \ldots, n_k\}$, then $v_p(b_n) \geq k \cdot \frac{p-2}{p-1}$ for each b_n. Theorem 3.1 follows, since now k can be chosen arbitrarily large, and therefore $v_p(b_n) = \infty$ for each member of the sequence b_n.

Choose $n = \max\{n_1, \ldots, n_k\}$ and recall the definition of $R(x)$ in (29).

$$R(x) = \sum_{i=0}^{n} d_i p^i \frac{x(x-1) \cdot \ldots \cdot (x-i+1)}{i!} \in \mathbb{Q}[x].$$

It is clear that for each $t \leq n$ we have

$$R(t) = \sum_{i=0}^{n} \binom{t}{i} p^i d_i = \sum_{i=0}^{t} \binom{t}{i} p^i d_i = b_t,$$

and since $R(x)$ has integer zeros n_1, …, n_k,

$$R(x) = (x - n_1) \cdot \ldots \cdot (x - n_k) Q(x)$$

for some polynomial $Q(x)$. Therefore also $v_p(R(t)) \geq v_p(Q(t))$, and we can estimate by Remark 3.4 and by the previous lemmata as follows:

$$v_p(b_t) = v_b(R(t)) \geq v_p(Q(t)) \geq \omega_0(Q) \geq \omega_k(R) \geq k \cdot \frac{p-2}{p-1}.$$

□

The following theorem is a special case of Skolem-Mahler-Lech theorem. We will not give a general case of the proof here, but instead we refer to [13], where it is explained how the proof extends into the most general case.

Theorem 3.5. *Let u_n be a linear recurrent sequence. Then the set $Z(u_n) = \{i \mid u_i = 0\}$ is a union of a finite set F and finitely many arithmetic progressions. That is, Z admits representation*

$$Z = F \cup (a_1 + N\mathbb{Z}) \cup \ldots \cup (a_r + N\mathbb{Z}).$$

Moreover, numbers N, a_1, …, a_r can be found algorithmically.

Remark 3.6. *The "constructive part" of the above theorem means that one can decide, given a linear recurrence, whether there are infinitely many n such that $u_n = 0$.*

Proof. Let each u_n be given as

$$u_n = a_{k-1}u_{n-1} + \ldots + a_1 u_{n-k+1} + a_0 u_{n-k}.$$

Without loss of generality we can assume that $a_0 \neq 0$, for otherwise the sequence u_n would satisfy a shorter recursion. Then we can define vectors $\boldsymbol{v} \in \mathbb{Z}^k$, $\boldsymbol{w} \in \mathbb{Z}^k$, and a matrix $M \in \mathbb{Z}^{k \times k}$ as in (3) such that

$$u_n = \boldsymbol{v} M^n \boldsymbol{w}^T,$$

and that $\det(M) = \pm a_0 \neq 0$.

Choose then a prime number p not dividing a_0 and consider the image M_p of matrix M in $\mathbb{F}_p^{k \times k}$ under the canonical projection $\mathbb{Z} \to \mathbb{F}_p$. Now that $p \nmid a_0$, we have that $\det(M_p) \neq 0$. Moreover, the group of invertible $k \times k$-matrices over $\mathbb{F}_p$ has cardinality at most p^{k^2}, so it follows that there exists a number $N \leq p^{k^2}$ such that $M_p^N = I$ in the group of invertible matrices over $\mathbb{F}_p$. Lifting the equation $M_p^N = I$ back to matrices over $\mathbb{Z}$ we learn that there is a number $N \leq p^{k^2}$ and a matrix $M_1 \in \mathbb{Z}^{k \times k}$ such that

$$M^N = I + pM_1.$$

Notice that numbers p and N, as well as the matrix M_1 can be found algorithmically.

Now, for any number n we can write $n = mN + r$, where $0 \leq r < N$ and see that

$$M^n = M^{mN+r} = M^{Nm}M^r = (I + pM_1)^m M^r.$$

Then we have

$$u_n = \boldsymbol{v} M^k \boldsymbol{w}^T = \boldsymbol{v}(I + pM_1)^m M^r \boldsymbol{w}^T,$$

i.e.,

$$u_{mN+r} = \boldsymbol{v}(I + pM_1)^m \boldsymbol{w}_r^T,$$

where $\boldsymbol{w}_r^T = M^r \boldsymbol{w}^T$.

Now we can split the sequence u_n into N different linearly recurrent sequences $u_m^{(r)}$ for each $r \in \{0, 1, \ldots, N-1\}$ by setting

$$u_m^{(r)} = u_{mN+r} = \boldsymbol{v}(I + pM_1)^m \boldsymbol{w}_r^T = \sum_{i=0}^{m} \binom{m}{i} p^i \boldsymbol{v} M_1^i \boldsymbol{w}_r^T.$$

By Theorem 3.1, sequence $u_m^{(r)}$ either vanishes identically or contains only finitely many zeros. To check whether $u_m^{(r)}$ vanishes identically is an easy task: it suffices to compute k first members of the sequence. The claim follows immediately. □

4 Decidable Cases

4.1 Dominant Roots

Before the main theorem of this section, we must introduce several lemmata. The following useful reduction was mentioned in [5].

Lemma 4.1. *Let*

$$u_n = a_{k-1}u_{n-1} + \ldots + a_1 u_{n-k+1} + a_0 u_{n-k} \tag{30}$$

be a linear recurrent sequence and

$$u_n = p_1(n)\lambda_1^n + p_2(n)\lambda_2^n + \ldots + p_r(n)\lambda_r^n \tag{31}$$

it representation as in (18). Sequence u_n can be algorithmically reduced to a finite number of linear recurrent sequences $u_m^{(0)}$, $u_m^{(1)}$ …, $u_m^{(N-1)}$ such that λ_i/λ_j is not a root of unity for $i \neq j$ in the representation (31) for sequences $u_m^{(i)}$.

Proof. If some λ_i/λ_j is a root of unity, let

$$N = \mathrm{lcm}\{\mathrm{ord}(\lambda_i/\lambda_j) \mid i \neq j \text{ and } \lambda_i/\lambda_j \text{ is a root of unity}\}.$$

We can split the original sequence u_n into M distinct subsequences: For each $j \in \{0, 1, \ldots, N-1\}$ we define $u_m^{(j)} = u_{j+mN}$ and notice that

$$u_m^{(j)} = u_{j+mN} = \sum_{i=1}^{r} p_i(j+mN)\lambda_i^{j+mN} = \sum_{i=1}^{r} p_i(j+mN)\lambda_i^{j}(\lambda_i^N)^m \tag{32}$$

If $\{\mu_1, \ldots, \mu_{r'}\}$ are the distinct elements of set $\{\lambda_1^N, \ldots, \lambda_r^N\}$, we can write (32) as

$$u_m^{(j)} = \sum_{i=1}^{r'} q_i(m)\mu_i^m \tag{33}$$

If $i \neq j$, then μ_i/μ_j is not a root of unity, since $(\mu_i/\mu_j)^t = 1$ is equivalent to $(\lambda_i/\lambda_j)^{Nt} = 1$ (we choose the notations such that $\mu_i = \lambda_i^N$ and $\mu_j = \lambda_j^N$), which implies $(\lambda_i/\lambda_j)^N = 1$. This contradicts the assumption that $\mu_i \neq \mu_j$.

We will yet demonstrate, that the property $\lambda_i = \overline{\lambda}_j \implies p_i(x) = \overline{p}_j(x)$ is preserved in this reduction. But this is very straightforward: Assume that $\mu_1 = \lambda_{i_1}^N = \ldots = \lambda_{i_k}^N$. Then trivially $\overline{\mu}_1 = \overline{\lambda}_{i_1}^N = \ldots = \overline{\lambda}_{i_k}^N$ and $q_1(x) = p_{i_1}(j + xN)\lambda_{i_1}^j + \ldots + p_{i_k}(j + xN)\lambda_{i_k}^j$ is the coefficient polynomial of μ_1, whereas the coefficient polynomial of $\overline{\mu}_1$ is $\overline{p}_{i_1}(j+xN)\overline{\lambda}_{i_1}^j + \ldots + \overline{p}_{i_k}(j+xN)\overline{\lambda}_{i_k}^j = \overline{q}_1(x)$. □

We will also need several height estimations.

Lemma 4.2. *Let α be an algebraic number and denote $q(x) = q_n x^n + q_{n-1}x^{n-1} + \ldots + q_1 x + q_0 \in \mathbb{Z}[x]$ and $H(q) = \max\{|q_0|, |q_1|, \ldots, |q_n|\}$. If $q(\alpha) = 0$, then there exists a (computable) number C_n depending on n only such that $H(\alpha) \leq C_n H(q)$.*

Proof. If $q(x)$ is irreducible, there is an integer C such that $q(x) = Cp(x)$, where $p(x)$ is the defining polynomial of α. Hence $H(q) = |C|\, H(\alpha) \geq H(\alpha)$, and we can choose $C_n = 1$. If $q(x)$ is not irreducible, then $q(x) = p(x)r(x)$, where $p(x), r(x) \in \mathbb{Z}[x]$ and $p(x)$ is the defining polynomial of α. We must show that the coefficients cannot cancel too much when computing the product $p(x)r(x)$. For this purpose, we will employ ***Bombieri's norm*** [6]: If $q(x)$ is as above, then the Bombieri's 2-norm is defined as

$$[q] = \Big(\sum_{i=0}^{n} \binom{n}{i}^{-1} |q_i|^2 \Big)^{\frac{1}{2}}.$$

Bombieri's inequality [6] states that if $p(x)r(x) = q(x)$, and $\deg(p) = m$, then

$$[p][r] \leq \binom{n}{m}^{\frac{1}{2}} [q].$$

Since $r(x) \in \mathbb{Z}[x]$, we have $[r] \geq 1$, and consequently

$$[p] \leq \binom{n}{m}^{\frac{1}{2}} [q]. \tag{34}$$

It can be shown [30] that

$$\sum_{i=0}^{n} \binom{n}{i}^{-1} = \frac{n+1}{2^n} \sum_{i=0}^{n} \frac{2^i}{i+1},$$

which implies that

$$\sum_{i=0}^{n} \binom{n}{i}^{-1} \leq \frac{8}{3} < 4,$$

and hence

$$[q] = \Big(\sum_{i=0}^{n} \binom{n}{i}^{-1} |q_i|^2 \Big)^{\frac{1}{2}} \leq \Big(\sum_{i=0}^{n} \binom{n}{i}^{-1} H(q)^2 \Big)^{\frac{1}{2}} < 2H(q). \tag{35}$$

Since $p(x)$ is the defining polynomial of α, there must be a coefficient p_j of $p(x)$ such that $H(\alpha) = |p_j|$. Then

$$[p] = \Big(\sum_{i=0}^{m} \binom{m}{i}^{-1} |p_i|^2 \Big)^{\frac{1}{2}} \geq \Big(\binom{m}{j}^{-1} H(\alpha)^2 \Big)^{\frac{1}{2}} = H(\alpha) \binom{m}{j}^{-\frac{1}{2}}. \tag{36}$$

Combining (34), (35) and (36) we see that

$$H(\alpha) \leq \binom{m}{j}^{\frac{1}{2}} \binom{n}{m}^{\frac{1}{2}} 2H(q) \leq \binom{m}{\lfloor m/2 \rfloor}^{\frac{1}{2}} \binom{n}{\lfloor n/2 \rfloor}^{\frac{1}{2}} 2H(q) \leq 2 \binom{n}{\lfloor n/2 \rfloor} H(q).$$

□

Lemma 4.3. *If α is an algebraic number of degree d and height H, and $\beta^n = \alpha$, then $\deg(\beta) \leq nd$ and $H(\beta) \leq C_{n,d}H(\alpha)$, where $C_{n,d}$ is a (computable) number depending on n and d alone.*

Proof. Let $p(x) = p_d x^d + p_{d-1}x^{d-1} + \ldots + p_1 x + p_0$ be the defining polynomial of α. Clearly polynomial

$$q(x) = p_d x^{nd} + p_{d-1}x^{n(d-1)} + \ldots + p_1 x^n + p_0$$

has β as a root, and the first claim follows directly. The second claim follows by Lemma 4.2, we can take $C_{n,d} = 2\binom{nd}{\lfloor nd/2 \rfloor}$. □

Lemma 4.4. *Let $p(x) \in \mathbb{C}[x]$. For each natural number n, let $p(n) = |p(n)|\, e^{i\phi_n}$. Then there are algorithmically computable numbers B and C such that $H(e^{i\phi_n}) \leq Bn^C$. Numbers B and C depend only on the polynomial $p(x)$. Moreover, $e^{i\phi_n} \to e^{i\phi}$ as $n \to \infty$, where $e^{i\phi}$ is an algebraic number which can be found algorithmically.*

Proof. Let $p(x) = p_m x^m + \ldots + p_1 x + p_0$. First we will estimate the height of $p(n)$. Let F be a field containing all the coefficients of $p(x)$ and denote $d = [F : \mathbb{Q}]$. Then clearly $p(n) \in F$ for each $n \in \mathbb{N}$, which implies that $\deg(p(n)) \leq d$ for each n. Let $H = \max\{H(p_0), \ldots, H(p_{m-1})\}$. Now if $\deg(\alpha) \leq d$, it is easy to see that $H(n\alpha) \leq n^d H(\alpha)$ for each $n \in \mathbb{N}$. In fact, if $\deg(\alpha) = d_1$ and

$$q(x) = a_{d_1} x^{d_1} + \ldots + a_1 x + a_0$$

is the defining polynomial of α, then polynomial

$$q_1(x) = a_{d_1} x^{d_1} + \ldots + a_1 n^{d_1 - 1} x + a_0 n^{d_1} \tag{37}$$

has $n\alpha$ as a zero. Moreover, (37) is irreducible, since otherwise $\deg(n\alpha) < d_1$ which would imply that also $\alpha = \frac{1}{n} \cdot n\alpha$ has degree less than d_1. Therefore (37) is the defining polynomial of $n\alpha$, and estimation $H(n\alpha) \leq n^d H(\alpha)$ follows.

As a consequence, we get that $H(p_i n^i) \leq n^{id} H \leq n^{md} H = H_1$ for each $i \in \{0, 1, \ldots, m\}$. For estimating the height of the sum of two algebraic numbers, we can use the following fact [1, p. 24]: If $\deg(\alpha)$, $\deg(\beta) \leq d$ and $H(\alpha)$, $H(\beta) \leq H$, then there exists a computable constant C_1 (depending on d only) such that $H(\alpha + \beta)$, $H(\alpha\beta) \leq H^{C_1}$. Thus

$$H(p_m n^m + p_{m-1} n^{m-1}) \leq H_1^{C_1}.$$

Similarly

$$H(p_m n^m + p_{m-1} n^{m-1} + p_{m-2} n^{m-2}) \leq (H_1^{C_1})^{C_1} = H_1^{C_1^2},$$

and continuing this way we see that

$$H(p(n)) = H(p_m n^m + p_{m-1} n^{m-1} + \ldots + p_1 n + p_0) \leq H_1^{C_1^m},$$

and that

$$H(p(n)^2) \leq (H_1^{C_1^m})^{C_1} = H_1^{C_1^{m+1}}$$

as well as

$$H(|p(n)|^2) = H(p(n)\overline{p}(n)) \leq (H_1^{C_1^m})^{C_1} = H_1^{C_1^{m+1}}.$$

By Lemma 4.3

$$H(|p(n)|) \leq C_2 H_1^{C_1^{m+1}},$$

where C_2 is a computable constant depending on d only, and $\deg(|p(n)|) \leq 2d$.

It is plain that $H(|p(n)|) = H(|p(n)|^{-1})$, which finally implies that

$$\begin{aligned} H(e^{i\phi_n}) &= H(|p(n)|^{-1} \cdot |p(n)| \, e^{i\phi_n}) \\ &= H(|p(n)|^{-1} \, p(n)) \leq (C_2 H_1^{C_1^{m+1}})^{C_3}, \end{aligned}$$

where C_3 depends on d only. Denoting $C_4 = C_2^{C_3} H^{C_1^{m+1} C_3}$, $C_5 = mdC_1^{m+1}C_3$, and recalling that $H_1 = n^{md}H$, we see that

$$H(e^{i\phi_n}) \leq C_4 n^{C_5},$$

where C_4 and C_5 are (computable) constants that depend only on polynomial $p(x)$.

The latter claim is trivial, since

$$\frac{p(n)}{|p(n)|} = \frac{p_m n^m}{|p(n)|} + O(\frac{1}{n}) = \frac{p_m}{|p(n)/n^m|} + O(\frac{1}{n}),$$

and

$$e^{i\phi_n} = \frac{p(n)}{|p(n)|} \to e^{i\phi},$$

where ϕ is the phase of $p_m = |p_m| \, e^{i\phi}$. Another obvious but important fact is that for any $\epsilon > 0$ it is possible to find algorithmically an integer M such that

$$\left| e^{i\phi} - e^{i\phi_n} \right| \leq \epsilon$$

whenever $n \geq M$. □

Lemma 4.5. *If $p_1(x), p_2(x) \in \mathbb{C}[x]$ are both of degree d, then $\left|\frac{p_1(x)}{p_2(x)}\right|$ tends to a finite limit as $x \to \infty$ along the real axis, and the convergence to the limit is ultimately monotonous. Moreover, if $|p_1(x)| \neq |p_2(x)|$, then $\left|\frac{p_1(x)}{p_2(x)}\right| = L \pm \Theta(x^{-k})$, where L, k and the coefficients belonging to Θ-notation can be algorithmically found.*

Proof. In this proof, we assume that $x \in \mathbb{R}$. Then $|p_1(x)|^2 = p_1(x)\overline{p}_1(x) \in \mathbb{R}[x]$ is a polynomial of degree $2d$. Let the leading coefficients of $p_1(x)$ and $p_2(x)$ be c_1 and c_2, respectively. Then it is easy to see that the leading coefficients of $|p_1(x)|^2$ and $|p_1(x)|^2$ are $|c_1|^2$ and $|c_2|^2$, respectively, and as a consequence

$$|p_1(x)|^2 - \frac{|c_1|^2}{|c_2|^2} |p_2(x)|^2 = q(x),$$

where $q(x) \in \mathbb{R}[x]$ is a nonzero polynomial of degree less than $2d$.

Polynomial $q(x)$ can be constructed from polynomials $p_1(x)$ and $p_2(x)$, and when the coefficients of $q(x)$ are known, it is easy to find M_1 such that $q(x) \neq 0$ when $x \geq M_1$. If $\deg(q(x)) = d_1 < 2d$, and $q(x)$ is ultimately positive (resp. negative), one can use standard techniques for finding constant C_1, $C_2 > 0$ such that $C_1 x^{d_1} \leq q(x) \leq C_2 x^{d_1}$ (resp. $-C_1 x^{d_1} \leq q(x) \leq -C_2 x^{d_1}$), when $x \geq M_2$. Similarly we can find positive constants C_3, C_4, and M_3 such that $C_3 x^{2d} \leq |p_2(x)|^2 \leq C_4 x^{2d}$, when $x \geq M_3$. If now $q(x)$ is ultimately positive, then

$$\left|\frac{p_1(x)}{p_2(x)}\right|^2 = \left|\frac{c_1}{c_2}\right|^2 + \frac{q(x)}{|p_2(x)|^2},$$

where the remainder satisfies

$$\frac{C_1 x^{d_1}}{C_4 x^{2d}} \leq \frac{q(x)}{|p_2(x)|^2} \leq \frac{C_2 x^{d_1}}{C_3 x^{2d}}$$

for $x \geq M = \max\{M_1, M_2, M_3\}$. The case that $q(x)$ is ultimately negative is treated in the same way.

The claim follows now from the fact that

$$\left|\frac{p_1(x)}{p_2(x)}\right|^2 - \left|\frac{c_1}{c_2}\right|^2 = \Big(\left|\frac{p_1(x)}{p_2(x)}\right| - \left|\frac{c_1}{c_2}\right|\Big)\Big(\left|\frac{p_1(x)}{p_2(x)}\right| + \left|\frac{c_1}{c_2}\right|\Big)$$

and that the latter factor of the product is in interval $(2\,|c_1/c_2| - \epsilon, 2\,|c_1/c_2| + \epsilon)$ for any $x \geq M_\epsilon$, where M_ϵ can be found algorithmically when ϵ is given. □

Proposition 4.1. *Let*

$$u_n = a_{k-1}u_{n-1} + \ldots + a_1 u_{n-k+1} + a_0 u_{n-k} \tag{38}$$

be a linear recurrent sequence and

$$u_n = p_1(n)\lambda_1^n + p_2(n)\lambda_2^n + \ldots + p_r(n)\lambda_r^n \tag{39}$$

it representation as in (18). Without loss of generality, we can assume that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_r|$. Problem SKOLEM *is decidable if one of the following cases hold:*

1. $|\lambda_1| > |\lambda_2|$.
2. $\lambda_2 = \overline{\lambda_1}$, $|\lambda_1| = |\lambda_2| > |\lambda_3|$.
3. $|\lambda_1| = |\lambda_2| = |\lambda_3| > |\lambda_4|$ *and* $\lambda_1 \in \mathbb{R}$, $\lambda_2 = \overline{\lambda_3}$.

Remark 4.2. *It is decidable whether one of the above conditions holds. In the continuation will also assume that the roots of the characteristic polynomial are always enumerated such that $|\lambda_1| \geq |\lambda_2| \geq \ldots$. The proof also shows that the above the claim holds if actually $r = 1$ in case (1) or if $r = 2$ in case (2), or if $r = 3$ in case (3).*

Remark 4.3. *It must be emphasized that even a more general result than this proposition has been reached in [22]. Here we get a restricted version of [22], by using somewhat lighter mathematical machinery.*

Proof. We prove only the last claim, the other ones follow easily from that. It is worth noticing that by Lemma 4.1, we can assume that λ_i/λ_j is not a root of unity, if $i \neq j$.

We denote by F the field which contains all the roots $\lambda_1, \ldots, \lambda_r$, as well as their absolute values $|\lambda_1|, \ldots, |\lambda_r|$. Furthermore, we assume that F is normal (see Remark 2.4) and denote $d = [F : \mathbb{Q}]$.

By the assumption and Proposition 2.13, we can write

$$u_n = p_1(n)\lambda_1^n + p_2(n)\lambda_2^n + \overline{p}_2(n)\overline{\lambda}_2^n + R_1(n), \tag{40}$$

where $R_1(n) = p_4(n)\lambda_4^n + \ldots + p_r(n)\lambda_r^n$, so $R_1(n)/\lambda_1^n$ tends exponentially fast to 0 as n tends to infinity. We denote the main term by $R_0(n) = u_n - R_1(n)$.

We prove the claim only in the case $\lambda_1 < 0$, the case $\lambda_1 > 0$ is even easier. Denote $\lambda_1 = -\lambda$, $\lambda_2 = \lambda e^{i\theta}$, $\lambda_3 = \lambda e^{-i\theta}$. We write also $p_1(n) = A_n$, $p_2(n) = B_n e^{i\phi_n}$, $\overline{p}_2(n) = B_n e^{-i\phi_n}$, where $B_n \in \mathbb{R}$, $\phi_n \in (-\pi, \pi]$. By proposition 2.13, $A_n \in \mathbb{R}$ anyway.

To examine the leading term of (40), we write

$$\begin{aligned} R_0(n) &= p_1(n)\lambda_1^n + p_2(n)\lambda_2^n + \overline{p}_2(n)\overline{\lambda}_2^n \\ &= A_n(-\lambda)^n + B_n e^{i\phi_n}\lambda^n e^{in\theta} + B_n e^{-i\phi_n}\lambda^n e^{-in\theta} \\ &= \lambda^n B_n\Big(\frac{A_n}{B_n}(-1)^n + e^{i(\phi_n+n\theta)} + e^{-i(\phi_n+n\theta)}\Big) \\ &= 2\lambda^n B_n\Big((-1)^n\frac{A_n}{2B_n} + \cos(\phi_n + n\theta)\Big). \end{aligned}$$

If $\deg(p_1) > \deg(p_2)$, then clearly $\left|\frac{A_n}{2B_n}\right|$ grows unrestricted when n tends to infinity. It is clearly possible to find (algorithmically) a number M_1 such that $\left|\frac{A_n}{2B_n}\right| \geq 2$ whenever $n \geq M_1$. For such values of n, we have

$$\left|(-1)^n\frac{A_n}{2B_n} + \cos(\phi_n + n\theta)\right| \geq \left|(-1)^n\frac{A_n}{2B_n}\right| - |\cos(\phi_n + n\theta)| \geq 2 - 1 = 1,$$

and hence $|R_0(n)| \geq 2\lambda^n B_n$ when $n \geq M_1$. Because B_n is an absolute value of a given nonzero polynomial, we can find (algorithmically) a number M_2 such that $B_n \geq 1/2$ when $n \geq M_2$. Thus

$$|u_n| = |R_0(n) + R_1(n)| \geq |R_0(n)| - |R_1(n)| \geq \lambda^n - |R_1(n)| = \lambda^n(1 - \frac{|R_1(n)|}{\lambda^n}).$$

It is now possible to find (algorithmically) a number M such that $\frac{|R_1(n)|}{\lambda^n} < 1$ whenever $n \geq M$. Then also $u_n \neq 0$ when $n \geq M$.

If $\deg(p_1) = \deg(p_2)$, then $\lim_{n\to\infty} \left|\frac{A_n}{2B_n}\right|$ is a finite number which can be straightforwardly recovered from the leading coefficients of p_1 and p_2. If $\deg(p_1) < \deg(p_2)$, then of course the aforementioned limit equals to 0. If $\lim_{n\to\infty} \left|\frac{A_n}{2B_n}\right| > 1$, one can find algorithmically a number M such that $u_n \neq 0$ whenever $n \geq M$, as easily as in the case $\deg(p_1) > \deg(p_2)$.

Lemma 4.5 becomes helpful if $\lim_{n\to\infty} \left|\frac{A_n}{2B_n}\right| = 1$. In this case, the convergence to the limit is ultimately monotonous, and following the proof of Lemma 4.5, we can effectively find a constant M and decide whether $\left|\frac{A_n}{2B_n}\right|$ approaches 1 from above or from below (for $n \geq M$). In the case that the convergence happens from above, we can utilize the proof of Lemma 4.5 to find positive constants M, C and k such that

$$\left|\frac{A_n}{2B_n}\right| - 1 \geq \frac{C}{n^k},$$

when $n \geq M$. Then

$$\left|(-1)^n \frac{A_n}{2B_n} + \cos(\phi_n + n\theta)\right| \geq \left|\frac{A_n}{2B_n}\right| - 1 \geq \frac{C}{n^k}$$

and

$$|u_n| \geq \left|\lambda^n B_n \cdot \frac{C}{n^k}\right| - |R_1(n)| = \lambda^n \Big(B_n \cdot \frac{C}{n^k} - \frac{|R_1(n)|}{\lambda^n}\Big),$$

when $n \geq M$. If the term $\frac{CB_n}{n^k}$ vanishes, it does so at most polynomially, but the latter term $\frac{|R_1(n)|}{\lambda^n}$ vanishes exponentially. It follows that one can find a number M_1 such that $u_n \neq 0$ whenever $n \geq M_1$.

Finally we assume that either $\left|\frac{A_n}{2B_n}\right|$ tends to 1 (ultimately) from below, or that it tends to a limit less than 1. In both cases, we choose first M_1 such that $\left|\frac{A_n}{2B_n}\right| \leq 1$, if $n \geq M_1$. We will extend Mignotte's result [20] to estimate the expression

$$\left|(-1)^n \frac{A_n}{2B_n} + \cos(\phi_n + n\theta)\right| \tag{41}$$

from below by using Baker's theorem. If we can find for (41) a lower bound of form $\frac{C}{n^k}$ for $n \geq M$, then the proof is complete, since we can argue as in the previous case (of course numbers k, C, and M must be found algorithmically).

Now that $\left|\frac{A_n}{2B_n}\right| \leq 1$ for each n under consideration, we can choose $\psi_n \in (-\pi, \pi]$ such that $\cos\psi_n = \frac{A_n}{2B_n}$. In fact, the description of $e^{i\psi_n}$ can be found by using equation $e^{i\psi_n} + e^{-i\psi_n} = \frac{A_n}{B_n}$, which also shows that $e^{i\psi_n}$ is an algebraic number.

Expression (41) takes now form

$$\left|\cos(\phi_n + n\theta) + (-1)^n \cos\psi_n\right|, \tag{42}$$

and we treat cases n even and n odd separately. For even n, (42) is of form

$$|\cos(\phi_n+n\theta)+\cos\psi_n| = \left|\cos\frac{\phi_n+n\theta+\psi_n}{2}\right|\left|\cos\frac{\phi_n+n\theta-\psi_n}{2}\right|, \qquad (43)$$

whereas for odd n, we have

$$|\cos(\phi_n+n\theta)-\cos\psi_n| = \left|\sin\frac{\phi_n+n\theta+\psi_n}{2}\right|\left|\sin\frac{\phi_n+n\theta-\psi_n}{2}\right|. \qquad (44)$$

We handle only the even case, (44) is even easier. In the right hand side of (43), one of the factors is minimal, so

$$|\cos(\phi_n+n\theta)+\cos\psi_n| \geq \left|\cos\frac{\phi_n+n\theta\pm\psi_n}{2}\right|^2, \qquad (45)$$

and we may concentrate on expressions

$$\left|\cos\frac{\phi_n+n\theta\pm\psi_n}{2}\right| = \left|\sin\frac{\phi_n+n\theta\pm\psi_n-\pi}{2}\right|. \qquad (46)$$

It is of course possible to choose $m_n\in\mathbb{Z}$ such that

$$\frac{\phi_n+n\theta\pm\psi_n-\pi}{2}+m_n\pi\in[-\frac{\pi}{2},\frac{\pi}{2}].$$

Moreover, since ϕ_n, ψ_n, $\theta\in[-\pi,\pi]$, such m_n satisfies

$$-\frac{n+2}{2}\leq m_n\leq\frac{n+4}{2}.$$

In particular, m_n satisfies $|m_n|\leq n$, whenever $n\geq 4$. We will then use inequality $|\sin x|\geq\frac{2}{\pi}|x|$ (valid for $x\in[-\frac{\pi}{2},\frac{\pi}{2}]$) to see that

$$\begin{aligned} & \left|\cos\frac{\phi_n+n\theta\pm\psi_n}{2}\right| \\ = & \left|\sin\left(\frac{\phi_n+n\theta\pm\psi_n-\pi}{2}+m_n\pi\right)\right| \\ \geq & \frac{2}{\pi}\left|\frac{\phi_n+n\theta\pm\psi_n-\pi}{2}+m_n\pi\right| \\ = & \frac{1}{\pi}\left|i\phi_n+ni\theta\pm i\psi_n+(2m_n-1)i\pi\right|. \qquad (47)\end{aligned}$$

Numbers $i\phi_n$, $i\theta$, $i\psi_n$, and $i\pi$ occurring in the latest expression are all logarithms of algebraic numbers. To apply Baker's theorem, we must estimate the heights of these algebraic numbers and to find out when (47) may become zero.

By Lemma 4.4 there are constants C_1 and C_2 such that $H(e^{i\phi_n})\leq C_1n^{C_2}$. Clearly $e^{i\phi_n}$ belongs to F, which shows that $d(e^{i\phi_n})\leq d$. It is plain that $H(e^{i\theta})$ is independent of n, and can be estimated using equation $\lambda_2 = \lambda e^{i\theta}$ as in the proof of Lemma 4.4. Moreover, $d(e^{i\theta}) \leq 2d$. Also $H(e^{i\pi}) = H(-1) = 1$ and $d(e^{i\pi}) = 1$.

It remains to estimate the height and the degree of $e^{i\psi_n}$. For that, equation

$$e^{i\psi_n} + e^{-i\psi_n} = \frac{A_n}{B_n} \tag{48}$$

offers a starting point. Both numbers A_n^2 and B_n^2 belong to F, so (48) shows that $e^{i\psi_n}$ satisfies a quadratic equation over a quadratic extension of F. This implies that $\deg(e^{i\psi_n}) \leq 4d$. Solution of (48) is

$$e^{i\psi_n} = \frac{1}{2}\left(\frac{A_n}{B_n} \pm \sqrt{\left(\frac{A_n}{B_n}\right)^2 - 4}\right),$$

which, using [1, p. 24] and Lemma 4.3 shows that it is possible to find algorithmically constants C_3 and C_4 such that $H(e^{i\psi_n}) \leq C_3 n^{C_4}$.

Before applying Baker's theorem we must still find out the values of n for which (47) becomes zero. Formula (47) is zero if and only if

$$e^{i\phi_n} e^{ni\theta} e^{\pm i\psi_n} e^{(2m_n-1)i\pi} = 1,$$

which is implies that

$$(e^{i2\theta})^n = e^{-i2(\phi_n \pm \psi_n)}. \tag{49}$$

We will show how to extend Proposition 2.7 to find out all solutions of (49). This is indeed rather straightforward: by Lemma 4.1 we can assume that $e^{2i\theta} = \lambda_2/\overline{\lambda}_2$ is not a root of unity, which implies that the powers $(e^{2i\theta})^n$ are all distinct. By Lemma 4.4 and construction of numbers ψ_n it is clear that there are limits $e^{i\phi_n} \to e^{i\phi}$ and $e^{i\psi_n} \to e^{i\psi}$, which both are algebraic numbers that can be found algorithmically. Moreover, for each $\epsilon > 0$ one can algorithmically find a number M_ϵ such that $\left|e^{i\phi_n} - e^{i\phi}\right|$, $\left|e^{i\psi_n} - e^{i\psi}\right| \leq \epsilon$ whenever $n \geq M_\epsilon$.

Now $\left|e^{i\theta}\right| = 1$, so we have to apply the latter part of the proof of Proposition 2.7. Recall from the that proof, that since $\lambda_2/\overline{\lambda}_2 = e^{2i\theta}$ is not a root of unity, there are two choices: either $e^{2i\theta}$ is an algebraic integer or not. If $e^{2i\theta}$ is an algebraic integer, there is a (computable) isomorphism $\sigma_i : \mathbb{C} \to \mathbb{C}$ such that $\left|\sigma_i(e^{2i\theta})\right| > 1$. Equation (49) gives then

$$\log\left|\sigma_i(e^{-2i(\phi_n \pm \psi_n)})\right| = n \log\left|\sigma_i(e^{2i\theta})\right|. \tag{50}$$

Now that $e^{-2i\phi_n \pm \psi_n}$ tends to a limit, we can find all potential numbers n for which (50) holds.

If $e^{2i\theta}$ is not an algebraic integer, there is a prime ideal P of the integer ring of field F for which $v_P(e^{2i\theta}) \neq 0$. It follows that

$$n v_P(e^{2i\theta}) = v_P(e^{-2i(\phi_n \pm \psi_n)})$$

and

$$|n| \leq \left|v_P(e^{-2i(\phi_n \pm \psi_n)})\right| \leq \log_2\left|N(e^{-2i(\phi_n \pm \psi_n)})\right|$$

(recall the proof of Proposition 2.7). Again since $e^{i(\phi_n \pm \psi_n)}$ tends to a limit, we can find all numbers n for which (47) is zero.

Notice that (47) is zero if and only if (46) is zero. It follows that also the version $r = 3$ (1 real root and a pair of complex roots with the same absolute value) is covered.

Hereafter we assume that n is chosen so large that (47) is not zero. We can then use Baker's theorem by choosing $B = n$ (recall that $|m_n| \leq n$) and $A = Dn^E$, where D and E are chosen such that $A \geq \max\{C_1 n^{C_2}, C_3 n^{C_4}\}$ for each n. Therefore

$$|i\phi_n + ni\theta \pm i\psi_n + (2m_n - 1)i\pi| \geq n^{-C' \log(Dn^E)^\kappa}, \tag{51}$$

κ is a (computable) constant, and C' is a (computable) number that depends only on d. Lower bound (51) is not exactly of the required form $\frac{C}{n^k}$, but (51) is good enough: by combining (42), (43), (45), (47) and (51) we see that

$$|R_0(n)| \geq 2\lambda^n B_n \frac{1}{n^{2C' \log(Dn^E)^\kappa}},$$

whereas $|R_1(n)| = O(\lambda'^n)$ with $|\lambda'| < \lambda$. It is straightforward to see that expression $|R_0(n)/R_1(n)|$ tends to ∞ as $n \to \infty$, and consequently one can find a limit M such that $u_n \neq 0$ when $n \geq M$. □

4.2 Restricted Depths

Proposition 4.4. *Problem* SKOLEM *is decidable if the recurrence has depth two, i.e.*

$$u_n = a_1 u_{n-1} + a_0 u_{n-2}.$$

Proof. Let λ_1 and λ_2 be the roots of the characteristic polynomial $p(x) = x^2 - a_1 x - a_0$. If $\lambda_1 = \lambda_2$, then the case 1) of Proposition 4.1 applies. If $\lambda_1 \neq \lambda_2$, but $|\lambda_1| = |\lambda_2|$, and $\lambda_1, \lambda_2 \in \mathbb{R}$, then necessarily $\lambda_2 = -\lambda_1$, so $\lambda_1/\lambda_2 = -1$ is a root of unity, and we can use the reduction of Lemma 4.1.

In the remaining case, we have $\lambda_2 = \overline{\lambda_1}$, which leads us to the case 2) of Proposition 4.1. We can write

$$u_n = 2A_1 \lambda_1^n \cos(\phi + n\theta),$$

where $\lambda_1 = \lambda e^{i\theta}$, and $A = A_1 e^{i\phi}$ are algebraic numbers. It remains to decide whether $\cos(\phi + n\theta) = 0$ for some n, but this is equivalent to deciding whether $e^{-2i\phi} = (e^{2i\theta})^n$ for some n. This is decidable by Proposition 2.7. □

Proposition 4.5. *Problem* SKOLEM *is decidable if the recurrence has depth three, i.e.*

$$u_n = a_2 u_{n-1} + a_1 u_{n-2} + a_0 u_{n-3}.$$

Proof. Let $p(x) = x^3 - a_2x^2 - a_1x - a_0$ be the characteristic polynomial of the recurrence and λ_1, λ_2, and λ_3 its roots with $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$. There are now two possibilities: 1) $p(x)$ has three real roots, or 2) $p(x)$ has only one real root.

The first case divides into three subcases: 1a) $|\lambda_1| > |\lambda_2|$ 1b) $|\lambda_1| = |\lambda_2| > |\lambda_3|$, and 1c) $|\lambda_1| = |\lambda_2| = |\lambda_3|$. Case 1a) is a subcase of Proposition 4.1, whereas in case 1b) we have either $\lambda_1 = \lambda_2$ or $\lambda_2 = -\lambda_1$. If $\lambda_1 = \lambda_2$, we have a multiple dominant root, again a subcase of Proposition 4.1 If $\lambda_2 = -\lambda_1$, then $\lambda_1/\lambda_2 = -1$ is a root of unity, and we can use the reduction of Lemma 4.1. In case 1c), we have either a triple root λ_1 (which is a subcase of Proposition 4.1), or a double root. Without loss of generality, we can assume $\lambda_1 = \lambda_2 = -\lambda_3$, so λ_2/λ_3 is a root of unity and we can again reduce as in the proof of Proposition 4.1.

In the second case, necessarily $\lambda_2 = \overline{\lambda_3}$, and this is a special case of Proposition 4.1: We can write

$$u_n = 2A_1\lambda^n \cos(\phi + n\theta) + C\lambda_3^n, \tag{52}$$

where $\lambda_1 = \lambda e^{i\theta} = \overline{\lambda_2}$, and $\lambda_3 \in \mathbb{R}$. If $|\lambda_3| > \lambda$, then we have case 1) of Proposition 4.1. If $|\lambda_3| < \lambda_1$, we have case 2) of the aforementioned proposition, whereas $|\lambda_3| = \lambda$ is a an instance of case 3) of Proposition 4.1. □

Proposition 4.6. *Problem* SKOLEM *is decidable if the recurrence has depth four.*

Proof. Now that the characteristic polynomial has degree four, we can divide into three cases: the number of the real roots is either 4, 2, or 0.

If there are four real roots, then either one of them is dominating (which is a special case of Proposition 4.1), or $\lambda_1 = \lambda$, $\lambda_2 = -\lambda$. But in the latter case, again λ_1/λ_2 is a root of unity, and we can use Lemma 4.1.

If there are exactly two real roots, one having modulus strictly greater than the complex roots, the decision is again easy (if the other real root has the same modulus, we can use again Lemma 4.1). On the other hand, if the real roots have smaller absolute value than the complex roots, then we have again a special case of Proposition 4.1. Therefore we can assume that $\lambda_1 = \lambda e^{i\theta}$, $\lambda_2 = \lambda e^{-i\theta}$, $\lambda_3 = \pm\lambda$, and that $|\lambda_4| \leq \lambda$. We can assume that $\lambda_4 \neq \pm\lambda_3$, since then λ_3/λ_4 would be a root of unity. Then $|\lambda_4| < |\lambda_3|$ and we have again a special case of Proposition 4.1.

In the last case, there are no real roots. If one of the complex roots is a double root, so is its conjugate, and

$$u_n \;=\; (A + Bn)\lambda_1^n + (\overline{A} + \overline{B}n)\overline{\lambda_1}^n,$$

which is a special case (2) of Proposition 4.1.

Finally we consider the case that all the complex roots are disjoint. Then

$$u_n = A\lambda_1^n + \overline{A\lambda_1}^n + C\lambda_3^n + \overline{C\lambda_3}^n, \tag{53}$$

and if $|\lambda_1| > |\lambda_3|$, we have again a special case of Proposition 4.1.

Therefore we can assume that $|\lambda_1| = |\lambda_3| = \lambda$. In this case we can rewrite (53) as

$$\begin{aligned} u_n &= A\lambda^n e^{in\theta_1} + \overline{A}\lambda^n e^{-in\theta_1} + \lambda^n e^{in\theta_2} + \overline{C}\lambda^n e^{-in\theta_2} \\ &= \lambda^n \Big(A(e^{i\theta_1})^n + \overline{A}(e^{i\theta_1})^{-n} + C(e^{i\theta_2})^n + \overline{C}(e^{i\theta_2})^{-n} \Big). \end{aligned} \tag{54}$$

To find out if (54) is zero for some $n \in \mathbb{N}$ is a special case of Proposition 2.10, a decidable task.

□

Proposition 4.7. *Problem* SKOLEM *is decidable if the recurrence has depth five.*

Proof. In this case, the characteristic polynomial has degree 5. It is possible to have 1, 3, or 5 real roots.

If there are 5 real roots, then either one of them is dominating (which is a subcase of Proposition 4.1), or some of them have the same absolute value (which we can exclude by Lemma 4.1).

If there are 3 real roots, then there are two complex roots consisting of a pair of complex conjugates. In this case, either the complex roots have greater absolute value than the real roots (which is a subcase of Proposition 4.1), or there is a real root which has exactly the same absolute value as the complex roots. If, in the latter case, there are more than one real roots with the same absolute value, we can apply Lemma 4.1. If there is exactly one real root sharing the absolute value with the complex roots, we have again a subcase of Lemma 4.1.

In the last case, there is only one real root and 4 complex roots. If the roots do not share the same absolute value, we have one of the cases discussed before. If one of the complex roots has multiplicity greater than 1, then so does its complex conjugate, and we have a subcase of Proposition 4.1. If all the roots have the same absolute value, then

$$\begin{aligned} u_n &= A\lambda_1^n + B\lambda_2 + \overline{B\lambda_2}^n + D\lambda_4 + \overline{D\lambda_4}^n \\ &= A(\pm\lambda)^n + B\lambda^n e^{in\theta_1} + \overline{B}\lambda^n e^{in\theta_1} + De^{in\theta_2}\lambda^n + \overline{D}e^{-in\theta_2} \\ &= \lambda^n \big(A(\pm 1)^n + B(e^{i\theta_1})^n + \overline{B}(e^{i\theta_1})^{-n} + D(e^{i\theta_2})^n + \overline{D}(e^{i\theta_2})^{-n} \big). \end{aligned}$$

We can now take separately the cases for odd and even n, and use Proposition 2.10 to decide whether $u_n = 0$ for some n. □

References

[1] A. Baker: *Transcendental Number Theory*. Cambridge University Press (1975).

[2] A. Baker: *The Theory of Linear Forms in Logarithms*. In A. Baker and D. W. Masser (eds.): *Transcendence Theory: Advances and Applications*, pp. 1–27. Academic Press (1977).

[3] A. Baker and G. Wüstholz: *Logarithmic forms and group varietis*. Journal für die reine und angewandte Mathematik 442, pp. 19–62 (1993).

[4] A. Baker, personal communication.

[5] J. Berstel and M. Mignotte: *Deux propriétés décidables des suites récurrentes linéaires*. Bull. Soc. math. France, 104, pp. 175–184 (1976).

[6] P. Borwein and T. Erdélyi: *Polynomials and Polynomial Inequalities*. Springer (1995).

[7] J. Cassaigne, T. Harju and J. Karhumäki: *On the decidability of the freeness of matrix semigroups.* Internat. J. of Algebra Comp. 9, 295 - 305 (1999).

[8] J. Cassaigne and J. Karhumäki: *Examples of undecidable problems for 2-generator matrix semigroups*. Theor. Comput. Sci. 204, No.1-2, 29-34 (1998).

[9] H. Cohen: *A Course in Computational Algebraic Number Theory*. Springer (1993).

[10] P. M. Cohn: ~~*Algebra, Vol 2*.~~ John Wiley and Sons (1977).

[11] V. Halava: *Decidable and Undecidable Problems in Matrix Theory*. TUCS Tech. Report 127 (1997).

[12] V. Halava and T. Harju: *Mortality in matrix semigroups*. Amer. Math. Monthly 108, 649 - 653 (2001).

[13] G. Hansel: *Une démonstration simple du théorème de Skolem-Mahler-Lech*. Theoret. Comput. Sci. 43, pp. 1–10 (1986).

[14] T. Harju and J. Karhumäki, *Morphisms*. In *Handbook of Formal Languages* (G. Rozenberg and A. Salomaa, eds.), vol. 1, Springer-Verlag, 1997.

[15] R. Kannan and R. J. Lipton: *Polynomial-Time Algorithm for the Orbit Problem*. Journal of the Association for Computing Machinery, Vol 33, No. 4, pp. 808–821 (1986).

[16] D.A. Klarner, J.-C. Birget and W. Satterfield: *On the undecidability of the freeness of the integer matrix semigroups*. Int. J. Algebra Comp. 1, 223-226 (1991).

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

[18] Z. Manna: *Mathematical theory of computation*. McGraw-Hill Book Comp, (1974).

[19] Y. Matiyasevich and G. Sénizergues: *Decision problems for semi-Thue systems with a few rules*. Theor. Comput. Sci. 330(1), 145-169, (2005).

[20] M. Mignotte: *A note on linear recursive sequences*. Journal of Australian Mathematical Society 20, pp- 242–244 (1975).

[21] M. Mignotte: *Some useful bounds*. In B. Buchberger, G. E. Collings, and R. Loos (eds.): *Computer Algebra*, pp. 259–263, Springer (1982).

[22] 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, pp. 63–76 (1984).

[23] C. Lech: *A note on recurring series*. Ark. Mat. 2, 417-421 (1953).

[24] M.S. Paterson: *Unsolvability in* 3×3 *matrices*. Studies in Appl. Math. **49**, 105-107 (1970).

[25] A. J. van der Poorten: *Linear Forms in Logarithms in the* p*-adic Case*. In A. Baker and D. W. Masser (eds.): *Transcendence Theory: Advances and Applications*, pp. 29–57. Academic Press (1977).

[26] P. Bell and I. Potapov: *On the Reachability of Invertible Diagonal Matrices*. Manuscript.

[27] J. B. Rosser and L. Schoenfeld: *Approximate formulas for some functions of prime numbers*, Illinois Journal of Mathematics 6:1, 64–94 (1962).

[28] A. Salomaa and M. Soittola: *Automata-theoretic aspects of formal power series*. Springer-Verlag, (1978).

[29] T. Skolem: *Ein Verfahren zur Behandlung gewisser exponentialer Gleichungen und diophantischer Gleichungen*. 8. Skand. Mat. Kongr., Stockhohn, 1934, 163-188. (1934).

[30] B. Sury: *Sum of the reciprocals of the binomial coefficients*. European Journal of Combinatorics 14, pp. 351–353 (1993).

[31] N. K. Vereshchagin: *The problem of the appearance of a zero in a linear recursive sequence* (Russian), Mat. Zametki 38, no. 2, pp. 177–189, 347 (1985). English translation *Occurrence of zero in a linear recursive sequence*. Math. Notes 38, nos 1–2, pp. 609–615, (1985).

Turku Centre *for* Computer Science

Lemminkäisenkatu 14 A, 20520 Turku, Finland | www.tucs.fi

University of Turku

- Department of Information Technology
- Department of Mathematics

Åbo Akademi University

- Department of Computer Science
- Institute for Advanced Management Systems Research

Turku School of Economics and Business Administration

- Institute of Information Systems Sciences

ISBN 952-12-1534-8
ISSN 1239-1891