On the Decidability of Membership in Matrix-exponential Semigroups

JOËL OUAKNINE and AMAURY POULY, Max Planck Institute for Software Systems
JOÃO SOUSA-PINTO and JAMES WORRELL, University of Oxford

We consider the decidability of the membership problem for matrix-exponential semigroups: Given $k \in \mathbb{N}$ and square matrices $A_1, \ldots, A_k, C$, all of the same dimension and with real algebraic entries, decide whether C is contained in the semigroup generated by the matrix exponentials $\exp(A_i t)$, where $i \in \{1, \ldots, k\}$ and $t \geq 0$. This problem can be seen as a continuous analog of Babai et al.'s and Cai et al.'s problem of solving multiplicative matrix equations and has applications to reachability analysis of linear hybrid automata and switching systems. Our main results are that the semigroup membership problem is undecidable in general, but decidable if we assume that $A_1, \ldots, A_k$ commute. The decidability proof is by reduction to a version of integer programming that has transcendental constants. We give a decision procedure for the latter using Baker's theorem on linear forms in logarithms of algebraic numbers, among other tools. The undecidability result is shown by reduction from Hilbert's Tenth Problem.

CCS Concepts: • **Theory of computation → Timed and hybrid models**;

Additional Key Words and Phrases: Linear forms in logarithms, matrix exponential, matrix reachability, matrix logarithms, commuting matrices, hybrid automata

ACM Reference format:
Joël Ouaknine, Amaury Pouly, João Sousa-Pinto, and James Worrell. 2019. On the Decidability of Membership in Matrix-exponential Semigroups. *J. ACM* 66, 3, Article 15 (May 2019), 24 pages.
https://doi.org/10.1145/3286487

1 INTRODUCTION

Reachability problems are a staple of theoretical computer science and algorithmic verification. In this article, our motivation stems from systems with both continuous variables and discrete control modes. Systems of this type permeate engineering and the physical sciences: examples include linear hybrid automata, switching systems, continuous-time Markov chains, and cyber-physical systems—see References [1, 2, 15, 16].

We focus on systems consisting of a finite number of discrete control modes (or states), having the property that the continuous variables of interest evolve in each mode according to some linear differential equation of the form $\dot{\boldsymbol{x}} = A\boldsymbol{x}$. Here $\boldsymbol{x}$ is a vector of continuous variables, and A is a square "rate" matrix of appropriate dimension. As is well known, in each mode the closed

Supported by ERC grant AVS-ISS (648701), and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — Projektnummer 389792660 — TRR 248.
Supported by EPSRC Fellowship EP/N008197/1.
Authors' addresses: J. Ouaknine and A. Pouly, Max Planck Institute for Software Systems, Saarland Informatics Campus, Campus E1 5, 66123 Saarbrücken, Germany; emails: {joel, pamaury}@mpi-sws.org; J. Sousa-Pinto and J. Worrell, Department of Computer Science, University of Oxford, 15 Parks Rd, Oxford OX1 3QD, UK; emails: {jspinto, jbw}@cs.ox.ac.uk.

0004-5411/2019/05-ART15
https://doi.org/10.1145/3286487

form solution $\boldsymbol{x}(t)$ to the differential equation admits a matrix-exponential representation of the form $\boldsymbol{x}(t) = \exp(At)\boldsymbol{x}(0)$. Thus, if such a system evolves through a series of k modes, where the ith mode has rate matrix A_i, then the overall effect on the initial configuration $\boldsymbol{x}(0)$ is determined by multiplying by the matrix $\prod_{i=1}^{k} \exp(A_i t_i)$, where $t_i \geq 0$ denotes the amount of time spent in the ith mode for $i = 1, \ldots, k$. A particularly interesting situation arises when the matrices A_i commute; in such cases the order in which the modes are visited (or, indeed, whether they are visited only once or several times) is immaterial, the only relevant data being the total time spent in each mode. Natural questions then arise as to what kinds of linear transformations can thus be achieved by such systems.

The following decision problem captures a fundamental reachability question in the setting described above. The *Matrix-Exponential Semigroup Membership Problem (MSMP)* asks, given square matrices $A_1, \ldots, A_k, C$, all of the same dimension and with real-algebraic entries, whether C is a member of the matrix semigroup generated by the set of matrices $\exp(A_i t)$ for $i \in \{1, \ldots, k\}$ and $t \geq 0$. The main result of this article shows decidability of this problem when the matrices $A_1, \ldots, A_k$ commute. We also show that the problem is undecidable in general.

1.1 Related Work

In the case of matrix-exponential semigroups with a single generator, the membership problem was shown to be decidable in Reference [12], and a polynomial-time procedure was subsequently given in Reference [9]. Further reachability problems in the setting of a single generator are considered in Reference [6].

There is a rich literature on membership problems for matrix semigroups, which can be seen as discrete analogs of the matrix-exponential semigroups studied in this article. Given square matrices $A_1, \ldots, A_k, C$, all of the same dimension, and with algebraic entries, the *Matrix Semigroup Membership Problem* consists in deciding whether the matrix C belongs to the multiplicative semigroup generated by $\{A_1, \ldots, A_k\}$. A related problem is to determine, given the multiplicative matrix equation $\prod_{i=1}^{k} A_i^{n_i} = C$, whether there is a solution $n_1, \ldots, n_k \in \mathbb{N}$. Both these problems have been shown to be undecidable in general [5, 23]. When the matrices $A_1, \ldots, A_k$ commute, the two problems are equivalent and known to be decidable [3]. Prior to Reference [3], the case $k = 1$ was shown to be decidable in Reference [17], and the case $k = 2$ was first shown to be decidable in Reference [8]. The case $k = 2$ without commutativity assumptions was shown to be decidable in Reference [5]. See Reference [13] for a relevant survey and Reference [10] for some interesting related problems.

1.2 Structure of the Paper

To prove decidability of the Matrix-Exponential Semigroup Membership Problem in case the generating matrices $A_1, \ldots, A_k$ all commute, we successively reformulate the problem until we eventually arrive at a version of integer programming that has transcendental constants. We show how to solve the latter problem using Baker's theorem on linear forms in logarithms of algebraic numbers and results in simultaneous Diophantine approximation.

If matrices A and B commute, then so do their exponentials $\exp(A)$ and $\exp(B)$. Using this fact, it immediately follows that in the case of commuting matrices $A_1, \ldots, A_k$, the Matrix-Exponential Semigroup Membership Problem is equivalent to the following problem:

Definition 1.1. Given square matrices $A_1, \ldots, A_k$ and C, all of the same dimension and with real algebraic entries, the *Matrix-Exponential Equation Problem (MEP)* consists in determining whether there exist real numbers $t_1, \ldots, t_k \geq 0$ such that $\prod_{i=1}^{k} \exp(A_i t_i) = C$.

The Matrix-Exponential Equation Problem corresponds to reachability in a switching system in which the number of mode switches is fixed *a priori*. Using results in linear algebra about commuting matrices and basic facts about matrix logarithms, the special case of the Matrix-Exponential Equation Problem for commuting matrices can be reduced to the following problem concerning systems of linear-exponential equations:

Definition 1.2. An instance of the *Linear-Exponential Equation Problem* (LEP) consists of a system of linear-exponential equations in nonnegative real variables $t_1, \ldots, t_k$:

$$\begin{gathered} \exp\left(\sum_{i=1}^{k} \lambda_i^{(j)} t_i\right) = c_j \exp(d_j) \quad (j = 1, \ldots, m) \\ A\boldsymbol{t} = \boldsymbol{b}, \end{gathered} \tag{1}$$

where $k, m \in \mathbb{N}$ and the constants $\lambda_i^{(j)}, c_j, d_j$ and the entries of the matrix A and vector $\boldsymbol{b}$ are (possibly complex) algebraic numbers. The problem asks to determine whether there exist nonnegative real numbers $t_1, \ldots, t_k \geq 0$ that satisfy the system (1).

By taking logarithms we can reduce the Linear-Exponential Equation Problem to a version of integer programming with transcendental constants. More precisely, we consider real constants that can be written as *linear forms in logarithms of algebraic numbers*, that is, numbers of the form $\alpha_0 + \sum_{i=1}^{m} \alpha_i \log(\beta_i)$, where $\alpha_0, \ldots, \alpha_m$ are algebraic numbers and $\beta_1, \ldots, \beta_m$ are non-zero algebraic numbers (with both the α_i and β_i being possibly complex) and log is a fixed branch of the complex logarithm function.

Definition 1.3. An instance of the *Algebraic-Logarithmic Integer Programming Problem* (ALIP) consists of a finite system of inequalities of the form $\pi\Lambda\boldsymbol{x} \leq \boldsymbol{e}$, where Λ is a matrix with real algebraic entries and where the coordinates of $\boldsymbol{e}$ are real linear forms in logarithms of algebraic numbers. The problem asks to determine whether such a system admits a solution $\boldsymbol{x}$ in integers.

Our strategy to decide the Matrix-Exponential Semigroup Membership Problem in the commutative case is to establish the following chain of reductions:

$$\text{MSMP} \equiv \text{MEP} \leq \text{LEP} \leq \text{ALIP}$$

and, finally, to show decidability of ALIP using results in transcendence theory and Diophantine approximation.

In the general setting (i.e., without assuming commutativity) we show that both the Matrix-Exponential Semigroup Membership Problem and the Matrix-Exponential Equation Problem are undecidable. The proof is by reduction from Hilbert's Tenth Problem. We also show undecidability of variants of these problems that involve vector reachability and hyperplane reachability.

Definition 1.4. Given square matrices $A_1, \ldots, A_k$ and vectors $\boldsymbol{x}, \boldsymbol{y}$, all of the same dimension and with real algebraic entries, the *Matrix-Exponential Vector Reachability Problem* (MVRP) consists of deciding whether there exists a matrix C in the semigroup generated by the set $\{\exp(A_i t) : t \geq 0, i = 1, \ldots, k\}$ such that $C\boldsymbol{x} = \boldsymbol{y}$. The *Matrix-Exponential Hyperplane Reachability Problem* (MHRP) asks whether there exists such a matrix C satisfying $\boldsymbol{x}^T C \boldsymbol{y} = 0$.

Our undecidability results will be established by the following chain of reductions (where HTP refers to Hilbert's Tenth Problem):

$$\text{HTP} \leq \text{MEP} \leq \text{MSMP} \leq \text{MVRP} \leq \text{MHRP}.$$

2 MATHEMATICAL BACKGROUND

In this section, we review some results in linear algebra, convex geometry, and number theory. We also introduce some specialised mathematical notation that will be needed in the subsequent development.

2.1 Linear Algebra

2.1.1 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 of A. The set $\sigma(A)$ of eigenvalues is the set of zeros of m, also known as the *spectrum* of A. The *index* of an eigenvalue λ, denoted by $\nu(\lambda)$, is its multiplicity as a zero of m. We use $\nu(A)$ to denote $\max_{\lambda\in\sigma(A)} \nu(\lambda)$: the maximum index over all eigenvalues of A. An eigenvalue λ is said to be *simple* if $\nu(\lambda) = 1$ and *repeated* otherwise. Given an eigenvalue $\lambda \in \sigma(A)$, we say that $\boldsymbol{v} \in \mathbb{C}^d$ is a *generalised eigenvector* of A if $\boldsymbol{v} \in \ker (A - \lambda I)^m$, for some $m \in \mathbb{N}$.

For each eigenvalue λ of A, we denote the subspace of $\mathbb{C}^d$ spanned by the set of generalised eigenvectors associated with λ by $\mathcal{V}_\lambda$. 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 with some non-real eigenvalue by $\mathcal{V}^c$.

Based on the Jordan decomposition of A, as described later on in this subsection, each vector $\boldsymbol{v} \in \mathbb{C}^d$ can be written uniquely as

$$\boldsymbol{v} = \sum_{\lambda\in\sigma(A)} \boldsymbol{v}_\lambda, \tag{2}$$

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 will need the following result:

Proposition 2.1. *Suppose that $\boldsymbol{v} \in \mathbb{R}^d$ and that $\boldsymbol{v} = \sum_{\lambda\in\sigma(A)} \boldsymbol{v}_\lambda$, where $\boldsymbol{v}_\lambda \in \mathcal{V}_\lambda$. For all $\lambda \in \sigma(A)$, it holds that $\boldsymbol{v}_{\overline{\lambda}}$ and $\boldsymbol{v}_\lambda$ are component-wise complex conjugates.*

Proof. Since A is real, $\boldsymbol{v}_\lambda \in \ker (A - \lambda I)^m$ implies that $\overline{\boldsymbol{v}_\lambda} \in \ker (A - \overline{\lambda} I)^m$ and hence that $\overline{\boldsymbol{v}_\lambda} \in \ker (A - \lambda I)^m$. In other words, both $\boldsymbol{v}_\lambda$ and $\overline{\boldsymbol{v}_\lambda}$ lie in $\mathcal{V}_\lambda$. The result now follows from the fact that

$$\mathbf{0} = \boldsymbol{v} - \overline{\boldsymbol{v}} = \sum_{\lambda\in\sigma(A)} (\boldsymbol{v}_\lambda - \overline{\boldsymbol{v}_{\overline{\lambda}}})$$

and from uniqueness of the decomposition (2). □

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

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

Moreover, given a rational matrix A, its Jordan Normal Form $A = Q^{-1}JQ$ can be computed in polynomial time, as shown in Reference [7].

Note that each vector $\boldsymbol{v}$ appearing as a column of the matrix Q^{-1} is a generalised eigenvector and that the index $\nu(\lambda)$ of each 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 the powers of a general matrix A. In fact, if J_i is a $l \times l$ Jordan block associated with an eigenvalue λ, then

$$J_i^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-1} & \cdots & \binom{n}{l-1}\lambda^{n-l+1} \\ 0 & \lambda^n & n\lambda^{n-1} & \cdots & \binom{n}{l-2}\lambda^{n-l+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n\lambda^{n-1} \\ 0 & 0 & 0 & \cdots & \lambda^n \end{pmatrix}, \tag{3}$$

where $\binom{n}{j}$ is defined to be 0 when $n < j$.

2.1.2 Matrix Exponentials. The exponential of a matrix $A \in \mathbb{C}^{d\times d}$ is defined by the power series

$$\exp(A) := \sum_{i=0}^{\infty} \frac{A^i}{i!} \in \mathbb{C}^{d\times d}.$$

The series above always converges, and so the exponential of a matrix is always well defined. Given $t \in \mathbb{C}$, one can obtain a closed-form representation of $\exp(At)$ as follows. Find $Q \in GL_d(\mathbb{C})$ such that $A = Q^{-1}JQ$ and $J = \operatorname{diag}(J_1, \ldots, J_N)$ is a block diagonal Jordan matrix. Then $\exp(At) = Q^{-1}\exp(Jt)Q$, where $\exp(Jt) = \operatorname{diag}(\exp(J_1t), \ldots, \exp(J_Nt))$. Note that, due to Equation (3), if

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

is a Jordan block associated with an eigenvalue λ, then

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

If A and B commute, then so do $\exp(A)$ and $\exp(B)$ and in this case we have $\exp(A)\exp(B) = \exp(A+B)$.

Proposition 2.2. *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 $t^n \exp(\lambda t)$, $n \in \mathbb{N}$.*

Proof. Note that if $A = Q^{-1}JQ$ and $J = \operatorname{diag}(J_1, \ldots, J_N)$ is a block diagonal Jordan matrix, then $\exp(At) = Q^{-1}\exp(Jt)Q$ and $\exp(Jt) = \operatorname{diag}(\exp(J_1t), \ldots, \exp(J_Nt))$. The result follows by observing that $Q\boldsymbol{v}$ is zero in every component other than those pertaining the block corresponding to the eigenspace $\mathcal{V}_\lambda$. □

2.1.3 Matrix Logarithms. Given a matrix $A \in \mathbb{C}^{d\times d}$, define the matrix logarithm

$$\log(A) := \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} \in \mathbb{C}^{d\times d}$$

whenever the above series converges. In particular, the matrix logarithm that is well defined in case A is unipotent (e.g., when A is a upper unitriangular matrix), since in this case the above series becomes a polynomial expression in A. In fact, it is well-known that the matrix exponential function and matrix logarithm functions yield a bijection between nilpotent and unipotent matrices in $\mathbb{C}^{d\times d}$ (see, e.g., Reference [14, Chapter 2]). Thus we have the following.

Proposition 2.3. *Given an upper unitriangular matrix $M \in \mathbb{C}^{d\times d}$, there exists a unique strictly upper triangular matrix L such that $\exp(L) = M$. Moreover, the entries of L lie in the field $\mathbb{Q}(M_{i,j} : 1 \le i, j \le d)$ generated over $\mathbb{Q}$ by the entries of M.*

2.1.4 Properties of Commuting Matrices. We present a useful decomposition of $\mathbb{C}^d$ induced by the commuting matrices $A_1, \ldots, A_k \in \mathbb{C}^{d\times d}$ (cf. Reference [8, Section 2]). Recall that $\sigma(A_i)$ denotes the spectrum of the matrix A_i and fix

$$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k) \in \sigma(A_1) \times \cdots \times \sigma(A_k).$$

Note that the generalised eigenspace of λ_i of A_i is equal to $\ker (A_i - \lambda_i I)^d$. (This is because the increasing sequence $(\ker (A_i - \lambda_i I)^n)_{n\in\mathbb{N}}$ of subspaces of $\mathbb{C}^d$ must converge in at most d steps.) With this in mind, we define the following subspace of $\mathbb{C}^d$:

$$\mathcal{V}_{\boldsymbol{\lambda}} = \bigcap_{i=1}^{k} \ker (A_i - \lambda_i I)^d.$$

Let $\Sigma = \{\boldsymbol{\lambda} \in \sigma(A_1) \times \cdots \times \sigma(A_k) : \mathcal{V}_{\boldsymbol{\lambda}} \neq \{\mathbf{0}\}\}$. Below, $A_i \restriction_{\mathcal{V}_{\boldsymbol{\lambda}}}$ denotes the restriction of the linear operator A_i to the linear subspace $\mathcal{V}_{\boldsymbol{\lambda}}$, which is invariant under A_i.

Theorem 2.4. *For all $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k) \in \Sigma$ and all $i \in \{1, \ldots, k\}$ the following properties hold:*

(1) $\mathcal{V}_{\boldsymbol{\lambda}}$ is invariant under A_i.
(2) $\sigma(A_i \restriction_{\mathcal{V}_{\boldsymbol{\lambda}}}) = \{\lambda_i\}$.
(3) $\mathbb{C}^d = \bigoplus_{\boldsymbol{\lambda}\in\Sigma} \mathcal{V}_{\boldsymbol{\lambda}}$.

Proof. We show by induction on k that the subspaces $\mathcal{V}_{\boldsymbol{\lambda}}$ satisfy the properties above.

When $k = 1$, the result follows from the existence of Jordan Canonical Forms. When $k > 1$, suppose that $\sigma(A_k) = \{\mu_1, \ldots, \mu_m\}$, and let $\mathcal{U}_j = \ker (A_k - \mu_j I)^d$, for $j \in \{1, \ldots, m\}$. Again, it follows from the existence of Jordan Canonical Forms that

$$\mathbb{C}^d = \bigoplus_{j=1}^{m} \mathcal{U}_m.$$

Pick $i \in \{1, \ldots, k-1\}$ and $j \in \{1, \ldots, m\}$. Now, as A_k and A_i commute, so do $(A_k - \mu_j I)$ and A_i. Therefore, for all $\boldsymbol{v} \in \mathcal{U}_j$, $(A_k - \mu_j I)^d A_i \boldsymbol{v} = A_i (A - \mu_j I)^d \boldsymbol{v} = \mathbf{0}$, so $A_i \boldsymbol{v} \in \mathcal{U}_j$; that is, $\mathcal{U}_j$ is invariant under A_i. The result follows from applying the induction hypothesis to the commuting operators $A_i \restriction_{\mathcal{U}_j}$. □

We will also make use of the following well-known result on simultaneous triangularisation of commuting matrices. See, for example, Reference [22].

Theorem 2.5. *Given commuting matrices $A_1, \ldots, A_k \in \mathbb{C}^{d\times d}$, there exists a matrix $P \in GL_d(\mathbb{C})$ such that $P^{-1}A_iP$ is upper triangular for all $i \in \{1, \ldots, k\}$. Moreover, if the entries of every matrix A_i are algebraic numbers then P may be chosen to have algebraic entries also.*

Note that Theorem 2.5 will allow us to reduce general instances of MSMP with commuting matrices to the subcase where all those matrices are in upper triangular form.

2.2 Number Theory

2.2.1 Linear Diophantine Equations with Algebraic Coefficients. A complex number α is said to be *algebraic* if it is a root of some non-zero polynomial with integer coefficients. The set of algebraic numbers forms a field, denoted $\overline{\mathbb{Q}}$. A complex number that is not algebraic is said to be *transcendental*.

For computational purposes, we represent an algebraic number by a polynomial P with rational coefficients such that $P(\alpha) = 0$, together with a numerical approximation $p + qi$, where $p, q \in \mathbb{Q}$, of sufficient accuracy to distinguish α from the other roots of P [11, Section 4.2.1]. Under this representation, arithmetic operations and equality testing can be carried out effectively, as can sign testing of real algebraic numbers

The subfield $\mathbb{K}$ of $\mathbb{C}$ generated by given algebraic numbers $\alpha_1, \ldots, \alpha_k$ is finite dimensional as a vector space over $\mathbb{Q}$. Moreover, given representations of $\alpha_1, \ldots, \alpha_k$, one can compute a basis of $\mathbb{K}$ over $\mathbb{Q}$ and the (rational) coefficients of each α_i with respect to this basis (see Reference [11, Section 4.5]).

Proposition 2.6. *Let $S = \{\boldsymbol{x} \in \mathbb{Z}^n : A\boldsymbol{x} = \boldsymbol{b}\}$, where matrix A and vector $\boldsymbol{b}$ have algebraic coefficients. Then we can decide whether S is non-empty and, if so, can compute $\boldsymbol{x}_0 \in \mathbb{Z}^d$ and $M \in \mathbb{Z}^{d\times s}$, for some $s \leq d$, such that $S = \{\boldsymbol{x}_0 + M\boldsymbol{y} : \boldsymbol{y} \in \mathbb{Z}^s\}$.*

Proof. Compute a basis $\alpha_1, \ldots, \alpha_k$ of the subfield of $\mathbb{C}$ that is generated by the entries of A and $\boldsymbol{b}$. Then we can write $A = \sum_{i=1}^{k} A_i\alpha_i$ and $\boldsymbol{b} = \sum_{i=1}^{k} \boldsymbol{b}_i\alpha_i$, where A_i is a matrix of rational numbers of the same dimension as A and $\boldsymbol{b}_i$ is a rational vector of the same dimension as $\boldsymbol{b}$. Now we have that

$$\begin{aligned} A\boldsymbol{x} = \boldsymbol{b} &\Leftrightarrow \left(\sum_{i=1}^{k} A_i\alpha_i\right)\boldsymbol{x} = \sum_{i=1}^{k} \boldsymbol{b}_i\alpha_i \\ &\Leftrightarrow A_i\boldsymbol{x} = \boldsymbol{b}_i, \forall i \in \{1, \ldots, k\}. \end{aligned}$$

Thus we have characterised S as the set of integer solutions of a system of equations with rational coefficients. Now the desired representation of S can be computed using, e.g., the procedure described in Reference [11, Chapter X]. □

2.2.2 Linear Forms in Logarithms. Let log denote the principal branch of the complex logarithm function, i.e., the imaginary part of $\log(z)$ lies in the interval $(-\pi, \pi]$ for all $z \neq 0$.

Given $k \in \mathbb{N}$, a number Λ of the form

$$\Lambda := \alpha_0 + \alpha_1 \log(\beta_1) + \cdots + \alpha_k \log(\beta_k), \tag{4}$$

where $\alpha_0, \ldots, \alpha_n, \beta_1, \ldots, \beta_k$ are algebraic numbers, with the β_i non-zero, is said to be a *linear form in logarithms of algebraic numbers*. Note that the collection of such linear forms is closed under addition, complex conjugation, and under multiplication by algebraic numbers.

To effectively manipulate linear forms, we will need the following result of Baker [4].

Theorem 2.7 (Baker). *Let $\beta_1, \ldots, \beta_k$ be non-zero algebraic numbers. If $\log(\beta_1), \ldots, \log(\beta_k)$ are linearly independent over $\mathbb{Q}$, then $1, \log(\beta_1), \ldots, \log(\beta_k)$ are linearly independent over $\overline{\mathbb{Q}}$.*

Proposition 2.8. *Let $\Lambda = \alpha_0 + \sum_{i=1}^{k} \alpha_i \log(\beta_i) \in \mathbb{R}$ be a linear form in logarithms of algebraic numbers for some nonnegative integer k. Then we can effectively determine the sign of Λ (zero, positive, or negative) and whether $\frac{\Lambda}{\pi}$ is algebraic.*

Proof. Clearly, we may assume without loss of generality that $\alpha_1, \ldots, \alpha_k$ are non-zero. We say that Λ is *reduced* if the set of terms $\mathcal{T}_\Lambda := \{\log(-1), \log(\beta_1), \ldots, \log(\beta_k)\}$ is linearly independent over $\mathbb{Q}$. Rational linear independence of $\mathcal{T}_\Lambda$ is equivalent to the requirement that the set of multiplicative relations $\mathcal{L} := \{(n_1, \ldots, n_k) \in \mathbb{Z}^k : \beta_1^{n_1} \cdots \beta_k^{n_k} = 1\}$ contain only the zero vector. Now $\mathcal{L}$ is a subgroup of $\mathbb{Z}^k$, and by a deep result of Masser [21] it is possible to compute a basis of $\mathcal{L}$ (see, e.g., Reference [8]). In particular, if Λ is not reduced, then we can compute an integer linear relation among $\log(-1), \log(\beta_1), \ldots, \log(\beta_k)$ and rewrite Λ so as to eliminate one of the terms $\log(\beta_i)$ in $\mathcal{T}_\Lambda$. Continuing in this way, we eventually reach an equivalent reduced form of Λ.

Suppose that Λ is reduced. Then by Theorem 2.7 the set $\{1, \log(-1), \log(\beta_1), \ldots, \log(\beta_k)\}$ is linearly independent over $\overline{\mathbb{Q}}$. It follows that $\Lambda = 0$ if and only if $k = 0$ and $\alpha_0 = 0$. Hence it is decidable whether or not $\Lambda = 0$. If $\Lambda \neq 0$, then we can determine the sign of Λ by computing a rational approximation of Λ of sufficient precision. Furthermore, we have that $\Lambda/\pi \in \overline{\mathbb{Q}}$ if and only if $k = 1$, $\alpha_0 = 0$, and $\beta_1 = -1$. □

2.2.3 Diophantine Approximation. Given a vector $\boldsymbol{a} \in \mathbb{R}^d$ and a set $S \subseteq \mathbb{R}^d$, write $\mathrm{dist}(\boldsymbol{a}, S)$ for the ℓ_1-distance between $\boldsymbol{a}$ and S, i.e., $\inf_{\boldsymbol{b} \in S} \|\boldsymbol{a} - \boldsymbol{b}\|_1$. We will need the following consequence of Kronecker's theorem on simultaneous inhomogeneous diophantine approximation (see Reference [18, Corollary 2.8]).

Theorem 2.9. *Suppose that $C = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k\} \subseteq \mathbb{R}^d$ and that no integer vector is orthogonal to C. Then for any $\boldsymbol{q} \in \mathbb{R}^d$ and for any $\varepsilon > 0$ there exist non-negative real numbers $\lambda_1, \ldots, \lambda_k$ such that*

$$\mathrm{dist}\left(\boldsymbol{q} + \sum_{i=1}^{k} \lambda_i \boldsymbol{c}_i, \mathbb{Z}^d\right) \leq \varepsilon.$$

2.2.4 Schanuel's Conjecture. Schanuel's Conjecture [19] is a unifying conjecture in transcendence theory that generalises many of the classical results in the field (including Theorem 2.7). The conjecture states that if $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$ are rationally linearly independent, then some k-element subset of $\{\alpha_1, \ldots, \alpha_k, e^{\alpha_1}, \ldots, e^{\alpha_k}\}$ is algebraically independent.

Assuming Schanuel's Conjecture, MacIntyre and Wilkie [20] have shown decidability of the first-order theory of the expansion of the real field with the exponentiation function and the sin and cos functions restricted to bounded intervals.

Theorem 2.10 (Wilkie and MacIntyre). *If Schanuel's conjecture is true, then, for each $n \in \mathbb{N}$, the first-order theory of the strucure $(\mathbb{R}, +, \cdot, \exp, \cos \restriction_{[0,n]}, \sin \restriction_{[0,n]})$ is decidable.*

2.3 Convex Geometry

2.3.1 Convex Polytopes. A *convex polytope* is a subset of $\mathbb{R}^n$ of the form

$$\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b}\},$$

where $A \in \mathbb{R}^{d \times n}$ and $\boldsymbol{b} \in \mathbb{R}^d$ for some $d \in \mathbb{N}$. When all the entries of A and $\boldsymbol{b}$ are algebraic numbers, $\mathcal{P}$ is said to have an algebraic description. In this case, we can decide non-emptiness of $\mathcal{P}$, e.g., by reduction to the satisfiability problem for the existential theory of real closed fields.

We will need the following result:

Theorem 2.11 (Minkowski-Weyl). *Any polytope $\mathcal{P} \subseteq \mathbb{R}^d$ can be written as the sum of two sets $\mathcal{H} \subseteq \mathbb{R}^d$ and $C \subseteq \mathbb{R}^d$, where $\mathcal{H}$ is a finitely generated convex hull and C is a finitely generated cone.*

2.3.2 Fourier-Motzkin Elimination. Fourier-Motzkin elimination is a simple method for eliminating variables from systems of linear inequalities. The procedure consists in isolating one variable at a time and comparing all its lower and upper bounds. Note that this method preserves the set of solutions on the remaining variables, so a solution of the reduced system can always be extended to a solution of the original one.

Proposition 2.12. *It is decidable whether a given convex polytope $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : \pi A\boldsymbol{x} < \boldsymbol{b}\}$ is empty, where the entries of A are all real algebraic numbers and those of $\boldsymbol{b}$ are real linear forms in logarithms of algebraic numbers. Moreover, if $\mathcal{P}$ is non-empty, then one can compute a rational vector $\boldsymbol{q} \in \mathcal{P}$.*

Proof. Note that $\mathcal{P}$ is non-empty iff there exists $\boldsymbol{y} \in \mathbb{R}^n$ with $A\boldsymbol{y} < \boldsymbol{b}$. The existence of such a vector $\boldsymbol{y}$ can be decided using Fourier-Motzkin elimination. The linear inequalities that are generated by the elimination process are such that each coefficient of a variable y_i is real algebraic and the constant term is a real linear form in logarithms of algebraic numbers. Having eliminated all variables, we can decide emptiness, since we can effectively decide whether a real linear form in logarithms is strictly positive by Proposition 2.8. □

3 EXAMPLE

The following example illustrates some elements of our approach to establishing decidability of the Matrix-Exponential Semigroup Membership Problem.

Let λ_1, λ_2 be real algebraic numbers such that $\lambda_1 > \lambda_2$ and consider the following commuting matrices A_1, A_2:

$$A_i = \begin{pmatrix} \lambda_i & 1 \\ 0 & \lambda_i \end{pmatrix}, i \in \{1, 2\}.$$

Given nonnegative real variables t_1, t_2, we have that (see Section 2.1.2)

$$\exp(A_i t_i) = \exp(\lambda_i t_i) \begin{pmatrix} 1 & t_i \\ 0 & 1 \end{pmatrix}, i \in \{1, 2\}. \tag{5}$$

Let c_1, c_2 be real algebraic numbers such that $c_1, c_2 > 0$, and let

$$C = \begin{pmatrix} c_1 & c_2 \\ 0 & c_1 \end{pmatrix}.$$

We would like to determine whether there exist $t_1, t_2 \in \mathbb{R}$, $t_1, t_2 \geq 0$ such that

$$\exp(A_1 t_1) \exp(A_2 t_2) = C. \tag{6}$$

Using the closed-form expression for matrix exponentials in Equation (5), Equation (6) is equivalent to the following pair of equations:

$$\exp(\lambda_1 t_1 + \lambda_2 t_2) = c_1$$
$$(t_1 + t_2) \exp(\lambda_1 t_1 + \lambda_2 t_2) = c_2.$$

Solving directly, we have

$$t_1 = \frac{\log(c_1) - \frac{c_2}{c_1}\lambda_2}{\lambda_1 - \lambda_2} \quad \text{and} \quad t_2 = \frac{\frac{c_2}{c_1}\lambda_1 - \log(c_1)}{\lambda_1 - \lambda_2}.$$

Now $t_1, t_2 \geq 0$ holds if and only if

$$\lambda_2 \leq \frac{c_1}{c_2} \log(c_1) \leq \lambda_1.$$

Whether these inequalities hold amounts to comparing linear forms in logarithms of algebraic numbers, which is decidable by Proposition 2.8.

4 DECIDABILITY IN THE COMMUTATIVE CASE

We first reduce the MEP with commuting matrices to the LEP, as defined in Section 1.2. The idea is to perform a change of basis so that the matrices $A_1, \ldots, A_k$, and C in an instance of MEP all become block-diagonal, with each block being upper triangular; we then analyse the problem blockwise, making use of Proposition 2.3 concerning logarithms of unipotent matrices.

Proposition 4.1. *MEP with commuting matrices reduces to LEP.*

Proof. Consider an instance of MEP, as given in Definition 1.1, with commuting $n \times n$ matrices $A_1, \ldots, A_k$ and target matrix C.

We first show how to define a matrix P such that each matrix $P^{-1}A_iP$ is block diagonal for $i = 1, \ldots, k$, with each block being upper triangular.

By Theorem 2.4, we can write $\mathbb{C}^n$ as a direct sum of subspaces $\mathbb{C}^n = \oplus_{j=1}^{m} \mathcal{V}_j$ such that for every subspace $\mathcal{V}_j$ and matrix A_i, $\mathcal{V}_j$ is an invariant subspace of A_i on which A_i has a single eigenvalue $\lambda_i^{(j)}$.

Define a matrix Q by picking an algebraic basis for each $\mathcal{V}_j$ and taking the columns of Q to be the successive vectors of each basis. Then, for $i = 1, \ldots, k$, the matrix $Q^{-1}A_iQ$ is block-diagonal, where the jth block is a matrix $B_i^{(j)}$ that represents $A_i \restriction \mathcal{V}_j$, $j = 1, \ldots, m$.

Fixing the index of a block $j \in \{1, \ldots, m\}$, note that the matrices $B_1^{(j)}, \ldots, B_k^{(j)}$ all commute. Thus we may apply Theorem 2.5 to obtain an algebraic matrix M_j such that each matrix $M_j^{-1}B_i^{(j)}M_j$ is upper triangular, $i = 1, \ldots, k$. Thus, we can write

$$M_j^{-1}B_i^{(j)}M_j = \lambda_i^{(j)}I + N_i^{(j)}$$

for some strictly upper triangular matrix $N_i^{(j)}$.

We define M to be the block-diagonal matrix $M = \mathrm{diag}(M_1, \ldots, M_m)$. Letting $P = QM$, we have that $P^{-1}A_iP$ is block-diagonal, with the jth block being $\lambda_i^{(j)}I + N_i^{(j)}$ for $j = 1, \ldots, m$. Now for all $t_1, \ldots, t - K \geq 0$, we have

$$\prod_{i=1}^{k} \exp(A_i t_i) = C \Leftrightarrow \prod_{i=1}^{k} \exp\left(P^{-1}A_iPt_i\right) = P^{-1}CP. \tag{7}$$

If $P^{-1}CP$ is not block-diagonal, with each block being upper triangular and with the same entries along the diagonal, then Equation (7) has no solution, and the problem instance must be negative. Otherwise, denoting the blocks of $P^{-1}CP$ by $D^{(j)}$ for $j \in \{1, \ldots, m\}$, our problem amounts to simultaneously solving the system of matrix equations

$$\prod_{i=1}^{k} \exp\left(\left(\lambda_i^{(j)}I + N_i^{(j)}\right)t_i\right) = D^{(j)}, \quad j \in \{1, \ldots, m\}, \tag{8}$$

where there is one equation for each block.

For each fixed j, the matrices $N_i^{(j)}$ inherit commutativity from the matrices $B_i^{(j)}$, so we have

$$\begin{aligned}\prod_{i=1}^{k} \exp\left((\lambda_i^{(j)}I + N_i^{(j)})t_i\right) &= \exp\left(\sum_{i=1}^{k}(\lambda_i^{(j)}I + N_i^{(j)})t_i\right) \\ &= \exp\left(\sum_{i=1}^{k}\lambda_i^{(j)}t_i\right) \cdot \exp\left(\sum_{i=1}^{k}N_i^{(j)}t_i\right).\end{aligned}$$

Hence the system (8) is equivalent to

$$\exp\left(\sum_{i=1}^{k}\lambda_i^{(j)}t_i\right)\cdot\exp\left(\sum_{i=1}^{k}N_i^{(j)}t_i\right)=D^{(j)}\quad j\in\{1,\ldots,m\}. \tag{9}$$

By assumption, the diagonal entries of each matrix $D^{(j)}$ are all equal to a single value—say, $c^{(j)}$. Since the diagonal entries of $\exp(\sum_{i=1}^{k}N^{(j)}t_i)$ are all 1, the system (9) is equivalent to

$$\exp\left(\sum_{i=1}^{k}\lambda_i^{(j)}t_i\right)=c^{(j)}\text{ and }\exp\left(\sum_{i=1}^{k}N_i^{(j)}t_i\right)=\frac{1}{c^{(j)}}D^{(j)}$$

for $j=1,\ldots,m$.

Applying Proposition 2.3, the above system can equivalently be written

$$\exp\left(\sum_{i=1}^{k}\lambda_i^{(j)}t_i\right)=c^{(j)}\text{ and }\sum_{i=1}^{k}N_i^{(j)}t_i=S^{(j)} \tag{10}$$

for some effectively computable matrices $S^{(j)}$ with algebraic entries, $j=1,\ldots,m$. But the above system of equations is an instance of LEP. □

Next we reduce the Linear-Exponential Equation Problem to the ALIP, as defined in Section 1.2.

Proposition 4.2. *LEP reduces to ALIP.*

Proof. Consider an instance of LEP, comprising a system of equations

$$\exp\left(\sum_{\ell=1}^{k}\lambda_\ell^{(j)}t_\ell\right)=c_j\exp(d_j)\quad j\in\{1,\ldots,m\}, \tag{11}$$

and linear equations $A\boldsymbol{t}=\boldsymbol{b}$.

Recall that log denotes the principal branch of the complex logarithm function. Note that if any $c_j=0$ for some j, then Equation (11) has no solution. Otherwise, by applying log to each equation in Equation (11), we get

$$\sum_{\ell=1}^{k}\lambda_\ell^{(j)}t_\ell=d_j+\log(c_j)+2\pi i n_j\quad j\in\{1,\ldots,m\}, \tag{12}$$

where $n_j\in\mathbb{Z}$.

The system of Equations (12) can be written in matrix form as

$$\Lambda\boldsymbol{t}\in\boldsymbol{d}+\log(\boldsymbol{c})+2\pi i\mathbb{Z}^m, \tag{13}$$

where Λ is the $m\times k$ matrix with $\Lambda_{j,\ell}=\lambda_\ell^{(j)}$ and log is applied pointwise to vectors. A solution to the given instance of LEP comprises a solution $\boldsymbol{t}\geq 0$ to Equation (13) that furthermore satisfies $A\boldsymbol{t}=\boldsymbol{b}$. Such a solution exists if and only if the convex polytope

$$\mathcal{T}:=\{\boldsymbol{x}\in\mathbb{R}^m:\exists\boldsymbol{t}\geq 0\,(\boldsymbol{d}+\log(\boldsymbol{c})+2\pi i\boldsymbol{x}=\Lambda\boldsymbol{t}\text{ and }A\boldsymbol{t}=\boldsymbol{b})\}$$

contains an integer point.

The linear constraints in the definition of $\mathcal{T}$ are such that the coefficients of $t_1,\ldots,t_k$ are algebraic and the constant terms are linear forms in logarithms of algebraic numbers. By dividing the equational constraints in the definition of $\mathcal{T}$ into real and imaginary parts, we can use Fourier-Motzkin elimination to eliminate the existentially quantified variables $\boldsymbol{t}$ and arrive at a characterisation $\mathcal{T}=\{\boldsymbol{x}:\pi B\boldsymbol{x}\leq\boldsymbol{e}\}$, for some effectively computable matrix B and vector $\boldsymbol{e}$ such that the

entries of B are real algebraic and the entries of $\boldsymbol{e}$ are real linear forms in logarithms of algebraic numbers. But this is the form of an instance of ALIP. □

We are left with the task of showing that ALIP is decidable. The key tools here are Theorem 2.9 and Proposition 2.12.

Proposition 4.3. *ALIP is decidable.*

Proof. We are given a convex polytope $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^d : \pi A\boldsymbol{x} \leq \boldsymbol{b}\}$, where the entries of A are real algebraic and the entries of $\boldsymbol{b}$ are real linear forms in logarithms of algebraic numbers. We wish to decide whether $\mathcal{P}$ contains an integer point.

First, we show how to reduce the problem of deciding whether $\mathcal{P}$ contains an integer point to finitely many instances of the problem of finding an integer point in a convex polytope $\mathcal{Q}$ of the form

$$\mathcal{Q} = \{\boldsymbol{x} \in \mathbb{R}^d : C\boldsymbol{x} = \boldsymbol{d},\ \pi E\boldsymbol{x} < \boldsymbol{f}\}, \tag{14}$$

where the entries of $C, E, \boldsymbol{d}$ are real algebraic and the entries of $\boldsymbol{f}$ are real linear forms in logarithms of algebraic numbers. We will obtain a collection of such polytopes $\mathcal{Q}$ from $\mathcal{P}$ by replacing the non-strict inequality "$\leq$" in every constraint of $\mathcal{P}$ with either strict inequality "$<$" or equality "$=$". To this end, note that when $\boldsymbol{x} \in \mathbb{Z}^d$, $A\boldsymbol{x}$ is a vector with algebraic coefficients; so whenever b_i/π is transcendental (which can be effectively checked by Proposition 2.8), we may replace the inequality $\pi(A\boldsymbol{x})_i \leq b_i$ in the definition of $\mathcal{P}$ with a strict inequality $\pi(A\boldsymbol{x})_i < b_i$ without affecting $\mathcal{P} \cap \mathbb{Z}^d$. However, whenever b_i/π is algebraic, we split our problem into two cases: In the first case, we replace the inequality $\pi(A\boldsymbol{x})_i \leq b_i$ with an equality $\pi(A\boldsymbol{x})_i = b_i$, and, in the second case, we replace the inequality $\pi(A\boldsymbol{x})_i \leq b_i$ with the strict inequality $\pi(A\boldsymbol{x})_i < b_i$. Thus, we obtain a finite collection of polytopes of the form (14).

Next we eliminate the equality constraints in the definition of $\mathcal{Q}$ so that only strict inequality constraints remain. Indeed, by Proposition 2.6, we can compute an integer matrix M and vector $\boldsymbol{x}_0$ such that

$$\{\boldsymbol{x} \in \mathbb{Z}^d : C\boldsymbol{x} = \boldsymbol{d}\} = \{\boldsymbol{x}_0 + M\boldsymbol{y} : \boldsymbol{y} \in \mathbb{Z}^{d'}\}$$

for some $d' < d$. Then the polytope $\mathcal{Q}' := \{\boldsymbol{y} \in \mathbb{R}^{d'} : \pi E(\boldsymbol{x}_0 + M\boldsymbol{y}) < \boldsymbol{f}\}$ contains an integer point iff $\mathcal{Q}$ contains an integer point. Now determining the existence of an integer point in a polytope of the form $\mathcal{Q}'$ is simply a version of ALIP in which all inequality constraints are strict.

It remains to show how to decide whether a polytope $\mathcal{P} := \{\boldsymbol{x} \in \mathbb{R}^d : \pi A\boldsymbol{x} < \boldsymbol{b}\}$, defined using only strict constraints, contains an integer point. To do this, we first use Theorem 2.12 to decide whether $\mathcal{P}$ is non-empty. In case $\mathcal{P}$ is non-empty, we can use Fourier-Motzkin elimination to find a vector $\boldsymbol{q} \in \mathbb{Q}^d$ and $\varepsilon > 0$ such that the closed l_1-ball $\mathcal{B}$ with centre $\boldsymbol{q}$ and radius ε with respect to the l_1 norm is contained in $\mathcal{P}$.

The next step is to consider the Minkowski-Weyl decomposition of $\mathcal{P}$, namely $\mathcal{P} = \mathcal{H} + \mathcal{C}$, where $\mathcal{H}$ is the convex hull of finitely many points of $\mathcal{P}$ and $\mathcal{C} = \{\boldsymbol{x} \in \mathbb{R}^d : A\boldsymbol{x} \leq \boldsymbol{0}\}$ is a cone with an algebraic description. Note that $\mathcal{P}$ is bounded if and only if $\mathcal{C} = \{\boldsymbol{0}\}$, in which case the problem at hand is simple: Consider the polytope $\mathcal{Q}$ with an algebraic description obtained by rounding up each coordinate of $\boldsymbol{b}/\pi$, which has the same conic part as $\mathcal{P}$ and which contains $\mathcal{P}$, and therefore is bounded; finally, compute a bound on $\mathcal{Q}$ (such a bound can be defined in the first-order theory of the reals), which is also a bound on $\mathcal{P}$, and test the integer points within that bound for membership in $\mathcal{P}$. Otherwise,

$$\mathcal{C} = \{\alpha_1\boldsymbol{c}_1 + \cdots + \alpha_k\boldsymbol{c}_k : \alpha_1, \ldots, \alpha_k \geq 0\},$$

where $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k \in \overline{\mathbb{Q}}^d$ are the extremal rays of $\mathcal{C}$. Note that $\boldsymbol{q} + \mathcal{C} \subseteq \mathcal{P}$ and that $\mathcal{B} + \mathcal{C} \subseteq \mathcal{P}$.

Now we consider a variation of an argument that appears in Reference [18]. Consider the computable set

$$\mathcal{L} = C^{\perp} \cap \mathbb{Z}^d.$$

If $\mathcal{L} = \{\mathbf{0}\}$, then, due to Theorem 2.9, it must be the case that there exist non-negative reals $\lambda_1, \ldots, \lambda_k$ such that

$$\operatorname{dist}\left(\boldsymbol{q} + \sum_{i=1}^{k} \lambda_i \boldsymbol{c}_i, \mathbb{Z}^d\right) \leq \varepsilon,$$

and we know that $\mathcal{P} \cap \mathbb{Z}^d \neq \emptyset$ from the fact that the closed ball $\mathcal{B}$ of centre $\boldsymbol{q}$ and radius ε with respect to the l_1 norm is contained in $\mathcal{P}$.

However, if $\mathcal{L} \neq \{\mathbf{0}\}$, then let $\boldsymbol{z} \in \mathcal{L} \setminus \{\mathbf{0}\}$. Since $\mathcal{H}$ is a bounded subset of $\mathbb{R}^n$, the set

$$\{\boldsymbol{z}^T \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\} = \{\boldsymbol{z}^T \boldsymbol{x} : \boldsymbol{x} \in \mathcal{H}\}$$

is a bounded subset of $\mathbb{R}$. Therefore, there exist $a, b \in \mathbb{Z}$ such that

$$\forall \boldsymbol{x} \in \mathcal{P}, a \leq \boldsymbol{z}^T \boldsymbol{x} \leq b,$$

so we can reduce our problem to $b - a + 1$ lower-dimensional instances by finding the integer points of $\{\boldsymbol{x} \in \mathcal{P} : \boldsymbol{z}^T \boldsymbol{x} = i\}$ for $i \in \{a, \ldots, b\}$. Note that we have seen earlier in the proof how to reduce the dimension of the ambient space when the polytope $\mathcal{P}$ is contained in an affine hyperplane. □

Putting together Proposition 4.1, Proposition 4.2, and Proposition 4.3, we obtain the following.

Theorem 4.4. *The Matrix-Exponential Semigroup Membership Problem (as given in Definition 1.1) is decidable in case the generating matrices $A_1, \ldots, A_k$ commute.*

5 UNDECIDABILITY OF THE NON-COMMUTATIVE CASE

In this section, we show that the Matrix Exponential Equation Problem and the Matrix-Exponential Semigroup Membership Problem are both undecidable in general.

5.1 Matrix-Exponential Equation Problem with Constraints

The proof of undecidability of MEP in the non-commutative case is by reduction from Hilbert's Tenth Problem. The reduction proceeds via several intermediate problems. These problems are obtained by augmenting MEP with various classes of arithmetic constraints on the real variables that appear in the statement of the problem.

Definition 5.1. We consider the following three classes of arithmetic constraints over real variables $t_1, t_2, \ldots$:

- $\mathcal{E}_{\pi\mathbb{Z}}$ comprises constraints of the form $t_i \in \alpha + \beta\pi\mathbb{Z}$, where α and $\beta \neq 0$ are real-valued constants such that $\cos(2\alpha\beta^{-1})$, β are both algebraic numbers.
- $\mathcal{E}_{+}$ comprises linear equations of the form $\alpha_1 t_1 + \cdots + \alpha_n t_n = \alpha_0$ for $\alpha_0, \ldots, \alpha_n$ real algebraic constants.
- $\mathcal{E}_{\times}$ comprises equations of the form $t_\ell = t_i t_j$.

These constraints will be useful in defining polynomial relations, as well as in enforcing that certain variables be integer multiples of π. They will play a crucial role in the reduction from Hilbert's Tenth Problem.

A class of constraints $\mathcal{E} \subseteq \mathcal{E}_{\pi\mathbb{Z}} \cup \mathcal{E}_{+} \cup \mathcal{E}_{\times}$ induces a generalisation of the MEP problem as follows:

Definition 5.2 (MEP with Constraints). Given a class of constraints $\mathcal{E} \subseteq \mathcal{E}_{\pi\mathbb{Z}} \cup \mathcal{E}_+ \cup \mathcal{E}_\times$, the problem MEP($\mathcal{E}$) is as follows. An instance consists of real algebraic matrices $A_1, \ldots, A_k, C$ and a finite set of constraints $E \subseteq \mathcal{E}$ on real variables $t_1, \ldots, t_k$. The question is whether there exist non-negative real values for $t_1, \ldots, t_k$ such that $\prod_{i=1}^{k} e^{A_i t_i} = C$ and the constraints E are all satisfied.

Note that in the above definition of MEP($\mathcal{E}$), the set of constraints E only utilises real variables $t_1, \ldots, t_k$ appearing in the matrix equation $\prod_{i=1}^{k} e^{A_i t_i} = C$. However, without loss of generality, we can allow constraints to utilise fresh variables t_i, for $i > k$, since we can always define a corresponding matrix $A_i = 0$ for such variables, for then $e^{A_i t_i} = I$ has no effect on the matrix product. In other words, we can without loss of generality have constraints in $\mathcal{E}$ with existentially quantified variables. In particular, we have the following useful observations:

- We can express inequality constraints of the form $t_i \neq \alpha$ in $\mathcal{E}_+ \cup \mathcal{E}_\times$ by using fresh variables t_j, t_ℓ. Indeed, $t_i \neq \alpha$ is satisfied whenever there exist values of t_j and t_ℓ such that $t_i = t_j + \alpha$ and $t_j t_\ell = 1$.
- By using fresh variables, $\mathcal{E}_+ \cup \mathcal{E}_\times$ can express polynomial constraints of the form $P(t_1, \ldots, t_n) = t$ for P a polynomial with integer coefficients.

We illustrate the above two observations in an example.

Example 5.3. Consider the problem, given matrices A_1, A_2 and C, to decide whether there exist $t_1, t_2 \geq 0$ such that

$$e^{A_1 t_1} e^{A_2 t_2} = C \text{ and } t_1^2 - 1 = t_2, t_2 \neq 0.$$

This is equivalent to the following instance of MEP($\mathcal{E}_+ \cup \mathcal{E}_\times$): Decide whether there exist $t_1, \ldots, t_5 \geq 0$ such that

$$\prod_{i=1}^{5} e^{A_i t_i} = C \text{ and } t_1 t_1 = t_3, t_3 - 1 = t_2, t_2 t_4 = t_5, t_5 = 1,$$

where A_1, A_2, and C are as above and $A_3 = A_4 = A_5 = 0$.

We will make heavy use of the following proposition to combine several instances of the constrained MEP into a single instance by combining matrices blockwise.

Proposition 5.4. *Given real algebraic matrices $A_1, \ldots, A_k, C$ and $A'_1, \ldots, A'_k, C'$, there exist real algebraic matrices $A''_1, \ldots, A''_k, C''$ such that for all $t_1, \ldots, t_k$:*

$$\prod_{i=1}^{k} e^{A''_i t_i} = C'' \qquad \Leftrightarrow \qquad \left(\prod_{i=1}^{k} e^{A_i t_i} = C\right) \wedge \left(\prod_{i=1}^{k} e^{A'_i t_i} = C'\right).$$

Proof. For any $i \in \{1, \ldots, k\}$, define

$$A''_i = \begin{bmatrix} A_i & 0 \\ 0 & A'_i \end{bmatrix}, \qquad C'' = \begin{bmatrix} C & 0 \\ 0 & C' \end{bmatrix}.$$

The result follows, because the matrix exponential can be computed blockwise (as is clear from its power series definition):

$$\prod_{i=1}^{k} e^{A''_i t_i} = \prod_{i=1}^{k} \begin{bmatrix} e^{A_i t_i} & 0 \\ 0 & e^{A'_i t_i} \end{bmatrix} = \begin{bmatrix} \prod_{i=1}^{k} e^{A_i t_i} & 0 \\ 0 & \prod_{i=1}^{k} e^{A'_i t_i} \end{bmatrix}. \qquad \square$$

We remark that in the statement of Proposition 5.4 the two matrix equations that are combined are over the same set of variables. However, we can clearly combine any two matrix equations for which the common variables appear in the same order in the respective products.

The core of the reduction is to show that the constraints in $\mathcal{E}_{\pi\mathbb{Z}}, \mathcal{E}_+$, and $\mathcal{E}_\times$ do not make the MEP problem harder.

Proposition 5.5. *MEP*$(\mathcal{E}_{\pi\mathbb{Z}} \cup \mathcal{E}_+ \cup \mathcal{E}_\times)$ *reduces to MEP*$(\mathcal{E}_+ \cup \mathcal{E}_\times)$.

Proof. Let $A_1, \ldots, A_k, C$ be real algebraic matrices and $E \subseteq \mathcal{E}_{\pi\mathbb{Z}} \cup \mathcal{E}_+ \cup \mathcal{E}_\times$ a finite set of constraints on real variables $t_1, \ldots, t_k$. Since E is finite, it suffices to show how to eliminate a single constraint in $\mathcal{E}_{\pi\mathbb{Z}} \cap E$ from E.

Let $t_j \in \alpha + \beta\pi\mathbb{Z}$ be a constraint in E. By definition of $\mathcal{E}_{\pi\mathbb{Z}}$, we have that $\beta \neq 0$ and that both β and $\cos(2\alpha\beta^{-1})$ are real algebraic. The Pythagorean theorem implies that $\sin(2\alpha\beta^{-1})$ is also real algebraic. Now define the following extra matrices:

$$A'_j = \begin{bmatrix} 0 & 2\beta^{-1} \\ -2\beta^{-1} & 0 \end{bmatrix}, C' = \begin{bmatrix} \cos(2\alpha\beta^{-1}) & \sin(2\alpha\beta^{-1}) \\ -\sin(2\alpha\beta^{-1}) & \cos(2\alpha\beta^{-1}) \end{bmatrix}.$$

Our assumptions ensure that A'_j and C' are both real algebraic.

We now have the following chain of equivalences:

$$\begin{aligned} e^{A'_j t_j} = C' &\Leftrightarrow \begin{bmatrix} \cos(2t_j\beta^{-1}) & \sin(2t_j\beta^{-1}) \\ -\sin(2t_j\beta^{-1}) & \cos(2t_j\beta^{-1}) \end{bmatrix} = C' \\ &\Leftrightarrow \cos(2t_j\beta^{-1}) = \cos(2\alpha\beta^{-1}) \\ &\qquad \wedge \sin(2t_j\beta^{-1}) = \sin(2\alpha\beta^{-1}) \\ &\Leftrightarrow 2\beta^{-1}t_j \equiv 2\alpha\beta^{-1} \pmod{2\pi} \\ &\Leftrightarrow t_j \in \alpha + \beta\pi\mathbb{Z}. \end{aligned}$$

Thus, the additional matrix equation $e^{A'_j t_j} = C'$ is equivalent to the constraint $t_j \in \alpha + \beta\pi\mathbb{Z}$. Applying Proposition 5.4, we can thus eliminate this constraint. □

Proposition 5.6. *MEP*$(\mathcal{E}_+ \cup \mathcal{E}_\times)$ *reduces to MEP*$(\mathcal{E}_+)$.

Proof. Let $A_1, \ldots, A_k, C$ be real algebraic matrices and $E \subseteq \mathcal{E}_+ \cup \mathcal{E}_\times$ a finite set of constraints on variables $t_1, \ldots, t_k$. We proceed as above, showing how to remove each constraint in $\mathcal{E}_\times$ from E. In so doing, we potentially increase the number and the dimension of matrices and add new constraints from $\mathcal{E}_+$.

Let $t_l = t_i t_j$ be an equation in E. To eliminate this equation, the first step is to introduce fresh variables x, x', y, y', z and add the constraints

$$t_i = x,\ t_j = y,\ t_\ell = z,$$

which are all in $\mathcal{E}_+$. We now add a new matrix equation over the fresh variables x, x', y, y', z that is equivalent to the constraint $xy = z$. Since this matrix equation involves a new set of variables, we are free to the set the order of the matrix products, which is crucial to express the desired constraint.

The key gadget is the following matrix product equation, which holds for any $x, x', y, y', z \geqslant 0$:

$$\begin{bmatrix} 1 & 0 & -z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -y' \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -x' & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & x - x' & z - xy \\ 0 & 1 & y - y' \\ 0 & 0 & 1 \end{bmatrix}.$$

Notice that each of the matrices on the left-hand side of the above equation has a single non-zero off-diagonal entry. Crucially, each matrix of this form can be expressed as an exponential. Indeed, we can write the above equation as a matrix-exponential product

$$e^{B_1 z} e^{B_2 y'} e^{B_3 x} e^{B_4 y} e^{B_5 x'} = \begin{bmatrix} 1 & x - x' & z - xy \\ 0 & 1 & y - y' \\ 0 & 0 & 1 \end{bmatrix}$$

for matrices

$$B_1 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad B_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad B_5 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \qquad B_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus the constraint $xy = z$ can be expressed as

$$e^{B_1 z} e^{B_2 y'} e^{B_3 x} e^{B_4 y} e^{B_5 x'} = I. \tag{15}$$

Again, we can apply Proposition 5.4 to combine Equation (15) with the matrix equation from the original problem instance and thus encode the constraint $z = xy$. □

Proposition 5.7. *MEP($\mathcal{E}_+$) reduces to MEP.*

Proof. Let $A_1, \ldots, A_k, C$ be real algebraic matrices and $E \subseteq \mathcal{E}_+$ a set of constraints. We proceed as above, showing how to eliminate each constraint from E that lies in $\mathcal{E}_+$, while preserving the set of solutions of the instance.

Let $\beta + \sum_{i=1}^{k} \alpha_i t_i = 0$ be an equation in E. Recall that $\beta, \alpha_1, \ldots, \alpha_k$ are real algebraic. Define the extra matrices $A'_1, \ldots, A'_k$ and C' as follows:

$$A'_i = \begin{bmatrix} 0 & \alpha_i \\ 0 & 0 \end{bmatrix}, \qquad C' = \begin{bmatrix} 1 & -\beta \\ 0 & 1 \end{bmatrix}.$$

Our assumptions ensure that $A'_1, \ldots, A'_k$ and C' are all real algebraic. Furthermore, the following extra product equation becomes

$$\prod_{i=1}^{k} e^{A'_i t_i} = C \Leftrightarrow \prod_{i=1}^{k} \begin{bmatrix} 1 & \alpha_i t_i \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\beta \\ 0 & 1 \end{bmatrix}$$
$$\Leftrightarrow \sum_{i=1}^{k} \alpha_i t_i = -\beta.$$

The result follows by applying Proposition 5.4. □

Combining Proposition 5.5, Proposition 5.6, and Proposition 5.7, we have the following.

Proposition 5.8. *MEP($\mathcal{E}_{\pi\mathbb{Z}} \cup \mathcal{E}_+ \cup \mathcal{E}_\times$) reduces to MEP.*

5.2 Reduction from Hilbert's Tenth Problem

Theorem 5.9. *MEP is undecidable in the non-commutative case.*

Proof. We have seen in the previous section that the problem MEP($\mathcal{E}_{\pi\mathbb{Z}} \cup \mathcal{E}_+ \cup \mathcal{E}_\times$) reduces to MEP without constraints. Thus, it suffices to reduce Hilbert's Tenth Problem to MEP($\mathcal{E}_{\pi\mathbb{Z}} \cup \mathcal{E}_+ \cup \mathcal{E}_\times$). In fact, the matrix equation will not play a role in the target of this reduction, only the additional constraints.

Let P be a polynomial of total degree d in k variables with integer coefficients. From P, we build a homogeneous polynomial Q by adding a new variable λ:

$$Q(\mathbf{x}, \lambda) = \lambda^d P\left(\frac{x_1}{\lambda}, \ldots, \frac{x_k}{\lambda}\right).$$

Note that Q still has integer coefficients. Furthermore, we have the relationship

$$Q(\mathbf{x}, 1) = P(\mathbf{x}).$$

As we have seen previously, it is easy to encode Q with constraints, in the sense that we can compute a finite set of constraints $E_Q \subseteq \mathcal{E}_+ \cup \mathcal{E}_\times$ utilising variables $t_0, \ldots, t_m, \lambda$ such that E is satisfied if and only if $t_0 = Q(t_1, \ldots, t_k, \lambda)$. Note that E_Q may need to utilise variables other than $t_1, \ldots, t_k$ to do that. Another finite set of equations $E_{\pi\mathbb{Z}} \subseteq \mathcal{E}_{\pi\mathbb{Z}}$ is used to encode that $t_1, \ldots, t_k, \lambda \in \pi\mathbb{Z}$. Finally, $E_= \subseteq \mathcal{E}_+ \cup \mathcal{E}_\times$ is used to encode $t_0 = 0$ and $1 \leqslant \lambda \leqslant 4$. The latter is done by adding the polynomial equations $\lambda = 1 + \alpha^2$ and $\lambda = 4 - \beta^2$ for some α and β. Finally, we have the following chain of equivalences:

$$\begin{aligned}
&\exists t_0, \ldots, \lambda \geqslant 0 \text{ s.t. } E_Q \cup E_{\pi\mathbb{Z}} \cup E_= \text{ is satisfied} \\
&\quad \Leftrightarrow \exists t_1, \ldots, \lambda \geqslant 0 \text{ s.t. } 0 = Q(t_1, \ldots, t_k, \lambda) \\
&\qquad \wedge\, t_1, \ldots, t_k, \lambda \in \pi\mathbb{Z} \wedge 1 \leqslant \lambda \leqslant 4 \\
&\quad \Leftrightarrow \exists n_1, \ldots, n_k \in \mathbb{N} \text{ s.t. } 0 = Q(\pi n_1, \ldots, \pi n_k, \pi) \\
&\quad \Leftrightarrow \exists n_1, \ldots, n_k \in \mathbb{N} \text{ s.t. } 0 = \pi^d Q(n_1, \ldots, n_k, 1) \\
&\quad \Leftrightarrow \exists n_1, \ldots, n_k \in \mathbb{N} \text{ s.t. } 0 = P(n_1, \ldots, n_k).
\end{aligned}$$

□

5.3 Enforcing a Matrix Product Order

In this section, we will present a gadget matrix-exponential semigroup that can enforce a certain partial order on the matrices reaching a particular target. This will be useful to establish the reduction MEP $\leq$ MSMP. More precisely, we will exhibit five matrices W, X, Y, Z, and G such that any product $G = \prod_{i=1}^{p} e^{A_i t_i}$, where $t_i > 0$ and $A_i \in \{W, X, Y, Z\}$ is such that all the "X" appear before the "Y." Define the following matrices:

$$W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \qquad X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad Z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix},$$

$$Y = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad G = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

One easily computes the exponentials W, X, Y, Z:

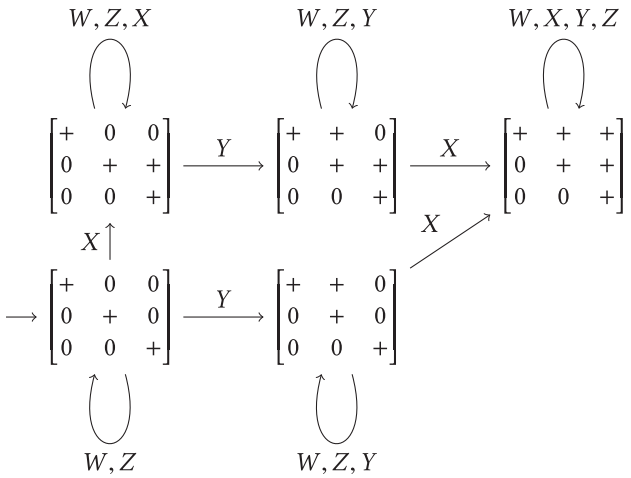
$$e^{Wt} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{2t} \end{bmatrix}, e^{Xt} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}, e^{Yt} = \begin{bmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, e^{Zt} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{-2t} \end{bmatrix}.$$

The crux of the proof will be based on the following asymmetry between X and Y, which leaves the top right corner zero in one case but nonzero in the other. As we will later observe, once the top right corner is nonzero, it cannot be cleared,

$$e^{Xt}e^{Yu} = \begin{bmatrix} 1 & u & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}, \qquad e^{Yt}e^{Xu} = \begin{bmatrix} 1 & u & tu \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}.$$

Proposition 5.10. *If there exist $p \in \mathbb{N}$, $A_i \in \{W, X, Y, Z\}$ and $t_i > 0$, for $i \in \{1, \ldots, p\}$, such that $\prod_{i=1}^{p} e^{A_i t_i} = G$, then the product contains at least one*[1] *"X" and one "Y," and all the "X" appear before the "Y." Formally, there exist i and j such that $A_i = X$ and $A_j = Y$, and for any such choice we have $i < j$.*

Proof. First, observe that any such product is always an upper triangular matrix with non-negative entries (because all the matrices have non-negative entries) and positive entries on the diagonal. Let M be such a matrix; we denote its coefficients by 0 if they are zero and + if they are positive. The following automaton should be read as follows: An arrow from M to M' annotated with A means that any product of a matrix of the shape of M by e^{At} with $t > 0$ will give a matrix with the shape of M'. Note that the empty product gives the identity. One easily checks that the following transitions hold.



Starting from the identity and applying the different products $e^{A_i t_i}$ in the automaton, it is clear that the only way to reach a matrix of the shape of G is to have all the "X" before "Y." Formally, by contradiction, if there were $i < j$ such that $A_i = Y$ and $A_j = X$, then by the automaton, we would end up with a matrix where the top right corner is nonzero, which contradicts $G = \prod_{i=1}^{p} e^{A_i t_i}$. □

The previous lemma shows that this semigroup enforces a partial order on the matrices in products that reach the matrix G. The next lemma shows that G can indeed be reached using these kinds of products, essentially proving that G belongs to this semigroup.

Proposition 5.11. *For any positive real t, there exists a non-negative real u such that*

$$e^{Wu}e^{Xt}e^{Yt}e^{Zu} = G \quad \text{or} \quad e^{Zu}e^{Xt}e^{Yt}e^{Wu} = G.$$

[1] Note that this is not entirely trivial, because we required only positive t_i in the product.

Proof. Consider the following products for an arbitrary $u \geqslant 0$:

$$e^{Wu}e^{Xt}e^{Yt}e^{Zu} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^u & 0 \\ 0 & 0 & e^{2u} \end{bmatrix} \begin{bmatrix} 1 & t & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-u} & 0 \\ 0 & 0 & e^{-2u} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^u & 0 \\ 0 & 0 & e^{2u} \end{bmatrix} \begin{bmatrix} 1 & te^{-u} & 0 \\ 0 & e^{-u} & te^{-2u} \\ 0 & 0 & e^{-2u} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & te^{-u} & 0 \\ 0 & 1 & te^{-u} \\ 0 & 0 & 1 \end{bmatrix},$$

and

$$e^{Zu}e^{Xt}e^{Yt}e^{Wu} = \begin{bmatrix} 1 & te^u & 0 \\ 0 & 1 & te^u \\ 0 & 0 & 1 \end{bmatrix}.$$

If $t \geqslant 1$, then choosing $u = \ln t \geqslant 0$ in the first product gives G; otherwise, choosing $u = \ln \frac{1}{t} \geqslant 0$ in the second product gives G. □

5.4 Undecidability of the Semigroup Membership Problem

We will now show the undecidability of MSMP.

In short, the difference is that we do not fix the order of the matrices in the product and that each matrix may be used more than once. We will now show the following key result.

Theorem 5.12. *MSMP is undecidable in the non-commutative case.*

Proof. We have seen in the previous section that MEP is undecidable in the non-commutative case. Thus, it suffices to reduce MEP to MSMP to show undecidability.

Let $A_1, \ldots, A_k, C \in \overline{\mathbb{Q}}^{n \times n}$ be an instance of MEP. Denote by I_m the identity of size m and 0_m the zero matrix of size m. Recall the 3×3 matrices W, X, Y, Z, G, defined in Section 5.3. For every $i \in \{2, \ldots, k-1\}$, define the following matrices:

$$B_i = \begin{bmatrix} A_i & & & & & \\ & 0_{3(i-2)} & & & & \\ & & Y & & & \\ & & & X & & \\ & & & & 0_{3(k-1-i)} & \end{bmatrix}, \qquad B'_i = \begin{bmatrix} 0_n & & & & & \\ & 0_{3(i-2)} & & & & \\ & & Y & & & \\ & & & X & & \\ & & & & 0_{3(k-1-i)} & \end{bmatrix}.$$

We also define the matrices

$$B_1 = \begin{bmatrix} A_1 & & \\ & X & \\ & & 0_{3(k-2)} \end{bmatrix}, \qquad B'_1 = \begin{bmatrix} 0_n & & \\ & X & \\ & & 0_{3(k-2)} \end{bmatrix},$$

$$B_k = \begin{bmatrix} A_1 & & \\ & 0_{3(k-2)} & \\ & & Y \end{bmatrix}, \qquad B'_k = \begin{bmatrix} 0_n & & \\ & 0_{3(k-2)} & \\ & & Y \end{bmatrix},$$

and, for every $i \in \{1, \ldots, k-1\}$,

$$W_i = \begin{bmatrix} 0_n & & & \\ & 0_{3(i-1)} & & \\ & & W & \\ & & & 0_{3(k-1-i)} \end{bmatrix}, \qquad Z_i = \begin{bmatrix} 0_n & & & \\ & 0_{3(i-1)} & & \\ & & Z & \\ & & & 0_{3(k-1-i)} \end{bmatrix}.$$

Finally, we define the target matrix:

$$C' = \begin{bmatrix} C & & & \\ & G & & \\ & & \ddots & \\ & & & G \end{bmatrix}.$$

We can now define our MSMP instance as follows; the target matrix is C' and the semigroup $\mathcal{G}$ is generated by

$$\left\{ e^{B_i t_i}, e^{B'_i t_i} : t_i \geqslant 0, i = 1, \ldots, k \right\}$$
$$\cup \left\{ e^{W_i t_i}, e^{Z_i t_i} : t_i \geqslant 0, i = 1, \ldots, k-1 \right\}.$$

We claim the original instance of MEP is satisfiable if and only if $C' \in \mathcal{G}$. Let us examine both directions independently.

Assume that the MEP instance is satisfiable. Then there exist $t_1, \ldots, t_k \geqslant 0$ such that

$$\prod_{i=1}^{k} e^{A_i t_i} = C.$$

Define $\tau = \max\{t_1, \ldots, t_k\} + 1$ (note that $\tau > 0$) and $t'_i = \tau - t_i \geqslant 0$ for every $i \in \{1, \ldots, k\}$. A straightforward calculation shows that:

$$\prod_{i=1}^{k} \left(e^{B_i t_i} e^{B'_i t'_i} \right) = \begin{bmatrix} \prod_{i=1}^{k} e^{A_i t_i} & & & \\ & e^{X\tau} e^{Y\tau} & & \\ & & \ddots & \\ & & & e^{X\tau} e^{Y\tau} \end{bmatrix}$$
$$= \begin{bmatrix} C & & & \\ & U & & \\ & & \ddots & \\ & & & U \end{bmatrix},$$

where $U = e^{X\tau} e^{Y\tau}$. Apply Proposition 5.11 to get $\lambda \geqslant 0$ such that either $e^{W\lambda} U e^{Z\lambda} = G$ or $e^{Z\lambda} U e^{W\lambda} = G$. In the first case, conclude by checking that

$$\prod_{i=1}^{k-1} e^{W_i \lambda} \prod_{i=1}^{k} \left(e^{B_i t_i} e^{B'_i t'_i} \right) \prod_{i=1}^{k-1} e^{Z_i \lambda} = \begin{bmatrix} C & & & \\ & G & & \\ & & \ddots & \\ & & & G \end{bmatrix} = C'.$$

In the second case, exchange the W_i and Z_i to get the same result. This concludes the proof that the MSMP instance is satisfiable, since all the products belong to $\mathcal{G}$.

Assume that the MSMP instance is satisfiable. Then there exist $t_1, \ldots, t_m > 0$ (we can always take them positive) and $M_1, \ldots, M_m \in \left\{B_i, B_i' : i = 1, \ldots, k\right\} \cup \left\{W_i, Z_i : i = 1, \ldots, k-1\right\}$ such that

$$\prod_{j=1}^{m} e^{M_j t_j} = C'. \tag{16}$$

Observe that, by construction, this product has the following form:

$$\prod_{j=1}^{m} e^{M_j t_j} = \begin{bmatrix} V & & & \\ & U_1 & & \\ & & \ddots & \\ & & & U_{k-1} \end{bmatrix},$$

where V belongs to the semigroup generated by $\{e^{A_i t} : t \geqslant 0\}$ and U_i belongs to the semigroup generated by $\{e^{Wt}, e^{Xt}, e^{Yt}, e^{Zt} : t \geqslant 0\}$. Since Equation (16) implies that $U_i = G$, we can apply Proposition 5.10 to get each product producing U_i must have all its "X" before its "Y." Furthermore, each U_i must contain at least one X and one Y in its product. For any $i \in \{1, \ldots, k\}$, let k_i (respectively, k_i') denote the first (respectively, last) index j such that $M_j = B_i$ or B_i'. Those indices exist because of the proposition, since at least one B_i or B_i' must appear for every i to get both an X and a Y in each product giving U_i. Obviously, $k_i \leqslant k_i'$ by definition. We now claim that the proposition implies that:

$$k_1 \leqslant k_1' < k_2 \leqslant k_2' < k_3 \cdots < k_{k-1} \leqslant k_{k-1}'.$$

Indeed, Proposition 5.10 ensures that in the product giving U_1, all the "X" appear before the "Y," but the only matrices that contribute some X to U_1 are B_1 and B_1', and the only matrices that contribute some Y to U_1 are B_2 and B_2'. Thus, $k_1' < k_2$, i.e., the last "X" coming from B_1 or B_1' is before the first "Y" coming from B_2 or B_2'. A similar reasoning ensures that $k_2' < k_3$ and so on. This shows that for any $i \in \{1, \ldots, k\}$, if $M_j = B_i$, then $j \in \{k_i, \ldots, k_i'\}$. Thus, all the B_1 appear before the B_2, which appear before the B_3, and so on. However, since the B_i are the only ones to contribute to V, then V must be of the form:

$$V = \prod_{i=1}^{k} e^{A_i t_i'},$$

where $t_i' \geqslant 0$ is the sum of all t_j such that $M_j = B_i$. Finally, $V = C$, so the instance of MEP is satisfiable. □

6 VECTOR AND HYPERPLANE REACHABILITY

In this section, we show that the vector and halfspace reachabililty problems for matrix-exponential semigroups, as given in Definition 1.4, are both undecidable.

Theorem 6.1. *The Matrix-Exponential Vector Reachability Problem is undecidable.*

Proof. This can be shown by reduction from the membership problem for matrix-exponential semigroups. In particular, given square matrices $B_1, \ldots, B_k, C$, we construct matrices $A_1, \ldots, A_k$ and vectors $\boldsymbol{x}, \boldsymbol{y}$ for which

$$\prod_{j=1}^{m} \exp(B_{i_j} t_j) = C \Leftrightarrow \prod_{j=1}^{m} \exp(A_{i_j} t_j)\boldsymbol{x} = \boldsymbol{y}. \tag{17}$$

Let $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ be the columns of C, from left to right, and let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ denote the canonical basis of $\mathbb{R}^n$. Then Equation (17) can be achieved by setting, for each $i \in \{1, \ldots, n\}$,

$$A_i = \begin{pmatrix} B_i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & B_i \end{pmatrix} \in \mathbb{R}^{n^2 \times n^2},$$

as well as

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_n \end{pmatrix} \text{ and } \boldsymbol{y} = \begin{pmatrix} \boldsymbol{c}_1 \\ \vdots \\ \boldsymbol{c}_n \end{pmatrix}. \qquad \square$$

Theorem 6.2. *The Matrix-Exponential Hyperplane Reachability Problem is undecidable.*

Proof. This can be shown by reduction from the vector reachability problem. Similarly to what we did in the previous proof, we will define matrices $C_1, \ldots, C_k$ and vectors $\boldsymbol{w}, \boldsymbol{z}$ such that

$$\prod_{i=1}^{k} \exp(A_i t_i)\boldsymbol{x} = \boldsymbol{y} \Leftrightarrow \boldsymbol{w}^T \prod_{i=1}^{k} \exp(C_i t_i)\boldsymbol{z} = 0.$$

Let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ denote the canonical basis of $\mathbb{R}^n$. Moreover, let

$$B_i = \begin{pmatrix} A_i & 0 \\ 0 & 0 \end{pmatrix}, \boldsymbol{u}_j = \begin{pmatrix} \boldsymbol{e}_j \\ -\boldsymbol{e}_j \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}$$

and

$$C_i = \begin{pmatrix} B_i \otimes I + I \otimes B_i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & B_i \otimes I + I \otimes B_i \end{pmatrix}.$$

Then

$$\begin{aligned}
&\prod_{i=1}^{k} \exp(A_i t_i)\boldsymbol{x} = \boldsymbol{y} \\
\Leftrightarrow &\sum_{j=1}^{n} \left(\boldsymbol{u}_j^T \prod_{i=1}^{k} \exp(B_i t_i)\boldsymbol{v} \right)^2 = 0 \\
\Leftrightarrow &\sum_{j=1}^{n} \left((\boldsymbol{u}_j \otimes \boldsymbol{u}_j)^T \prod_{i=1}^{k} (\exp(B_i t_i) \otimes \exp(B_i t_i)) (\boldsymbol{v} \otimes \boldsymbol{v}) \right) = 0 \\
\Leftrightarrow &\sum_{j=1}^{n} \left((\boldsymbol{u}_j \otimes \boldsymbol{u}_j)^T \prod_{i=1}^{k} \exp ((B_i \otimes I + I \otimes B_i) t_i) (\boldsymbol{v} \otimes \boldsymbol{v}) \right) = 0 \\
\Leftrightarrow &\begin{pmatrix} \boldsymbol{u}_1 \otimes \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_n \otimes \boldsymbol{u}_n \end{pmatrix}^T \prod_{i=1}^{k} \exp(C_i t_i) \begin{pmatrix} \boldsymbol{v} \otimes \boldsymbol{v} \\ \vdots \\ \boldsymbol{v} \otimes \boldsymbol{v} \end{pmatrix} = 0.
\end{aligned}$$

The result then follows by taking

$$\boldsymbol{w} = \begin{pmatrix} \boldsymbol{u}_1 \otimes \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_n \otimes \boldsymbol{u}_n \end{pmatrix} \text{ and } \boldsymbol{z} = \begin{pmatrix} \boldsymbol{v} \otimes \boldsymbol{v} \\ \vdots \\ \boldsymbol{v} \otimes \boldsymbol{v} \end{pmatrix}. \qquad \square$$

7 TURING-DEGREE OF MSMP

We are interested in classifying the Turing-degree of MSMP. Recall that the first level of the arithmetical hierarchy, denoted Σ_1, consists of the recursively enumerable languages, while the second level of the arithmetical hierarchy, denoted Σ_2, consists of the languages that can be enumerated by a Turing machine with an oracle to a language in Σ_1.

Theorem 7.1. *MSMP lies in* Σ_2.

Proof. Given an instance of MSMP, determined by generating matrices $A_1, \ldots, A_k$ and target C, consider the functions $f_{\boldsymbol{w}} : \mathbb{R}^{|\boldsymbol{w}|} \to \mathbb{R}$, defined for each $\boldsymbol{w} \in \{1, \ldots, k\}^*$ by

$$f_{\boldsymbol{w}}(\boldsymbol{t}) = \left\| \prod_{i=1}^{|\boldsymbol{w}|} \exp(A_{w_i} t_i) - C \right\|_F^2,$$

where $\| \cdot \|_F$ denotes the Frobenius norm of a matrix (obtained as the positive square root of the sum of the squares of the matrix entries). Note that each $f_{\boldsymbol{w}}$ is an exponential-polynomial that takes values in the nonnegative reals. Clearly, the given instance of MSMP is positive if and only if some $f_{\boldsymbol{w}}$ has a (necessarily tangential) zero.

Before proceeding, note that exponential-polynomials are closed under differentiation and that they are computable functions.

Let $\boldsymbol{b} : \mathbb{N} \to \{1, \ldots, k\}^*$ be any computable surjection. For each $n \in \mathbb{N}$ and $\boldsymbol{w} \in \{\boldsymbol{b}(1), \ldots, \boldsymbol{b}(n)\}$, consider the Turing machine $A_{\boldsymbol{w},n}$ that does the following: For each $m \in \mathbb{N}$, partition $[0, n]^{|\boldsymbol{w}|}$ in a uniform grid, with mesh size m^{-1}, and compute the approximate value of $f_{\boldsymbol{w}}$ with error at most m^{-1} at each grid point; if it is ever the case that all the approximate values of $f_{\boldsymbol{w}}$ are greater than $(1 + \frac{L_{\boldsymbol{w},n}\sqrt{|\boldsymbol{w}|}}{2})m^{-1}$ (where $L_{\boldsymbol{w},n}$ is an upper bound on $\|\nabla f_{\boldsymbol{w}}\| \restriction_{[0,n]}$, which we can compute by using the triangle inequality and the monotonicity of the exponential function), $A_{\boldsymbol{w},n}$ halts. Due to the Mean Value Theorem and to the compactness of $[0, n]$, $A_{\boldsymbol{w},n}$ halts if and only if $f_{\boldsymbol{w}} \restriction_{[0,n]}$ does not have a zero. Thus, the instance of MSMP is positive if and only if some $A_{\boldsymbol{w},n}$ does not halt.

Now consider the Turing machine B with access to an oracle for the Halting Problem that, for each $n \in \mathbb{N}$ and $\boldsymbol{w} \in \{\boldsymbol{b}(1), \ldots, \boldsymbol{b}(n)\}$, uses the oracle to decide whether $A_{\boldsymbol{w},n}$ halts and B only if halts if some oracle call determines that $A_{\boldsymbol{w},n}$ runs forever. Then B halts if and only if the MSMP instance in consideration is positive. □

Moreover, the following result holds.

Theorem 7.2. *If Schanuel's conjecture is true, then MSMP lies in* Σ_1.

Proof. Let $f_{\boldsymbol{w}}, \boldsymbol{w} \in \{1, \ldots, k\}^\star$ be as in the proof of Theorem 7.1. Consider the Turing Machine T, which, for each $n \in \mathbb{N}$ and $\boldsymbol{w} \in \{\boldsymbol{b}(1), \ldots, \boldsymbol{b}(n)\}$, uses Theorem 2.10 to decide whether $f_{\boldsymbol{w}}$ admits a zero in the region $[0, n]^{|\boldsymbol{w}|}$ and halts when such a zero is found. Then T halts if and only if the instance of MSMP under consideration is positive. □

8 CONCLUSION

We have shown that the Matrix-Exponential Semigroup Membership Problem is undecidable in general but decidable when the generating matrices $A_1, \ldots, A_k$ commute. Our results are analogous to what is known for the discrete version of this problem—the Matrix Semigroup Membership Problem—namely decidability in the commutative case and undecidability in general (see Section 1.1). Finally, we have shown that the Matrix-Exponential Semigroup Membership Problem is in Σ_1 if Schanuel's conjecture is true and in Σ_2 unconditionally. Note that for the Matrix Semigroup Membership Problem, membership in Σ_1 follows trivially from the fact that a finitely generated matrix semigroup is recursively enumerable.

It would be interesting to look at possibly decidable restrictions of Matrix-Exponential Semigroup Membership Problem: For example, the case with two non-commuting generators, which was shown to be decidable in the case of the analogous discrete problem in Reference [5]. Bounding the dimension of the ambient vector space might also yield decidability, along the lines of results in the discrete case in Reference [24]. Finally, deriving upper and lower bounds for the computational complexity of the commutative case remains to be addressed.

REFERENCES

[1] R. Alur. 2015. *Principles of Cyber-Physical Systems*. MIT Press.
[2] E. Asarin, V. Mysore, A. Pnueli, and G. Schneider. 2012. Low dimensional hybrid systems—Decidable, undecidable, don't know. *Inf. Comput.* 211 (2012), 138–159.
[3] L. Babai, R. Beals, J.-Y. Cai, G. Ivanyos, and E. M. Luks. 1996. Multiplicative equations over commuting matrices. In *SODA*. ACM/SIAM, 498–507.
[4] A. Baker. 1975. *Transcendental Number Theory*. Cambridge University Press.
[5] P. Bell, V. Halava, T. Harju, J. Karhumäki, and I. Potapov. 2008. Matrix equations and Hilbert's Tenth problem. *Int. J. Algebr. Comput.* 18, 8 (2008), 1231–1241.
[6] P. C. Bell, J.-C. Delvenne, R. M. Jungers, and V. D. Blondel. 2010. The continuous Skolem-Pisot problem. *Theor. Comput. Sci.* 411, 40–42 (2010), 3625–3634.
[7] J.-Y. Cai. 1994. Computing Jordan normal forms exactly for commuting matrices in polynomial time. *International J. Found. Comput. Sci.* 5, 3,4 (1994), 293–302.
[8] J.-Y. Cai, R. J. Lipton, and Y. Zalcstein. 2000. The complexity of the A B C problem. *SIAM J. Comput.* 29, 6 (2000), 1878–1888.
[9] T. Chen, N. Yu, and T. Han. 2015. Continuous-time orbit problems are decidable in polynomial-time. *Inform. Process. Lett.* 115, 1 (2015), 11–14.
[10] C. Choffrut and J. Karhumäki. 2005. Some decision problems on integer matrices. *Theor. Inf. Appl.* 39, 1 (2005), 125–131.
[11] H. Cohen. 1993. *A Course in Computational Algebraic Number Theory*. Springer-Verlag.
[12] E. Hainry. 2008. Reachability in linear dynamical systems. In *CiE (LNCS)*, Vol. 5028. Springer, 241–250.
[13] V. Halava. 1997. *Decidable and Undecidable Problems in Matrix Theory*. Technical Report 127. Turku Centre for Computer Science.
[14] B. Hall. 2015. *Lie Groups, Lie Algebras, and Representations*. Springer.
[15] T. A. Henzinger. 1996. The theory of hybrid automata. In *LICS*. IEEE Computer Society, 278–292.
[16] T. A. Henzinger, P. W. Kopke, A. Puri, and P. Varaiya. 1995. What's decidable about hybrid automata? In *STOC*. ACM, 373–382.
[17] R. Kannan and R. J. Lipton. 1986. Polynomial-time algorithm for the orbit problem. *J. ACM* 33, 4 (1986), 808–821.
[18] L. Khachiyan and L. Porkolab. 2000. Integer optimization on convex semialgebraic sets. *Discr. Comput. Geom.* 23, 2 (2000), 207–224.
[19] S. Lang. 1966. *Introduction to Transcendental Number Theory*.
[20] A. Macintyre and A. J. Wilkie. 1996. On the decidability of the real exponential field. In *Kreiseliana: About and Around Georg Kreisel*. A K Peters, 441–467.
[21] D. W. Masser. 1988. Linear relations on algebraic groups. In *New Advances in Transcendence Theory*. Cambridge University Press.
[22] M. Newman. 1967. Two classical theorems on commuting matrices. *J. Res. Nat. Bur. Stand.* 71 B, 2, 3 (1967), 69–71.
[23] M. S. Paterson. 1970. Undecidability in 3×3 matrices. *Journal of Mathematics and Physics* 49, 1 (1970), 105–107.
[24] I. Potapov and P. Semukhin. 2017. Decidability of the membership problem for 2×2 integer matrices. In *SODA*. SIAM, 170–186.

Received October 2017; accepted October 2018