The Polytope-Collision Problem

Shaull Almagor[1], Joël Ouaknine[2], and James Worrell[3]

1 **Department of Computer Science, Oxford University, UK**
`shaull.almagor@cs.ox.ac.uk`

2 **Department of Computer Science, Oxford University, UK**
`joel@cs.ox.ac.uk`

3 **Department of Computer Science, Oxford University, UK**
`jbw@cs.ox.ac.uk`

Abstract

The *Orbit Problem* consists of determining, given a matrix $\mathcal{A} \in \mathbb{R}^{d\times d}$ and vectors $x, y \in \mathbb{R}^d$, whether there exists $n \in \mathbb{N}$ such that $\mathcal{A}^n = y$. This problem was shown to be decidable in a seminal work of Kannan and Lipton in the 1980s. Subsequently, Kannan and Lipton noted that the Orbit Problem becomes considerably harder when the *target* y is replaced with a subspace of $\mathbb{R}^d$. Recently, it was shown that the problem is decidable for vector-space targets of dimension at most three, followed by another development showing that the problem is in **PSPACE** for polytope targets of dimension at most three.

In this work, we take a dual look at the problem, and consider the case where the *initial* vector x is replaced with a polytope P_1, and the target is a polytope P_2. Then, the question is whether there exists $n \in \mathbb{N}$ such that $\mathcal{A}^n P_1 \cap P_2 \neq \emptyset$. We show that the problem can be decided in **PSPACE** for dimension at most three. As in previous works, decidability in the case of higher dimensions is left open, as the problem is known to be hard for long-standing number-theoretic open problems.

Our proof begins by formulating the problem as the satisfiability of a parametrized family of sentences in the existential first-order theory of real-closed fields. Then, after removing quantifiers, we are left with instances of simultaneous positivity of sums of exponentials. Using techniques from transcendental number theory, and separation bounds on algebraic numbers, we are able to solve such instances in **PSPACE**.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems, I.1.2 Algorithms

Keywords and phrases linear dynamical systems, orbit problem, algebraic algorithms

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

42nd Conference on Very Important Topics (CVIT 2016).
Editors: John Q. Open and Joan R. Acces; Article No. 23; pp. 23:1–23:21
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Given a linear transformation $\mathcal{A}$ over the vector space $\mathbb{R}^d$, together with a starting point x, the *orbit* of x under $\mathcal{A}$ is the infinite sequence $x, \mathcal{A}x, \mathcal{A}^2x, \ldots$. A natural decision problem in discrete linear dynamical systems is whether the orbit of x ever hits a particular target set V (assuming suitable, effective representations of $\mathcal{A}$, x, and V). An early instance of this problem was raised by Harrison in 1969 [12] for the special case in which V is simply a point in $\mathbb{R}^d$. Decidability remained open for over ten years, and was finally settled in a seminal paper of Kannan and Lipton, who moreover gave a polynomial-time decision procedure [13]. In subsequent work [14], Kannan and Lipton noted that the Orbit Problem becomes considerably harder when the target V is replaced by a subspace of $\mathbb{R}^d$: indeed, if V has dimension $d-1$, the problem is equivalent to the *Skolem Problem*, known to be **NP-Hard** but whose decidability has remained open for over 80 years [21]. However, for low-dimensional target spaces, the Orbit Problem becomes more tractable. Indeed, it was recently shown in [7] that the problem is decidable for vector-space targets of dimension at most three, with polynomial-time complexity for one-dimensional targets, and complexity in $\mathbf{NP}^{\mathbf{RP}}$ for two- and three-dimensional targets. Another development followed in [8], where the authors consider more intricate target sets, namely polytopes. It is shown in [8] that up to dimension three, the problem can be solved in **PSPACE**. In addition, it is shown that for higher dimensions, the problem becomes hard with respect to long-standing number-theoretic open problems.

A key motivation for studying the Orbit Problem comes from program verification, particularly the problem of determining whether a simple while loop with affine assignments and guards will terminate or not. Similar reachability questions were considered and left open by Lee and Yannakakis in [15] for what they termed “real affine transition systems”. Similarly, decidability for the case of a single-halfspace target was mentioned as an open problem by Braverman in [5].

An important aspect of termination problems for linear loops is the quantification of the initial point. Traditionally, the ‘Termination problem’ in the program-verification literature (see, e.g. [4]) refers to termination of while loops for *all* possible initial starting points. In [17] the traditional Termination Problem is solved over the integers for while loops, assuming diagonalisability of the associated linear transformation. To our knowledge, very little else is known on the general problem of universally quantified inputs. In contrast, the works in [7, 8] study the termination problem where the input is fixed (but the target space is complicated). This corresponds to verifying the termination of a concrete run of a linear loop. It should be noted that the techniques used for analyzing the latter differ significantly from the former.

In this work, we take a dual look at the problem, and study the case where the input is existentially quantified. Thus, we are given a set $P_1 \subseteq \mathbb{R}^d$, and a target set P_2, and the problem is to decide whether there exists $x \in P_1$ and $n \in \mathbb{N}$ such that $\mathcal{A}^n x \in P_2$. In practice, this corresponds to deciding *safety* properties of linear loops: we think of P_2 as some error set, and the problem is to decide whether there exists an input that would cause the program to reach the error set.

Specifically, the focus of this paper is the **3D Polytope-Collision Problem** (**3DPCP**, for short): Given two polytopes P_1 and P_2 in $\mathbb{R}^3$ (represented as an intersection of halfspaces) and a matrix with real-algebraic entries[1] $\mathcal{A} \in (\mathbb{A} \cap \mathbb{R})^{3\times 3}$, determine whether there exists a

[1] We denote by $\mathbb{A}$ the set of algebraic numbers.

point $x \in P_1$ and a natural number n such that $\mathcal{A}^n x \in P_2$.

We present the following effectiveness result on the 3D Polytope-Collision Problem.

▶ **Theorem 1.** *3DPCP is decidable in* **PSPACE**.

Note that as proved in [8], when the dimension is at least four, the polytope-collision problem becomes hard with respect to number-theoretic open problems.

Before describing our approach, we explain why this result is somewhat surprising. Consider a simplification of 3DPCP, where the initial polytope P is a segment between points x and y, and we wish to decide whether the orbit of P under the matrix $\mathcal{A}$ collides with another polytope R. We can represent P as the single point (x, y) in $\mathbb{R}^6$, and extend $\mathcal{A}$ to a matrix $\mathcal{B} \in \mathbb{R}^{6\times 6}$ that has two copies of $\mathcal{A}$ on its diagonal. Then, the orbit of P under $\mathcal{A}$ corresponds to the orbit of (x, y) under $\mathcal{B}$. However, the respective target space in $\mathbb{R}^6$ becomes the set of all points (u, v) such that the line between u and v in $\mathbb{R}^3$ intersects R. While this is a semi-algebraic set, it is quite complicated, and recall that the polytope hitting problem is already hard in dimension four. Thus, this approach suggests that the problem may be as hard as the hitting problem in $\mathbb{R}^6$.

Technically, the above intricacy prevents us from using the techniques previously employed on fixed-input orbit problems, e.g. [8]. There, describing the dominant behavior of the orbit is relatively straightforward, and the difficulty is reasoning about hitting the target. In our setting, merely describing the orbit involves symbolic quantifier elimination, as described next, and reasoning about hitting the target therefore involves symbolic analysis.

Our approach to proving Theorem 1 is as follows. Observe that 3DPCP can be formulated as the problem of deciding whether there exists $n \in \mathbb{N}$ such that $\mathcal{A}^n P_1$ intersects P_2 (where $\mathcal{A}^n P_1 = \{\mathcal{A}^n x : x \in P_1\}$). In Section 3 we reduce this formulation of 3DPCP to the problem of solving a system of inequalities, as we now describe.

In Section 3.1 we identify two types of intersection of 3D polytopes, namely (1) where a vertex of one polytope lies in the other polytope, and (2) where an edge of one polytope intersects a face of the other polytope. We show that under a certain representation, an intersection of polytopes is always of one of these types. Note that while each of these types seems symmetric with respect to the two polytopes, in our setting the polytopes have an inherent asymmetry, as $\mathcal{A}^n P_1$ is dependent on n whereas P_2 is not.

In order to overcome this asymmetry, in Section 3.2 we reduce 3DPCP to the case where the matrix $\mathcal{A}$ is invertible. Then, considering $\mathcal{A}^n P_1$ and P_2 is symmetric to considering P_1 and $(\mathcal{A}^{-1})^n P_2$.

Next, in Section 3.3 we observe that intersections of Type (1) can be decided using the work in [8], and we are left to address intersections of Type (2). We formulate this type of intersection as $\exists n \in \mathbb{N}\ \Phi(\alpha^n, \overline{\alpha}^n, \rho^n)$, where Φ is a sentence in the existential first-order theory of real-closed fields, and $\alpha, \overline{\alpha}$, and ρ are the eigenvalues of the matrix $\mathcal{A}$, with $\alpha \in \mathbb{A}\setminus\mathbb{R}$ and $\rho \in \mathbb{A} \cap \mathbb{R}$ (the case where $\mathcal{A}$ has only real eigenvalues is simpler, and we handle it in Appendix C). Moreover, Φ contains only linear expressions (with respect to its variables, where n is treated as a constant), and at most three real variables. We proceed by eliminating the quantifiers from Φ. We use the fact that the expressions in $\Phi(n)$ are linear to apply the simple Fourier-Motzkin quantifier-elimination algorithm [11]. We note that while other quantifier-elimination algorithms (e.g., [20]) offer better asymptotic complexity, since the number of variables in Φ is constant, Fourier-Motzkin elimination takes polynomial time. Moreover, its simplicity allows us to keep track of the expressions in the quantifier free equivalent of $\Phi(n)$. Specifically, we show that this output consists of a disjunction of systems, where each system is a conjunction of expressions of the form

$$A\alpha^{2n} + \overline{A}\overline{\alpha}^{2n} + B\alpha^n\rho^n + \overline{B}\overline{\alpha}^n\rho^n + C\rho^{2n} + D|\alpha|^{2n} + E\alpha^n + \overline{E}\overline{\alpha}^n + F\rho^n + G \bowtie 0 \qquad (1)$$

where $\bowtie \in \{>, =\}$.

Finally, Section 4 is the heart of our technical contribution, in which we show how to solve such systems. Intuitively, we normalize Expression (1) such that the maximal modulus of its terms is 1, thus obtaining an expression of the form $A\gamma^{2n} + \overline{A}\overline{\gamma}^{2n} + B\gamma^n + \overline{B}\overline{\gamma}^n + C + r(n) \bowtie 0$ with $|\gamma| = 1$ and $r(n)$ tending exponentially fast to 0. We then consider two cases, depending on whether γ is a root of unity or not. If γ is a root of unity, we show that it is enough to consider polynomially many expressions with only real elements, which can be handled using relatively standard techniques. If γ is not a root of unity, things are more involved. Then, by utilizing consequences of the Baker-Wüstholz theorem [2], we are able to show that the expression $|A\gamma^{2n} + \overline{A}\overline{\gamma}^{2n} + B\gamma^n + \overline{B}\overline{\gamma}^n + C|$ is bounded away from 0 by an inverse polynomial in n. Then, using a separation bound due to Mignotte [16], we show that $r(n)$ decays fast enough to obtain a bound $N \in \mathbb{N}$ such that $r(n)$ does not affect the sign of $A\gamma^{2n} + \overline{A}\overline{\gamma}^{2n} + B\gamma^n + \overline{B}\overline{\gamma}^n + C$ for all $n > \mathbb{N}$. Finally, since γ is not a root of unity, it is dense in the unit circle, and we can replace the analysis of the former expression by analysis of the simpler function $f(z) = Az^2 + \overline{A}\overline{z}^2 + Bz + \overline{B}\overline{z} + C$ on the unity circle, from which we obtain our main result.

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 algebraic numbers (denoted $\mathbb{A}$) and their (efficient) manipulation. The main references include [3, 9, 20]. 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[2] for algebraic numbers is to encode α as a tuple comprising its defining polynomial together with rational approximations of its real and imaginary parts of sufficient precision to distinguish α from the other roots of p_α. More precisely, α can be represented by $(p_\alpha, a, b, r) \in \mathbb{Z}[x] \times \mathbb{Q}^3$ provided that α is the unique root of p_α inside the circle in $\mathbb{C}$ of radius r centred at $a + bi$. A separation bound due to Mignotte [16] asserts that for roots $\alpha \neq \beta$ of a polynomial $p \in \mathbb{Z}[x]$, we have

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

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.

[2] Note that this representation is not unique.

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, 9, 19]. 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 2 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, 9].

2.2 First-order theory of the reals

Let $\overrightarrow{x} = x_1, \ldots, x_m$ be a list of m real-valued variables, and let $\sigma(\overrightarrow{x})$ be a Boolean combination of atomic predicates of the form $g(\overrightarrow{x}) \bowtie 0$, where each $g(\overrightarrow{x}) \in \mathbb{Z}[x]$ is a polynomial with integer coefficients over these variables, and $\bowtie \in \{>, =\}$. A *sentence of the first-order theory of the reals* is of the form $Q_1x_1Q_2x_2 \cdots Q_mx_m\sigma(\overrightarrow{x})$, 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 [22]. His procedure, however, has non-elementary complexity. Many substantial improvements followed over the years, starting with Collins' technique of cylindrical algebraic decomposition [10], and culminating with the fine-grained analysis of Renegar [20]. In this paper, we focus exclusively on the situation in which the number of variables is uniformly bounded.

▶ **Theorem 2** (Renegar). *Let $M \in \mathbb{N}$ be fixed, let τ be of the form $Q_1x_1Q_2x_2 \cdots Q_mx_m\sigma(\overrightarrow{x})$. Assume that the number of variables in τ is bounded by M (i.e., $m \leq M$). Then the truth value of τ can be determined in time polynomial in $\|\tau\|$.*

An important property of the first-order theory of the reals is that it admits *quantifier elimination*. That is, consider two lists of variables $\overrightarrow{x}, \overrightarrow{y}$ and a sentence $Q_1x_1 \cdots Q_mx_m\sigma(\overrightarrow{x}, \overrightarrow{y})$ with the variables of $\overrightarrow{y}$ being free, then there exists an (unquantified) sentence $\sigma'(\overrightarrow{y})$ such that for every assignment π to the variables in $\overrightarrow{y}$ it holds that $\sigma'(\pi)$ is true iff $Q_1x_1 \cdots Q_mx_m\sigma(\overrightarrow{x}, \pi)$ is true.

When the polynomials in σ are all linear and the quantifiers are all existential, then quantifier elimination can be performed using the Fourier-Motzkin quantifier-elimination algorithm [11] (see Appendix B for details). The benefit of this algorithm is its simplicity, which allows us to remove quantifiers symbolically.

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

2.3 Polytopes and their representation

A *polytope* P in $\mathbb{R}^3$ is an intersection of finitely many halfspaces in $\mathbb{R}^3$: $P = \{x \in \mathbb{R}^3 : v_1^T x \geq c_1 \wedge \ldots \wedge v_k^T x \geq c_k\}$ for vectors $v_1, \ldots, v_k \in \mathbb{R}^3$ and numbers $c_1, \ldots, c_k \in \mathbb{R}$. The *halfspace* *description* of P is then $(v_1, c_1), \dots, (v_k, c_k)$. When all entries are algebraic, we denote by $\|P\|$ the description length.

The *dimension* of a polytope P, denoted $\dim(P)$, is the dimension of the subspace of $\mathbb{R}^3$ spanned by P. The dimension of P can be computed in time polynomial in $\|P\|$ by solving polynomially many linear programs. In $\mathbb{R}^3$, the dimension of a polytope is in $\{0, \dots, 3\}$. A 2D boundary of a 3D polytope is a 2D polytope called a *face*. Similarly, the boundries of 2D polytopes (and in particular of faces) are called *edges*, and the boundries of edges are *vertices*. Every 3D polytope, except the trivial $\mathbb{R}^3$ and $\emptyset$, has at least one face (but not necessarily edges or vertices). Since vertices and edges are crucial for our algorithms, we present the following lemma from [8]

▶ **Lemma 3** ([8] Lemma A.1). *Suppose $P \subseteq \mathbb{R}^3$ is a 2D polytope. Then $P = \bigcup_{i=1}^m A_i$, where m is finite and each A_i is of the form $A_i = \{u_i + \alpha v_i + \beta w_i : T_i(\alpha, \beta)\}$ where $u_i, v_i, w_i \in \mathbb{R}^3$ and the predicates $T_i(\alpha, \beta)$ are from the following:*

- $T_i(\alpha, \beta) \equiv \alpha \geq 0 \wedge \beta \geq 0$ *(A_i is an infinite cone)*
- $T_i(\alpha, \beta) \equiv \alpha \geq 0 \wedge \beta \geq 0 \wedge \alpha + \beta \leq 1$ *(A_i is a triangle)*
- $T_i(\alpha, \beta) \equiv \alpha \geq 0 \wedge \beta \geq 0 \wedge \beta \leq 1$ *(A_i is an infinite strip)*

Furthermore, if we are given a halfspace description of P with length $\|P\|$, the size of the representation of each vector u_i, v_i, w_i is at most $\|P\|^{O(1)}$.

Note that since the representation of u_i, v_i, and w_i is polynomial, it follows that m is at most exponential in $\|P\|$, and moreover, that iterating over the sets A_i can be done in **PSPACE**.

3 From 3DPCP to a System of Inequalities

In this section we reduce 3DPCP to the problem of solving a system of inequalities. More precisely, we show how to solve 3DPCP by solving an exponential number of systems of equalities and inequalities, and that iterating over these systems can be done in **PSPACE**. In Section 4 we tackle the main technical challenge of solving each such system in **PSPACE**, thus concluding the proof of Theorem 1.

As mentioned in Section 1, we start by studying the intersection of polytopes.

3.1 Intersection of polytopes

Consider two intersecting polytopes Q_1 and Q_2 in $\mathbb{R}^3$. In this section, we characterize the intersection of Q_1 and Q_2, which would later simplify the solution of 3DPCP. To illustrate the idea, assume that both Q_1 and Q_2 are bounded 3D polytopes. In this case, we can assume w.l.o.g. that Q_1 and Q_2 are both tetrahedra. Indeed, every bounded 3D polytope with d vertices can be decomposed into a union of at most $\binom{d}{4}$ tetrahedra, and two such decompositions intersect iff two of the tetrahedra in the respective decompositions intersect. Under this assumption, there are two possible "types" of intersections: either Q_1 is contained in Q_2 (or vice-versa), or an edge of Q_1 intersects a face of Q_2 (or vice-versa). When the polytopes are bounded, we can relax the first requirement, and require instead that a vertex of Q_1 lies in Q_2 (or vice-versa).

In general, however, Q_1 or Q_2 may be unbounded. In this case we need to be slightly more careful. Indeed, as stated in Section 2.3, unbounded polytopes might have no vertices or edges, but only faces (unless the polytope is $\mathbb{R}^3$ or $\emptyset$, in which case the problem is trivial). For example, consider the case where Q_1 and Q_2 are infinite prisms. Then, it is possible that $Q_1 \cap Q_2 \neq \emptyset$ and neither are contained in each other, but no edge of Q_1 intersects a face of Q_2 (and vice-versa).

Therefore, to get the above characterization for unbounded polytopes, we need to add "fictive" edges. Since we assume the input polytopes are non trivial, then each of them has at least one face, and recall that the faces of a 3D polytope are 2D polytopes. By employing Lemma 3 on the faces of the polytopes, we get that each face of Q_1 and of Q_2 can be written as $\bigcup_{i=1}^{m} A_i$ as per Lemma 3. Observe that every set A_i in the decomposition of Lemma 3 has at least two edges and one vertex, and that a non-empty intersection $A_i \cap A'_j$ in such decompositions also intersects an edge of at least one of the two sets (the only involved case is the intersection of two infinite strips, where one should notice that the strips are only infinite to one side).

We conclude that the above characterization of the intersection of polytopes is correct also for unbounded ones. In the following, when we refer to a vertex/edge of an unbounded polytope, we mean the vertices and edges of the sets in the decomposition of Lemma 3.

Thus, we have that Q_1 intersects Q_2 if at least one of the following holds:

1. There exists a vertex of Q_1 that is in Q_2.
2. There exists a vertex of Q_2 that is in Q_1.
3. An edge of Q_1 intersects a face of Q_2.
4. An edge of Q_2 intersects a face of Q_1.

3.2 Reduction to the invertible case

In the notations of Section 3.1, we wish to check the intersection of $Q_1 = \mathcal{A}^n P_1$ and $Q_2 = P_2$ for an existentially quantified $n \in \mathbb{N}$. As mentioned in Section 1, if $\mathcal{A}$ is invertible, then the problem is symmetric with respect to Q_1 and Q_2. Indeed, $\mathcal{A}^n P_1$ intersects P_2 iff P_1 intersects $(\mathcal{A}^{-1})^n P_2$. However, if $\mathcal{A}$ is not invertible, the problem is not clearly symmetric. In this section, we reduce 3DPCP to the case where $\mathcal{A}$ is an invertible matrix.

Consider polytopes $P, R \subseteq \mathbb{R}^3$, and let $\mathcal{A} \in (\mathbb{A} \cap \mathbb{R})^{3\times 3}$ be a singular matrix, so 0 is an eigenvalue of $\mathcal{A}$. Consider first the case where the multiplicity of 0 is 1. Thus, we can write $\mathcal{A} = D^{-1}\begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix} D$ where D is an invertible matrix with real-algebraic entries, and $B \in (\mathbb{A} \cap \mathbb{R})^{2\times 2}$. Indeed, if $\mathcal{A}$ has only real eigenvalues then this is achieved by converting $\mathcal{A}$ to Jordan form, and if $\mathcal{A}$ has complex eigenvalues α and $\overline{\alpha}$, then this is achieved by setting $D = (v, u, w)$ where v is an eigenvector corresponding to 0, and $u + iw$ is an eigenvector corresponding to α. In addition, B is invertible, since its eigenvalues are the nonzero eigenvalues of $\mathcal{A}$.

In Appendix A, we show that in this case, there exist polytopes $P', R' \subseteq \mathbb{R}^2$ such that for every $n \geq 2$ the following holds: there exists $x \in P$ such that $\mathcal{A}^n x \in R$ iff there exists $x' \in P'$ such that $B^{n-1}x' \in R'$. Thus, it is enough to consider the polytopes P', R' and the invertible matrix B. Moreover, we show that computing P' and R' can be done in polynomial time. We also show a similar approach can be taken when 0 has multiplicity 2 or 3 (with the latter being trivial, since $\mathcal{A}$ is then nilpotent).

It should be noted that in the reduction above, even if the input had only rational entries, the output may still require a real-algebraic description. However, the degree and height of the algebraic numbers involved in the description of the output polytopes remain polynomial in the size of the input.

Finally, we note that we can always *increase* the dimension of the problem while maintaining an invertible matrix. Indeed, Given a invertible matrix $B \in (\mathbb{A} \cap \mathbb{R})^{2\times 2}$, we can consider the invertible matrix $\begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$, and change $P, R \subseteq \mathbb{R}^2$ to $\{1\} \times P, \{1\} \times R \subseteq \mathbb{R}^3$ (and a similar approach when $B \in (\mathbb{A} \cap \mathbb{R})^{1\times 1}$). Thus, it is enough to solve the problem in the invertible case in dimension 3.

3.3 From the invertible case to an equation system

In this section we focus on solving 3DPCP in the invertible case.

Let P_1, P_2 be the input polytopes (whose description may contain algebraic numbers, as per the reduction of Section 3.2), and let $\mathcal{A} \in (\mathbb{A} \cap \mathbb{R})^{3\times 3}$ be an invertible matrix. By Section 3.1, and since $\mathcal{A}$ is invertible, it suffices to decide whether there exists a number $n \in \mathbb{N}$ such that either there exists a vertex x of P_1 with $\mathcal{A}^n x \in P_2$, or there exists an edge e of P_1 such that $\mathcal{A}^n e$ intersects a face of P_2. Note that we may need to reverse the roles of P_1 and P_2, and use $\mathcal{A}^{-1}$ instead of $\mathcal{A}$. We remark that $\left\|\mathcal{A}^{-1}\right\|$ is polynomial in $\|\mathcal{A}\|$, and moreover — since the eigenvalues of $\mathcal{A}^{-1}$ are inverses of those of $\mathcal{A}$ — the description length of the eigenvalues of $\mathcal{A}^{-1}$ is equal to that of $\mathcal{A}$.

In [8], the authors show that the problem of deciding, given a polyhedron P in $\mathbb{R}^3$, a vector $x \in \mathbb{R}^3$, and a matrix $\mathcal{A} \in (\mathbb{A}\cap\mathbb{R})^{3\times 3}$, whether there exists $n \in \mathbb{N}$ such that $\mathcal{A}^n x \in P$ is solvable in **PSPACE**. This solves the former case. It remains to solve the latter.

We thus assume that we are given as input a matrix $\mathcal{A} \in (\mathbb{A} \cap \mathbb{R})^{3\times 3}$, an edge $E = \{u + \lambda v : \lambda \in J\}$ where $u, v \in \mathbb{R}^3$ and J is either $[0, 1]$ or $[0, \infty)$, and a face $F = \{s + \mu t + \nu r : T(\mu, \nu)\}$, where $s, t, r \in \mathbb{R}^3$ and $T(\mu, \nu)$ is one of the following predicates (as per Lemma 3):

- $T(\mu, \nu) \equiv \mu \geq 0 \wedge \nu \geq 0$
- $T(\mu, \nu) \equiv \mu \geq 0 \wedge \nu \geq 0 \wedge \mu + \nu \leq 1$
- $T(\mu, \nu) \equiv \mu \geq 0 \wedge \nu \geq 0 \wedge \nu \leq 1$

We wish to determine whether there exists a number n and $x \in E$ such that $\mathcal{A}^n x \in F$. In the following, we will treat the case where $E = \{u + \lambda v : \lambda \in [0, 1]\}$ and $F = \{s + \mu t + \nu r : \mu \geq 0 \wedge \nu \geq 0 \wedge \mu + \nu \leq 1\}$. The other cases are slightly simpler, and can be solved *mutatis-mutandis*.

Consider the eigenvalues of $\mathcal{A}$. Since $\mathcal{A}$ is a 3×3 invertible matrix, either all the eigenvalues are real, or there is one real eigenvalue ρ, and two complex, conjugate eigenvalues, α and $\overline{\alpha}$. In the latter case, $\mathcal{A}$ is also diagonalizable. We consider here the latter case. In Appendix C we show how to handle the former case, which is easier.

Thus, let us assume that the eigenvalues of $\mathcal{A}$ are $\rho \in \mathbb{A}\cap\mathbb{R}$ and $\alpha, \overline{\alpha} \in \mathbb{A}$. We can compute an invertible matrix $B \in \mathbb{A}^{3\times 3}$ such that $A = B^{-1} \begin{pmatrix} \rho & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \overline{\alpha} \end{pmatrix} B$, and the rows of B are the respective eigenvectors. Note that if w_α is an eigenvector of α, then $\overline{w_\alpha}$ is eigenvector of $\overline{\alpha}$, so we can write $B = \begin{pmatrix} w_\rho & w_\alpha & \overline{w_\alpha} \end{pmatrix}^T$. We now have that $\mathcal{A}^n = B^{-1} \begin{pmatrix} \rho^n & 0 & 0 \\ 0 & \alpha^n & 0 \\ 0 & 0 & \overline{\alpha}^n \end{pmatrix} B$ for every $n \in \mathbb{N}$. By analyzing the structure of B and B^{-1}, it is not hard to verify that every entry of $\mathcal{A}^n$ is a linear combination of $\alpha^n, \overline{\alpha}^n$ and ρ^n such that the coefficients of α^n and $\overline{\alpha}^n$ are conjugates, and the coefficient of ρ^n is real. That is, for every $1 \leq i, j \leq 3$ it holds that $(\mathcal{A}^n)_{i,j} = c_{i,j}\alpha^n + \overline{c_{i,j}}\,\overline{\alpha}^n + d_{i,j}\rho^n$ for coefficients $c_{i,j} \in \mathbb{A}$ and $d_{i,j} \in \mathbb{A} \cap \mathbb{R}$ (independent of n).

Consider a vector $x = u + \lambda v \in E$. We can write $\mathcal{A}^n x = \mathcal{A}^n u + \lambda \mathcal{A}^n v$, and observe that for $1 \leq i \leq 3$ we have $(\mathcal{A}^n u)_i = (c_{i,1}u_1 + c_{i,2}u_2 + c_{i,3}u_3)\alpha^n + \overline{(c_{i,1}u_1 + c_{i,2}u_2 + c_{i,3}u_3)}\,\overline{\alpha}^n + (d_{i,1}u_1 + d_{i,2}u_2 + d_{i,3}u_3)\rho^n$, and a similar structure holds for $\mathcal{A}^n v$. By renaming the coefficients, we can write $(\mathcal{A}^n u + \lambda \mathcal{A}^n v)_i = f_i\alpha^n + \overline{f_i}\,\overline{\alpha}^n + g_i\rho^n + \lambda(h_i\alpha^n + \overline{h_i}\,\overline{\alpha}^n + k_i\rho^n)$ where $f_i, h_i \in \mathbb{A}$ and $g_i, k_i \in \mathbb{A} \cap \mathbb{R}$ for $1 \leq i \leq 3$.

We can now formulate the problem as follows: does there exist a number $n \in \mathbb{N}$ such that the following first-order sentence is true: $\exists \lambda, \mu, \nu : 0 \le \lambda, \mu, \nu \le 1 \wedge \mu + \nu \le 1 \wedge$

$$\bigwedge_{i=1}^{3} \big(f_i \alpha^n + \overline{f_i}\, \overline{\alpha}^n + g_i \rho^n + \lambda (h_i \alpha^n + \overline{h_i}\, \overline{\alpha}^n + k_i \rho^n) = s_i + \mu t_i + \nu r_i \big) \tag{3}$$

As mentioned in Section 2.2, we can convert (3) to an equivalent, quantifier-free sentence. Since our reasoning requires this equivalent sentence to have a special structure, we must explicitly remove the quantifiers. This is done in Appendix B using Fourier-Motzkin quantifier elimination [11], where we conclude the following.

▶ **Theorem 4.** *There exist constants M, M' such that the sentence (3) is equivalent to a disjunction $\bigvee_{i=1}^{M} \mathrm{Sys}_i$ where for every $1 \le i \le M$, $\mathrm{Sys_i}$ is a conjunction of at most M' expressions of the form*

$$A\alpha^{2n} + \overline{A}\overline{\alpha}^{2n} + B\alpha^n \rho^n + \overline{B}\overline{\alpha}^n \rho^n + C\rho^{2n} + D|\alpha|^{2n} + E\alpha^n + \overline{E}\overline{\alpha}^n + F\rho^n + G \bowtie 0 \tag{4}$$

where $\bowtie \in \{>, =\}$, $A, B, E \in \mathbb{A}$, and $C, D, F, G \in \mathbb{A} \cap \mathbb{R}$. Moreover, the description of Sys *is polynomial in $\|I\|$ (the description length of the input).*

4 Solving the System

This section constitutes the main technical challenge of the paper, namely to decide whether there exists $n \in \mathbb{N}$ such that the disjunction presented in Theorem 4 is true. We refer to such an n as a *solution* for the disjunction.

We first note that it is enough to consider each system in the disjunction separately. Indeed, since the number of systems is bounded, independent of the input, we can try to solve each one separately. Our goal is then to decide, given a system Sys of expressions as per Theorem 4, whether there exists a solution $n \in \mathbb{N}$ that satisfies all the expressions simultaneously.

We divide our analysis to two cases. First we handle the (straightforward) case where $\frac{\alpha}{|\alpha|}$ is a root of unity. We then proceed to consider the more involved case, where $\frac{\alpha}{|\alpha|}$ is not a root of unity.

4.1 The case where $\frac{\alpha}{|\alpha|}$ is a root of unity

Suppose that $\frac{\alpha}{|\alpha|}$, denoted γ, is a root of unity. We can now treat (4) as

$$\begin{aligned} &|\alpha|^{2n} A\gamma^{2n} + |\alpha|^{2n}\overline{A}\overline{\gamma}^{2n} + |\alpha|^n B\gamma^n \rho^n + |\alpha|^n \overline{B}\overline{\gamma}^n \rho^n + C\rho^{2n} + D|\alpha|^{2n} \\ &+ |\alpha|^n E\gamma^n + |\alpha|^n \overline{E}\overline{\gamma}^n + F\rho^n + G \bowtie 0 \end{aligned}$$

Let d be the order of γ, then γ^2 is also a root of unity of order at most d. Thus, there are at most d^2 possible values for (γ^n, γ^{2n}), determined by the pair $(n \bmod d, 2n \bmod d)$. We can now treat the expression as d^2 expressions of real-algebraic sums of exponentials. We show that $d \le \deg(\gamma)^2$, so these can be solved in **PSPACE** using standard techniques of asymptotic analysis, by considering the coefficients and the moduli of α and ρ (see Appendix D for details).

4.2 The case where $\frac{\alpha}{|\alpha|}$ is not a root of unity

When $\gamma = \frac{\alpha}{|\alpha|}$ is not a root of unity, things are more involved. Nonetheless, we prove the following theorem.

▶ **Theorem 5.** *The problem of deciding whether a system* Sys *of expressions of the form* (4) *has a solution, is in* **PSPACE***.*

Before proving the theorem, we need some definitions. In the following, we assume w.l.o.g. that $\rho > 0$. Indeed, if $\rho < 0$ then we can divide into two cases according to the parity of n, and solve each separately (note that $\rho \neq 0$ since the matrix $\mathcal{A}$ is invertible).

For an expression of the form (4), we obtain its *normalized expression* by dividing it by $(\max\{|\alpha|^2, |\alpha|\rho, \rho^2, |\alpha|, |\rho|\})^n$ (and such that the coefficient of the element we divide by is nonzero). Thus, the normalized expression is of the form

$$A\gamma^{2n} + \overline{A}\overline{\gamma}^{2n} + B\gamma^n + \overline{B}\overline{\gamma}^n + C + r(n) \bowtie 0, \tag{5}$$

with $\gamma \in \mathbb{A}$ such that $|\gamma| = 1$ and γ is not a root of unity, $A, B \in \mathbb{A}$ and $C \in \mathbb{A} \cap \mathbb{R}$ are not all 0, and $r(n) = \sum_{l=1}^{m} D_l\beta_l^n + \overline{D_l\beta_l}^n$, where $|\beta_l| < 1$ for every $1 \leq l \leq m$, and $0 \leq m \leq 4$ (note that for uniformity we treat real numbers in $r(n)$ as a sum of complex conjugates). For every $1 \leq l \leq m$, β_l is a quotient of two elements from the set $\{\alpha, \alpha^2, \rho, \rho^2, \alpha\rho\}$. Since α and ρ are eigenvalues of $\mathcal{A}$, $\deg(\alpha), \deg(\rho)$ are $\|\mathcal{A}\|^{O(1)}$. Thus, by Section 2.1, $\deg(\beta_l) = \|\mathcal{A}\|^{O(1)}$, and $H(\beta_l) = 2^{\|\mathcal{A}\|^{O(1)}}$.

Since γ is not a root of unity, then $\{\gamma^n : n \in \mathbb{N}\}$ is dense in the unit circle. With this motivation in mind, we define, for a normalized expression, its *dominant function* $f : \mathbb{C} \to \mathbb{R}$ as $f(z) = Az^2 + \overline{A}\overline{z}^2 + Bz + \overline{B}\overline{z} + C$. Observe that (5) is now equivalent to $f(\gamma^n) + r(n) \bowtie 0$.

The following lemma is our main technical tool in proving Theorem 5.

▶ **Lemma 6.** *Consider a normalized expression as in* (5)*. Let* $\|I\|$ *be its encoding length, and let* f *be its dominant function. Then there exists* $N \in \mathbb{N}$ *computable in polynomial time in* $\|I\|$ *with* $N = 2^{\|I\|^{O(1)}}$ *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 $f(n) + r(n) = 0$, then $n \leq N$. The proof of Lemma 6 relies on the following lemma from [18], which is itself a consequence of the Baker-Wüstholz Theorem [2].

▶ **Lemma 7** ([18])**.** *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* $|\zeta^n - \xi| > \frac{1}{n^{(\|\zeta\| + \|\xi\|)^D}}$*.*

We now turn to prove Lemma 6. The following synopsis contains the main ideas. The full proof can be found in Appendix E.

Proof (Synopsis): Since $\{\gamma^n : n \in \mathbb{N}\}$ is dense on the unit circle, we consider $f(z)$ for z in the unit circle. In the full proof, we show that $\{z : f(z) = 0 \wedge |z| = 1\}$ contains at most four points $\{z_1, \ldots, z_4\}$, whose coordinates are algebraic. 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 $N_1 \in \mathbb{N}$ such that $\gamma^n \notin \{z_1, \ldots, z_4\}$ for every $n > N_1$. Moreover, by Lemma D.1 in [6], we have that $N_1 = k^{O(1)}$, where $k = \|\gamma\| + \sum_{j=1}^{4} \|z_j\|$, and N_1 can be computed in polynomial time in k. Then, by Lemma 7, there exists a constant $D \in \mathbb{N}$ such that for every $n \geq N_1$ and $1 \leq j \leq 4$ we have that $|\gamma^n - z_j| > \frac{1}{n^{(k^D)}}$. 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. While proving that this holds in general is not very difficult, note that we also need the bound on N in the statement of the Lemma to be effectively computable and to be $2^{\|I\|^{O(1)}}$, which complicates things significantly.

We consider the function $g : (-\pi, \pi] \to \mathbb{R}$ defined by $g(x) = f(e^{ix})$. Explicitly, we have $g(x) = 2|A|\cos(2x+\theta_A)+2|B|\cos(x+\theta_B)+C$ where $\theta_A = \arg(A)$ and $\theta_B = \arg(B)$. By the above, g has at most four roots, denoted $\varphi_1, \dots, \varphi_4$. We now show that there exist $N_2 \in \mathbb{N}$ and a non-negative polynomial $p(n)$ such that $f(\gamma^n) = g(\arg(\gamma^n)) > \frac{1}{p(n)}$ for every $n > N_2$. For every $1 \le j \le 4$ consider the first non-zero Taylor polynomial T_j of g around φ_j. In Lemma 9 we show that the degree of such approximations is at most 3. We show that there exists $\epsilon_1 > 0$ such that for every $x \in (\varphi_j - \epsilon_1, \varphi_j + \epsilon)$ it holds that (1) $|g(x) - T_j(x)| \le \frac{1}{2}|T_j(x)|$, (2) g is monotone on either side of φ_j, and (3) T is monotone with the same tendency of g (see Figure 1 for an illustration). In Lemma 10 we also show that crucially, we can require ϵ_1 to be efficiently computable and $\frac{1}{\epsilon} = 2^{n^{O(1)}}$.

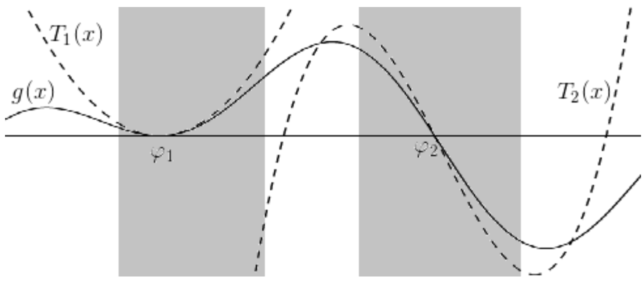

Figure 1 $g(x)$ and two Taylor polynomials: $T_1(x)$ around φ_1 and $T_2(x)$ around φ_2. The shaded regions show where requirements (1)–(3) hold, which determine ϵ_1. Observe that for T_1, the most restrictive requirement is $|g(x) - T_1(x)| \le \frac{1}{2}T_1(x)$, whereas for T_2 the restriction is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

Consider $n \in \mathbb{N}$ such that $\gamma^n \in \bigcup_{j=1}^4 (\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$ and such that $n > N_1$, then as we have seen above, $\frac{1}{n^{(k^D)}} < |\gamma^n - z_j|$. But $|\gamma^n - z_j| < |\arg(\gamma^n) - \varphi_j|$ (since the euclidean distance is smaller than the arc length), so $|\arg(\gamma^n) - \varphi_j| > \frac{1}{n^{(k^D)}}$. From requirements (1) and (2) of ϵ_1, we get that $|g(\arg(\gamma^n))| \ge \frac{1}{2}|T_j(\gamma^n)|$ and from the monotonicity of T_j in the neighbourhood of φ_j (requirement (3)), we have that $\frac{1}{2}|T_j(\gamma^n)| > \frac{1}{2}\min\left\{|T_j(\varphi_j + \frac{1}{n^{(k^D)}})|, |T_j(\varphi_j - \frac{1}{n^{(k^D)}})|\right\}$, from which we conclude that $|g(\arg(\gamma^n))| > \frac{1}{p(n)}$ for some non-negative polynomial p. Moreover, we can compute the representation of p in polynomial time.

Finally, for $x \notin \bigcup_{j=1}^4 (\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$, we have that $|g(x)|$ is bounded from below by a constant. Our careful accounting of $\|\epsilon_1\|$ in Lemma 10 allows us to compute this bound, and show that it is not too small.

The last step in the proof is to show that $r(n)$ decreases fast enough such that $r(n) < \frac{1}{p(n)}$ for every $n > N_3$ for some large enough $N_3 \in \mathbb{N}$. Clearly this holds eventually, since $r(n)$ decreases exponentially. However, we also need a bound on the size of N_3, which requires more effort. Recall that $r(n) = \sum_{l=1}^m D_l\beta_l^n + \overline{D_l\beta_l}^n$. By applying The root separation bound (2) from Section 2.1 to $1 - |\beta_l|$, we compute $\epsilon \in (0,1)$ and $N_3 \in \mathbb{N}$ such that $\frac{1}{\epsilon}$ and N_3 are $2^{\|I\|^{O(1)}}$, and for every $n > N_3$ it holds that $|r(n)| < (1-\epsilon)^n$. Using this, we can find $N_4 \in \mathbb{N}$ such that $N_4 = 2^{\|I\|^{O(1)}}$ and $|r(n)| < \frac{1}{p(n)}$ for all $n > N_4$, from which we can conclude the proof. ◀

We are now ready to prove Theorem 5

Proof. For every expression in Sys, let f be the corresponding function as per Lemma 6, and compute its respective bound N. If $\bowtie$ is "=", then by Lemma 6, if the equation is satisfiable for $n \in \mathbb{N}$, then $n < N$.

If all the $\bowtie$ are "$>$", then for each such inequality compute $\{z : f(z) > 0\}$. 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 in polynomial time using Theorem 2.

Thus, it remains to check whether there exists a solution $n < N$. Recall that $N = 2^{\|I\|^{O(1)}}$. Thus, in order to check whether the system is solved for $n < N$, we need to compute, e.g., α^{2n}, whose representation is exponential in $\|I\|$, so a naive implementation would take exponential space.

Instead, we take a similar approach to [8]: by representing numbers as arithmetic circuits, deciding the positivity (or testing for 0 equality) can be done using an oracle to **PosSLP**, which by [1] is in the counting hierarchy. By first guessing $n < N$, the problem can be solved in $\mathbf{NP}^{\mathbf{PosSLP}}$, which is contained in **PSPACE**. ◀

5 Conclusions

5.1 Proof of Theorem 1

We conclude by giving an explicit proof of Theorem 1: Given polytopes P_1 and P_2 and a matrix $\mathcal{A}$, if $\mathcal{A}$ is singular, we first apply (in polynomial time) the reduction in Section 3.2. Thus, we can assume $\mathcal{A}$ is invertible. Next, if P_1 or P_2 are unbounded, for each unbounded face F we proceed as follows: decompose F as per Lemma 3, so $F = \bigcup_{i=1}^{m} A_i$, and recall that iterating over the A_i's can be done in **PSPACE**. In each iteration, consider an edge E of P_1 and a face F of P_2 (both of which may belong to sets A_i as above). Formulate the first-order sentence (3) in Section 3.3, and apply Theorem 4 to obtain an equivalent disjunction of systems $\bigvee_{i=1}^{M} \mathrm{Sys_i}$, where M is constant. Then, for each system $\mathrm{Sys_i}$, check in **PSPACE** whether it has a solution, using either Section 4.1 or Theorem 5. If no solution was found, check in **PSPACE** whether a vertex of P_1 collides with P_2, using the algorithm in [8]. Then, if still no solution is found, repeat the same procedure by interchanging the roles of P_1 and P_2, and considering the matrix $\mathcal{A}^{-1}$ instead of $\mathcal{A}$. The correctness and complexity of this procedure follow from the proofs of the respective theorems.

5.2 Discussion

This paper studies an extension of the Orbit Problem, in which the input is existentially quantified over a polytope, and the target is a polytope. The importance of this work is twofold: from a practical perspective, we provide an algorithm for deciding the termination of linear while loops with affine guards, up to dimension three, when the input is not fixed. From a more theoretical perspective, and as already pointed out by Kannan and Lipton in [14], the Orbit Problem and its variants are closely related to long-standing open problems such as the *Skolem Problem*, and various number-theoretic problems. It is therefore useful and compelling to push the borders of decidability, in order to identify the core of the remaining difficulties, and to eventually hopefully overcome them.

Finally, as discussed in Section 1, the problem at hand can be viewed as a particular case of the Orbit Problem in dimension six where the target is a semi-algebraic set. As the general problem is known to be hard even in dimension four, our work here suggests that interesting and useful fragments are tractable even in high dimensions.

References

1 Eric Allender, Peter Bürgisser, Johan Kjeldgaard-Pedersen, and Peter Bro Miltersen. On the complexity of numerical analysis. *SIAM Journal on Computing*, 38(5):1987–2006, 2009.

2 Alan Baker and Gisbert Wüstholz. Logarithmic forms and group varieties. *J. reine angew. Math*, 442(19-62):3, 1993.

3 Saugata Basu, Richard Pollack, and Marie-Francoise Roy. *Algorithms in real algebraic geometry*, volume 20033. Springer, 2005.

4 Amir M Ben-Amram and Samir Genaim. Ranking functions for linear-constraint loops. *Journal of the ACM (JACM)*, 61(4):26, 2014.

5 Mark Braverman. Termination of integer linear programs. In *International Conference on Computer Aided Verification*, pages 372–385. Springer, 2006.

6 Ventsislav Chonev, Joël Ouaknine, and James Worrell. On the complexity of the orbit problem. *arXiv preprint arXiv:1303.2981*, 2013.

7 Ventsislav Chonev, Joël Ouaknine, and James Worrell. The orbit problem in higher dimensions. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 941–950. ACM, 2013.

8 Ventsislav Chonev, Joël Ouaknine, and James Worrell. The polyhedron-hitting problem. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 940–956. SIAM, 2015.

9 Henri Cohen. *A course in computational algebraic number theory*, volume 138. Springer Science & Business Media, 2013.

10 George 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.

11 Jean Baptiste Joseph Fourier. Solution d'une question particuliere du calcul des inégalités. *Nouveau Bulletin des Sciences par la Société philomatique de Paris*, 99:100, 1826.

12 Michael A Harrison. Lectures on linear sequential machines. Technical report, DTIC Document, 1969.

13 Ravindran Kannan and Richard J Lipton. The orbit problem is decidable. In *Proceedings of the twelfth annual ACM symposium on Theory of computing*, pages 252–261. ACM, 1980.

14 Ravindran Kannan and Richard J Lipton. Polynomial-time algorithm for the orbit problem. *Journal of the ACM (JACM)*, 33(4):808–821, 1986.

15 David Lee and Mihalis Yannakakis. Online minimization of transition systems. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 264–274. ACM, 1992.

16 Maurice Mignotte. Some useful bounds. In *Computer algebra*, pages 259–263. Springer, 1983.

17 Joël Ouaknine, João Sousa Pinto, and James Worrell. On termination of integer linear loops. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 957–969. SIAM, 2015.

18 Joël Ouaknine and James Worrell. Ultimate positivity is decidable for simple linear recurrence sequences. In *International Colloquium on Automata, Languages, and Programming*, pages 330–341. Springer, 2014.

19 Victor Y Pan. Optimal and nearly optimal algorithms for approximating polynomial zeros. *Computers & Mathematics with Applications*, 31(12):97–138, 1996.

20 James 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.

21 Terence Tao. *Structure and randomness: pages from year one of a mathematical blog*. American Mathematical Soc., 2008.

22 Alfred Tarski. A decision method for elementary algebra and geometry. 1951.

A Reduction to the Invertible Case

Recall that we are given polytopes $P, R \subseteq \mathbb{R}^3$ and a matrix $\mathcal{A} \in (\mathbb{A} \cap \mathbb{R})^{3\times 3}$, where 0 is an eigenvalue of $\mathcal{A}$ with multiplicity 1. As discussed in Section 3.2, we can write $\mathcal{A} = D^{-1}\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix} D$ where D is an invertible matrix with real-algebraic entries, and $B \in (\mathbb{A} \cap \mathbb{R})^{2\times 2}$ is also invertible.

Then, for every $x \in P$ and $n \in \mathbb{N}$ it holds that $\mathcal{A}^n x \in R$ iff $D^{-1}\begin{pmatrix}0 & 0\\ 0 & B^n\end{pmatrix} Dx \in R$ iff $\begin{pmatrix}0 & 0\\ 0 & B^n\end{pmatrix} Dx \in DR$ (where $DR = \{Dv : v \in R\}$). Observe that the first coordinate of $\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix} Dx$ is 0 for every vector x, and Consider the set $\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix} DP = \left\{\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix} Dx : x \in P\right\}$. We can write

$$\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix} DP = \left\{\begin{pmatrix}0\\ x_1\\ x_2\end{pmatrix} : \begin{pmatrix}x_1\\ x_2\end{pmatrix} \in P'\right\}$$

where $P' \subseteq \mathbb{R}^2$ is the intersection of $\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix} DP$ with the $[yz]$ plane, in the standard basis $\left\{\begin{pmatrix}1\\ 0\end{pmatrix}, \begin{pmatrix}0\\ 1\end{pmatrix}\right\}$.

Now, for $n \geq 1$, we get that there exists $x \in P$ such that $\mathcal{A}^n x \in R$ iff there exists $x \in P$ such that

$$\begin{pmatrix}0 & 0\\ 0 & B^{n-1}\end{pmatrix}\begin{pmatrix}0 & 0\\ 0 & B\end{pmatrix} Dx \in DR$$

iff there exists $x' \in P'$ such that $B^{n-1}x' \in \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\left(DR \cap sp(\left\{\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}, \begin{pmatrix}0\\ 0\\ 1\end{pmatrix}\right\})\right)$.

Since all the intersections and matrices above can be computed in polynomial time, and since the intersections above are polytopes, we conclude that if $\mathcal{A}$ is singular, we can reduce the dimension of the problem.

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

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

B Quantifier Elimination

In this section we eliminate quantifiers from the expression

$$\exists \lambda, \mu, \nu : 0 \leq \lambda, \mu, \nu \leq 1 \wedge \mu + \nu \leq 1 \wedge \bigwedge_{i=1}^{3} \left(f_i \alpha^n + \overline{f_i}\, \overline{\alpha}^n + g_i \rho^n + \lambda (h_i \alpha^n + \overline{h_i}\, \overline{\alpha}^n + k_i \rho^n) = s_i + \mu t_i + \nu r_i \right) \tag{6}$$

using the Fourier-Motzkin quantifier-elimination algorithm.

We start by recalling the Fourier-Motzkin algorithm. Given a set of linear inequalities in the variables x (which we want to eliminate), isolate x in each equation. Then, for each pair of equations of the form $x \leq \text{expression}_1$ and $x \geq \text{expression}_2$, add the inequality $\text{expression}_1 \geq \text{expression}_2$ (with analogous rules for strict inequalities). After doing so for every relevant pair of inequalities, remove all original inequalities involving x. The Fourier-Motzkin Theorem states that the new system is satisfiable iff the original system is satisfiable. Note that the new system is also a system of linear inequalities in the original variables.

By repeating this process for all variables, we end up with an equivalent, variable-free system of inequalities.

For the purpose of proving Theorem 4, we need some assumptions on the coefficients of the resulting inequalities, in order to have the form described in Theorem 4. We thus analyze in some detail the specific application of Fourier-Motzkin elimination to our setting.

B.1 Proof of Theorem 4

We start by explicitly writing down the expressions we consider in 6. We think of "$=$" as a pair of "$\geq$" and "$\leq$" inequalities.

$$\begin{aligned}
&\lambda \leq 1\\
&\lambda \geq 0\\
&\mu \leq 1\\
&\mu \geq 0\\
&\nu \leq 1\\
&\nu \geq 0\\
&\mu + \nu \leq 1\\
&\overline{\alpha}^n \overline{f_1} + \alpha^n f_1 + \rho^n g_1 + \lambda \left(\overline{\alpha}^n \overline{h_1} + \alpha^n h_1 + \rho^n k_1 \right) - (\nu r_1 + s_1 + \mu t_1) = 0\\
&\overline{\alpha}^n \overline{f_2} + \alpha^n f_2 + \rho^n g_2 + \lambda \left(\overline{\alpha}^n \overline{h_2} + \alpha^n h_2 + \rho^n k_2 \right) - (\nu r_2 + s_2 + \mu t_2) = 0\\
&\overline{\alpha}^n \overline{f_3} + \alpha^n f_3 + \rho^n g_3 + \lambda \left(\overline{\alpha}^n \overline{h_3} + \alpha^n h_3 + \rho^n k_3 \right) - (\nu r_3 + s_3 + \mu t_3) = 0
\end{aligned}$$

We make the following observations on the structure of the system.

▶ Observation 8. The coefficients of the system above satisfy the following.

1. The coefficients of ν do not depend on α, ρ or n.
2. The coefficients of μ do not depend on α, ρ or n.
3. The coefficients of λ are either constant, or of the form $A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n$, for some $A \in \mathbb{A}$ and $B \in \mathbb{R} \cap \mathbb{A}$ (that is, the coefficients of α^n and $\overline{\alpha}^n$ are conjugates, and the coefficient of ρ^n is real)
4. The free coefficients of the form $A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n + C$, for some $A \in \mathbb{A}$ and $B, C \in \mathbb{R} \cap \mathbb{A}$

We eliminate ν first. By Observation 8.1, after isolating ν (which involves dividing by the coefficient of ν), Observations 8.2, 8.3, and 8.4 still hold. Thus, after eliminating ν and aggregating the coefficients of μ and λ, Observations 8.2, 8.3, and 8.4 still hold, and Observation 8.1 is irrelevant, since ν was eliminated.

By Observation 8.2, following the same reasoning for eliminating μ results in a system of inequalities in λ that satisfies Observations 8.3 and 8.4.

It now remains to eliminate λ. Note that here, even isolating λ is not trivial. Indeed, in order to divide by a coefficient $A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n$, we need to know its sign (and whether it is 0). Thus, at this point in the elimination, we split the system into a disjunction of systems, where in each system we add an assumption on the sign of $A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n$. Thus, an inequality of the form $(A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n)\lambda \leq \mathit{expression}$ will yield a disjunction of three systems:

- $\lambda \leq \frac{\mathit{expression}}{A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n} \wedge A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n > 0$
- $\lambda \geq \frac{\mathit{expression}}{A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n} \wedge A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n < 0$
- $0 \leq \mathit{expression} \wedge A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n = 0$

After constructing these systems and combining the inequalities according to the algorithm, we multiply by a common denominator to get a system of inequalities without variables. In these inequalities, we multiply expressions of the form of Observation 8.4 by either constants, or by expressions of the form of Observation 3. Thus, end up with expressions of either the form of Observation 8.4, or of the form

$$(A\alpha^n + \overline{A}\overline{\alpha}^n + B\rho^n)(A'\alpha^n + \overline{A'}\overline{\alpha}^n + B'\rho^n + C') = AA'\alpha^{2n} + \overline{AA'}\overline{\alpha}^{2n} + \\ (AB' + A'B)\alpha^n\rho^n + (\overline{A}B' + \overline{A'}B)\overline{\alpha^n}\rho^n + BB'\rho^{2n} + (A\overline{A'} + \overline{A}A')|\alpha|^{2n} + C'A\alpha^n + C'\overline{A}\overline{\alpha}^n + C'B\rho^n$$

Finally, by renaming the coefficients, and by adding a constant term, both the latter form and that of Observation 8.3 are as described in Theorem 4. Finally, we split every nonstrict inequality to a disjunction of an equality and a strict inequality, and distribute the conjunction over them.

We note that the numbers of systems and equations are bounded by constants, since the removal does not depend on the coefficients, but only on the form of the expressions. ◀

C The case of only real eigenvalues

In this section we consider the case where the matrix $\mathcal{A}$ has only real eigenvalues, denoted ρ_1, ρ_2, ρ_3. In this case, by converting $\mathcal{A}$ 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. $\mathcal{A} = B^{-1}\begin{pmatrix}\rho_1 & 0 & 0\\ 0 & \rho_2 & 0\\ 0 & 0 & \rho_3\end{pmatrix}B$, in which case $\mathcal{A}^n = B^{-1}\begin{pmatrix}\rho_1^n & 0 & 0\\ 0 & \rho_2^n & 0\\ 0 & 0 & \rho_3^n\end{pmatrix}B$.
2. $\mathcal{A} = B^{-1}\begin{pmatrix}\rho_1 & 1 & 0\\ 0 & \rho_2 & 0\\ 0 & 0 & \rho_3\end{pmatrix}B$ with $\rho_1 = \rho_2$, in which case $\mathcal{A}^n = B^{-1}\begin{pmatrix}\rho_1^n & n\rho_1^{n-1} & 0\\ 0 & \rho_1^n & 0\\ 0 & 0 & \rho_3^n\end{pmatrix}B$.
3. $\mathcal{A} = B^{-1}\begin{pmatrix}\rho_1 & 1 & 0\\ 0 & \rho_2 & 1\\ 0 & 0 & \rho_3\end{pmatrix}B$ with $\rho_1 = \rho_2 = \rho_3$, in which case $\mathcal{A}^n = B^{-1}$ $\begin{pmatrix}\rho_1^n & n\rho_1^{n-1} & \frac{1}{2}n(n-1)\rho_1^{n-2}\\ 0 & \rho_1^n & n\rho_1^{n-1}\\ 0 & 0 & \rho_1^n\end{pmatrix}B$.

We consider here the latter case, as the first two are similar and simpler. We start by following the lines of Section 3.3. That is, we formulate the problem as a first-order sentence, and proceed to remove the quantifiers as per Appendix B. Consider $n \in \mathbb{N}$ and a vector v, then we can write

$$\mathcal{A}^n v = B^{-1} \begin{pmatrix} \rho_1^n & n\rho_1^{n-1} & \frac{1}{2}n(n-1)\rho_1^{n-2} \\ 0 & \rho_1^n & n\rho_1^{n-1} \\ 0 & 0 & \rho_1^n \end{pmatrix} Bv = a\rho_1^n + bn\rho_1^{n-1} + cn(n-1)\rho_1^{n-2}$$

Thus, the formulation of the first-order sentence 3 in Section 3.3 takes a similar form in this case, and after applying quantifier elimination, we end up with a disjunction as per Theorem 4, where the expressions in each system are of the form

$$A\rho_1^{2n} + Bn\rho_1^{2n} + Cn^2\rho_1^{2n} + D\rho_1^n + En\rho_1^n + Fn^2\rho_1^n + G \bowtie 0$$

Assuming $\rho_1 > 0$ (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 different components, such that either for every $n > N$ the equation holds, or for every $n > N$ it does not hold. This is done in a similar manner to the proof of Theorem 5. Thus, either we determine that a solution exists since all the expressions are satisfied for large enough n, or we need to check the solutions up to N, which can be done in **PSPACE** (as in the proof of Theorem 5).

D The case where $\frac{\alpha}{|\alpha|}$ is a root of unity

Let $\gamma = \frac{\alpha}{|\alpha|}$. We assume that γ is a root of unity. Thus, there exists $d \in \mathbb{N}$ such that $\gamma^d = 1$. After obtaining the systems of expressions as per Theorem 4, each expression 4 can be written as

$$|\alpha|^{2n}A\gamma^{2n}+|\alpha|^{2n}\overline{A}\overline{\gamma}^{2n}+|\alpha|^n B\gamma^n\rho^n+|\alpha|^n\overline{B}\overline{\gamma}^n\rho^n+C\rho^{2n}+D|\alpha|^{2n}+|\alpha|^n E\gamma^n+|\alpha|^n\overline{E}\overline{\gamma}^n+F\rho^n+G \bowtie 0$$

Observe that γ^2 is also a root of unity of order at most d. Thus, for every $n \in \mathbb{N}$ it holds that $(\gamma^n, \gamma^{2n}) = (\gamma^{n \bmod d}, \gamma^{2n \bmod d})$. Consider the set $V = \{(n \bmod d, 2n \bmod d) : n \in \mathbb{N}\}$, and note that $|V| \leq d^2$. For every $(k, k') \in V$, let $N_{(k,k')}$ be the minimal number such that $(n \bmod d, 2n \bmod d) = (k, k')$. Observe that $\{n \in \mathbb{N} : (n \bmod d, 2n \bmod d) = (k, k')\} = \{N_{(k,k')} + m|V| : m \in \mathbb{N}\}$. For each system Sys of expressions, we construct $|V|$ systems $\left\{\mathrm{Sys}_{(k,k')}\right\}_{(k,k')\in V}$ such that $\mathrm{Sys}_{(k,k')}$ is obtained from Sys by replacing, in every expression, γ^n with γ^k, replacing γ^{2n} with $\gamma^{k'}$, and replacing n in the remaining powers by $N_{(k,k')}+m|V|$. By pushing constants into the coefficients and renaming $\alpha^|V| = \beta$ and $\rho^|V| = \delta$, the expression above can be written as

$$|\beta|^{2m}A\gamma^{k'}+|\beta|^{2m}\overline{A}\overline{\gamma}^{k'}+|\beta|^m B\gamma^k\delta^m+|\beta|^m\overline{B}\overline{\gamma}^k\delta^m+C\delta^{2m}+D|\beta|^{2m}+|\beta|^m E\gamma^k+|\beta|^m\overline{E}\overline{\gamma}^k+F\delta^m+G \bowtie 0$$

This becomes

$$2\mathrm{Re}(A\gamma^{k'})|\beta|^{2m} + 2\mathrm{Re}(B\gamma^k)|\beta|^m\delta^m + C\delta^{2m} + D|\beta|^{2m} + 2\mathrm{Re}(E\gamma^k)|\beta|^m + F\delta^m + G \bowtie 0$$

These expressions contain only real-algebraic constants, and thus the system can be solved in similar techniques as those of Appendix C.

Finally, we show that the number of systems is polynomial, by showing that $d \leq \deg(\gamma)^2$. The proof appears in [14], 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 $d \leq \deg(\gamma)^2$.

E Proof of Lemma 6

By identifying $\mathbb{C}$ with $\mathbb{R}^2$ (where $z = x + iy$ is identified with (x, y)) we identify f with the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\begin{aligned}
f(x,y) =& A(x+iy)^2 + \overline{A(x+iy)^2} + B(x+iy) + \overline{B(x+iy)} + C \\
=& 2\mathrm{Re}(A(x+iy)^2) + 2\mathrm{Re}(B(x+iy)) + C \\
=& 2\mathrm{Re}((\mathrm{Re}(A) + i\mathrm{Im}(A))(x^2 - y^2 + i2xy)) + 2\mathrm{Re}((\mathrm{Re}(B) + i\mathrm{Im}(B))(x+iy)) + C \\
=& 2(\mathrm{Re}(A)(x^2-y^2) - \mathrm{Im}(A)2xy) + 2(\mathrm{Re}(B)x - \mathrm{Im}(B)y) + C
\end{aligned}$$

Since $\{\gamma^n : n \in \mathbb{N}\}$ is dense on the unit circle, our interest in f is also about the unit circle. Since f is a polynomial with algebraic coefficients, we can find in polynomial time a description of the algebraic set $\{(x,y) : f(x,y) = 0 \wedge x^2 + y^2 = 1\}$. Note that since the coefficients of x^2 and y^2 in f are either both 0, or they differ in their sign, then this set is not the entire unit circle. Therefore, by Bézout's Theorem, this set is discrete and consists of at most 4 points. Indeed, this set is the intersection of distinct contours of bivariate quadratic polynomials, so it corresponds to the roots of a polynomial of degree at most 4. Let $\{(x_1,y_1),\ldots,(x_4,y_4)\}$ be these points, and let $z_1 = x_1 + iy_1, \ldots, z_4 = x_4 + iy_4$ be the respective complex numbers. Note that these points have algebraic coordinates, so $z_1, \ldots, z_4$ are algebraic numbers. Moreover, since these numbers are attained as the roots of a polynomial of degree 4 whose coefficients are polynomial in those of f, then we have that $\|z_1\|,\ldots,\|z_4\|$ are polynomial in $\|f\|$. Note that if $A = B = 0$, then $C \neq 0$ by our assumption, and there are no roots. Thus, we assume for now that A and B are not both 0. We remove this assumption after we are done handling the roots.

Since γ is not a root of unity, then in particular, for every $n_1 \neq n_2 \in \mathbb{N}$ it holds that $\gamma^{n_1} \neq \gamma^{n_2}$. Thus, there exists $N_1 \in \mathbb{N}$ such that $\gamma^n \notin \{z_1,\ldots,z_4\}$ for every $n > N_1$. Moreover, by [6], we have that $N_1 = k^{O(1)}$, where $k = \|\gamma\| + \sum_{j=1}^4 \|z_j\|$, and N_1 can be computed in polynomial time in k. Then, by Lemma 7, there exists a constant $D \in \mathbb{N}$ such that for every $n \geq N_1$ and $1 \leq j \leq 4$ we have that $|\gamma^n - z_j| > \frac{1}{n^{(k^D)}}$.

Let $\theta_A = \arg(A)$, $\theta_B = \arg(B)$, and $\varphi_j = \arg(z_j)$ for every $1 \leq j \leq 4$. We assume w.l.o.g. that the angles φ_j are all in $(-\pi,\pi)$. Otherwise (if one of the angles is exactly π), we shift the domain such that all the angles are in the interior. Define $g : (-\pi,\pi] \to \mathbb{R}$ by $g(x) = f(e^{ix})$, so that $g(x) = 2|A|\cos(2x+\theta_A) + 2|B|\cos(x+\theta_B) + C$. Our next step is to show that $|g|$ is bounded from below by a polynomial. More precisely, we will show that $|g|$ is bounded from below in neighbourhoods of the roots of g, and give a lower bound on the value of $|g|$ outside these neighbourhoods. Technically, we will use the Taylor polynomials of g to obtain these bounds. For every $1 \leq j \leq 4$, 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 $T_j(x) = \frac{g^{(d_j)}}{d_j!}(x-\varphi)^{d_j}$. We now show that in fact, the degrees of these polynomials are at most three.

▶ **Lemma 9.** *$d_j \leq 4$ for every $1 \leq j \leq 3$.*

Proof. It is enough to show that at every point where $g(x) = 0$, at least one of the first three derivatives of g is non-zero. Assume by way of contradiction that the first three derivatives are all 0 at x, and $g(x) = 0$, then we have

$$\begin{aligned}
g(x) =& 2|A|\cos(2x+\theta_A) + 2|B|\cos(x+\theta_B) = 0 \\
g'(x) =& -4|A|\sin(2x+\theta_A) - 2|B|\sin(x+\theta_B) = 0 \\
g''(x) =& -8|A|\cos(2x+\theta_A) - 2|B|\cos(x+\theta_B) = 0 \\
g^{(3)}(x) =& 16|A|\sin(2x+\theta_A) + 2|B|\sin(x+\theta_B) = 0
\end{aligned}$$

Pairing the odd and even derivatives, this can be written as

$$\begin{pmatrix} 2|A| & 2|B| \\ -8|A| & -2|B| \end{pmatrix}\begin{pmatrix} \cos(2x+\theta_A) \\ \cos(x+\theta_B) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} -4|A| & -2|B| \\ 16|A| & 2|B| \end{pmatrix}\begin{pmatrix} \sin(2x+\theta_A) \\ \sin(x+\theta_B) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

If either $|A|$ or $|B|$ are 0 (but not both, as per our assumption above), then clearly either the first or second derivatives are always nonzero (since this is a single trigonometric function). If $|A|, |B| \neq 0$, then the matrices are invertible, so it must hold that $\sin(2x+\theta_A) = \sin(x+\theta_B) = \cos(2x+\theta_A) = \cos(x+\theta_B) = 0$, which clearly has no solution. ◀

Thus, $T_j(x)$ is a polynomial of degree at most three, with $T_j(\varphi_j) = 0$. We remark that T_j is computable in polynomial time (in $\|f\|$), as the coefficients are polynomial in $\|A\|, \|B\|, \|\gamma\|$.

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 $M_j = \max_{x\in[-\pi,\pi]} \left\{g^{(d_j+1)}(x)\right\} \leq 32|A| + 2|B|$ (where g is extended naturally to the domain $[-\pi, \pi]$).

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

1. $|g(x) - T_j(x)| \leq \frac{1}{2}|T_j(x)|$ for every $x \in (\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$.
2. $\mathrm{sign}(g'(x))$ does not change in $(\varphi_j, \varphi_j + \epsilon_1)$ nor in $(\varphi_j - \epsilon_1, \varphi_j)$.
3. $\mathrm{sign}(g'(x)) = \mathrm{sign}(T_j'(x))$ 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 $\theta_j \in (-\pi, \pi)$ for all $1 \leq j \leq 4$.

An ϵ_1 as above exists since $T_j(x)$ is of degree d_j, whereas the $|g(x) - T(x)|$ is of degree $d_j + 1$, since there are only finitely many points where $g'(x) = 0$, and since $T'(x)$ is the Taylor polynomial of degree $d_j - 1$ of $g'(x)$ around φ_j, so by bounding the distance $|g'(x) - T'(x)|$ we can conclude the third requirement (see Figure 1 for an illustration). For the following, we need also to compute ϵ_1, we thus proceed with the following lemma.

▶ **Lemma 10.** *ϵ_1 can be computed in polynomial time in $\|f\|$, and $\frac{1}{\epsilon} = 2^{n^{O(1)}}$.*

Proof. We start with Condition 2, and compute $\delta_1 > 0$ such that $\mathrm{sign}(g'(x))$ 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 $g(x) = f(e^{ix})$, then we have $g'(x) = f'(e^{ix})ie^{ix}$. Since f' is a polynomial with algebraic coefficients, then $F = \{z : |z| = 1 \wedge f'(z)iz = 0\}$ consists of algebraic numbers whose degree and height are polynomial in those of A and B, and we have that $\{x : g'(x) = 0\} = \{\arg(z) : z \in F\}$. By similar arguments as those by which we found the roots of f on the unit circle, we can conclude that F contains at most four points. Thus, it is enough to set δ_1 such that $\left(\bigcup_{j=1}^{4}(\varphi_j - \delta_1, \varphi_j) \cup (\varphi_j, \varphi_j + \delta_1)\right) \cap F = \emptyset$. By Equation (2), 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 $f'(z)iz$. Thus, $1/|z - z'|$ is $2^{\|f\|^{O(1)}}$, and has a polynomial description. Since $|\arg(z) - \arg(z')| > |z - z'|$, we conclude that by setting $\delta_1 = \min |z - z'| : z \neq z' \in F/3$, and it holds that $\frac{1}{\delta_1}$ has a polynomial description in $\|f\|$, and δ_1 satisfies the required condition.

We now proceed to handle Condition 1, and 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 $T_j(x) = \frac{g^{(d_j)}}{d_j!}(x - \varphi_j)^{d_j}$. Note that this case is more challenging than Condition 2, as unlike $g(x) = f(e^{ix})$, the polynomial $T_j(x)$ has potentially transcendental coefficients (namely φ_j). In order to ignore the absolute value, assume $T_j(x) \geq g(x) \geq \frac{1}{2}T_j(x) > 0$ in an interval $(\varphi_j, \varphi_j + \xi)$ for some $\xi > 0$ (the other cases are treated similarly). Then, the inequality above becomes $g(x) - \frac{1}{2}T_j(x) \geq 0$. Since the degree of T_j is d_j, then by the definition of T_j, the first $d_j - 1$ derivatives of g in φ_j vanish. Define $h(x) = g(x) - \frac{1}{2}T_j(x)$, then we have $h(\varphi_j) = 0$, $h'(\varphi_j) = 0, \ldots, h^{(d_j-1)}(\varphi_j) = 0$ and

$h^{(d_j)}(\varphi_j) = g^{(d_j)}(\varphi_j) - \frac{1}{2}g^{(d_j)}(\varphi_j) = \frac{1}{2}g^{(d_j)}(\varphi_j)$. By our assumption, $T_j(x) \geq \frac{1}{2}T_j(x)$ for $x \in (\varphi, \varphi+\xi)$, so $h^{(d_j)} > 0$. In addition, recall that $|h^{(d_j+1)}(x)| = |g^{(d_j+1)}(x)| \leq M_j \leq 64|A| + 2|B|$ for every $x \in [-\pi, \pi]$. Thus, by writing the d_j-th Taylor expansion of $h(x)$ around φ, we have that $h(x) = h^{d_j}(\varphi)(x-\varphi)^{d_j} + R(x)$ where $|R(x)| \leq \frac{M_j}{(d_j+1)!}(x-\varphi)^{k+1}$. We thus have that for $x \in (\varphi, \varphi + \frac{g^{d_j}(\varphi)(d_j+1)}{2M_j})$ it holds that $h(x) \geq 0$. We can now set $\delta_2 = \frac{g^{d_j}(\varphi)(d_j+1)}{2M_j}$, which satisfies the required condition (or a similar δ_2 after analyzing the other cases).

Finally, we address Condition 3, and 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. 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 $x \in (\varphi_j - \delta_3, \varphi_j + \delta_3)$.

By setting $\epsilon_1 = \min\{\delta_1, \delta_2, \delta_3\}$, we conclude the proof. ◀

Conditions 1,2,3 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 $\bigcup_{j=1}^4(\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$ and for its complement. Then, taking their minimum and bounding it with a polynomial yields $p(n)$.

At this point we also drop the assumption that either A or B are nonzero. Indeed, if $A = B = 0$, then $C \neq 0$, and the above is trivial.

We start by considering the case where $\arg(\gamma^n) \in \bigcup_{j=1}^4(\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$. Recall that since γ is not a root of unity, then for every $n > N_1$ it holds that $\gamma^n \notin \{z_1, \ldots, z_4\}$. Then, by Lemma 7, for every $1 \leq j \leq 4$ and every $n \geq N_2 = \max\{N_1, 2\}$ we have $|\gamma^n - z_j| > \frac{1}{n^{(k^D)}}$. 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, $|\arg(\gamma^n) - \varphi_j| > \frac{1}{n^{(k^D)}}$, so either $\arg(\gamma^n) > \varphi_j + \frac{1}{n^{(k^D)}}$ or $\arg(\gamma^n) < \varphi_j - \frac{1}{n^{(k^D)}}$. Next, we have that if $\arg(\gamma^n) \in (\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$ for some $1 \leq j \leq 4$, then $|g(\arg(\gamma^n))| \geq \frac{1}{2}|T_j(\arg(\gamma^n))| \geq \frac{1}{2}\min\left\{|T_j(\varphi_j + \frac{1}{n^{(k^D)}})|, |T_j(\varphi_j - \frac{1}{n^{(k^D)}})|\right\}$, where the last inequality follows from condition 3 above, which implies that T_j is monotone with the same tendency as g.

Observe that $T_j(\varphi_j - \frac{1}{n^{(k^D)}}) = \frac{g^{(d_j)}(\varphi)}{d_j!}\frac{1}{n^{(k^D)}}$ and similarly $T_j(\varphi_j + \frac{1}{n^{(k^D)}}) = -\frac{g^{(d_j)}(\varphi)}{d_j!}\frac{1}{n^{(k^D)}}$ are both inverse polynomials. 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 $x \notin \bigcup_{j=1}^4(\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)$ we can compute in polynomial time a bound $B > 0$ such that $|g(x)| > B$. Indeed, $B = \min\left\{|g(x)| : x \in [-\pi, \pi] \setminus \bigcup_{j=1}^4(\varphi_j - \epsilon_1, \varphi_j + \epsilon_1)\right\}$ (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 $1 \leq j \leq 4$ (and by Lemma 10, $\|\epsilon_1\|$ 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.

Nest, 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 $r(n) = \sum_{l=1}^{m} D_l \beta_l^n + \overline{D_l \beta_l}^n$. Let $1 \leq l \leq m$, and consider β_l. Since β_l is algebraic, then so is $1 - |\beta_l|$. Indeed, $1 - |\beta_l| = 1 - \sqrt{\beta_l \overline{\beta_l}}$. Moreover, we get that $\deg(1 - |\beta_l|) \leq \deg(\beta_l)^4$ the root of a polynomial of degree at most $\deg(\beta_l)^4$, and of height polynomial in $H(\beta_l)$. Since $|\beta_l| < 1$, By applying Equation 2, we get $1 - |\beta_l| = |1 - |\beta_l|| > \frac{\sqrt{6}}{d^{(d+1)/2} H(\beta_l)^{d-1}}$ where $d = \deg(\beta_l)^{O(1)}$. Recall that $H(\beta_l) = 2^{\|I\|^{O(1)}}$. Thus, we can compute $\epsilon \in (0, 1)$ and $N_3 \in \mathbb{N}$ such that:

1. $\frac{1}{\epsilon} = 2^{\|I\|^{O(1)}}$
2. $N_3 = 2^{\|I\|^{O(1)}}$
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 $N_4 = 2^{\|I\|^{O(1)}}$) 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 sinc e $|g(\arg(\gamma^n))| > \frac{1}{p(n)}$ 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 $f(\gamma^n) < 0$.

Which concludes the proof of Lemma 6. ◀