The Semialgebraic Orbit Problem

Shaull Almagor
Department of Computer Science, Oxford University, UK
shaull.almagor@cs.ox.ac.uk

Joël Ouaknine
Max Planck Institute for Software Systems, Germany

Department of Computer Science, Oxford University, UK
joel@mpi-sws.org

James Worrell
Department of Computer Science, Oxford University, UK
jbw@cs.ox.ac.uk

Abstract

The *Semialgebraic Orbit Problem* is a fundamental reachability question that arises in the analysis of discrete-time linear dynamical systems such as automata, Markov chains, recurrence sequences, and linear while loops. An instance of the problem comprises a dimension $d \in \mathbb{N}$, a square matrix $A \in \mathbb{Q}^{d \times d}$, and semialgebraic source and target sets $S, T \subseteq \mathbb{R}^d$. The question is whether there exists $x \in S$ and $n \in \mathbb{N}$ such that $A^n x \in T$.

The main result of this paper is that the Semialgebraic Orbit Problem is decidable for dimension $d \leq 3$. Our decision procedure relies on separation bounds for algebraic numbers as well as a classical result of transcendental number theory—Baker's theorem on linear forms in logarithms of algebraic numbers. We moreover argue that our main result represents a natural limit to what can be decided (with respect to reachability) about the orbit of a single matrix. On the one hand, semialgebraic sets are arguably the largest general class of subsets of $\mathbb{R}^d$ for which membership is decidable. On the other hand, previous work has shown that in dimension $d = 4$, giving a decision procedure for the special case of the Orbit Problem with singleton source set S and polytope target set T would entail major breakthroughs in Diophantine approximation.

2012 ACM Subject Classification Computing methodologies → Algebraic algorithms; Theory of computation → Logic and verification

Keywords and phrases linear dynamical systems, Orbit Problem, first order theory of the reals

Digital Object Identifier 10.4230/LIPIcs.STACS.2019.6

Funding *Joël Ouaknine*: Supported by ERC grant AVS-ISS (648701), and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — Projektnummer 389792660 — TRR 248

James Worrell: Supported by EPSRC Fellowship EP/N008197/1

36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019).
Editors: Rolf Niedermeier and Christophe Paul; Article No. 6; pp. 6:1–6:20

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

This paper concerns decision problems of the following form: given $d \in \mathbb{N}$, a square matrix $A \in \mathbb{Q}^{d\times d}$, and respective *source* and *target* sets $S, T \subseteq \mathbb{R}^d$, does there exist $n \in \mathbb{N}$ and $x \in S$ such that $A^n x \in T$? One way to categorise such problems is according to the types of sets allowed for the source and target (e.g., polytopes or semialgebraic sets). We collectively refer to the various problems that arise in this way as *Orbit Problems*. Orbit Problems occur naturally in the reachability analysis of discrete-time linear dynamical systems, including Markov chains, automata, recurrence sequences, and linear loops in program analysis (see [5, 11, 9] and references therein).

In order to describe the main result of this paper in relation to existing work, we identify three successively more general types of Orbit Problems. In the *point-to-point* version both the source and target are singletons with rational coordinates; in the *Polytopic Orbit Problem* the source and target S and T are polytopes (i.e., sets defined by conjunctions of linear inequalities with rational coefficients); in the *Semialgebraic Orbit Problem* S and T are semialgebraic sets defined with rational parameters.

The question of the decidability of the point-to-point Orbit Problem was raised by Harrison in 1969 [10]. The problem remained open for ten years until it was finally resolved in a seminal paper of Kannan and Lipton [11], who in fact gave a polynomial-time decision procedure.

The Polytopic Orbit Problem is considerably more challenging than the point-to-point version, and its decidablity seems out of reach for now. Indeed the special case in which S is a singleton and T is a linear subspace of $\mathbb{R}^d$ of dimension $d-1$ is a well-known decision problem in its own right, called the *Skolem Problem*, whose decidability has been open for many decades [20]. In contrast to the point-to-point case the only positive decidability results for the Polytopic Orbit Problem are in the case of fixed dimension d. For the Skolem Problem, decidability is known for $d \leq 4$ [14, 22]. In case S and T are allowed to be arbitrary polytopes, decidability is known in case $d \leq 3$ [1] (see also [4]). While Kannan and Lipton's decision procedure in the point-to-point case mainly relied on algebraic number theory (e.g., separation bounds between algebraic numbers and prime factorisation of ideals in rings of algebraic integers), the decision procedures for the Skolem Problem and the Polytopic Orbit Problem additionally use results about transcendental numbers (specifically Baker's theorem about linear forms in logarithms of algebraic numbers). It was shown in [4] that the existence of a decision procedure for the Polytopic Orbit Problem in dimension $d = 4$ would entail computability of the Diophantine approximation types of a general class of transcendental numbers (a problem considered intractable at present). Not only does this suggest that the use of transcendental number theory is unavoidable in analysing the Polytopic Orbit Problem, it also indicates that further progress beyond the case $d = 3$ is contingent upon significant advances in the field of Diophantine approximation.

In this paper we remain in dimension $d = 3$ and consider a generalisation of previous work by allowing the source and target sets to be semialgebraic, that is, defined by Boolean combinations of polynomial equalities and inequalities. This allows us to handle three-dimensional source and target sets in much greater geometrical generality than polytopes. In applications to program analysis and dynamical systems, semialgebraic sets are indispensable in formulating sufficiently expressive models (e.g., to describe initial conditions and transition guards) and in model analysis (e.g., in synthesising invariants and barrier certificates and approximating sets of reachable states) [15, 12].

The Semialgebraic Orbit Problem could be reduced to the polytopic case in a fairly

straightforward fashion by increasing the dimension d according to the degree of the polynomials appearing in the semialgebraic constraints. However such a general approach is doomed to failure in view of the obstacles to obtaining decidability in the polytopic case beyond dimension 3 and instead we develop specific techniques for the semialgebraic case that are considerably more challenging than in the Polytopic Problem. As in previous work on the Skolem Problem and on the Polytopic Orbit Problem, Baker's Theorem plays a crucial role in the present development. The main difficulty in generalising from the polytopic case to the semialgebraic case lies in the delicate analytic arguments that are required to bring Baker's Theorem to bear. More precisely: (i) we need to resort to symbolic quantifier elimination (in lieu of explicit Fourier-Motzkin elimination, which had been used in the Polytopic Orbit Problem), since we are now dealing with non-linear constraints; (ii) we also need to perform spectral calculations symbolically, via the use of Vandermonde methods, instead of the explicit direct approach possible in our earlier work; and (iii) we replace triangulation of polytopes by cylindrical algebraic decomposition of semialgebraic sets into cells, which again necessitates a new symbolic treatment along with a substantially refined analysis based on Taylor approximation of the attendant functions.

In summary, this paper provides a decision procedure for the Orbit Problem in dimension $d = 3$ with semialgebraic source and target sets. The latter appear to be a natural limit to the positive decidability results that can be obtained for this problem, barring major new advances in Diophantine approximation.

At a technical level, our contributions are twofold: in Section 3 we start by analysing the case of the Orbit Problem in which S is a singleton and T a semialgebraic set. We then reduce this problem in Section 3.1 to solving certain systems of polynomial-exponential equalities and inequalities, and in Section 3.2 we show precisely how to solve such systems. The second technical contribution consists in handling the general case of the Semialgebraic Orbit Problem, in Section 4. There, we show how to circumvent problems that arise when quantifying over the set S, and arrive at a system that can ultimately be solved using the techniques and results developed in Section 3.2.

2 Mathematical Tools

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

2.1 Algebraic numbers

For $p \in \mathbb{Z}[x]$ a polynomial with integer coefficients, we denote by $\|p\|$ the bit length of its representation as a list of coefficients encoded in binary. Note that the *degree* of p, denoted $\deg(p)$ is at most $\|p\|$, and the *height* of p — i.e., the maximum of the absolute values of its coefficients, denoted $H(p)$ — is at most $2^{\|p\|}$.

We begin by summarising some basic facts about the field of algebraic numbers (denoted $\mathbb{A}$) and (efficient) arithmetic therein. The main references include [3, 6, 19]. 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, and are denoted $\deg(\alpha)$ and $H(\alpha)$. A standard representation[1] for algebraic numbers is to encode α as a tuple comprising its defining polynomial together with

[1] Note that this representation is not unique.

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 [13] asserts that for roots $\alpha \neq \beta$ of a polynomial $p \in \mathbb{Z}[x]$, we have

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

where $d = \deg(p)$ and $H = H(p)$. 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 $\|p\|$ [3, 6, 17]. Thus given an algebraic number α 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.

Note that Equation (1) 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 $\deg(\alpha) + \deg(\beta)$, and of height at most $H(\alpha)H(\beta)$. Given algebraic numbers α and β, one can compute $\alpha + \beta$, $\alpha\beta$, $1/\alpha$ (for $\alpha \neq 0$), $\overline{\alpha}$, 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 [3, 6].

2.2 First-order theory of the reals

Let $\vec{x} = x_1, \ldots, x_m$ be a list of m real-valued variables, and let $\sigma(\vec{x})$ be a Boolean combination of atomic predicates of the form $g(\vec{x}) \sim 0$, where each $g(\vec{x}) \in \mathbb{Z}[x]$ is a polynomial with integer coefficients over these variables, and $\sim \in \{>, =\}$. A *formula of the first-order theory of the reals* is of the form $Q_1 x_1 Q_2 x_2 \cdots Q_m x_m \sigma(\vec{x})$, where each Q_i is one of the quantifiers $\exists$ or $\forall$. Let us denote the above formula by τ , and write $\|\tau\|$ to denote the bit length of its syntactic representation. Tarski famously showed that the first-order theory of the reals is decidable [21]. His procedure, however, has non-elementary complexity. Many substantial improvements followed over the years, starting with Collins's technique of cylindrical algebraic decomposition [7], and culminating with the fine-grained analysis of Renegar [19]. In this paper, we will use the following theorems [18, 19].

▶ **Theorem 1** (Renegar [18]). *The problem of deciding whether a closed formula τ of the form above holds over the reals is in **2EXP**, and in **PSPACE** if τ has only existential quantifiers.*

▶ **Theorem 2** (Renegar [19]). *There is an algorithm that, given a formula $\tau(x_1, \ldots, x_m)$ where $x_1, \ldots, x_m$ are free variables, computes an equivalent quantifier-free formula in* disjunctive normal form *(DNF) $\Phi(x_1, \ldots, x_m) = \bigvee_I \bigwedge_J R_{I,J}(x_1, \ldots, x_m) \sim_{I,J} 0$ where $R_{I,J}$ is a polynomial* [2] *and $\sim_{I,J} \in \{>, =\}$. Moreover, the algorithm runs in time $2^{2^{\mathcal{O}(\|\tau\|)}}$, and in particular, $\|\Phi\| = 2^{2^{\mathcal{O}(\|\tau\|)}}$.*

A set $S \subseteq \mathbb{R}^d$ is *semialgebraic* if there exists a formula $\Phi(x_1, \ldots, x_d)$ in the first-order theory of the reals with free variables $x_1, \ldots, x_d$ such that $S = \{(c_1, \ldots, c_d) : \Phi(c_1, \ldots, c_d) \text{ is true}\}$.

[2] Technically, the indices should be I, J_I, but we omit the dependency of J on I for simplicity.

We remark that algebraic constants can also be incorporated as coefficients in the first-order theory of the reals (and in particular, in the definition of semialgebraic sets), as follows. Consider a polynomial $g(x_1, \ldots, x_m)$ with algebraic coefficients $c_1, \ldots, c_k$. We replace every c_j with a new, existentially-quantified variable y_j, and add to the sentence the predicates $p_{c_j}(y_j) = 0$ and $(y_j - (a + bi))^2 < r^2$, where (p_{c_j}, a, b, r) is the representation of c_j. Then, in any evaluation of this formula to True, it must hold that y_j is assigned value c_j.

3 Almost Self-Conjugate Systems of Inequalities

In this section we lay the groundwork for solving the Semialgebraic Orbit Problem. We do so by initially treating the case where the set S of initial points is a singleton.

3.1 Analysis of the Point-to-Semialgebraic Orbit Problem

The *point-to-semialgebraic Orbit Problem* is to decide, given a matrix $A \in \mathbb{Q}^{3\times 3}$, an initial point $s \in \mathbb{Q}^3$ and a semialgebraic target $T \subseteq \mathbb{R}^3$, whether there exists $n \in \mathbb{N}$ such that $A^n s \in T$.

By Theorem 2, we can compute a quantifier-free representation of T. That is, we can write $T = \{(x, y, z) : \bigvee_I \bigwedge_J R_{I,J}(x, y, z) \sim_{I,J} 0\}$ where $R_{I,J}$ are polynomials with integer coefficients, and $\sim_{I,J} \in \{>, =\}$. For the purpose of solving the point-to-semialgebraic Orbit Problem, we note that it is enough to consider each disjunct separately. Thus, we can assume $T = \{(x, y, z) : \bigwedge_J R_J(x, y, z) \sim_J 0\}$, and it remains to decide whether there exists $n \in \mathbb{N}$ such that $\bigwedge_J R_J(A^n s) \sim_J 0$.

Note that, as per Theorem 2, we have that $\|R_J\| = 2^{2^{\mathcal{O}(\|T\|)}}$. Moreover, the number of terms in the DNF formula above can itself be doubly-exponential in $\|T\|$. Complexity wise, this is the most expensive part of our algorithm.

Consider the eigenvalues of A. Since A is a 3×3 matrix, then either it has only real eigenvalues, or it has one real eigenvalue and two conjugate complex eigenvalues. In particular, if A has complex eigenvalues, then it is diagonalisable.

The case where A has only real eigenvalues is treated in Appendix A for the general case of the Semialgebraic Orbit Problem, and is considerably simpler.

Henceforth, we assume A has complex eigenvalues, so that $A = PDP^{-1}$ with $D = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \overline{\lambda} & 0 \\ 0 & 0 & \rho \end{pmatrix}$, where λ is a complex eigenvalue, $\rho \in \mathbb{R}$, and P an invertible matrix.

Observe that $A^n = PD^nP^{-1}$. By carefully analysing the structure of P, it is not hard to show that $A^n s = \begin{pmatrix} a_1\lambda^n + \overline{a_1}\overline{\lambda}^n + b_1\rho^n \\ a_2\lambda^n + \overline{a_2}\overline{\lambda}^n + b_2\rho^n \\ a_3\lambda^n + \overline{a_3}\overline{\lambda}^n + b_3\rho^n \end{pmatrix}$ where the a_i and b_i are algebraic and the b_i are also real (see Appendix E for a detailed analysis).

Thus, we want to decide whether there exists $n \in \mathbb{N}$ such that $R_J(a_1\lambda^n + \overline{a_1}\overline{\lambda}^n + b_1\rho^n, a_2\lambda^n + \overline{a_2}\overline{\lambda}^n + b_2\rho^n, a_3\lambda^n + \overline{a_3}\overline{\lambda}^n + b_3\rho^n) \sim_J 0$ for every J. Since R_J is a polynomial, then by aggregating coefficients we can write

$$R_J(A^n s) = \sum_{0 \le p_1, p_2, p_3 \le k} \alpha_{p_1,p_2,p_3} \lambda^{np_1} \overline{\lambda}^{np_2} \rho^{np_3} + \overline{\alpha_{p_1,p_2,p_3}} \overline{\lambda}^{np_1} \lambda^{np_2} \rho^{np_3} \tag{2}$$

for some $k \in \mathbb{N}$. Note that we treat the (real) coefficients of ρ as a sum of complex conjugate coefficients, but this can easily be achieved by writing e.g., $c\rho^{np} = \frac{c}{2}\rho^{np} + \frac{c}{2}\rho^{np}$.

We notice that for every J, the polynomial $R_J(A^n s)$, consists of conjugate summands. More precisely, $R_J(A^n s)$, when viewed as a polynomial in $\lambda^n, \overline{\lambda}^n$, and ρ^n, has the following property.

▷ Property 3 (Almost Self-Conjugate Polynomial). A complex polynomial $Q(z_1, z_2, z_3)$ over $\mathbb{C}^3$ is *almost self-conjugate* if

$$Q(z_1, z_2, z_3) = \sum_{0 \leq t_1, t_2, t_3 \leq \ell} \delta_{t_1,t_2,t_3} z_1^{t_1} z_2^{t_2} z_3^{t_3} + \overline{\delta_{t_1,t_2,t_3}} z_2^{t_1} z_1^{t_2} z_3^{t_3}.$$

That is, if $z_2 = \overline{z_1}$ and z_3 is a real variable, then the monomials in Q appear in conjugate pairs with conjugate coefficients.

We refer to the conjunction $\bigwedge_J R_J(A^n s) \sim_J 0$ as an *almost self-conjugate system*. It remains to show that we can decide whether there exists $n \in \mathbb{N}$ that solves the system.

3.2 Solving Almost Self-Conjugate Systems

Our starting point is now an almost self-conjugate system as described above. In the following, we will consider a single conjunct $R_J(A^n s) \sim_J 0$.

We start by normalising the expression $R_J(A^n s) \sim_J 0$ in the form of (2), as follows. Let $\Lambda = \max\left\{|\lambda^{p_1}\overline{\lambda}^{p_2}\rho^{p_3}| : \alpha_{p_1,p_2,p_3} \neq \emptyset\right\}$, we divide the expression in (2) by Λ^n, and get that $R_J(A^n s) \sim_J 0$ iff

$$\sum_{m=0}^{k} \beta_m \gamma^{nm} + \overline{\beta_m}\overline{\gamma}^{nm} + r(n) \sim_J 0 \tag{3}$$

where the β_m are algebraic coefficients, $\gamma = \frac{\lambda}{|\lambda|}$ satisfies $|\gamma| = 1$ and $r(n) = \sum_{l=1}^{k'} \chi_l \mu_l^n + \overline{\chi_l \mu_l}^n$ with χ_l being algebraic coefficients, and $|\mu_l| < 1$ for every $1 \leq l \leq k'$. Moreover, every μ_l is a quotient of two elements of the form $\lambda^{p_1}\overline{\lambda}^{p_2}\rho^{p_3}$, and thus, by Section 2.1, $\deg(\mu_l) = \|R_J\|^{\mathcal{O}(1)}$ and $H(\mu_l) = 2^{\|R_J\|^{\mathcal{O}(1)}}$. Note that for simplicity, we reuse the number k, although it may differ from k in (2). We refer to Equation (3) as the *normalised expression*.

In the following, we assume that at least one of the β_j is nonzero for $j \geq 1$. Indeed, otherwise we can recast our analysis on $r(n)$, which is of lower order.

We now split our analysis according to whether or not γ is a root of unity. That is, whether $\gamma^d = 1$ for some $d \in \mathbb{N}$.

3.2.1 The case where γ is a root of unity

Suppose that γ is a root of unity. Then, the set $\{\gamma^n : n \in \mathbb{N}\}$ is a finite set $\left\{\gamma^0, \ldots, \gamma^{d-1}\right\}$. Thus, by splitting the analysis of $A^n s$ according to n mod d, we can reduce the problem to d instances which involve only real numbers. In Appendix B we detail how to handle this case, and comment on its complexity.

3.2.2 The case where γ is not a root of unity

When γ is not a root of unity, the set $\{\gamma^n : n \in \mathbb{N}\}$ is dense in the unit circle. With this motivation in mind, we define, for a normalised expression, its *dominant function* $f : \mathbb{C} \to \mathbb{R}$ as $f(z) = \sum_{m=0}^{k} \beta_m z^m + \overline{\beta_m}\overline{z}^m$. Observe that (3) is now equivalent to $f(\gamma^n) + r(n) \sim_J 0$.

Our main technical tool in handling (3) is the following lemma.

▶ **Lemma 4.** *Consider a normalised expression as in* (3)*. Let* $\|\mathcal{I}\|$ *be its encoding length, and let* f *be its dominant function. Then there exists* $N \in \mathbb{N}$ *computable in polynomial time in* $\|\mathcal{I}\|$ *with* $N = 2^{\|\mathcal{I}\|^{\mathcal{O}(1)}}$ *such that for every* $n > N$ *it holds that*

1. $f(\gamma^n) \neq 0$,
2. $f(\gamma^n) > 0$ *iff* $f(\gamma^n) + r(n) > 0$,
3. $f(\gamma^n) < 0$ *iff* $f(\gamma^n) + r(n) < 0$.

In particular, the lemma implies that if $f(\gamma^n) + r(n) = 0$, then $n \leq N$. That is, if $\sim_J$ is "=", then there is a bound on n that solves the system.

▶ Remark 5. In the formulation of Lemma 4, we measure the complexity with respect to $\|\mathcal{I}\|$. However, recall that when the input is T, we actually have $\|\mathcal{I}\| = 2^{2^{\mathcal{O}(\|T\|)}}$. The analysis in Lemma 4 thus allows us to separate the blowup required for analysing the semialgebraic target from our algorithmic contribution. In particular, when the target has bounded description length, we can obtain better complexity bounds.

We prove Lemma 4 in the remainder of this section.

Since $\{\gamma^n : n \in \mathbb{N}\}$ is dense on the unit circle, our interest in f is also about the unit circle. By identifying $\mathbb{C}$ with $\mathbb{R}^2$, we can think of f as a function of two real variables. In this view, $f(x, y)$ is a polynomial with algebraic coefficients, and we can therefore compute a description of the algebraic set $Z_f = \{(x, y) : f(x, y) = 0 \wedge x^2 + y^2 = 1\}$. We start by showing that this set is finite. Define $g : (-\pi, \pi] \to \mathbb{R}$ by $g(x) = f(e^{ix})$. Explicitly, we have $g(x) = \sum_{m=0}^{k} 2|\beta_m| \cos(mx + \theta_m)$ where $\theta_m = \arg(\beta_m)$. Clearly there is a one-to-one correspondence between Z_f and the roots of g.

We present the following proposition, which will be reused later in the proof.

▶ **Proposition 6.** *For every* $x \in (-\pi, \pi]$ *there exists* $1 \leq i \leq 4k$ *such that* $g^{(i)}(x) \neq 0$*, where* $g^{(i)}$ *is the* i*-th derivative of* g*.*

Proof. Assume by way of contradiction that $g'(x) = \ldots = g^{4k}(x) = 0$. For every $1 \leq i \leq 4k$ we have that

$$g^{(i)}(x) = \begin{cases} \sum_{m=1}^{k} m^i 2|\beta_m| \cos(mx + \theta_m) & i \equiv_4 0 \\ \sum_{m=1}^{k} -m^i 2|\beta_m| \sin(mx + \theta_m) & i \equiv_4 1 \\ \sum_{m=1}^{k} -m^i 2|\beta_m| \cos(mx + \theta_m) & i \equiv_4 2 \\ \sum_{m=1}^{k} m^i 2|\beta_m| \sin(mx + \theta_m) & i \equiv_4 3 \end{cases}$$

(note that the summand that corresponds to $m = 0$ is dropped out in the derivative, as it is constant).

Splitting according $i \mod 4$, we rewrite the equations $g^{(i)}(x) = 0$ in matrix form as follows.[3]

$$\text{for } i \equiv_4 0: \quad \begin{pmatrix} 1^4 & 2^4 & \cdots & k^4 \\ 1^8 & 2^8 & \cdots & k^8 \\ \vdots & \vdots & \vdots & \vdots \\ 1^{4k} & 2^{4k} & \cdots & k^{4k} \end{pmatrix} \begin{pmatrix} 2|\beta_1| \cos(x + \theta_1) \\ 2|\beta_2| \cos(2x + \theta_2) \\ \vdots \\ 2|\beta_k| \cos(kx + \theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

[3] By splitting modulo 2, we could actually improve the bound in the proposition from $4k$ to $2k$, but this further complicates the proof.

$$\text{for } i \equiv_4 1: \quad \begin{pmatrix} -1^1 & -2^1 & \cdots & -k^1 \\ -1^5 & -2^5 & \cdots & -k^5 \\ \vdots & \vdots & \vdots & \vdots \\ -1^{4k-3} & -2^{4k-3} & \cdots & -k^{4k-3} \end{pmatrix} \begin{pmatrix} 2|\beta_1|\sin(x+\theta_1) \\ 2|\beta_2|\sin(2x+\theta_2) \\ \vdots \\ 2|\beta_k|\sin(kx+\theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{for } i \equiv_4 2: \quad \begin{pmatrix} -1^2 & -2^2 & \cdots & -k^2 \\ -1^6 & -2^6 & \cdots & -k^6 \\ \vdots & \vdots & \vdots & \vdots \\ -1^{4k-2} & -2^{4k-2} & \cdots & -k^{4k-2} \end{pmatrix} \begin{pmatrix} 2|\beta_1|\cos(x+\theta_1) \\ 2|\beta_2|\cos(2x+\theta_2) \\ \vdots \\ 2|\beta_k|\cos(kx+\theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{for } i \equiv_4 3: \quad \begin{pmatrix} 1^3 & 2^3 & \cdots & k^3 \\ 1^7 & 2^7 & \cdots & k^7 \\ \vdots & \vdots & \vdots & \vdots \\ 1^{4k-1} & 2^{4k-1} & \cdots & k^{4k-1} \end{pmatrix} \begin{pmatrix} 2|\beta_1|\sin(x+\theta_1) \\ 2|\beta_2|\sin(2x+\theta_2) \\ \vdots \\ 2|\beta_k|\sin(kx+\theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Observe that the matrices we obtain are minors of Vandermonde Matrices (up to their sign), and as such are non-singular [8]. It follows that

$$\begin{pmatrix} 2|\beta_1|\sin(x+\theta_1) \\ 2|\beta_2|\sin(2x+\theta_2) \\ \vdots \\ 2|\beta_k|\sin(kx+\theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 2|\beta_1|\cos(x+\theta_1) \\ 2|\beta_2|\cos(2x+\theta_2) \\ \vdots \\ 2|\beta_k|\cos(kx+\theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Recall that we assume at least one β_j is nonzero for some $1 \leq j \leq k$, so we have $\cos(jx+\theta_j) = \sin(jx+\theta_j) = 0$, which is clearly a contradiction. We thus conclude the proof. ◀

By Proposition 6, it follows that g is not constant, and therefore $f(x,y)$ is not constant on the curve $x^2+y^2=1$. By Bezout's Theorem, we have that Z_f is finite, and consists of at most $4k$ points. Moreover, f is a semialgebraic function (that is, its graph $\{(x,y,f(x,y)) : x,y \in \mathbb{R}\}$ is semialgebraic set in $\mathbb{R}^3$). Thus, the points in Z_f have semialgebraic coordinates, and we can compute them. By identifying $\mathbb{R}^2$ with $\mathbb{C}$, denote $Z_f = \{z_1, \ldots, z_{4k}\}$.

▶ Remark 7. Since the polynomial f has algebraic coefficients, it is not immediately clear how the degree and height of the points in Z_f relate to $\|f\|$. However, recall that the algebraic coefficients in f are polynomials in the entries of $A^n s$, which are, in turn, algebraic numbers of degree at most 3 whose description is polynomial in that of A and s.

Thus, we can define Z_f with a formula in the first-order theory of the reals with a fixed number of variables. Using results of Renegar [19], we show in Appendix F that the points in Z_f have semialgebraic coordinates with description length polynomial in $\|f\|$.

We now employ the following lemma from [16], which is itself a consequence of the Baker-Wüstholz Theorem [2].

▶ **Lemma 8** ([16]). *There exists $D \in \mathbb{N}$ such that for all algebraic numbers ζ, ξ of modulus 1, and for every $n \geq 2$, if $\zeta^n \neq \xi$, then $|\zeta^n - \xi| > \frac{1}{n^{(\|\zeta\|+\|\xi\|)^D}}$.*

Since γ is not a root of unity, it holds that $\gamma^{n_1} \neq \gamma^{n_2}$ for every $n_1 \neq n_2 \in \mathbb{N}$. Thus, there exists a computable $N_1 \in \mathbb{N}$ such that $\gamma^n \notin Z_f$ for every $n > N_1$. Moreover, by [5, Lemma D.1], we have that $N_1 = \|f\|^{\mathcal{O}(1)}$. By Lemma 8, there exists a constant $D \in \mathbb{N}$ such that for every $n \geq N_1$ and $1 \leq j \leq 4k$ we have that $|\gamma^n - z_j| > \frac{1}{n^{(\|f\|^D)}}$ (since $\|z_j\| + \|\gamma\| = \mathcal{O}(\|f\|)$). Intuitively, for $n > N_1$ we have that γ^n does not get close to any z_i "too quickly" as a function of n. In particular, for $n > N_1$ we have $f(\gamma^n) \neq 0$. It thus remains to show that for

large enough n, $r(n)$ does not affect the sign of $f(\gamma^n) + r(n)$. Intuitively, this is the case because $r(n)$ decreases exponentially, while $|f(\gamma^n)|$ is bounded from below by an inverse polynomial.

For every $z_j \in Z_f$, let $\varphi_j = \arg z_j$, so that $f(z) = 0$ iff $g(\varphi_j) = 0$. We assume w.l.o.g. that $\varphi_j \in (-\pi, \pi)$ for every $1 \leq j \leq 4k$. Indeed, if $\varphi_j = \pi$ for some j, then we can shift the domain of g slightly so that all zeros are in the interior.

For every $1 \leq j \leq 4k$, let T_j be the Taylor polynomial of g around φ_j such that the degree d_j of T_j is minimal and T_j is not identically 0. Thus, we have $T_j(x) = \frac{g^{(d_j)}(\varphi_j)}{d_j!}(x - \varphi_j)^{d_j}$. By Proposition 6 we have that $d_j \leq 4k$ for every j. In addition, the description of T_j is computable from that of $\|f\|$.

By Taylor's inequality, we have that for every $x \in [-\pi, \pi]$ it holds that $|g(x) - T_j(x)| \leq \frac{M_j |x - \varphi_j|^{d_j+1}}{(d_j+1)!}$ where $M_j = \max_{x \in [-\pi,\pi]} \left\{ g^{(d_j+1)}(x) \right\}$ (where g is extended naturally to the domain $[-\pi, \pi]$). By our description of $g^{(d_j+1)}(x)$, we see that M_j is bounded by $M = 4k \max_{1 \leq i \leq k} \{|\beta_i|\} k^{4k+1}$.

Let $\epsilon_1 > 0$ be such that the following conditions hold for every $1 \leq j \leq 4k$.

1. $\text{sign}(g'(x))$ does not change in $(\varphi_j, \varphi_j + \epsilon_1)$ nor in $(\varphi_j - \epsilon_1, \varphi_j)$.
2. $|g(x) - T_j(x)| \leq \frac{1}{2}|T_j(x)|$ for every $x \in (\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$.
3. $\text{sign}(g'(x)) = \text{sign}(T_j'(x))$ for every $x \in (\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$.

Note that we can assume $(\varphi_j - \epsilon_1, \varphi_j + \epsilon_1) \subseteq (-\pi, \pi)$, since by our assumption $\varphi_j \in (-\pi, \pi)$ for all $1 \leq j \leq 4k$.

An ϵ_1 as above exists due to the following properties (see Figure 1 for an illustration):

- There are only finitely many points where $g'(x) = 0$,
- $T_j(x)$ is of degree d_j, whereas $|g(x) - T_j(x)|$ is upper-bounded by a polynomial of degree $d_j + 1$, and
- $T_j'(x)$ is the Taylor polynomial of degree $d_j - 1$ of $g'(x)$ around φ_j, so by bounding the distance $|g'(x) - T_j'(x)|$ we can conclude the third requirement.

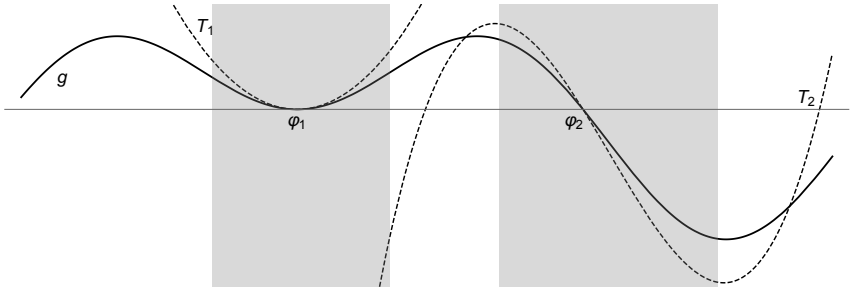

Figure 1 $g(x)$ and two Taylor polynomials: $T_1(x)$ around φ_1 and $T_2(x)$ around φ_2. The shaded regions show where requirements (1)–(3) hold, which determine ϵ_1. Observe that for T_1, the most restrictive requirement is $|g(x) - T_1(x)| \leq \frac{1}{2}T_1(x)$, whereas for T_2 the restriction is the requirement that $T_2(x)$ is monotone.

In order to establish Lemma 4, we must be able to effectively compute ϵ_1. We thus proceed with the following lemma.

▶ **Lemma 9.** *ϵ_1 can be computed in polynomial time in $\|f\|$, and $\frac{1}{\epsilon_1} = 2^{\|f\|^{\mathcal{O}(1)}}$.*

Proof. We compute $\delta_1, \delta_2, \delta_3$ that satisfy requirements 1,2, and 3, respectively. Then, taking $\epsilon_1 = \min\{\delta_1, \delta_2, \delta_3\}$ will conclude the proof.

Condition 1: We compute $\delta_1 > 0$ such that $\text{sign}(g'(x))$ does not change in $(\varphi_j - \delta_1, \varphi_j)$ nor in $(\varphi_j, \varphi_j + \delta_1)$. This is done as follows. Recall that $g(x) = f(e^{ix}) = \sum_{m=0}^{k} \beta_m e^{imx} + \overline{\beta_m} e^{imx}$. It is not hard to check that $g'(x) = \sum_{m=0}^{k} im\beta_m e^{imx} + \overline{im\beta_m} e^{imx}$. Let $\widehat{f}(z) : \mathbb{C} \to \mathbb{R}$ be the function $\widehat{f}(z) = \sum_{m=0}^{k} im\beta_m z + \overline{im\beta_m}\overline{z}$, then $g'(x) = \widehat{f}(e^{ix})$ and $\left\|\widehat{f}\right\| = \mathcal{O}(\|f\|)$.

Consider the algebraic set $F = \left\{z : |z| = 1 \wedge \widehat{f}(z) = 0\right\}$, then $\{x : g'(x) = 0\} = \{\arg(z) : z \in F\}$. By similar arguments as those by which we found the roots of f on the unit circle, namely by adapting Proposition 6 to $\widehat{f}$, we can conclude that F contains at most $4k$ points. Thus, it is enough to set δ_1 such that $\left(\bigcup_{j=1}^{4k}(\varphi_j - \delta_1, \varphi_j) \cup (\varphi_j, \varphi_j + \delta_1)\right) \cap F = \emptyset$.

By Equation (1), we have that for $z \neq z' \in F$ it holds that $|z - z'| > \frac{\sqrt{6}}{d^{\frac{d+1}{2}} \cdot H^{d-1}}$ where d and H are the degree and height of the roots of $\widehat{f}(z)$ (see Remark 7). Thus, $1/|z - z'|$ is $2^{\|f\|^{\mathcal{O}(1)}}$, and has a polynomial description. Since $|\arg(z) - \arg(z')| > |z - z'|$, we conclude that by setting $\delta_1 = \min\{|z - z'| : z \neq z' \in F\}/3$, it holds that $\frac{1}{\delta_1}$ has a polynomial description in $\|f\|$, and δ_1 satisfies the required condition.

Condition 2: Next, we compute $\delta_2 > 0$ such that $|g(x) - T_j(x)| \leq \frac{1}{2}|T_j(x)|$ for every $x \in (\varphi_j - \delta_2, \varphi_j + \delta_2)$. Recall that $T_j(x) = \frac{g^{(d_j)}(\varphi_j)}{d_j!}(x - \varphi_j)^{d_j}$. Note that this case is more challenging than Condition 1, as unlike $g(x) = f(e^{ix})$, the polynomial $T_j(x)$ has potentially transcendental coefficients (namely φ_j). For clarity, we omit the index j in the following. Thus, we write T, d, φ instead of T_j, d_j, φ_j, etc.

In order to ignore the absolute value, assume $T(x) \geq g(x) > 0$ in an interval $(\varphi, \varphi + \xi)$ for some $\xi > 0$ (the other cases are treated similarly). Then, the inequality above becomes $g(x) - \frac{1}{2}T(x) \geq 0$. Since the degree of T is d, then by the definition of T, the first $d - 1$ derivatives of g in φ vanish. Define $h(x) = g(x) - \frac{1}{2}T(x)$, then we have $h(\varphi) = 0$, $h'(\varphi) = 0, \ldots, h^{(d-1)}(\varphi) = 0$ and $h^{(d)}(\varphi) = g^{(d)}(\varphi) - \frac{1}{2}g^{(d)}(\varphi) = \frac{1}{2}g^{(d)}(\varphi)$. By our assumption, $T(x) \geq \frac{1}{2}T(x)$ for $x \in (\varphi, \varphi + \xi)$, so $h^{(d)}(\varphi) > 0$. In addition, recall that $|h^{(d+1)}(x)| = |g^{(d+1)}(x)| \leq M$ for every $x \in [-\pi, \pi]$. Thus, by writing the d-th Taylor expansion of $h(x)$ around φ, we have that $h(x) = \frac{h^{(d)}(\varphi)}{d!}(x - \varphi)^d + E(x)$ where $|E(x)| \leq \frac{M}{(d+1)!}(x - \varphi)^{d+1}$. We now have that

$$h(x) \geq \frac{1}{2}\frac{g^{(d)}(\varphi)}{d!}(x - \varphi)^d - \frac{M}{(d+1)!}(x - \varphi)^{d+1}.$$

Taking $x \in (\varphi, \varphi + \frac{g^{(d)}(\varphi)(d+1)}{2M})$, it is easy to check that $h(x) \geq 0$. We can now set $\delta_2 = \frac{g^{(d)}(\varphi)(d+1)}{2M}$, which satisfies the required condition.

Condition 3: Finally, we compute $\delta_3 > 0$ such that $\text{sign}(g'(x)) = \text{sign}(T_j'(x))$ for every $x \in (\varphi_j - \delta_3, \varphi_j + \delta_3)$. Observe that $T_j'(x)$ is the $d_j - 1$-th Taylor polynomial of $g'(x)$ around φ^j. Thus, by following the reasoning used to find δ_2, we can find δ_3 such that $|g'(x) - T_j'(x)| \leq \frac{1}{2}|T_j'(x)|$ for every $x \in (\varphi - \delta_3, \varphi + \delta_3)$, and in particular it holds that $\text{sign}(g'(x)) = \text{sign}(T_j'(x))$ for every $x \in (\varphi_j - \delta_3, \varphi_j + \delta_3)$.

As mentioned above, by setting $\epsilon_1 = \min\{\delta_1, \delta_2, \delta_3\}$, we conclude the proof. ◀

Conditions 1,2, and 3 above imply that within the intervals $(\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$ we have that $|g(x)| \geq \frac{1}{2}|T_j(x)|$, that $g(x)$ and $T_j(x)$ have the same sign, and that they are both decreasing/increasing together.

We now claim that there exists a polynomial $p(n)$ and a number $N_2 \in \mathbb{N}$ such that for every $n > N_2$ it holds that $|g(\arg(\gamma^n))| > \frac{1}{p(n)}$. In order to compute $p(n)$, we compute

separate polynomials for the domain $\bigcup_{j=1}^{4k}(\varphi_j-\epsilon_1,\varphi_j+\epsilon_1)$ and for its complement. Then, taking their minimum and bounding it from below by another polynomial yields $p(n)$.

We start by considering the case where $\arg(\gamma^n)\in\bigcup_{j=1}^{4k}(\varphi_j-\epsilon_1,\varphi_j+\epsilon_1)$. Recall that since γ is not a root of unity, then for every $n>N_1$ it holds that $\gamma^n\notin Z_f=\{z_1,\ldots,z_{4k}\}$. Then, by Lemma 8, for every $1\leq j\leq 4k$ and every $n\geq N_2=\max\{N_1,2\}$ we have $|\gamma^n-z_j|>\frac{1}{n^{(\|f\|^D)}}$. In addition, $|\gamma^n-z_j|\leq|\arg(\gamma^n)-\varphi_j|$ (since the LHS is the Euclidean distance and the RHS is the spherical distance). Therefore, $|\arg(\gamma^n)-\varphi_j|>\frac{1}{n^{(\|f\|^D)}}$, so either $\arg(\gamma^n)>\varphi_j+\frac{1}{n^{(\|f\|^D)}}$ or $\arg(\gamma^n)<\varphi_j-\frac{1}{n^{(\|f\|^D)}}$. Next, we have that if $\arg(\gamma^n)\in(\varphi_j-\epsilon_1,\varphi_j+\epsilon_1)$ for some $1\leq j\leq 4k$, then $|g(\arg(\gamma^n))|\geq\frac{1}{2}|T_j(\arg(\gamma^n))|\geq\frac{1}{2}\min\left\{|T_j(\varphi_j+\frac{1}{n^{(\|f\|^D)}})|,|T_j(\varphi_j-\frac{1}{n^{(\|f\|^D)}})|\right\}$, where the last inequality follows from condition 3 above, which implies that T_j is monotone with the same tendency as g.

Observe that $T_j(\varphi_j-\frac{1}{n^{(\|f\|^D)}})=\frac{g^{(d_j)}(\varphi)}{d_j!}\frac{1}{n^{(\|f\|^D)}}$ and that similarly $T_j(\varphi_j+\frac{1}{n^{(\|f\|^D)}})=-\frac{g^{(d_j)}(\varphi)}{d_j!}\frac{1}{n^{(\|f\|^D)}}$ are both inverse polynomials (in n). Thus, $|g(\arg(\gamma^n))|$ is bounded from below by an inverse polynomial. Moreover, these polynomials can be easily computed in time polynomial in $\|f\|$.

Finally, we note that for $x\notin\bigcup_{j=1}^{4k}(\varphi_j-\epsilon_1,\varphi_j+\epsilon_1)$ we can compute in polynomial time a bound $B>0$ such that $|g(x)|>B$. Indeed, $B=\min\{|g(x)|:x\in[-\pi,\pi]\setminus\bigcup_{j=1}^{4k}(\varphi_j-\epsilon_1,\varphi_j+\epsilon_1)\}$ (where $g(-\pi)$ is defined naturally by extending the domain), and we have that $|B|>0$ since we assumed non of the φ_j are exactly at π (in which case we would have had $g(-\pi)=0$). In particular, we can combine the two domains and compute a polynomial p as required. We remark that we can compute $\|B\|$ in polynomial time, since it is either at least $\frac{1}{2}|T_j(\varphi_j\pm\epsilon_1)|$ for some $1\leq j\leq 4k$ (and by Lemma 9, $\|\epsilon_1\|$ can be computed in polynomial time), or it is the value of one of the extrema of g, and the latter can be computed by finding the extrema of the (algebraic) function f on the unit circle.

To recap, for every $n>N_2$ it holds that $|g(\arg(\gamma^n))|>\frac{1}{p(n)}$ for a non-negative polynomial p, and both N_2 and p can be computed in polynomial time in the description of the input.

Next, we wish to find $N_3\in\mathbb{N}$ such that for every $n>N_3$ it holds that $r(n)<\frac{1}{p(n)}$. Recall that $r(n)=\sum_{l=1}^{k'}\chi_l\mu_l^n+\overline{\chi_l\mu_l}^n$ where for every $1\leq l\leq k'$ we have that μ_l is algebraic with $\deg(\mu_l)=\|f\|^{\mathcal{O}(1)}$ and $H(\mu_l)=2^{\|f\|^{\mathcal{O}(1)}}$. Observe that $1-|\mu_l|$ is also an algebraic number. Indeed, $1-|\mu_l|=1-\sqrt{\mu_l\overline{\mu_l}}$. Moreover, we get that $\deg(1-|\mu_l|)\leq\deg(\mu_l)^4$, as it is the root of a polynomial of degree at most $\deg(\mu_l)^4$, and that $H(1-|\mu_l|)$ is polynomial in $H(\mu_l)$. Since $|\mu_l|<1$, by applying Equation (1), we get $1-|\mu_l|=|1-|\mu_l||>\frac{\sqrt{6}}{d^{(d+1)/2}H(\mu_l)^{d-1}}$ where $d=\deg(\mu_l)^{\mathcal{O}(1)}$ and $H(\mu_l)=2^{\|I\|^{\mathcal{O}(1)}}$. It follows that we can compute $\delta\in(0,1)$ with $\frac{1}{\delta}=2^{\|I\|^{\mathcal{O}(1)}}$ such that $1-|\mu_l|>\delta$, and hence $|\mu|^n<1-\delta$. Thus,

$$|r(n)|\leq\sum_{l=1}^{k'}2|\chi_l||\mu_l|^{mn}\leq\sum_{l=1}^{k'}2|\chi_l|(1-\delta)^{mn}\leq 2k'\max_{1\leq l\leq k'}|\chi_l|(1-\delta)^n$$

We can now compute $\epsilon\in(0,1)$ and $N_3\in\mathbb{N}$ such that:

1. $\frac{1}{\epsilon}=2^{\|I\|^{\mathcal{O}(1)}}$
2. $N_3=2^{\|I\|^{\mathcal{O}(1)}}$
3. For every $n>N_3$ it holds that $|r(n)|<(1-\epsilon)^n$

Finally, by taking $N_4\in\mathbb{N}$ such that $(1-\epsilon)^n<\frac{1}{p(n)}$ (which satisfies $N_4=2^{\|I\|^{\mathcal{O}(1)}}$) for all $n>N_4$, we can now conclude that for every $n>\max\{N_2,N_3,N_4\}$, the following hold.

1. $f(\gamma^n)=g(\arg(\gamma^n))\neq 0$.

2. If $f(\gamma^n) > 0$, then $g(\arg(\gamma^n)) > 0$, so $g(\arg(\gamma^n)) > \frac{1}{p(n)}$. Since $|r(n)| < \frac{1}{p(n)}$, it follows that $f(\gamma^n)+r(n) = g(\arg(\gamma^n))+r(n) > \frac{1}{p(n)} - |r(n)| > 0$. Conversely, if $f(\gamma^n)+r(n) > 0$, then $g(\arg(\gamma^n)) + r(n) > 0$, but since $|g(\arg(\gamma^n))| > \frac{1}{p(n)}$ and $|r(n)| < \frac{1}{p(n)}$, then it must hold that $g(\arg(\gamma^n)) > 0$, so $f(\gamma^n) > 0$.
3. If $f(\gamma^n) < 0$, then $g(\arg(\gamma^n)) < 0$, so $g(\arg(\gamma^n)) < -\frac{1}{p(n)}$. Since $|r(n)| < \frac{1}{p(n)}$, it follows that $f(\gamma^n)+r(n) = g(\arg(\gamma^n))+r(n) < -\frac{1}{p(n)}+|r(n)| < 0$. Conversely, if $f(\gamma^n)+r(n) < 0$, then $g(\arg(\gamma^n)) + r(n) < 0$, but since $|g(\arg(\gamma^n))| > \frac{1}{p(n)}$ and $|r(n)| < \frac{1}{p(n)}$, then it must hold that $g(\arg(\gamma^n)) < 0$, so $f(\gamma^n) < 0$.

This concludes the proof of Lemma 4. ◀

We are now ready to use Lemma 4 in order to solve the systems.

▶ **Theorem 10.** *The problem of deciding whether an almost self-conjugate system has a solution is decidable.*

Proof. Consider an almost self-conjugate system of the form $\bigwedge_J R_J(A^n s) \sim_J 0$. For each expression $R_J(A^n s) \sim_J 0$, let f be the corresponding dominant function, as per Lemma 4, and compute its respective bound N. If $\sim_J$ is "=", then by Lemma 4, if the equation is satisfiable for $n \in \mathbb{N}$, then $n < N$.

If all the $\sim_J$ are ">", then for each such inequality compute $\{z : f(z) > 0\}$, which is a semialgebraic set. If the intersection of these sets is empty, then if n is a solution for the system, it must hold that $n < N$. If the intersection is non-empty, then it is an open set. Since γ is not a root of unity, then $\{\gamma^n : n \in \mathbb{N}\}$ is dense in the unit circle. Thus, there exists $n > N$ such that γ^n is in the above intersection, so the system has a solution. Checking the emptiness of the intersection can be done using Theorem 1.

Thus, it remains to check whether there exists a solution $n < N$, which is clearly decidable. ◀

Observe that from Theorem 10, combined with Section 3.1, we can conclude the decidability of the point-to-semialgebraic Orbit Problem. However, as it turns out, we can reuse Theorem 10 to obtain a much stronger result, namely the decidability of the Semialgebraic Orbit Problem.

4 The Semialgebraic Orbit Problem

In [1], we proved that the following problem is decidable: given two polytopes $S, T \subseteq \mathbb{R}^3$ and a matrix $A \in \mathbb{Q}^{3\times 3}$, does there exist $n \in \mathbb{N}$ such that $A^n S \cap T \neq \emptyset$. We now show that the techniques developed here can be used as an alternative solution for this problem, and in fact solve a much stronger variant, where S and T are replaced by semialgebraic sets. That is, given two semialgebraic sets $S, T \subseteq \mathbb{R}^3$ and a matrix $A \in \mathbb{Q}^{3\times 3}$, does there exist $n \in \mathbb{N}$ such that $A^n S \cap T \neq \emptyset$.

▶ **Theorem 11.** *The Semialgebraic Orbit Problem is decidable.*

Proof. Consider semialgebraic sets $S, T \subseteq \mathbb{R}^3$ and a matrix $A \in \mathbb{Q}^{3\times 3}$, as described above. Recall that we can write $S = \{\vec{x} : \bigvee_I \bigwedge_J R_{I,J}(\vec{x}) \sim_{I,J} 0\}$ and similarly for T. Since we want to decide whether some point in S hits T, we can consider each disjunct in the description of S separately. Thus, we henceforth assume $S = \{\vec{x} : \bigwedge_J R_J(\vec{x}) \sim_J 0\}$.

We now turn to characterise the set $A^n S$ for every $n \in \mathbb{N}$. For this purpose, we assume A is invertible. The case where A is not invertible can be reduced to analysis in a lower

dimension, and is handled in Appendix D. For every $n \in \mathbb{N}$, we now have

$$A^n S = \left\{ A^n \vec{x} : \bigwedge_J R_J(\vec{x}) \sim_J 0 \right\} = \left\{ \vec{x} : \bigwedge_J R_J((A^{-1})^n \vec{x}) \sim_J 0 \right\}.$$

We further assume that A has a complex eigenvalue. As in Section 3, the case where all eigenvalues are real is simpler (even if A is not diagonalisable), and is handled in Appendix A. We can now write $A = PDP^{-1}$ with $D = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \overline{\lambda} & 0 \\ 0 & 0 & \rho \end{pmatrix}$, where λ is a complex eigenvalue, $\rho \in \mathbb{R}$, and P an invertible matrix. We thus have $A^{-1} = PD^{-1}P^{-1}$ where $D^{-1} = \begin{pmatrix} \frac{\overline{\lambda}}{|\lambda|^2} & 0 & 0 \\ 0 & \frac{\lambda}{|\lambda|^2} & 0 \\ 0 & 0 & \rho^{-1} \end{pmatrix}$. We denote $\zeta = \frac{\overline{\lambda}}{|\lambda|^2}$ and $\eta = \rho^{-1}$, so $D^{-1} = \begin{pmatrix} \zeta & 0 & 0 \\ 0 & \overline{\zeta} & 0 \\ 0 & 0 & \eta \end{pmatrix}$. As in Section 3, by analysing the structure of P and P^{-1}, we have that for $\vec{x} = (x_1, x_2, x_3)$, $(A^{-1})^n(\vec{x})_i = \sum_{j=1}^{3} (a_{i,j}\zeta^n + \overline{a_{i,j}}\overline{\zeta}^n + b_{i,j}\eta^n)x_j$ with $a_{i,j} \in \mathbb{A}$ and $b_{i,j} \in \mathbb{A} \cap \mathbb{R}$. That is, each coordinate $1 \leq i \leq 3$, is a linear combination of x_1, x_2, x_3 where the coefficients are of the form above. In particular, the coefficient of every x_j is an almost self-conjugate polynomial (see Appendix E for a complete analysis).

Consider a monomial of the form $x_1^{s_1}x_2^{s_2}x_3^{s_3}$ in $R_J(\vec{x})$. Replacing $\vec{x}$ with $(A^{-1})^n\vec{x}$, the monomial then becomes $Q(\zeta^n, \overline{\zeta}^n, \eta^n)x_1^{s_1}x_2^{s_2}x_3^{s_3}$, where $Q(z_1, z_2, z_3)$ is an almost self-conjugate polynomial. Indeed, this follows since the coordinates of $(A^{-1})^n\vec{x}$ above are almost self-conjugate, and products of almost self-conjugate polynomials remain almost self-conjugate.

Recall that the polynomials R_J in the description of S have integer (and in particular, real) coefficients. By lifting the discussion about monomials to R_J, we can write

$$R_J((A^{-1})^n(\vec{x})) = \sum_{0 \leq s_1, s_2, s_3 \leq k} Q^J_{s_1,s_2,s_3}(\zeta^n, \overline{\zeta}^n, \eta^n)x_1^{s_1}x_2^{s_2}x_3^{s_3}$$

where $k \in \mathbb{N}$ and the coefficients $Q^J_{s_1,s_2,s_3}$ are almost self-conjugate.

Observe that now, there exists $n \in \mathbb{N}$ such that $A^n S \cap T \neq \emptyset$ iff there exists $n \in \mathbb{N}$ and $\vec{x} \in \mathbb{R}^3$ such that $\vec{x} \in T$ and

$$\bigwedge_J \sum_{0 \leq s_1, s_2, s_3 \leq k} Q^J_{s_1,s_2,s_3}(\zeta^n, \overline{\zeta}^n, \eta^n)x_1^{s_1}x_2^{s_2}x_3^{s_3} \sim_J 0. \tag{4}$$

Intuitively, we now want to eliminate the quantifiers on $\vec{x}$ in the expression above. However, we cannot readily do so, as the expression is also quantified by $n \in \mathbb{N}$. Nonetheless, in the following we manage to circumvent this problem by increasing the dimension of the problem.

Let K be the number of polynomials $Q^J_{s_1,s_2,s_3}$ that appear in the conjunction (4) above, indexed by J, s_1, s_2, s_3. Consider the set

$$U = \left\{ (y_1, \ldots, y_K) \in \mathbb{R}^K : \begin{array}{l} \exists \vec{x} \in \mathbb{R}^3,\ x \in T \wedge \\ \bigwedge_J \sum_{0 \leq s_1, s_2, s_3 \leq k} y^J_{s_1,s_2,s_3} x_1^{s_1}x_2^{s_2}x_3^{s_3} \sim_J 0 \end{array} \right\}$$

That is, U is obtained by replacing each polynomial $Q^J_{s_1,s_2,s_3}$ with a "placeholder" real variable $y^J_{s_1,s_2,s_3}$. U is clearly a semialgebraic set, so by Theorem 2, we can eliminate the quantifier on $\vec{x}$, and write

$$U = \left\{ (y_1, \ldots, y_K) \in \mathbb{R}^K : \bigwedge_J S_J(y_1, \ldots, y_K) \sim_J 0 \right\}$$

where S_J are polynomials with integer coefficients. It is now the case that there exists $n \in \mathbb{N}$ such that $A^n S \cap T \neq \emptyset$ iff there exists $n \in \mathbb{N}$ such that $(Q_1(\zeta^n, \overline{\zeta}^n, \eta^n), \ldots, Q_K(\zeta^n, \overline{\zeta}^n, \eta^n)) \in U$. That is, we need to decide whether there exists $n \in \mathbb{N}$ such that $S_J(Q_1(\zeta^n, \overline{\zeta}^n, \eta^n), \ldots, Q_K(\zeta^n, \overline{\zeta}^n, \eta^n)) \sim_J 0$ for every J.

It is easy to see that since the polynomials Q_i are almost self-conjugate, then so is $S_J(Q_1(\zeta^n, \overline{\zeta}^n, \eta^n), \ldots, Q_K(\zeta^n, \overline{\zeta}^n, \eta^n))$, (when viewed as a polynomial in $\zeta^n, \overline{\zeta}^n, \eta^n$).

Thus, the conjunction

$$\bigwedge_J S_J(Q_1(\zeta^n, \overline{\zeta}^n, \eta^n), \ldots, Q_K(\zeta^n, \overline{\zeta}^n, \eta^n))$$

is an almost self-conjugate system, and by Theorem 10, it is decidable whether it has a solution. This concludes the proof. ◀

5 Discussion

This paper establishes the decidability of the Semialgebraic Orbit Problem in dimension at most three. The class of semialgebraic sets is arguably the largest natural class for which membership is decidable. Thus, our results reach the limit of what can be decided about the orbit of a single matrix. Moreover, our techniques shed light on the decidability (or hardness) of orbit problems in higher dimensions: the techniques we develop for analysing orbits can be applied to any matrix (in any dimension) whose eigenvalues have arguments that are pairwise linearly dependent over $\mathbb{Q}$ (i.e., the arguments of all the eigenvalues are rational multiples of some angle θ). Indeed, it is easy to see that the orbits generated by such matrices can be reduced to solving almost self-conjugate systems (see Section 3). This can be put in contrast to known hardness results [4] in dimension $d \geq 4$, which require a single pair of eigenvalues whose arguments do not satisfy the above property. Thus, we significantly sharpen the border of known decidability, and allow future research to focus on hard instances.

Technically, our contribution uncovers two interesting tools. First, the identification of almost self-conjugate polynomials, and their amenability to analysis (Section 3), and second, the ability to abstract away integral exponents in order to perform quantifier elimination, by increasing the dimension (Section 4). The former arises naturally in the context of matrix exponentiation, while the latter is an obstacle that is often encountered when quantifying over semialgebraic sets in the presence of a discrete operator (e.g., matrix exponentiation). In the future, we plan to further investigate the applications of these directions.

References

1. S. Almagor, J. Ouaknine, and J. Worrell. The polytope-collision problem. In *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, pages 24:1–24:14, 2017.
2. A. Baker and G. Wüstholz. Logarithmic forms and group varieties. *J. reine angew. Math*, 442(19-62):3, 1993.
3. S. Basu, R. Pollack, and M-F. Roy. *Algorithms in real algebraic geometry*, volume 20033. Springer, 2005.
4. V. Chonev, J. Ouaknine, and J. Worrell. The polyhedron-hitting problem. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 940–956. SIAM, 2015.
5. V. Chonev, J. Ouaknine, and J. Worrell. On the complexity of the orbit problem. *J. ACM*, 63(3):23:1–23:18, 2016.
6. H. Cohen. *A course in computational algebraic number theory*, volume 138. Springer Science & Business Media, 2013.
7. G. E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decompostion. In *Automata Theory and Formal Languages 2nd GI Conference Kaiserslautern, May 20–23, 1975*, pages 134–183. Springer, 1975.
8. F.R. Gantmacher. *The Theory of Matrices*. Number v. 2 in The Theory of Matrices. Chelsea Publishing Company, 1959.
9. 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.
10. M. A Harrison. Lectures on linear sequential machines. Technical report, DTIC Document, 1969.
11. R. Kannan and R. J. Lipton. Polynomial-time algorithm for the orbit problem. *Journal of the ACM (JACM)*, 33(4):808–821, 1986.
12. G. Lafferriere, G. J. Pappas, and S. Yovine. Symbolic reachability computation for families of linear vector fields. *J. Symb. Comput.*, 32(3):231–253, 2001.
13. M. Mignotte. Some useful bounds. In *Computer algebra*, pages 259–263. Springer, 1983.
14. M. Mignotte, T. Shorey, and R. Tijdeman. The distance between terms of an algebraic recurrence sequence. *J. für die reine und angewandte Math.*, 349, 1984.
15. M. Müller-Olm and H. Seidl. Computing polynomial program invariants. *Inf. Process. Lett.*, 91(5):233–244, 2004.
16. J. Ouaknine and J. Worrell. Ultimate positivity is decidable for simple linear recurrence sequences. In *International Colloquium on Automata, Languages, and Programming*, pages 330–341. Springer, 2014.
17. V. Y. Pan. Optimal and nearly optimal algorithms for approximating polynomial zeros. *Computers & Mathematics with Applications*, 31(12):97–138, 1996.
18. J. Renegar. A faster PSPACE algorithm for deciding the existential theory of the reals. In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 291–295, 1988.
19. 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. *Journal of symbolic computation*, 13(3):255–299, 1992.
20. T. Tao. *Structure and randomness: pages from year one of a mathematical blog*. American Mathematical Soc., 2008.
21. A. Tarski. A decision method for elementary algebra and geometry. 1951.
22. N. K. Vereshchagin. Occurrence of zero in a linear recursive sequence. *Mathematical notes of the Academy of Sciences of the USSR*, 38(2):609–615, 1985.

A The case of only real eigenvalues

In this section we consider the Semialgebraic Orbit Problem in the case where the matrix A has only real eigenvalues, denoted ρ_1, ρ_2, ρ_3. In this case, by converting A to Jordan normal form, there exists an invertible matrix $B \in (\mathbb{A} \cap \mathbb{R})^{3\times 3}$ such that one of the following holds:

1. $A = B^{-1}\begin{pmatrix}\rho_1 & 0 & 0\\ 0 & \rho_2 & 0\\ 0 & 0 & \rho_3\end{pmatrix}B$, in which case $A^n = B^{-1}\begin{pmatrix}\rho_1^n & 0 & 0\\ 0 & \rho_2^n & 0\\ 0 & 0 & \rho_3^n\end{pmatrix}B$.
2. $A = B^{-1}\begin{pmatrix}\rho_1 & 1 & 0\\ 0 & \rho_2 & 0\\ 0 & 0 & \rho_3\end{pmatrix}B$ with $\rho_1 = \rho_2$, in which case $A^n = B^{-1}\begin{pmatrix}\rho_1^n & n\rho_1^{n-1} & 0\\ 0 & \rho_1^n & 0\\ 0 & 0 & \rho_3^n\end{pmatrix}B$.
3. $A = B^{-1}\begin{pmatrix}\rho_1 & 1 & 0\\ 0 & \rho_2 & 1\\ 0 & 0 & \rho_3\end{pmatrix}B$ with $\rho_1 = \rho_2 = \rho_3$, in which case $A^n = B^{-1}$ $\begin{pmatrix}\rho_1^n & n\rho_1^{n-1} & \frac{1}{2}n(n-1)\rho_1^{n-2}\\ 0 & \rho_1^n & n\rho_1^{n-1}\\ 0 & 0 & \rho_1^n\end{pmatrix}B$.

In any of the forms above, we can write

$$A^n s = \begin{pmatrix}A_1(n)\rho_1^n + B_1(n)\rho_2^n + C_1(n)\rho_3^n\\ A_2(n)\rho_1^n + B_2(n)\rho_2^n + C_2(n)\rho_3^n\\ A_3(n)\rho_1^n + B_3(n)\rho_2^n + C_3(n)\rho_3^n\end{pmatrix}$$

where the A_i, B_i, and C_i are polynomials whose degree is less than the multiplicity of their corresponding eigenvalue.

In Sections 3 and 4, we reduce the problem to finding a solution to an almost self-conjugate system. In the case of real eigenvalues, the notion of almost self-conjugate is meaningless, as there are no complex numbers involved. Thus, following the analysis thereof, and plugging the entries of $A^n s$, we reduce the problem to solving a system of expressions of the form $\bigwedge_J R_J(A^n s) \sim_J 0$, where

$$R_J(A^n s) = \sum_{0\le p_1,p_2,p_3\le k} \alpha^J_{p_1,p_2,p_3}(n)\rho_1^{p_1 n}\rho_2^{p_2 n}\rho_3^{p_3 n} \tag{5}$$

for some $k \in \mathbb{N}$, and $\alpha^J_{p_1,p_2,p_3}(n)$ are polynomials.

Assuming $\rho_1, \rho_2, \rho_3 > 0$ (otherwise we can split according to odd and even n), for each such expression we can compute a bound $N \in \mathbb{N}$ based on the rate of growth of the summands, such that either for every $n > N$ the equation holds, or for every $n > N$ it does not hold.

B The case where γ is a root of unity

We assume that $\gamma = \frac{\lambda}{|\lambda|}$ is a root of unity. That is, there exists $d \in \mathbb{N}$ such that $\gamma^d = 1$, so we have that $\{\gamma^n : n \in \mathbb{N}\} = \{\gamma^0, \ldots, \gamma^{d-1}\}$.

Let $n \in \mathbb{N}$ and write $m = (n \bmod d)$. We can now write

$$A^n s = \begin{pmatrix}a_1|\lambda|^n\gamma^m + \overline{a_1}|\lambda|^n\overline{\gamma}^m + b_1\rho^n\\ a_2|\lambda|^n\gamma^m + \overline{a_2}|\lambda|^n\overline{\gamma}^m + b_2\rho^n\\ a_3|\lambda|^n\gamma^m + \overline{a_3}|\lambda|^n\overline{\gamma}^m + b_3\rho^n\end{pmatrix} = \begin{pmatrix}2\mathrm{Re}(a_1\gamma^m)|\lambda|^n + b_1\rho^n\\ 2\mathrm{Re}(a_2\gamma^m)|\lambda|^n + b_2\rho^n\\ 2\mathrm{Re}(a_3\gamma^m)|\lambda|^n + b_3\rho^n\end{pmatrix}$$

Observe that there exists $n \in \mathbb{N}$ such that $A^n s \in T$ iff there exists $0 \leq m \leq d-1$ and $r \in \mathbb{N} \cup \{0\}$ such that $A^{rd+m} s \in T$. We can thus split our analysis according to $m \in \{0, \ldots, d-1\}$. For every such m, we need to decide whether there exists $r \in \mathbb{N} \cup \{0\}$ such that $\begin{pmatrix} 2\mathrm{Re}(a_1\gamma^m)|\lambda|^m(|\lambda|^d)^r + b_1\rho^m(\rho^d)^r \\ 2\mathrm{Re}(a_2\gamma^m)|\lambda|^m(|\lambda|^d)^r + b_2\rho^m(\rho^d)^r \\ 2\mathrm{Re}(a_3\gamma^m)|\lambda|^m(|\lambda|^d)^r + b_3\rho^m(\rho^d)^r \end{pmatrix}$ Note that γ^m, $|\lambda|^m$ and ρ^m are constants. Therefore, these expressions contain only realalgebraic constants, the system can be viewed as a case handled in the setting of all real eigenvalues. We can thus proceed with the analysis in Section A.

Finally, we remark that $d \leq \deg(\gamma)^2$. The proof appears in [11], and we bring it here for completeness. Since γ is a primitive root of unity of order d, then the defining polynomial p_γ of γ is the d-th Cyclotomic polynomial, so $\deg(\gamma) = \Phi(d)$, where Φ is Euler's totient function. Since $\Phi(d) \geq \sqrt{d}$, we get that $d \leq \deg(\gamma)^2$. Therefore, the number of cases we consider is polynomial in the original input, and does not involve a blowup in the complexity.

C Matrix Forms in Proposition 6

Recall that we have

$$g^{(i)}(x) = \begin{cases} \sum_{m=1}^{k} m^i 2|\beta_m| \cos(mx + \theta_m) & i \equiv_4 0 \\ \sum_{m=1}^{k} -m^i 2|\beta_m| \sin(mx + \theta_m) & i \equiv_4 1 \\ \sum_{m=1}^{k} -m^i 2|\beta_m| \cos(mx + \theta_m) & i \equiv_4 2 \\ \sum_{m=1}^{k} m^i 2|\beta_m| \sin(mx + \theta_m) & i \equiv_4 3 \end{cases}$$

Writing this in matrix form, split by $i \mod 4$, we have the following.

$$\text{for } i \equiv_4 0: \quad \begin{pmatrix} 1^4 & 2^4 & \cdots & k^4 \\ 1^8 & 2^8 & \cdots & k^8 \\ \vdots & \vdots & \vdots & \vdots \\ 1^{4k} & 2^{4k} & \cdots & k^{4k} \end{pmatrix} \begin{pmatrix} 2|\beta_1| \cos(x + \theta_1) \\ 2|\beta_2| \cos(2x + \theta_2) \\ \vdots \\ 2|\beta_k| \cos(kx + \theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{for } i \equiv_4 1: \quad \begin{pmatrix} -1^1 & -2^1 & \cdots & -k^1 \\ -1^5 & -2^5 & \cdots & -k^5 \\ \vdots & \vdots & \vdots & \vdots \\ -1^{4k-3} & -2^{4k-3} & \cdots & -k^{4k-3} \end{pmatrix} \begin{pmatrix} 2|\beta_1| \sin(x + \theta_1) \\ 2|\beta_2| \sin(2x + \theta_2) \\ \vdots \\ 2|\beta_k| \sin(kx + \theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{for } i \equiv_4 2: \quad \begin{pmatrix} -1^2 & -2^2 & \cdots & -k^2 \\ -1^6 & -2^6 & \cdots & -k^6 \\ \vdots & \vdots & \vdots & \vdots \\ -1^{4k-2} & -2^{4k-2} & \cdots & -k^{4k-2} \end{pmatrix} \begin{pmatrix} 2|\beta_1| \cos(x + \theta_1) \\ 2|\beta_2| \cos(2x + \theta_2) \\ \vdots \\ 2|\beta_k| \cos(kx + \theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{for } i \equiv_4 3: \quad \begin{pmatrix} 1^3 & 2^3 & \cdots & k^3 \\ 1^7 & 2^7 & \cdots & k^7 \\ \vdots & \vdots & \vdots & \vdots \\ 1^{4k-1} & 2^{4k-1} & \cdots & k^{4k-1} \end{pmatrix} \begin{pmatrix} 2|\beta_1| \sin(x + \theta_1) \\ 2|\beta_2| \sin(2x + \theta_2) \\ \vdots \\ 2|\beta_k| \sin(kx + \theta_k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

D The Case where A is Singular

In this section, we reduce the Semialgebraic Orbit Problem in the case where A is a singular matrix to the case where A is non-singular. Intuitively, we simply cast our analysis to a lower dimension by projecting A on its nonzero eigenvalues.

In this case, we are given semialgebraic sets $S, T \subseteq \mathbb{R}^3$ and a matrix $A \in \mathbb{Q}^{3\times 3}$, where 0 is an eigenvalue of A.

We start with the case where the multiplicity of the eigenvalue 0 is 1. Then, we can write $A = P\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}P^{-1}$ where P and B are invertible matrices with rational entries. Indeed, since $0 \in \mathbb{Q}$, then we can decompose $\mathbb{Q}^3$ as a direct sum $\mathbb{Q}^3 = V_0 \oplus V_0^\perp$ where V_0 has dimension 1 and $V_0^\perp$ has dimension 2. Let $u \in \mathbb{Q}^3$ be an eigenvector corresponding to 0, so that $\text{span}(u) = V_0$, and let $v, w \in \mathbb{Q}^3$ such that $\text{span}(v, w) = V_0^\perp$. We now have that $Av, Aw \in V_0^\perp$, so we can write $Av = c_1v + c_2w$ and $Aw = d_1v + d_2w$ for some $c_1, c_2, d_1, d_2 \in \mathbb{Q}$. Let $P = (u, v, w)$ and $B = \begin{pmatrix}c_1 & d_1\\ c_2 & d_2\end{pmatrix}$, then it is easy to verify that P and B are invertible, and that $AP = P\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}$, so $A = P\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}P^{-1}$, as we wanted.

We now observe the following:

$$\begin{aligned}
&\exists n \in \mathbb{N}\ \exists x \in S : A^n x \in T && \iff\\
&\exists n \in \mathbb{N}\ \exists x \in S : P\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}P^{-1}x \in T && \iff\\
&\exists n \in \mathbb{N}\ \exists x' \in P^{-1}S : P\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}x' \in P^{-1}T
\end{aligned}$$

Denote $S' = P^{-1}S$ and $T' = P^{-1}T$, we proceed[4]:

$$\begin{aligned}
&\exists n \in \mathbb{N}\ \exists x' \in S' : \begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}x' \in T' && \iff\\
&\exists n \in \mathbb{N}\ \exists x' \in S' : \begin{pmatrix}0 & 0\\ 0 & B^{n-1}\end{pmatrix}\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}x' \in T' && \iff\\
&\exists n \in \mathbb{N}\ \exists x'' \in \begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}S' : \begin{pmatrix}0 & 0\\ 0 & B^{n-1}\end{pmatrix}x'' \in T' && \iff \qquad (6)
\end{aligned}$$

Denote $S'' = \begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}S'$, and observe that

$$S'' = \left\{\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix}\begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix} : \begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix} \in S'\right\} = \left\{\begin{pmatrix}0\\ B\begin{pmatrix}y_2\\ y_3\end{pmatrix}\end{pmatrix} : \begin{pmatrix}0\\ y_2\\ y_3\end{pmatrix} \in S'\right\} = \left\{\begin{pmatrix}0\\ z_2\\ z_3\end{pmatrix} : \begin{pmatrix}0\\ B^{-1}\begin{pmatrix}z_2\\ z_3\end{pmatrix}\end{pmatrix} \in S'\right\}$$

Thus, the vectors in S'' have 0 in their first coordinate. For such vectors, we have the

[4] In the following we ignore the case where $n = 0$, as this can be checked initially by deciding whether $S \cap T \neq \emptyset$.

following:

$$\begin{pmatrix} 0 & 0 \\ 0 & B^{n-1} \end{pmatrix} \begin{pmatrix} 0 \\ z_2 \\ z_3 \end{pmatrix} \in T' \iff B^{n-1} \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} \in \left\{ \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} : \begin{pmatrix} 0 \\ x_2 \\ x_3 \end{pmatrix} \in T' \right\}$$

Let $S_2'' = \left\{ \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} : \begin{pmatrix} 0 \\ z_2 \\ z_3 \end{pmatrix} \in S'' \right\}$ and $T_2' = \left\{ \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} : \begin{pmatrix} 0 \\ z_2 \\ z_3 \end{pmatrix} \in T' \right\}$, then the condition in (6) holds iff

$$\exists n \in \mathbb{N}\ \exists \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} \in S_2'' : B^{n-1} \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} \in T_2'$$

Since S_2'' and T_2' are semialgebraic (and are in fact easily computable from S and T), we conclude that we can reduce the dimension of the problem.

Next, if the multiplicity of 0 is 2, then we can write $A = P \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho \end{pmatrix} P^{-1}$ where ρ is a real eigenvalue. Then $A^n = P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho^n \end{pmatrix} P^{-1}$ for every $n \geq 2$, and the same approach as above can be taken.

Finally, if the multiplicity of 0 is 3, then $A^3 = 0$, so the problem becomes trivial.

E Change of Basis Matrices in the 3×3 case

In this section we consider a diagonalisable matrix $A \in \mathbb{Q}^{3\times 3}$ with complex eigenvalues. Thus, we can write $A = PDP^{-1}$ with $D = \mathrm{diag}(\lambda, \overline{\lambda}, \rho)$ with $\lambda \in \mathbb{A}$ and $\rho \in \mathbb{A} \cap \mathbb{R}$.

Note that the columns of the matrix P are eigenvectors of A, and moreover, conjugate eigenvalues have conjugate eigenvectors and real eigenvalues have real eigenvectors. We can therefore assume

$$P = \begin{pmatrix} a & \overline{a} & d \\ b & \overline{b} & e \\ c & \overline{c} & f \end{pmatrix}$$

for $a, b, c \in \mathbb{A}$ and $d, e, f \in \mathbb{R} \cap \mathbb{A}$.

▶ **Lemma 12.** *Let $E = diag(\delta_1, \delta_2, \delta_3)$ be a diagonal matrix, then every coordinate of PEP^{-1} is of the form $\alpha\delta_1 + \overline{\alpha}\delta_2 + \beta\delta_3$, where $\alpha \in \mathbb{A}$ and $\beta \in \mathbb{A} \cap \mathbb{R}$.*

Proof. The proof is straightforward: we compute the matrix P^{-1}, and then the product PEP^{-1}.

We leave it to the reader to verify the following: first, the determinant of P is pure-imaginary, i.e., $\det(P) = mi$ for $m \in \mathbb{R} \cap \mathbb{A}$. Second, we have

$$P^{-1} = \frac{1}{mi} \begin{pmatrix} f\overline{b} - e\overline{c} & d\overline{c} - f\overline{a} & e\overline{a} - d\overline{b} \\ ce - bf & af - cd & bd - ae \\ b\overline{c} - c\overline{b} & c\overline{a} - a\overline{c} & a\overline{b} - b\overline{a} \end{pmatrix}$$

Finally, it is very easy (yet tedious) to verify that PEP^{-1} satisfies the claim. We demonstrate by computing the coordinate $(PEP^{-1})_{1,2}$.

We have that the first row of PE is $(a\delta_1, \overline{a}\delta_2, d\delta_3)$, and hence

$$\begin{aligned}(PEP^{-1})_{1,2} &= (PE)_{1,1}P^{-1}_{1,2} + (PE)_{1,2}P^{-1}_{2,2} + (PE)_{1,3}P^{-1}_{3,2} \\ &= \frac{1}{mi}\left(a\delta_1(d\overline{c} - f\overline{a}) + \overline{a}\delta_2(af - cd) + d\delta_3(c\overline{a} - a\overline{c})\right) \\ &= \frac{1}{m}\left(-i\delta_1(ad\overline{c} - af\overline{a}) + i\delta_2(\overline{a}cd - af\overline{a}) - i\delta_3(dc\overline{a} - da\overline{c})\right)\end{aligned}$$

It is now easy to see that the coefficients of δ_1 and δ_2 are conjugates, and the coefficient of δ_3 is real, as desired. ◀

F Bounds on the Description Size of Points in Z_f

We complete the analysis of Remark 7.

Recall that $f(z) = \sum_{m=0}^{k} \beta_m z^m + \overline{\beta_m}\overline{z}^m$, and $Z_f = \{z : f(z) = 0 \wedge |z| = 1\}$. Further recall that for every $0 \le m \le k$, β_m is a polynomial in $a_1, a_2, a_3, \overline{a_1}, \overline{a_2}, \overline{a_3}, b_1, b_2, b_3$, where all the latter are linear combinations of roots of the characteristic polynomial of A, and are therefore algebraic numbers of degree at most 3 and description polynomial in $\|A\| + \|s\|$.

We can now express the condition $f(z) = 0$ using a quantified formula in the first order theory of the reals by replacing each of the constants above (i.e. a_1, etc.) by their corresponding description, as per Section 2.2. It follows that in this description, there are at most 9 variables. We now employ the following result due to Renegar [19].

▶ **Theorem 13** (Renegar). *Let $M \in \mathbb{N}$ be fixed. Let $\tau(\mathbf{y})$ be a formula of the first order theory of the reals. Assume that the number of (free and bound) variables in $\tau(\mathbf{y})$ is bounded by M. Denote the degree of $\tau(\mathbf{y})$ by d and the number of atomic predicates in $\tau(\mathbf{y})$ by n.*

There is a polynomial time (polynomial in $\|\tau(\mathbf{y})\|$) procedure which computes an equivalent quantifier-free formula

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

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

1. *Each of I and J_i (for $1 \le i \le I$) is bounded by $(n+d)^{O(1)}$.*
2. *The degree of $\chi(\mathbf{y})$ is bounded by $(n+d)^{O(1)}$.*
3. *The height of $\chi(\mathbf{y})$ is bounded by $2^{\|\tau(\mathbf{y})\|(n+d)^{O(1)}}$.*

We apply this theorem to the description of Z_f given above, where we identify $\mathbb{C}$ with $\mathbb{R}^2$ so that f is indeed a polynomial. Then, we obtain in polynomial time a description of Z_f. Moreover, the degrees of the entries is bounded by $\|f\|^{O(1)}$ and their height is bounded by $2^{\|f\|^{O(1)}}$.