On the Polytope Escape Problem for Continuous Linear Dynamical Systems

Joël Ouaknine
MPI-SWS and Oxford U.
joel@mpi-sws.org

João Sousa-Pinto
Oxford U.
jspinto@cs.ox.ac.uk

James Worrell
Oxford U.
jbw@cs.ox.ac.uk

ABSTRACT

The Polytope Escape Problem for continuous linear dynamical systems consists of deciding, given an affine function $f : \mathbb{R}^d \to \mathbb{R}^d$ and a convex polytope $\mathcal{P} \subseteq \mathbb{R}^d$, both with rational descriptions, whether there exists an initial point $\boldsymbol{x}_0$ in $\mathcal{P}$ such that the trajectory of the unique solution to the differential equation

$$\begin{cases} \dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t)) \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{cases}$$

is entirely contained in $\mathcal{P}$. We show that this problem is reducible in polynomial time to the decision version of linear programming with real algebraic coefficients. The latter is a special case of the decision problem for the existential theory of real closed fields, which is known to lie between **NP** and **PSPACE**. Our algorithm makes use of spectral techniques and relies, among others, on tools from Diophantine approximation.

CCS Concepts

•**Theory of computation** → **Timed and hybrid models**;

Keywords

Orbit Problem; Continuous Linear Dynamical Systems

1. INTRODUCTION

In ambient space $\mathbb{R}^d$, a *continuous linear dynamical system* is a trajectory $\boldsymbol{x}(t)$, where t ranges over the non-negative reals, defined by a differential equation $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ in which the function f is *affine* or *linear*. If the initial point $\boldsymbol{x}(0)$ is given, the differential equation uniquely defines the entire trajectory. (Linear) dynamical systems have been extensively studied in Mathematics, Physics, and Engineering, and more recently have played an increasingly important role in Computer Science, notably in the modelling and

analysis of cyber-physical systems; a recent and authoritative textbook on the matter is [2].

In the study of dynamical systems, particularly from the perspective of control theory, considerable attention has been given to the study of *invariant sets*, i.e., subsets of $\mathbb{R}^d$ from which no trajectory can escape; see, e.g., [10, 5, 3, 20]. Our focus in the present chapter is on sets with the dual property that *no trajectory remains trapped*. Such sets play a key role in analysing *liveness* properties in cyber-physical systems (see, for instance, [2]): discrete progress is ensured by guaranteeing that all trajectories (i.e., from any initial starting point) must eventually reach a point at which they 'escape' (temporarily or permanently) the set in question.

More precisely, given an affine function $f : \mathbb{R}^d \to \mathbb{R}^d$ and a convex polytope $\mathcal{P} \subseteq \mathbb{R}^d$, both specified using rational coefficients encoded in binary, we consider the *Polytope Escape Problem* which asks whether there is some point $\boldsymbol{x}_0$ in $\mathcal{P}$ for which the corresponding trajectory of the solution to the differential equation

$$\begin{cases} \dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t)) \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{cases}$$

is entirely contained in $\mathcal{P}$. Our main result is to show that this problem is decidable by reducing it in polynomial time to the decision version of linear programming with real algebraic coefficients, which itself reduces in polynomial time to deciding the truth of a sentence in the first-order theory of the reals: a problem whose complexity is known to lie between **NP** and **PSPACE** [9]. Our algorithm makes use of spectral techniques and relies among others on tools from Diophantine approximation.

It is interesting to note that a seemingly closely related problem, that of determining whether a given trajectory of a linear dynamical system ever hits a given hyperplane (also known as the *continuous Skolem Problem*), is not known to be decidable; see, in particular, [4, 12, 11]. When the target is instead taken to be a single point (rather than a hyperplane), the corresponding reachability question (known as the *continuous Orbit Problem*) can be decided in polynomial time [14].

2. MATHEMATICAL BACKGROUND

2.1 Kronecker's Theorem

Let $\mathbb{T}$ denote the group of complex numbers of modulus 1, with multiplication as group operation. Then the function $\phi : \mathbb{R} \to \mathbb{T}$ given by $\phi(x) = \exp(2\pi i x)$ is a homomorphism

from the additive group of real numbers to $\mathbb{T}$, with kernel the subgroup of integers.

Recall from [15] the following classical theorem of Kronecker on simultaneous inhomogeneous Diophantine approximation.

THEOREM 1 (KRONECKER). *Let* $\theta_1, \ldots, \theta_s$ *be real numbers such that the set* $\{\theta_1, \ldots, \theta_s, 1\}$ *is linearly independent over* $\mathbb{Q}$. *Then for all* $\psi_1, \ldots, \psi_s \in \mathbb{R}$ *and* $\varepsilon > 0$, *there exists a positive integer* n *and integers* $n_1, \ldots, n_s$ *such that*

$$|n\theta_1 - \psi_1 - n_1| < \varepsilon, \ldots, |n\theta_s - \psi_s - n_s| < \varepsilon.$$

We obtain the following simple corollary:

COROLLARY 2. *Let* $\theta_1, \ldots, \theta_s$ *be real numbers such that the set* $\{\theta_1, \ldots, \theta_s, 1\}$ *is linearly independent over* $\mathbb{Q}$. *Then*

$$\{(\phi(n\theta_1), \ldots, \phi(n\theta_s)) : n \in \mathbb{N}\}$$

is a dense subset of $\mathbb{T}^s$.

PROOF. Since ϕ is surjective, an arbitrary element of $\mathbb{T}^s$ can be written in the form $(\phi(\psi_1), \ldots, \phi(\psi_s))$ for some real numbers $\psi_1, \ldots, \psi_s$. Applying Kronecker's Theorem, we get that for all $\varepsilon > 0$, there exists a positive integer n and integers $n_1, \ldots, n_s$ such that

$$|n\theta_1 - \psi_1 - n_1| < \varepsilon, \ldots, |n\theta_s - \psi_s - n_s| < \varepsilon.$$

By continuity of ϕ it follows that $(\phi(\psi_1), \ldots, \phi(\psi_s))$ is a limit point of $\{(\phi(n\theta_1), \ldots, \phi(n\theta_s)) : n \in \mathbb{N}\}$. This establishes the result. $\square$

2.2 Laurent polynomials

A multivariate *Laurent polynomial* is a polynomial in positive and negative powers of variables $z_1, \ldots, z_s$ with complex coefficients. We are interested in Laurent polynomials of the special form

$$g = \sum_{j=1}^{k} \left(c_j z_1^{n_{1,j}} \ldots z_s^{n_{s,j}} + \overline{c_j} z_1^{-n_{1,j}} \ldots z_s^{-n_{s,j}} \right),$$

where $c_1, \ldots, c_k \in \mathbb{C}$ and $n_{1,1}, \ldots, n_{s,k} \in \mathbb{Z}$. We call such g *self-conjugate Laurent polynomials*. Notice that if $a_1, \ldots, a_s \in \mathbb{T}$ then $g(a_1, \ldots, a_s)$ is a real number, so we may regard g as a function from $\mathbb{T}^s$ to $\mathbb{R}$.

LEMMA 3. *Let* $g \in \mathbb{C}[z_1^{\pm 1}, \ldots, z_s^{\pm 1}]$ *be a self-conjugate Laurent polynomial that has no constant term. Given real numbers* $\theta_1, \ldots, \theta_s \in \mathbb{R}$ *such that* $\theta_1, \ldots, \theta_s, 1$ *are linearly independent over* $\mathbb{Q}$, *define a function* $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ *by*

$$f(t) = g(\phi(t\theta_1), \ldots, \phi(t\theta_s)).$$

Then either f *is identically zero, or*

$$\liminf_{n \to \infty} f(n) < 0,$$

where n *ranges over the nonnegative integers.*

PROOF. Recall that we may regard g as a function from $\mathbb{T}^s$ to $\mathbb{R}$. Now we consider the function $g \circ \phi^s : \mathbb{R}^s \to \mathbb{R}$,

$$(x_1, \ldots, x_s) \mapsto g(\phi(x_1), \ldots, \phi(x_s)).$$

We use an averaging argument to establish that either $g \circ \phi^s$ is identically zero on $\mathbb{R}^s$ or there exist $x_1^*, \ldots, x_s^* \in [0, 1]$ such that $g(\phi(x_1^*), \ldots, \phi(x_s^*)) < 0$.

Since $\int_0^1 \exp(2\pi i n x) dx = 0$ for all non-zero integers n, it holds that

$$\int_0^1 \ldots \int_0^1 g(\phi(x_1), \ldots, \phi(x_s)) dx_1 \ldots dx_s = 0.$$

Suppose that $g \circ \phi^s$ is not identically zero over $\mathbb{R}^s$ and hence not identically zero over $[0, 1]^s$. Then $g \circ \phi^s$ cannot be nonnegative on $[0, 1]^s$, since the integral over a set of positive measure of a continuous nonnegative function that is not identically zero must be strictly positive. We conclude that there must exist $(x_1^*, \ldots, x_s^*) \in [0, 1]^s$ such that $g(\phi(x_1^*), \ldots, \phi(x_s^*)) < 0$.

By assumption, $\theta_1, \ldots, \theta_s, 1$ are linearly independent over $\mathbb{Q}$. By Corollary 2 it follows that

$$\{(\phi(n\theta_1), \ldots, \phi(n\theta_s)) : n \in \mathbb{N}\}$$

is dense in $\mathbb{T}^s$ and hence has $(\phi(x_1^*), \ldots, \phi(x_s^*))$ as a limit point. Since $g \circ \phi^s$ is continuous, there are arbitrarily large $n \in \mathbb{N}$ for which

$$f(n) = g(\phi(n\theta_1), \ldots, \phi(n\theta_s)) \leq \tfrac{1}{2} g(\phi(x_1^*), \ldots, \phi(x_s^*)) < 0,$$

which proves the result. $\square$

Note that this proof could be made constructive by using an effective version of Kronecker's Theorem, as studied in [7] and [17], although we do not make use of this fact in the present paper.

We say that a self-conjugate Laurent polynomial g is *simple* if it has no constant term and each monomial mentions only a single variable. More precisely, g is simple if it can be written in the form

$$g = \sum_{j=1}^{k} c_j z_{i_j}^{n_j} + \overline{c_j} z_{i_j}^{-n_j},$$

where $c_1, \ldots, c_k \in \mathbb{C}$, $i_1, \ldots, i_k \in \{1, \ldots, s\}$, and $n_1, \ldots, n_k \in \mathbb{Z}$.

The following consequence of Lemma 3 will be key to proving decidability of the problem at hand. It is an extension of Lemma 4 from [6].

THEOREM 4. *Let* $g \in \mathbb{C}[z_1^{\pm 1}, \ldots, z_s^{\pm 1}]$ *be a simple self-conjugate Laurent polynomial and* $\theta_1, \ldots, \theta_s$ *non-zero real numbers. Then either*

$$g(\phi(t\theta_1), \ldots, \phi(t\theta_s)) = 0 \text{ for all } t \in \mathbb{R}$$

or

$$\liminf_{n \to \infty} g(\phi(n\theta_1), \ldots, \phi(n\theta_s)) < 0,$$

where n *ranges over the nonnegative integers.*

PROOF. Note that if $1, \theta_1, \ldots, \theta_s$ are linearly independent over $\mathbb{Q}$ then the result follows from Lemma 3. Otherwise, let $\{\theta_{i_1}, \ldots, \theta_{i_k}\}$ be a maximal subset of $\{\theta_1, \ldots, \theta_s\}$ such that $1, \theta_{i_1}, \ldots, \theta_{i_k}$ are linearly independent over $\mathbb{Q}$.

Then, for some $N \in \mathbb{N}$ and each j, one can write

$$N\theta_j = \left(m + \sum_{l=1}^{k} n_l \theta_{i_l} \right),$$

where $m, n_1, \ldots, n_k$ are integers that depend on j, whilst N

does not depend on j. It follows that for all j and $t \in \mathbb{R}$,

$$\phi(N\theta_j t) = \phi(mt) \cdot \prod_{l=1}^{k} \phi(n_l \theta_{i_l} t)$$

$$= \phi(t)^m \cdot \prod_{l=1}^{k} \phi(\theta_{i_l} t)^{n_l} \,.$$

In other words, for all $j \geq k+1$, $\phi(N\theta_j t)$ can be written as a product of positive and negative powers of the terms

$$\phi(t), \phi(\theta_{i_1} t), \dots, \phi(\theta_{i_k} t) \,.$$

It follows that there exists a self-conjugate Laurent polynomial $h \in \mathbb{C}[z_1^{\pm 1}, \dots, z_k^{\pm 1}]$, not necessarily simple, but with zero constant term, such that for all $t \in \mathbb{R}$,

$$g(\phi(N\theta_1 t), \dots, \phi(N\theta_s t)) = h(\phi(\theta_{i_1} t), \dots, \phi(\theta_{i_k} t)) \,.$$

Since $1, \theta_{i_1}, \dots, \theta_{i_k}$ are linearly independent over $\mathbb{Q}$, the result follows by applying Lemma 3 to h. $\quad\square$

2.3 Jordan Canonical Forms

Let $A \in \mathbb{Q}^{d \times d}$ be a square matrix with rational entries. The *minimal polynomial* of A is the unique monic polynomial $m(x) \in \mathbb{Q}[x]$ of least degree such that $m(A) = 0$. By the Cayley-Hamilton Theorem the degree of m is at most the dimension d of A. The set $\sigma(A)$ of eigenvalues is the set of roots of m. The *index* of an eigenvalue λ, denoted by $\nu(\lambda)$, is its multiplicity as a root of m. We use $\nu(A)$ to denote $\max_{\lambda \in \sigma(A)} \nu(\lambda)$: the maximum index over all eigenvalues of A. Given an eigenvalue $\lambda \in \sigma(A)$, we say that $\boldsymbol{v} \in \mathbb{C}^d$ is a *generalised eigenvector* of A if $\boldsymbol{v} \in \ker(A - \lambda I)^k$, for some $k \in \mathbb{N}$.

We denote by $\mathcal{V}_\lambda$ the subspace of $\mathbb{C}^d$ spanned by the set of generalised eigenvectors associated with some eigenvalue λ of A. We denote the subspace of $\mathbb{C}^d$ spanned by the set of generalised eigenvectors associated with some real eigenvalue by $\mathcal{V}^r$. We likewise denote the subspace of $\mathbb{C}^d$ spanned by the set of generalised eigenvectors associated to eigenvalues with non-zero imaginary part by $\mathcal{V}^c$.

It is well known that each vector $\boldsymbol{v} \in \mathbb{C}^d$ can be written uniquely as $\boldsymbol{v} = \sum_{\lambda \in \sigma(A)} \boldsymbol{v}_\lambda$, where $\boldsymbol{v}_\lambda \in \mathcal{V}_\lambda$. It follows that $\boldsymbol{v}$ can also be uniquely written as $\boldsymbol{v} = \boldsymbol{v}^r + \boldsymbol{v}^c$, where $\boldsymbol{v}^r \in \mathcal{V}^r$ and $\boldsymbol{v}^c \in \mathcal{V}^c$.

We can write any matrix $A \in \mathbb{C}^{d \times d}$ as $A = Q^{-1} J Q$ for some invertible matrix Q and block diagonal Jordan matrix $J = diag(J_1, \dots, J_N)$, with each block J_i having the following form:

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

Given a rational matrix A, its Jordan Normal Form $A = Q^{-1} J Q$ can be computed in polynomial time, as shown in [8].

Note that each vector $\boldsymbol{v}$ appearing as a column of the matrix Q^{-1} is a generalised eigenvector of A. We also note that the index $\nu(\lambda)$ of some eigenvalue λ corresponds to the dimension of the largest Jordan block associated with it.

One can obtain a closed-form expression for powers of block diagonal Jordan matrices, and use this to get a closed-form expression for exponential block diagonal Jordan matrices. In fact, if J_i is a $k \times k$ Jordan block associated with some eigenvalue λ, then

$$J_i^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-1} & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ 0 & \lambda^n & n\lambda^{n-1} & \cdots & \binom{n}{k-2}\lambda^{n-k+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n\lambda^{n-1} \\ 0 & 0 & 0 & \cdots & \lambda^n \end{pmatrix}$$

and

$$\exp(J_i t) = \exp(\lambda t) \begin{pmatrix} 1 & t & \cdots & \frac{t^{k-1}}{(k-1)!} \\ 0 & 1 & \cdots & \frac{t^{k-2}}{(k-2)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

In the above, $\binom{n}{j}$ is defined to be 0 when $n < j$.

PROPOSITION 5. *Let $\boldsymbol{v}$ lie in the generalised eigenspace $\mathcal{V}_\lambda$ for some $\lambda \in \sigma(A)$. Then $\boldsymbol{b}^T \exp(At)\boldsymbol{v}$ is a linear combination of terms of the form $t^n \exp(\lambda t)$.*

PROOF. Note that, if $A = Q^{-1} J Q$ and $J = diag(J_1, \dots, J_N)$ is a block diagonal Jordan matrix, then

$$\exp(At) = Q^{-1} \exp(Jt) Q$$

and

$$\exp(Jt) = diag(\exp(J_1 t), \dots, \exp(J_N t)) \,.$$

The result follows by observing that $Q\boldsymbol{v}$ is zero in every component other than those pertaining the block corresponding to the eigenspace $\mathcal{V}_\lambda$. $\quad\square$

In order to compare the asymptotic growth of expressions of the form $t^n \exp(\lambda t)$, for $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}_0$, we define $\prec$ to be the lexicographic order on $\mathbb{R} \times \mathbb{N}_0$, that is,

$$(\eta, j) \prec (\rho, m) \quad \text{iff} \quad \eta < \rho \text{ or } (\eta = \rho \text{ and } j < m) \,.$$

Clearly $\exp(\eta t) t^j = o(\exp(\rho t) t^m)$ as $t \to \infty$ if and only if $(\eta, j) \prec (\rho, m)$.

DEFINITION 1. *If $\boldsymbol{b}^T \exp(At)\boldsymbol{v}$ is not identically zero, the maximal $(\rho, m) \in \mathbb{R} \times \mathbb{N}_0$ (with respect to $\prec$) for which there is a term $t^m \exp(\lambda t)$ with $\Re(\lambda) = \rho$ in the closed-form expression for $\boldsymbol{b}^T \exp(At)\boldsymbol{v}$ is called* dominant *for $\boldsymbol{b}^T \exp(At)\boldsymbol{v}$.*

Before we can proceed, we shall need the following auxiliary result:

PROPOSITION 6. *Suppose that $\boldsymbol{v} \in \mathbb{R}^d$ and that*

$$\boldsymbol{v} = \sum_{\lambda \in \sigma(A)} \boldsymbol{v}_\lambda \,,$$

where $\boldsymbol{v}_\lambda \in \mathcal{V}_\lambda$. Then $\boldsymbol{v}_{\overline{\lambda}}$ and $\boldsymbol{v}_\lambda$ are component-wise complex conjugates.

PROOF. We start by observing that

$$\boldsymbol{0} = \boldsymbol{v} - \overline{\boldsymbol{v}} = \sum_{\lambda \in \sigma(A)} (\boldsymbol{v}_\lambda - \overline{\boldsymbol{v}_{\overline{\lambda}}}) \,. \tag{1}$$

But if $\boldsymbol{v}_\lambda \in \ker(A - \lambda I)^k$ then $\overline{\boldsymbol{v}_\lambda} \in \ker(A - \overline{\lambda} I)^k$, and hence $\overline{\boldsymbol{v}_{\overline{\lambda}}} \in \mathcal{V}_\lambda$. Thus each summand $\boldsymbol{v}_\lambda - \overline{\boldsymbol{v}_{\overline{\lambda}}}$ in (1) lies in $\mathcal{V}_\lambda$. Since $\mathbb{C}^d$ is a direct sum of the generalised eigenspaces of A, we must have $\boldsymbol{v}_\lambda = \overline{\boldsymbol{v}_{\overline{\lambda}}}$ for all $\lambda \in \sigma(A)$. $\square$

We now derive a corollary of Theorem 4.

COROLLARY 7. *Consider a function of the form* $h(t) = \boldsymbol{b}^T \exp(At)\boldsymbol{v}^c$, *where* $\boldsymbol{v}^c \in \mathcal{V}^c$, *with* $(\rho, m) \in \mathbb{R} \times \mathbb{N}_0$ *dominant. If* $h(t) \not\equiv 0$, *then we have*

$$-\infty < \liminf_{t \to \infty} \frac{h(t)}{\exp(\rho t) t^m} < 0 \,.$$

PROOF. Let

$$\Re(\sigma(A)) = \{\eta \in \mathbb{R} : \eta + i\theta \in \sigma(A), \text{ for some } \theta \in \mathbb{R}\} \,.$$

For each $\eta \in \Re(\sigma(A))$ define $\boldsymbol{\theta}_\eta = \{\theta \in \mathbb{R}_{>0} : \eta + i\theta \in \sigma(A)\}$. By abuse of notation, we also use $\boldsymbol{\theta}_\eta$ to refer to the vector whose coordinates are exactly the members of this set, ordered in an increasing way. We note that, due to Proposition 6 and Proposition 5, the following holds:

$$\boldsymbol{b}^T \exp(At)\boldsymbol{v}^c = \boldsymbol{b}^T \exp(At) \sum_{\eta \in \Re(\sigma(A))} \sum_{\theta \in \boldsymbol{\theta}_\eta} \boldsymbol{v}_{\eta+i\theta} + \boldsymbol{v}_{\eta-i\theta}$$
$$= \sum_{\eta \in \Re(\sigma(A))} \sum_{\theta \in \boldsymbol{\theta}_\eta} \boldsymbol{b}^T \exp(At)\boldsymbol{v}_{\eta+i\theta}$$
$$+ \overline{\boldsymbol{b}^T \exp(At)\boldsymbol{v}_{\eta+i\theta}}$$
$$= \sum_{\eta \in \Re(\sigma(A))} \sum_{j=0}^{\nu(A)-1} t^j \exp(\eta t) g_{(\eta,j)}(\exp(i\boldsymbol{\theta}_\eta t))$$

for some simple self-conjugate Laurent polynomials $g_{(\eta,j)}$. Note that

$$(\rho, m) = \max_{\prec} \{(\eta, j) \in \mathbb{R} \times \mathbb{N}_0 : g_{(\eta,j)}(\exp(i\boldsymbol{\theta}_\eta t)) \not\equiv 0\} \,.$$

The result then follows from Theorem 4 and the fact that

$$\liminf_{t \to \infty} \frac{h(t)}{\exp(\rho t) t^m} = \liminf_{t \to \infty} g_{(\rho,m)}(\exp(i\boldsymbol{\theta}_\rho t)) \,.$$

$\square$

2.4 Computation with Algebraic Numbers

In this section, we briefly explain how one can represent and manipulate algebraic numbers efficiently.

Any given algebraic number α can be represented as a tuple (p, a, ε), where $p \in \mathbb{Q}[x]$ is its minimal polynomial, $a = a_1 + a_2 i$, with $a_1, a_2 \in \mathbb{Q}$, is an approximation of α, and $\varepsilon \in \mathbb{Q}$ is sufficiently small that α is the unique root of p within distance ε of a. This is referred to as the standard or canonical representation of an algebraic number.

Let $f \in \mathbb{Z}[x]$ be a polynomial. The following root-separation bound, due to Mignotte [18], can be used to give a value of ε such that any disk of radius ε in the complex plane contains at most one root of f.

PROPOSITION 8. *Let* $f \in \mathbb{Z}[x]$. *If* α_1 *and* α_2 *are distinct roots of* f, *then*

$$|\alpha_1 - \alpha_2| > \frac{\sqrt{6}}{d^{(d+1)/2} H^{d-1}}$$

where d *and* H *are respectively the degree and height (maximum absolute value of the coefficients) of* f.

It follows that in the canonical representation (p, a, ε) of an algebraic number α, where p has degree d and height H, we may choose a_1, a_2, ϵ to have bit length polynomial in d and $\log H$.

Given canonical representations of two algebraic numbers α and β, one can compute canonical representations of $\alpha + \beta$, $\alpha\beta$, and α/β, all in polynomial time. More specifically, one can:

- factor an arbitrary polynomial with rational coefficients as a product of irreducible polynomials in polynomial time using the LLL algorithm, described in [16];

- compute an approximation of an arbitrary algebraic number accurate up to polynomially many bits in polynomial time, due to the work in [19];

- use the sub-resultant algorithm (see Algorithm 3.3.7 in [13]) and the two aforementioned procedures to compute canonical representations of sums, differences, multiplications, and quotient of two canonically represented algebraic numbers.

3. EXISTENTIAL FIRST-ORDER THEORY OF THE REALS

Let $\boldsymbol{x} = (x_1, \ldots, x_m)$ be a list of m real-valued variables, and let $\sigma(\boldsymbol{x})$ be a Boolean combination of atomic predicates of the form $g(\boldsymbol{x}) \sim 0$, where each $g(\boldsymbol{x})$ is a polynomial with integer coefficients in the variables $\boldsymbol{x}$, and $\sim$ is either $>$ or $=$. Tarski has famously shown that we can decide the truth over the field $\mathbb{R}$ of sentences of the form $\phi = Q_1 x_1 \cdots Q_m x_m \sigma(\boldsymbol{x})$, where Q_i is either $\exists$ or $\forall$. He did so by showing that this theory admits quantifier elimination (Tarski-Seidenberg Theorem [21]). The set of all true sentences of such form is called the first-order theory of the reals, and the set of all true sentences where only existential quantification is allowed is called the existential first-order theory of the reals. The complexity class $\exists\mathbb{R}$ is defined as the set of problems having a polynomial-time many-one reduction to the existential theory of the reals. It was shown in [9] that $\exists\mathbb{R} \subseteq PSPACE$.

We also remark that our standard representation of algebraic numbers allows us to write them explicitly in the first-order theory of the reals, that is, given $\alpha \in \mathbb{A}$, there exists a sentence $\sigma(x)$ such that $\sigma(x)$ is true if and only if $x = \alpha$. Thus, we allow their use when writing sentences in the first-order theory of the reals, for simplicity.

The decision version of linear programming with canonically-defined algebraic coefficients is in $\exists\mathbb{R}$, as the emptiness of a convex polytope can easily be described by a sentence of the form $\exists x_1 \cdots \exists x_n \sigma(\boldsymbol{x})$.

Finally, we note that even though the decision version of linear programming with rational coefficients is in P, allowing algebraic coefficients makes things more complicated. While it has been shown in [1] that this is solvable in time polynomial in the size of the problem instance and on the degree of the smallest number field containing all algebraic numbers in each instance, it turns out that in the problem at hand the degree of that extension can be exponential in the size of the input. In other words, the splitting field of the characteristic polynomial of a matrix can have a degree which is exponential in the degree of the characteristic polynomial.

4. THE POLYTOPE ESCAPE PROBLEM

The Polytope Escape Problem for continuous linear dynamical systems consists of deciding, given an affine function $f : \mathbb{R}^d \to \mathbb{R}^d$ and a convex polytope $\mathcal{P} \subseteq \mathbb{R}^d$, whether there exists an initial point $\boldsymbol{x}_0 \in \mathcal{P}$ for which the trajectory of the unique solution to the differential equation $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t)), \boldsymbol{x}(0) = \boldsymbol{x}_0, t \geq 0$, is entirely contained in $\mathcal{P}$. A starting point $\boldsymbol{x}_0 \in \mathcal{P}$ is said to be *trapped* if the trajectory of the corresponding solution is contained in $\mathcal{P}$, and *eventually trapped* if the trajectory of the corresponding solution contains a trapped point. Therefore, the Polytope Escape Problem amounts to deciding whether a trapped point exists, which in turn is equivalent to deciding whether an eventually trapped point exists.

The goal of this section is to prove the following result:

THEOREM 9. *The Polytope Escape Problem is polynomial-time reducible to the decision version of linear programming with algebraic coefficients.*

A d-dimensional instance of the Polytope Escape Problem is a pair $(f, \mathcal{P})$, where $f : \mathbb{R}^d \to \mathbb{R}^d$ is an affine function and $\mathcal{P} \subseteq \mathbb{R}^d$ is a convex polytope. In this formulation we assume that all numbers involved in the definition of f and $\mathcal{P}$ are rational.[1]

An instance $(f, \mathcal{P})$ of the Polytope Escape Problem is said to be *homogeneous* if f is a linear function and $\mathcal{P}$ is a convex polytope cone (in particular, $\boldsymbol{x} \in \mathcal{P}, \alpha > 0 \Rightarrow \alpha\boldsymbol{x} \in \mathcal{P}$).

The restriction of the Polytope Escape Problem to homogeneous instances is called the homogeneous Polytope Escape Problem.

LEMMA 10. *The Polytope Escape Problem is polynomial-time reducible to the homogeneous Polytope Escape Problem.*

PROOF. Let $(f, \mathcal{P})$ be an instance of the Polytope Escape Problem in $\mathbb{R}^d$, and write

$$f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{a} \text{ and } \mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^d : B_1\boldsymbol{x} > \boldsymbol{b}_1 \wedge B_2\boldsymbol{x} \geq \boldsymbol{b}_2\}.$$

Now define

$$A' = \begin{pmatrix} A & \boldsymbol{a} \\ \boldsymbol{0}^T & 0 \end{pmatrix}, B_1' = \begin{pmatrix} B_1 & -\boldsymbol{b}_1 \\ \boldsymbol{0}^T & 1 \end{pmatrix}, B_2' = \begin{pmatrix} B_2 & -\boldsymbol{b}_2 \end{pmatrix},$$

$$\mathcal{P}' = \left\{ \begin{pmatrix} \boldsymbol{x} \\ y \end{pmatrix} \in \mathbb{R}^{d+1} : B_1' \begin{pmatrix} \boldsymbol{x} \\ y \end{pmatrix} > \boldsymbol{0} \wedge B_2' \begin{pmatrix} \boldsymbol{x} \\ y \end{pmatrix} \geq \boldsymbol{0} \right\},$$

and

$$g\begin{pmatrix} \boldsymbol{x} \\ y \end{pmatrix} = A' \begin{pmatrix} \boldsymbol{x} \\ y \end{pmatrix}.$$

Then $(g, \mathcal{P}')$ is a homogeneous instance of the Polytope Escape Problem.

It is clear that $\boldsymbol{x}(t)$ satisfies the differential equation $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ if and only if $\begin{pmatrix} \boldsymbol{x}(t) \\ 1 \end{pmatrix}$ satisfies the differential equation $\begin{pmatrix} \dot{\boldsymbol{x}} \\ \dot{y} \end{pmatrix} = g\begin{pmatrix} \boldsymbol{x} \\ y \end{pmatrix} == \begin{pmatrix} A\boldsymbol{x} + y\boldsymbol{a} \\ 0 \end{pmatrix}$. In general, in any

---

[1] The assumption of rationality is required to justify some of our complexity claims (e.g., Jordan Canonical Forms are only known to be polynomial-time computable for matrices with rational coordinates). Nevertheless, our procedure remains valid in a more general setting, and in fact, the overall $\exists\mathbb{R}$ complexity of our algorithm would not be affected if one allowed real algebraic numbers when defining problem instances.

trajectory $\begin{pmatrix} \boldsymbol{x} \\ y \end{pmatrix}$ that satisfies this last differential equation, the y-component must be constant.

We claim that $(f, \mathcal{P})$ is a positive instance of the Polytope Escape Problem if and only if $(g, \mathcal{P}')$ is a positive instance. Indeed, if the point $\boldsymbol{x}_0 \in \mathbb{R}^d$ is trapped in $(f, \mathcal{P})$ then the point $\begin{pmatrix} \boldsymbol{x}_0 \\ 1 \end{pmatrix}$ is trapped in $(g, \mathcal{P}')$. Conversely, suppose that $\begin{pmatrix} \boldsymbol{x}_0 \\ y_0 \end{pmatrix}$ is trapped in $(g, \mathcal{P}')$. Then, since $B_1' \begin{pmatrix} \boldsymbol{x}_0 \\ y_0 \end{pmatrix} > \boldsymbol{0}$, we must have $y_0 > 0$. Scaling, it follows that $\begin{pmatrix} y_0^{-1}\boldsymbol{x}_0 \\ 1 \end{pmatrix}$ is also trapped in $(g, \mathcal{P}')$. This implies that $y_0^{-1}\boldsymbol{x}_0$ is trapped in $(f, \mathcal{P})$. $\square$

We remind the reader that the unique solution of the differential equation $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t)), \boldsymbol{x}(0) = \boldsymbol{x}_0, t \geq 0$, where $f(\boldsymbol{x}) = A\boldsymbol{x}$, is given by $\boldsymbol{x}(t) = \exp(At)\boldsymbol{x}_0$. In this setting, the sets of trapped and eventually trapped points are, respectively:

$$T = \{\boldsymbol{x}_0 \in \mathbb{R}^d : \forall t \geq 0, \exp(At)\boldsymbol{x}_0 \in \mathcal{P}\}$$

$$ET = \{\boldsymbol{x}_0 \in \mathbb{R}^d : \exists t \geq 0, \exp(At)\boldsymbol{x}_0 \in T\}$$

Note that both T and ET are convex subsets of $\mathbb{R}^d$.

LEMMA 11. *The homogeneous Polytope Escape Problem is polynomial-time reducible to the decision version of linear programming with algebraic coefficients.*

PROOF. Let $\boldsymbol{x}_0 = \boldsymbol{x}_0^r + \boldsymbol{x}_0^c$, where $\boldsymbol{x}_0^r \in \mathcal{V}^r$ and $\boldsymbol{x}_0^c \in \mathcal{V}^c$. We start by showing that if $\boldsymbol{x}_0$ lies in the set T of trapped points then its component $\boldsymbol{x}_0^r$ in the real eigenspace $\mathcal{V}^r$ lies in the set ET of eventually trapped points. Due to the fact that the intersection of finitely many convex polytopes is still a convex polytope, it suffices to prove this claim for the case when $\mathcal{P}$ is defined by a single inequality—say $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{b}^T\boldsymbol{x} \triangleright 0\}$, where $\triangleright$ is either $>$ or $\geq$.

We may assume that $\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^c$ is not identically zero, as in that case

$$\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0 \equiv \boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^r$$

and our claim holds trivially. Also, if $\boldsymbol{x}_0 \in T$, it cannot hold that

$$\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^r \equiv 0,$$

since $\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^c$ is negative infinitely often by Corollary 7.

Suppose that $\boldsymbol{x}_0 \in T$ and let (ρ, m) and (η, j) be the dominant indices for $\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^r$ and $\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^c$ respectively. Then by Proposition 5 we have

$$\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^r = \exp(\rho t)t^m(c + o(1)) \qquad (2)$$

as $t \to \infty$, where c is a non-zero real number. We will show that $c > 0$, from which it follows that $\boldsymbol{x}_0^r \in ET$.

It must hold that $(\eta, j) \preceq (\rho, m)$. Indeed, if $(\eta, j) \succ (\rho, m)$, then, as $t \to \infty$,

$$\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0 = \exp(\eta t)t^j \left(\underbrace{\frac{\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^c}{\exp(\eta t)t^j}}_{A} + o(1) \right),$$

but the limit inferior of the term A above is strictly negative by Corollary 7, contradicting the fact that $\boldsymbol{x}_0 \in T$.

If $(\eta, j) = (\rho, m)$, then, as $t \to \infty$,

$$\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0 = \exp(\rho t)t^m \left(c + \frac{\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^c}{\exp(\rho t)t^m} + o(1) \right),$$

and by invoking Corollary 7 as above, it follows that $c > 0$.

Finally, if $(\eta, j) \prec (\rho, m)$, then, as $t \to \infty$,

$$\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0^c = \exp(\rho t)t^m \cdot o(1), \qquad (3)$$

and hence, by (2) and (3), it follows that

$$\boldsymbol{b}^T \exp(At)\boldsymbol{x}_0 = \exp(\rho t)t^m \left(c + o(1) \right).$$

From the fact that $\boldsymbol{x}_0 \in T$ and that $c \neq 0$ we must have $c > 0$.

In all cases it holds that $c > 0$ and hence $\boldsymbol{x}_0^r \in ET$.

Having argued that $ET \neq \emptyset$ iff $ET \cap \mathcal{V}^r \neq \emptyset$, we will now show that the set $ET \cap \mathcal{V}^r$ is a convex polytope that we can efficiently compute. As before, it suffices to prove this claim for the case when $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{b}^T \boldsymbol{x} \triangleright 0\}$ (where $\triangleright$ is either $>$ or $\geq$).

In what follows, we let $[K]$ denote the set $\{0, \ldots, K-1\}$. We can write

$$\boldsymbol{b}^T \exp(At) = \sum_{(\eta,j) \in \sigma(A) \times [\nu(A)]} \exp(\eta t)t^j \boldsymbol{u}_{(\eta,j)}^T,$$

where $\boldsymbol{u}_{(\eta,j)}^T$ is a vector of coefficients.

Note that if $\boldsymbol{x} \in \mathcal{V}^r$ and $(\eta, j) \in (\sigma(A) \setminus \mathbb{R}) \times \mathbb{N}_0$, then $\boldsymbol{u}_{(\eta,j)}^T \boldsymbol{x} = 0$, as $\boldsymbol{u}_{(\eta,j)}^T \boldsymbol{x}$ is the coefficient of $t^j \exp(\eta t)$ in $\boldsymbol{b}^T \exp(At)\boldsymbol{x}$, and $\mathcal{V}^r$ is invariant under $\exp(At)$. Moreover,

$$ET \cap \mathcal{V}^r = (\mathcal{B} \cap \mathcal{C}) \cup \begin{cases} \{\boldsymbol{0}\} & \text{if } \triangleright \text{ is } \geq \\ \emptyset & \text{if } \triangleright \text{ is } > \end{cases}$$

where

$$\mathcal{B} = \bigcap_{(\eta,j) \in (\sigma(A) \setminus \mathbb{R}) \times [\nu(A)]} \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{u}_{(\eta,j)}^T \boldsymbol{x} = 0\}$$

$$\mathcal{C} = \bigcup_{(\eta,j) \in (\sigma(A) \cap \mathbb{R}) \times [\nu(A)]} \left[\{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{u}_{(\eta,j)}^T \boldsymbol{x} > 0\} \cap \right.$$

$$\left. \bigcap_{(\rho,m) \succ (\eta,j)} \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{u}_{(\rho,m)}^T \boldsymbol{x} = 0\} \right]$$

The set $ET \cap \mathcal{V}^r$ can be seen to be convex from the above characterisation. Alternatively, note that ET can be shown to be convex from its definition and that $\mathcal{V}^r$ is convex, therefore so must be their intersection. Thus $ET \cap \mathcal{V}^r$ must be a convex polytope whose definition possibly involves canonically-represented real algebraic numbers, and the Polytope Escape Problem reduces to testing this polytope for non-emptiness. $\square$

5. CONCLUSION

We have shown that the Polytope Escape Problem for continuous-time linear dynamical systems is decidable, and in fact, polynomial-time reducible to the decision problem for the existential theory of real closed fields. Given an instance of the problem $(f, \mathcal{P})$, with f an affine map, our decision procedure involves analysing the real eigenstructure of the linear operator $g(\boldsymbol{x}) := f(\boldsymbol{x}) - f(\boldsymbol{0})$. In fact, we showed that all complex eigenvalues could essentially be ignored for the purposes of deciding this problem.

Interestingly, the seemingly closely related question of whether a given single trajectory of a linear dynamical system remains trapped within a given polytope appears to be considerably more challenging and is not known to be decidable. In that instance, it seems that the influence of the complex eigenstructure cannot simply be discarded.

6. REFERENCES

[1] I. Adler and P. Beling. Polynomial algorithms for linear programming over the algebraic numbers. *Algorithmica*, 12(6):436–457, 1994.

[2] R. Alur. *Principles of Cyber-Physical Systems*. MIT Press, 2015.

[3] A. Bacciotti and L. Mazzi. Stability of dynamical polysystems via families of Lyapunov functions. *Jour. Nonlin. Analysis*, 67:2167–2179, 2007.

[4] P. C. Bell, J. Delvenne, R. M. Jungers, and V. D. Blondel. The continuous Skolem-Pisot problem. *Theor. Comput. Sci.*, 411(40-42):3625–3634, 2010.

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

[6] M. Braverman. Termination of integer linear programs. In *Proc. Intern. Conf. on Computer Aided Verification (CAV)*, volume 4144 of *LNCS*. Springer, 2006.

[7] D. Bridges and P. Schuster. A simple constructive proof of Kronecker's density theorem. *Elemente der Mathematik*, 61:152–154, 2006.

[8] J.-Y. Cai. Computing Jordan normal forms exactly for commuting matrices in polynomial time. *Int. J. Found. Comput. Sci.*, 5(3/4):293–302, 1994.

[9] J. Canny. Some algebraic and geometric computations in PSPACE. In *Proceedings of STOC'88*, pages 460–467. ACM, 1988.

[10] E. B. Castelan and J.-C. Hennet. On invariant polyhedra of continuous-time linear systems. *IEEE Transactions on Automatic Control*, 38(11):1680–85, 1993.

[11] V. Chonev, J. Ouaknine, and J. Worrell. On recurrent reachability for continuous linear dynamical systems. In *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, pages 515–524, 2016.

[12] V. Chonev, J. Ouaknine, and J. Worrell. On the Skolem Problem for continuous linear dynamical systems. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 100:1–100:13, 2016.

[13] H. Cohen. *A Course in Computational Algebraic Number Theory*. Springer-Verlag, 1993.

[14] E. Hainry. Reachability in linear dynamical systems. In *Logic and Theory of Algorithms, 4th Conference on Computability in Europe, CiE 2008, Athens, Greece, June 15-20, 2008, Proceedings*, pages 241–250, 2008.

[15] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, 1938.

[16] A. Lenstra, H. L. Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.

[17] G. Malajovich. An effective version of Kronecker's theorem on simultaneous diophantine approximation. Technical report, UFRJ, 1996.

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

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

[20] S. Sankaranarayanan, T. Dang, and F. Ivancic. A policy iteration technique for time elapse over template polyhedra. In *Proceedings of HSCC*, volume 4981 of *LNCS*. Springer, 2008.

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