On the Identity Problem for Unitriangular Matrices of Dimension Four

Ruiwen Dong ✉
Department of Computer Science, University of Oxford, UK

Abstract

We show that the Identity Problem is decidable in polynomial time for finitely generated subsemigroups of the group $\mathsf{UT}(4,\mathbb{Z})$ of 4×4 unitriangular integer matrices. As a byproduct of our proof, we also show the polynomial-time decidability of several subset reachability problems in $\mathsf{UT}(4,\mathbb{Z})$.

2012 ACM Subject Classification Computing methodologies → Symbolic and algebraic manipulation

Keywords and phrases identity problem, matrix semigroups, unitriangular matrices

Digital Object Identifier 10.4230/LIPIcs.MFCS.2022.43

Related Version *Full Version*: https://arxiv.org/abs/2202.05225

1 Introduction

Among the most prominent algorithmic problems for matrix semigroups are the *Identity Problem* and the *Membership Problem*. For the Membership Problem, the input is a finite set of square matrices $A_1,\ldots,A_k$ and a target matrix A. The problem is to decide whether A lies in the semigroup generated by $A_1,\ldots,A_k$. The Identity Problem is the Membership Problem restricted to the case where A is the identity matrix. These two problems are closely related to each other, and, as shown in many circumstances, studying the Identity Problem is usually the first step in studying the Membership Problem.

For general matrices, the Membership Problem is undecidable by a classical result of Markov [10]. Indeed, it is one of the earliest undecidability results on algorithmic problems in matrix semigroups. Most variants of the problem remain undecidable in low dimension. For example, the *Mortality Problem*, which is the Membership Problem in which the target matrix is 0, is undecidable in dimension three [12]. In dimension four, the Membership Problem is undecidable for matrices in $\mathrm{SL}(4,\mathbb{Z})$ (see [11]), while the Identity Problem is undecidable for the set of 4×4 integer matrices $\mathcal{M}_{4\times 4}(\mathbb{Z})$ (see [2]).

However, there has also been steady progress on the decidability side. The Membership Problem is shown to be decidable for $\mathrm{GL}(2,\mathbb{Z})$ in [4]. This decidability result is then extended to 2×2 integer matrices with nonzero determinant [13], and to 2×2 integer matrices with determinants equal to 0 and ± 1 [14]. It remains an intricate open problem whether the Membership Problem or the Identity Problem is decidable for $\mathrm{SL}(3,\mathbb{Z})$.

Recently, there has been more progress on closing the decidability gap by restricting consideration to the class of unitriangular matrices. It has long been known that the *Group Membership Problem* is decidable for $\mathsf{UT}(n,\mathbb{Z})$, the group of unitriangular integer matrices of dimension n. The Group Membership Problem asks to decide whether a matrix A lies in the *group* generated by given matrices $A_1,\ldots,A_k$. In fact, it is decidable for all finitely generated solvable matrix groups [8]. Later, Babai et al. [1] showed that the Group Membership Problem for *commuting matrices* can be computed in polynomial time (note that commuting matrices are simultaneously upper-triangularizable). However, there are significant differences between the group case and the semigroup case. In fact, for large enough n, the *Knapsack Problem* for $\mathsf{UT}(n,\mathbb{Z})$ is undecidable [7]. Given matrices $A_1,\ldots,A_k$

and A, the Knapsack Problem asks to decide whether there exist natural numbers $e_1, \dots, e_k$ such that $A_1^{e_1} \cdots A_k^{e_k} = A$. From the undecidability of the Knapsack Problem, one can deduce the undecidability of the semigroup Membership Problem for $\mathsf{UT}(n, \mathbb{Z})$ for large enough n [9].

Nevertheless, there have been some positive decidability results. The Identity Problem has been shown to be decidable for the group of 3×3 unitriangular integer matrices $\mathsf{UT}(3, \mathbb{Z})$ and the Heisenberg groups H_{2n+1} in [6]. Shortly after, the decidability result was extended to the Membership Problem [5]. Ko et al. left open the problem whether the Identity Problem in $\mathsf{UT}(n, \mathbb{Z})$ is decidable for $n \geq 4$, as well as finding the smallest n for which the Membership Problem for $\mathsf{UT}(n, \mathbb{Z})$ becomes undecidable.

The main result of this paper is that the Identity Problem is decidable in polynomial time for $\mathsf{UT}(4, \mathbb{Z})$. This further narrows the gap between decidability and undecidability and can be regarded as a first step towards the Membership Problem for $\mathsf{UT}(4, \mathbb{Z})$. The foundation of our method is the arguments developed in [5] for the Membership Problem of $\mathsf{UT}(3, \mathbb{Z})$. However, in order to pass from dimension three to four, we need to introduce additional methods from convex geometry, linear programming and even use the aid of computational algebraic geometry software. The proof for $\mathsf{UT}(3, \mathbb{Z})$ heavily relies on the fact that the subgroup generated by commutators of matrices from a given subset of $\{A_1, \dots, A_k\} \subset \mathsf{UT}(3, \mathbb{Z})$ is isomorphic to a subgroup of $\mathbb{Z}$. This is no longer the case for $\mathsf{UT}(4, \mathbb{Z})$. However, $\mathsf{UT}(4, \mathbb{Z})$ is still metabelian [15], and its derived subgroup is isomorphic to $\mathbb{Z}^3$. Given a finite set $\mathcal{G} \subseteq \mathsf{UT}(4, \mathbb{Z})$, we construct elements in $\langle \mathcal{G} \rangle$ that fall inside the derived subgroup of $\mathsf{UT}(4, \mathbb{Z})$. These elements then generate a cone in $\mathbb{Z}^3$ under the isomorphism between the derived subgroup and $\mathbb{Z}^3$. The possible shapes of this cone will determine the Identity Problem.

There is strong evidence that the new techniques introduced in this paper can help tackle the Identity Problem for $\mathsf{UT}(n, \mathbb{Z})$ with $n \geq 5$.

2 Preliminaries

Denote by $\mathsf{UT}(4, \mathbb{Z})$ the group of upper triangular integer matrices with ones on the diagonal:

$$\mathsf{UT}(4, \mathbb{Z}) \coloneqq \left\{ \begin{pmatrix} 1 & a & d & f \\ 0 & 1 & b & e \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{pmatrix} \middle| \, a, b, c, d, e, f \in \mathbb{Z} \right\}.$$

Denote its normal subgroups

$$\mathsf{U}_1 \coloneqq \left\{ \begin{pmatrix} 1 & 0 & d & f \\ 0 & 1 & 0 & e \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \middle| \, d, e, f \in \mathbb{Z} \right\}, \quad \mathsf{U}_2 \coloneqq \left\{ \begin{pmatrix} 1 & 0 & 0 & f \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \middle| \, f \in \mathbb{Z} \right\}$$

in the lower central series: $\mathsf{UT}(4, \mathbb{Z}) \trianglerighteq \mathsf{U}_1 = [\mathsf{UT}(4, \mathbb{Z}), \mathsf{UT}(4, \mathbb{Z})] \trianglerighteq \mathsf{U}_2 = [\mathsf{UT}(4, \mathbb{Z}), \mathsf{U}_1]$ (see [15, Chapter 5]). In particular, U_1 and U_2 are respectively the derived subgroup and the centre of $\mathsf{UT}(4, \mathbb{Z})$. For convenience, we introduce the following notations:

$$UT(a, b, c; d, e, f) \coloneqq \begin{pmatrix} 1 & a & d & f \\ 0 & 1 & b & e \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad U_1(d, e, f) \coloneqq UT(0, 0, 0; d, e, f).$$

There are surjective group homomorphisms $\varphi_0 \colon \mathsf{UT}(4, \mathbb{Z}) \to \mathbb{Z}^3$ defined by

$$\varphi_0(UT(a, b, c; d, e, f)) = (a, b, c),$$

with $\ker(\varphi_0) = \mathsf{U}_1$, and $\varphi_1 \colon \mathsf{U}_1 \to \mathbb{Z}^2$,

$$\varphi_1(U_1(d,e,f)) = (d,e),$$

with $\ker(\varphi_1) = \mathsf{U}_2$. Moreover, U_1 is itself abelian, with a natural isomorphism $\tau \colon \mathsf{U}_1 \xrightarrow{\sim} \mathbb{Z}^3$:

$$\tau(U_1(d,e,f)) = (d,e,f).$$

Denote by τ_d the projection $U_1(d,e,f) \mapsto d$, τ_e the projection $U_1(d,e,f) \mapsto e$, and τ_f the projection $U_1(d,e,f) \mapsto f$. Then, $\varphi_1 = (\tau_d, \tau_e)$ and $\tau = (\tau_d, \tau_e, \tau_f)$.

Finally, define the subgroup of $\mathsf{UT}(4,\mathbb{Z})$:

$$\mathsf{U}_{10} \coloneqq \{U_1(0,e,f) \mid e,f \in \mathbb{Z}\} \trianglelefteq \mathsf{U}_1.$$

For a finite set of matrices $\mathcal{G} = \{A_1, \ldots, A_k\}$, denote by $\langle \mathcal{G} \rangle$ the semigroup generated by $\mathcal{G}$. In this paper, we are concerned with the following problems.

▶ **Definition 1.** Let G be a monoid of matrices, and H a subset of G.

(i) The *Identity Problem* in G asks, given a finite set of matrices $\mathcal{G}$ in G, whether $I \in \langle \mathcal{G} \rangle$. If this is the case, we say that the identity matrix is *reachable*.

(ii) The *H-Reachability Problem* in G asks, given a finite set of matrices $\mathcal{G}$ in G, whether $H \cap \langle \mathcal{G} \rangle \neq \emptyset$. If this is the case, we say that H is *reachable*.

The main result of this paper is that the Identity Problem in $\mathsf{UT}(4,\mathbb{Z})$ is decidable in polynomial time, with respect to the number of bits required to encode all the entries of the matrices in $\mathcal{G}$ (each matrix $UT(a,b,c,d,e,f)$ is encoded by the entries a,b,c,d,e,f).

It turns out that the three problems: Identity Problem, U_2-Reachability and U_{10}-Reachability are interconnected and it is more convenient to devise algorithms that decide them simultaneously. A trivial observation is that, because $I \in \mathsf{U}_2 \subset \mathsf{U}_{10}$, a positive instance of the Identity Problem is also a positive instance of U_2-Reachability; and a positive instance of U_2-Reachability is also a positive instance of U_{10}-Reachability.

The following definitions will be used throughout this paper.

▶ **Definition 2** (String, product and Parikh vector)**.** Let $\mathcal{G} = \{A_1, \ldots, A_k\}$ be a fixed set of matrices in $\mathsf{UT}(4,\mathbb{Z})$. A *string* of $\mathcal{G}$ is an expression $B_1 B_2 \cdots B_m$ such that $B_i \in \mathcal{G}, i = 1, \ldots, m$. The *product* of a string $B_1 B_2 \cdots B_m$ is the matrix $P \in \mathsf{UT}(4,\mathbb{Z})$ such that $P = B_1 B_2 \cdots B_m$. The *Parikh vector* of a string $B_1 B_2 \cdots B_m$ is the vector $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_k) \in \mathbb{Z}_{\geq 0}$ where

$$\ell_j = \mathrm{card}(\{i \mid B_i = A_j\}), j = 1, \ldots, k.$$

When $\mathcal{G}$ is clear from the context, we simply use the term "string" instead of "string of $\mathcal{G}$".

For an integer $n \geq 1$, the *Heisenberg group* of dimension $2n+1$ is the group H_{2n+1} of $(n+2) \times (n+2)$ integer matrices of the form $H = \begin{pmatrix} 1 & \boldsymbol{a} & c \\ 0 & I_n & \boldsymbol{b}^\top \\ 0 & 0 & 1 \end{pmatrix}$, where $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{Z}^n$, $c \in \mathbb{Z}$.

The following result comes from [6] and [5].

▶ **Lemma 3** ([5, Theorem 7])**.** *The Identity Problem and the Membership Problem in* H_{2n+1} *are decidable for all* $n \geq 1$.

3 Identity problem, U_2- and U_{10}-Reachability in $\mathsf{UT}(4,\mathbb{Z})$

In this section, we construct algorithms that decide the Identity Problem, U_2-Reachability and U_{10}-Reachability in $\mathsf{UT}(4,\mathbb{Z})$.

3.1 Overview of decision strategy

For any set of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l \in \mathbb{R}^n$, denote by

$$\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l \rangle_{\mathbb{R}_{\geq 0}} \coloneqq \left\{ \sum_{i=1}^{l} r_i \boldsymbol{v}_i \middle| r_i \in \mathbb{R}_{\geq 0}, i = 1, \ldots, l \right\}$$

the $\mathbb{R}_{\geq 0}$-cone generated by $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l$, and by $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l \rangle_{\mathbb{R}}$ the $\mathbb{R}$-vector space spanned by $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l$.

Let $\mathcal{G} = \{A_1, \ldots, A_k\}$ be a set of matrices in $\mathsf{UT}(4, \mathbb{Z})$, for which we want to decide the Identity Problem, U_2-Reachability and U_{10}-Reachability. Define the $\mathbb{R}_{\geq 0}$-cone

$$\mathcal{C} \coloneqq \langle \varphi_0(A_1), \ldots, \varphi_0(A_k) \rangle_{\mathbb{R}_{\geq 0}}, \tag{1}$$

and denote by $\mathcal{C}^{lin}$ its lineality space, i.e. the largest linear subspace (by inclusion) contained in $\mathcal{C}$. In particular, $\mathcal{C}^{lin} = \mathcal{C} \cap -\mathcal{C}$. A basis of $\mathcal{C}^{lin}$ can be effectively computed in polynomial time [16]. For any matrix $A_i \in \mathcal{G}$, the projection $\varphi_0(A_i)$ can be either in $\mathcal{C}^{lin}$ or in $\mathcal{C} \setminus \mathcal{C}^{lin}$. However, in order to reach U_1, which contains the identity matrix, U_2 and U_{10}, one can only use matrices A_i with $\varphi_0(A_i) \in \mathcal{C}^{lin}$. This is formally stated by the following proposition.

▶ **Proposition 4.** *If the product of a string $B_1 \cdots B_m$ is in U_1, then every $B_j, j = 1, \ldots m$, must be in the set $\{A_i \in \mathcal{G} \mid \varphi_0(A_i) \in \mathcal{C}^{lin}\}$.*

Proof. Suppose on the contrary that some B_j satisfies $\varphi_0(B_j) \in \mathcal{C} \setminus \mathcal{C}^{lin}$.

Since φ_0 is a group homomorphism, we have

$$B_1 \cdots B_m \in \mathsf{U}_1 \iff \varphi_0(B_1 \cdots B_m) = \mathbf{0} \iff \sum_{i=1}^{m} \varphi_0(B_i) = \mathbf{0}.$$

Therefore, $-\varphi_0(B_j) = \sum_{i \neq j} \varphi_0(B_i) \in \mathcal{C}$.

Hence, the linear subspace $\varphi_0(B_j)\mathbb{R} = \langle \varphi_0(B_j), -\varphi_0(B_j) \rangle_{\mathbb{R}_{\geq 0}}$ is contained in $\mathcal{C}$. This yields $\varphi_0(B_j)\mathbb{R} \subseteq \mathcal{C}^{lin}$, a contradiction to $\varphi_0(B_j) \in \mathcal{C} \setminus \mathcal{C}^{lin}$. ◀

The overall strategy for constructing our algorithm is to use induction on $\mathrm{card}(\mathcal{G})$. If $\mathrm{card}(\mathcal{G}) = 0$, then the answers to the Identity Problem, U_2-Reachability and U_{10}-Reachability are all negative. Suppose now that we have an algorithm that decides all three problems for every set of at most $k-1$ matrices, we will construct an algorithm that decides them for a set of k matrices $\mathcal{G} = \{A_1, \ldots, A_k\}$. By Proposition 4, if some matrix A_i satisfies $\varphi_0(A_i) \in \mathcal{C} \setminus \mathcal{C}^{lin}$, then we can discard it without changing the answer to the Identity Problem or $\mathsf{U}_2, \mathsf{U}_{10}$-Reachability. This decreases the number of elements in $\mathcal{G}$, and an algorithm is available by the induction hypothesis on $\mathrm{card}(\mathcal{G})$. Hence, we can suppose that every $A_i \in \mathcal{G}$ satisfies $\varphi_0(A_i) \in \mathcal{C}^{lin}$, so $\mathcal{C} = \mathcal{C}^{lin}$ is a linear space.

Since $\varphi_0(A_i) \in \mathbb{Z}^3$, $\mathcal{C}$ is a linear subspace of $\mathbb{R}^3$. We identify cases according to the dimension of $\mathcal{C}$, with each of the following four subsections treating the case of dimension 3, 1, 0, 2. The pseudocode of the decision procedure for the Identity Problem is given here as a reference point for the detailed case analysis. The decision procedures for U_2-reachability and U_{10}-reachability follow similar patterns and their pseudocode is given in the appendix of the full version of this paper. Note that the decision procedure for the Identity Problem invokes the decision procedure for U_2-reachability as a subroutine. Similarly, the decision procedure for U_2-reachability will invoke the decision procedure for U_{10}-reachability as a subroutine.

Algorithm 1 IdentityProblem(): deciding the Identity Problem for a subset of $\mathsf{UT}(4,\mathbb{Z})$.

Input: A set $\mathcal{G} = \{A_1, \ldots, A_k\}$ of matrices in $\mathsf{UT}(4,\mathbb{Z})$.
Output: True or False.

Step 1: Compute the cone $\mathcal{C}$ and its lineality space $\mathcal{C}^{lin}$. For $i = 1, \ldots, k$, if some $\varphi_0(A_i)$ is not in $\mathcal{C}^{lin}$, return IdentityProblem($\mathcal{G} \setminus \{A_i\}$).
Step 2:
- **a.** If $\dim(\mathcal{C}) = 3$, return True.
- **b.** If $\dim(\mathcal{C}) = 1$, return True if the condition in Proposition 15(i) is satisfied, otherwise return False.
- **c.** If $\dim(\mathcal{C}) = 0$, return True if $\tau(A_i), i = 1, \ldots, m$ generate a semigroup containing $\mathbf{0}$, otherwise return False.
- **d.** If $\dim(\mathcal{C}) = 2$, compute a non-zero vector $(p, q, r) \in \mathbb{Q}^3$ orthogonal to $\mathcal{C}$.
  - **i.** If $p = 0$, but q, r are not zero, or $r = 0$, but q, p are not zero. Compute L_0, if $\mathrm{supp}(L_0) = \{1, \ldots, k\}$, return True, otherwise return IdentityProblem($\{A_i \mid i \in \mathrm{supp}(L_0)\}$).
  - **ii.** If $p = r = 0$, problem reduces to Identity Problem in H_5.
  - **iii.** If $p = q = 0, r \neq 0$ or $r = q = 0, p \neq 0$, compute A'_i as in (9). Return U2Reachability($A'_1, \ldots, A'_k$) (see full version of paper).

We now give an overview of the motivation behind classifying cases according to the dimension of $\mathcal{C}$. As a convention, we always use $A_i, i = 1, \ldots, k$ to denote elements of the fixed generating set $\mathcal{G}$, and Greek letters to denote their entries, i.e. $A_i = UT(\alpha_i, \beta_i, \kappa_i; \delta_i, \epsilon_i, \phi_i)$. We use $B_i, i = 1, \ldots, m$ to denote arbitrary elements in $\langle\mathcal{G}\rangle$ (when appearing in *strings*, they are elements in $\mathcal{G}$), and Latin letters to denote their entries, i.e. $B_i = UT(a_i, b_i, c_i, d_i, e_i, f_i)$. The variables B_i can depend on the context.

First of all, we need some results on the structure of products in $\mathsf{UT}(4,\mathbb{Z})$. For a positive integer m, denote by S_m the permutation group of the set $\{1, \ldots, m\}$. Throughout this paper, given some matrices $B_1, \ldots, B_m \in \langle\mathcal{G}\rangle$, we will often be computing the product of strings of the form $B^t_{\sigma(1)} \cdots B^t_{\sigma(m)}$, where $\sigma \in \mathrm{S}_m$ and $t \in \mathbb{Z}_{\geq 0}$. The overall idea is to find various strings $B^t_{\sigma(1)} \cdots B^t_{\sigma(m)}$ whose product is in $\mathsf{U}_1 \overset{\tau}{\cong} \mathbb{Z}^3$, then use them to generate an abelian semigroup containing the identity matrix. Let us define the following important values and abbreviations that will be used throughout this paper. These complicated formulas are related to the *logarithm* of the matrices B_i, and readers can for the time being ignore their exact form and treat them as black boxes.

▶ **Notation 5.** Given a series of matrices $B_1, \ldots, B_m$ where $B_i = UT(a_i, b_i, c_i; d_i, e_i, f_i), i = 1, \ldots, m$, we introduce the following notation:

(i) For $\sigma \in \mathrm{S}_m, t \in \mathbb{Z}_{\geq 0}$,

$$B(\sigma, t) \coloneqq B^t_{\sigma(1)} \cdots B^t_{\sigma(m)}. \tag{2}$$

(ii) For $\sigma \in \mathrm{S}_m$,

$$D_\sigma \coloneqq \sum_{i<j} a_{\sigma(i)} b_{\sigma(j)} + \frac{1}{2} \sum_{i=1}^{m} a_i b_i, \quad E_\sigma \coloneqq \sum_{i<j} b_{\sigma(i)} c_{\sigma(j)} + \frac{1}{2} \sum_{i=1}^{m} b_i c_i,$$

$$F_\sigma \coloneqq \sum_{i<j<k} a_{\sigma(i)} b_{\sigma(j)} c_{\sigma(k)} + \frac{1}{2} \sum_{i<j} (a_{\sigma(i)} b_{\sigma(i)} c_{\sigma(j)} + a_{\sigma(i)} b_{\sigma(j)} c_{\sigma(j)}) + \frac{1}{6} \sum_{i=1}^{m} a_i b_i c_i,$$

$$G_\sigma \coloneqq \sum_{i<j}(a_{\sigma(i)}e_{\sigma(j)} + d_{\sigma(i)}c_{\sigma(j)} - \frac{1}{2}a_{\sigma(i)}b_{\sigma(j)}c_{\sigma(j)} - \frac{1}{2}a_{\sigma(i)}b_{\sigma(i)}c_{\sigma(j)}) + \frac{1}{2}\sum_{i=1}^{m}(a_ie_i + d_ic_i - a_ib_ic_i). \quad (3)$$

(iii) For $i = 1, \ldots, m$,

$$D_i \coloneqq d_i - \frac{1}{2}a_ib_i, \quad E_i \coloneqq e_i - \frac{1}{2}b_ic_i, \quad F_i \coloneqq f_i - \frac{1}{2}(a_ie_i + d_ic_i) + \frac{1}{3}a_ib_ic_i. \quad (4)$$

The following proposition gives an exact expression for $B(\sigma, t)$. Because of the heavily computational nature of most of our propositions, their proofs are given in the appendix of the full version of this paper.

▶ **Proposition 6.** *Let $B_i = UT(a_i, b_i, c_i; d_i, e_i, f_i), i = 1, \ldots, m$, $\sigma \in \mathrm{S}_m, t \in \mathbb{Z}_{\geq 0}$, then*

$$B(\sigma, t) = UT\left(t\sum_{i=1}^{m}a_i, t\sum_{i=1}^{m}b_i, t\sum_{i=1}^{m}c_i; t^2D_\sigma + t\sum_{i=1}^{m}D_i, t^2E_\sigma + t\sum_{i=1}^{m}E_i, t^3F_\sigma + t^2G_\sigma + t\sum_{i=1}^{m}F_i\right). \quad (5)$$

Notice that $B(\sigma, t) \in \mathsf{U}_1$ if and only if $\sum_{i=1}^m a_i = \sum_{i=1}^m b_i = \sum_{i=1}^m c_i = 0$, a condition that does not depend on the value of t.

Proposition 6 shows that, if $B(\sigma, t)$ is in U_1, then as $t \to \infty$, the asymptotic behaviour of $\tau(B(\sigma, t))$ approaches the vector $(t^2D_\sigma, t^2E_\sigma, t^3F_\sigma)$, provided that $D_\sigma, E_\sigma, F_\sigma$ do not vanish. Therefore, the hope is that, as t, σ vary, the vectors $(t^2D_\sigma, t^2E_\sigma, t^3F_\sigma)$ can generate $\mathbb{R}^3$ as an $\mathbb{R}_{\geq 0}$-cone, barring a few degenerate cases. If these degenerate cases do not happen, then the different vectors $\tau(B(\sigma, t))$ will also generate $\mathbb{R}^3$ as an $\mathbb{R}_{\geq 0}$-cone. In particular, the identity element in $\mathbb{R}^3$ can be generated by $\tau(B(\sigma, t))$ as an additive semigroup, giving a positive answer to the Identity Problem. For the degenerate cases, they will be treated individually. As it will turn out, there are only two types of degeneracy (which may occur simultaneously):

(i) $F_\sigma = 0$ for all σ.

(ii) For some $p, r \in \mathbb{Q}$, possibly zero, we have $pD_\sigma = rE_\sigma$ for all σ.

When (i) occurs, the asymptotic behaviour of $\tau(B(\sigma, t))$ approaches the vector $(t^2D_\sigma, t^2E_\sigma, t^2G_\sigma)$, since G_σ is the second most dominant term after F_σ. This situation reminds us of the Identity Problem for H_3, and can be solved in a similar way. When (ii) occurs, the vectors $(t^2D_\sigma, t^2E_\sigma, t^3F_\sigma)$ are constrained to a strict linear subspace of $\mathbb{R}^3$. Hence, in order to describe the $\mathbb{R}_{\geq 0}$-cone generated by the vectors $\tau(B(\sigma, t))$, one needs to consider the sub-dominant terms as well, i.e. the terms $t\sum_{i=1}^m D_i, t\sum_{i=1}^m E_i$.

The rest of this paper aims to formalize this idea. We first exhibit a series of lemmas that characterise these degenerate cases. Our first lemma shows that, supposing $B(\sigma, t) \in \mathsf{U}_1$, then degenerate case (ii) happens if and only if $\langle \varphi_0(B_1), \ldots, \varphi_0(B_m)\rangle_{\mathbb{R}}$ is degenerate (i.e. of dimension at most 2).

▶ **Lemma 7.** *Given $p, r \in \mathbb{R}$ and $m \geq 2$. Suppose $\sum_{i=1}^m a_i = \sum_{i=1}^m b_i = \sum_{i=1}^m c_i = 0$. The two following statements are equivalent:*

(i) *For all $\sigma \in \mathrm{S}_m$, $pD_\sigma = rE_\sigma$.*

(ii) *Either $b_i = 0$ for all $i = 1, \ldots, m$, or there exist $q \in \mathbb{R}$, such that $pa_i + qb_i + rc_i = 0$ for all $i = 1, \ldots, m$.*

The next lemma shows that if $B(\sigma,t) \in \mathsf{U}_1$, then by "inverting" σ, we get a permutation σ' such that (D_σ, E_σ) and $(D_{\sigma'}, E_{\sigma'})$ are opposites of one another.

▶ **Lemma 8.** *Suppose* $\sum_{i=1}^m a_i = \sum_{i=1}^m b_i = \sum_{i=1}^m c_i = 0$, $m \geq 2$. *For every* $\sigma \in \mathrm{S}_m$, *there exists* $\sigma' \in \mathrm{S}_m$, *such that* $(D_{\sigma'}, E_{\sigma'}) = -(D_\sigma, E_\sigma)$.

We then show that, if $B(\sigma,t) \in \mathsf{U}_1$, then the value of F_σ for different $\sigma \in \mathrm{S}_m$ sums up to zero:

▶ **Lemma 9.** *Suppose* $\sum_{i=1}^m a_i = \sum_{i=1}^m b_i = \sum_{i=1}^m c_i = 0$, *where* $m \geq 3$. *Then we have* $\sum_{\sigma\in\mathrm{S}_m} F_\sigma = 0$.

The last lemma characterizes situations where the aforementioned degenerate case (i) happens. Its proof relies on the aid of a computational algebraic geometry software due to the complexity of the expressions F_σ.

▶ **Lemma 10.** *Let* $m = 4$. *Suppose* $\sum_{i=1}^4 a_i = \sum_{i=1}^4 b_i = \sum_{i=1}^4 c_i = 0$. *Then,* $F_\sigma = 0$ *for all* $\sigma \in \mathrm{S}_4$, *if and only if at least one of the following four conditions holds:*

(i) $a_1 = a_2 = a_3 = a_4 = 0$.

(ii) $b_1 = b_2 = b_3 = b_4 = 0$.

(iii) $c_1 = c_2 = c_3 = c_4 = 0$.

(iv) $\operatorname{rank}\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{pmatrix} \leq 1$.

A common idea of Lemma 7 and Lemma 10 is that the degeneracy of $(D_\sigma, E_\sigma, F_\sigma)$ is related to the degeneracy of $\varphi_0(B_1), \ldots, \varphi_0(B_m)$. Hence, it is natural to consider the degeneracy of the vectors $\varphi_0(A_i), i = 1, \ldots, k$, where $A_i \in \mathcal{G}$ are the elements of the generating set. This degeneracy is described by the dimension of the linear space $\mathcal{C}$ discussed at the beginning of the section. This justifies the classification according to $\dim(\mathcal{C})$. We now begin the case analysis.

3.2 $\mathcal{C}$ has dimension 3

The main idea of this case is that, for a well chosen set of matrices $B_1, B_2, B_3, B_4 \in \langle\mathcal{G}\rangle$, the vectors $(D_\sigma, E_\sigma, F_\sigma), \sigma \in \mathrm{S}_4$, are not degenerate and the asymptotic behaviour of $\tau(B(\sigma,t))$ approaches the vector $(t^2 D_\sigma, t^2 E_\sigma, t^3 F_\sigma)$, leading to a positive answer to the Identity Problem.

Let $B_1, B_2, B_3, B_4 \in \langle\mathcal{G}\rangle$ with $B_i = UT(a_i, b_i, c_i; d_i, e_i, f_i), i = 1, \ldots, 4$ be such that

$$\sum_{i=1}^4 \varphi_0(B_i) = 0 \tag{6}$$

and

$$\langle \varphi_0(B_1), \varphi_0(B_2), \varphi_0(B_3), \varphi_0(B_4) \rangle_{\mathbb{R}_{\geq 0}} = \mathcal{C} = \mathbb{R}^3. \tag{7}$$

Equation (6) shows that $B(\sigma,t) \in \mathsf{U}_1$ for all $\sigma \in \mathrm{S}_4, t \in \mathbb{Z}_{\geq 0}$.

The following lemma shows that the d, e-coordinates of different $\tau(B(\sigma,t))$ generate $\mathbb{R}^2$ as an $\mathbb{R}_{\geq 0}$-cone.

▶ **Lemma 11.** *Assuming* (6) *and* (7), *we have* $\langle \{\varphi_1(B(\sigma,t)) \mid \sigma \in \mathrm{S}_4, t \in \mathbb{Z}\} \rangle_{\mathbb{R}_{\geq 0}} = \mathbb{R}^2$.

Proof. First, we claim that $\langle\{(D_\sigma, E_\sigma) \mid \sigma \in \mathrm{S}_4\}\rangle_{\mathbb{R}} = \mathbb{R}^2$.

In fact, suppose to the contrary that $\langle\{(D_\sigma, E_\sigma) \mid \sigma \in \mathrm{S}_4\}\rangle_{\mathbb{R}}$ has dimension at most 1. Then there exist $p, r \in \mathbb{R}$, not both zero, such that for all $\sigma \in \mathrm{S}_4$, $pD_\sigma = rE_\sigma$. By Lemma 7, this means that either $b_i = 0$ for all i or there exists some $q \in \mathbb{R}$ such that $pa_i + qb_i + rc_i = 0$ for all i. In both cases, the $\mathbb{R}$-linear subspace spanned by $\varphi_0(B_1), \varphi_0(B_2), \varphi_0(B_3), \varphi_0(B_4)$ has dimension at most 2, contradicting Equation (7). This proves the claim. Hence, there exist $\sigma_1, \sigma_2 \in \mathrm{S}_4$ such that $(D_{\sigma_1}, E_{\sigma_1})$ and $(D_{\sigma_2}, E_{\sigma_2})$ span $\mathbb{R}^2$ as an $\mathbb{R}$-linear space.

Next, by Lemma 8, there exist $\sigma_1', \sigma_2' \in \mathrm{S}_4$ such that $(D_{\sigma_1'}, E_{\sigma_1'}) = -(D_{\sigma_1}, E_{\sigma_1})$ and $(D_{\sigma_2'}, E_{\sigma_2'}) = -(D_{\sigma_2}, E_{\sigma_2})$. It follows that $(D_{\sigma_1}, E_{\sigma_1}), (D_{\sigma_2}, E_{\sigma_2}), (D_{\sigma_1'}, E_{\sigma_1'}), (D_{\sigma_2'}, E_{\sigma_2'})$ generate $\mathbb{R}^2$ as an $\mathbb{R}_{\geq 0}$-cone, and all four vectors are non-zero.

Finally, consider the products $B(\sigma, t)$ with $\sigma \in \{\sigma_1, \sigma_2, \sigma_1', \sigma_2'\}$. By Proposition 6, when $t \to +\infty$, we have $\varphi_1(B(\sigma, t)) = (D_\sigma, E_\sigma)t^2 + O(t)$. Therefore, when t is large enough, the angle between $\varphi_1(B(\sigma, t))$ and (D_σ, E_σ) tends to zero, for all $\sigma \in \{\sigma_1, \sigma_2, \sigma_1', \sigma_2'\}$. Hence, for large enough t, $\varphi_1(B(\sigma_1, t)), \varphi_1(B(\sigma_2, t)), \varphi_1(B(\sigma_1', t)), \varphi_1(B(\sigma_2', t))$ generate $\mathbb{R}^2$ as an $\mathbb{R}_{\geq 0}$-cone. This proves the Lemma. ◀

The next proposition shows that as σ, t vary, the vectors $\tau(B(\sigma, t))$ generate $\mathbb{R}^3$ as an $\mathbb{R}_{\geq 0}$-cone.

▶ **Proposition 12.** *Assuming* (6) *and* (7)*, we have* $\langle\{\tau(B(\sigma, t)) \mid \sigma \in \mathrm{S}_4, t \in \mathbb{Z}\}\rangle_{\mathbb{R}_{\geq 0}} = \mathbb{R}^3$.

Proof. First, note that all $B(\sigma, t)$ have integer coefficients. By Lemma 11, there exist elements $P_1, P_2, P_3 \in \langle\{B(\sigma, t) \mid \sigma \in \mathrm{S}_4, t \in \mathbb{Z}\}\rangle$ such that $\varphi_1(P_i), i = 1, 2, 3$ generate $\mathbb{R}^2$ as an $\mathbb{R}_{\geq 0}$-cone (see Figure 1 for an illustration.).

Next, the idea is to find two additional matrices $P_+, P_- \in \{B(\sigma, t) \mid \sigma \in \mathrm{S}_4, t \in \mathbb{Z}\}$, whose images under τ are relatively close to the f-axis in $\mathbb{R}^3$. By Lemmas 9 and 10, there exist $\sigma_+, \sigma_- \in \mathrm{S}_4$ such that $F_{\sigma_+} > 0, F_{\sigma_-} < 0$. Indeed, by condition (7), none of the four conditions of Lemma 10 hold. Thus there exists $\sigma \in \mathrm{S}_4$ such that $F_\sigma \neq 0$. Then Lemma 9 shows we can find σ_+ and σ_- such that $F_{\sigma_+} > 0$ and $F_{\sigma_-} < 0$.

By Proposition 6, when $t \to +\infty$, we have $\tau_f(B(\sigma_+, t)) = F_{\sigma_+}t^3 + O(t^2)$ and $\tau_f(B(\sigma_-, t)) = F_{\sigma_-}t^3 + O(t^2)$, whereas $\tau_d(B(\sigma_\pm, t)) = O(t^2)$ and $\tau_e(B(\sigma_\pm, t)) = O(t^2)$. Therefore, when t is large enough, the angle between $\tau(B(\sigma_+, t))$ and $(0, 0, 1)$ tends to zero, as well as the angle between $\tau(B(\sigma_-, t))$ and $(0, 0, -1)$.

Finally, we claim that there exists t such that $\tau(P_1), \tau(P_2), \tau(P_3), \tau(B(\sigma_+, t)), \tau(B(\sigma_-, t))$ generate $\mathbb{R}^3$ as an $\mathbb{R}_{\geq 0}$-cone. See Figure 1 for an illustration. To justify this claim, suppose to the contrary that for every t, the $\mathbb{R}_{\geq 0}$-cone spanned by the five vectors $\tau(P_1), \tau(P_2), \tau(P_3)$, $\tau(B(\sigma_+, t))$, $\tau(B(\sigma_-, t))$ is a proper subset of $\mathbb{R}^3$. In other words, if we denote by $\langle \cdot, \cdot \rangle$ the canonical inner product of $\mathbb{R}^3$, then there exists a vector $\boldsymbol{v}_t$ with norm 1, such that $\langle \boldsymbol{v}_t, \tau(P_i) \rangle \geq 0, i = 1, 2, 3$ and $\langle \boldsymbol{v}_t, \tau(B(\sigma_\pm, t)) \rangle \geq 0$. For example, we can take $\boldsymbol{v}_t$ to be any normalized vector in the dual of the cone generated by these five vectors ([3, Chapter 2.6]). By the compactness of the unit sphere, $\{\boldsymbol{v}_t\}_{t \in \mathbb{N}}$ has a limit point $\boldsymbol{v}$. We have $\langle \boldsymbol{v}, \tau(P_i) \rangle \geq 0, i = 1, 2, 3$, so $\boldsymbol{v}$ is not orthogonal to the f-axis, otherwise $\tau(P_i), i = 1, 2, 3$ would all be on the same side of a hyperplane passing through the f-axis, contradicting the fact that their d, e-coordinates generate $\mathbb{R}^2$ as an $\mathbb{R}_{\geq 0}$-cone. Hence, $\langle \boldsymbol{v}, (0, 0, 1) \rangle \neq 0$. Without loss of generality, suppose $\langle \boldsymbol{v}, (0, 0, 1) \rangle < 0$. When $t \to \infty$, the angle between $(0, 0, 1)$ and $\tau(B(\sigma_+, t))$ tends to zero. Therefore, for all large enough t, we have $\langle \boldsymbol{v}, \tau(B(\sigma_+, t)) \rangle < 0$. Since $\boldsymbol{v}$ is a limit point of $\{\boldsymbol{v}_t\}_{t \in \mathbb{N}}$, there exists a large enough t such that $\langle \boldsymbol{v}_t, \tau(B(\sigma_+, t)) \rangle < 0$. This contradicts the fact that $\langle \boldsymbol{v}_t, \tau(B(\sigma_+, t)) \rangle \geq 0$ for all t. ◀

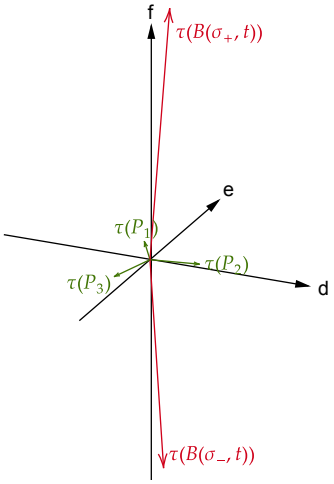

Figure 1 Illustration of the five vectors constructed in Proposition 12.

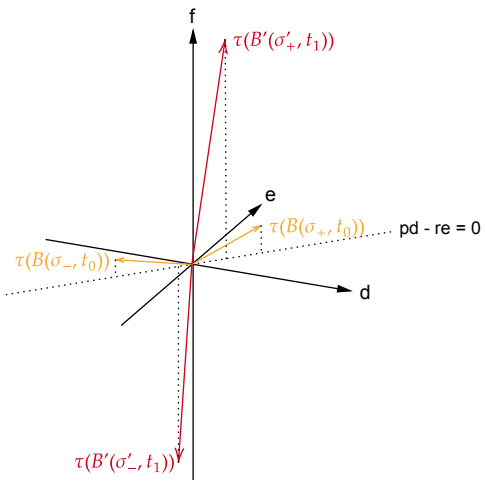

Figure 2 Illustration of the four vectors constructed in Proposition 19.

▶ **Corollary 13.** *When $\mathcal{C}$ has dimension 3, the identity matrix is reachable (and hence also U_2 and U_{10}).*

Proof. By Proposition 12, one can find $Q_1, Q_2, Q_3, Q_4 \in \langle \mathcal{G} \rangle \cap \mathsf{U}_1$ such that $\tau(Q_i), i = 1, \ldots, 4$ generate $\mathbb{R}^3$ as an $\mathbb{R}_{\geq 0}$-cone. In particular, $-\tau(Q_1) \in \langle \tau(Q_1), \tau(Q_2), \tau(Q_3), \tau(Q_4) \rangle_{\mathbb{R}_{\geq 0}}$. So there exist $x_i \in \mathbb{R}_{\geq 0}, i = 1, \ldots, 4$, not all zero, such that $\sum_{i=1}^4 x_i \tau(Q_i) = \mathbf{0}$. Since $\tau(Q_i), i = 1, \ldots, 4$ have integer entries, one can suppose $x_i \in \mathbb{Z}_{\geq 0}$. Hence, $\tau(\prod_{i=1}^4 Q_i^{x_i}) = \sum_{i=1}^4 x_i \tau(Q_i) = \mathbf{0}$, which yields $I = \prod_{i=1}^4 Q_i^{x_i} \in \langle \mathcal{G} \rangle$. ◀

3.3 $\mathcal{C}$ has dimension 1

Next, we consider the case where $\dim \mathcal{C} = 1$. The main idea of this case is that if the product of a string $B_1 \cdots B_m$ is in U_1, then all $D_\sigma, E_\sigma, F_\sigma$ vanish, so $\tau(B(\sigma, t))$ is determined by some linear terms as well as by G_σ. Recall that we write $A_i = UT(\alpha_i, \beta_i, \kappa_i; \delta_i, \epsilon_i, \phi_i)$, $i = 1, \ldots, k$. Similar to notation (4), we define the following quantities for convenience:

$$\Delta_i \coloneqq \delta_i - \frac{1}{2}\alpha_i\beta_i, \quad \mathcal{E}_i \coloneqq \epsilon_i - \frac{1}{2}\beta_i\kappa_i, \quad \Phi_i \coloneqq \alpha_i\beta_i\kappa_i - \frac{1}{2}(\alpha_i\epsilon_i + \delta_i\kappa_i) + \frac{1}{3}\phi_i. \tag{8}$$

Since $\mathcal{C}$ has dimension 1, there exist $\alpha, \beta, \kappa \in \mathbb{Z}$ such that $\varphi_0(A_i) = (\alpha, \beta, \kappa) \cdot \rho_i$ for $\rho_i \in \mathbb{Z}, i = 1, \ldots, k$.

▶ **Proposition 14.** *Suppose $\varphi_0(A_i) = (\alpha, \beta, \kappa) \cdot \rho_i$ for $\rho_i \in \mathbb{Z}, i = 1, \ldots, k$. Let $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_k)$ be the Parikh vector of a string $B_1 \cdots B_m$, with the product $P = B_1 \cdots B_m$. Then*

(i) *$P \in \mathsf{U}_{10}$ if and only if $\sum_{i=1}^k \ell_i \rho_i = 0$ and $\sum_{i=1}^k \ell_i \Delta_i = 0$.*

(ii) *$P \in \mathsf{U}_2$ if and only if $\sum_{i=1}^k \ell_i \rho_i = 0, \sum_{i=1}^k \ell_i \Delta_i = 0$ and $\sum_{i=1}^k \ell_i \mathcal{E}_i = 0$.*

The immediate consequence of Proposition 14 is that U_2-Reachability and U_{10}-Reachability are decidable using linear programming (LP). For example, U_{10}-Reachability has a positive answer if and only if the LP instance $\sum_{i=1}^k \ell_i \rho_i = 0$, $\sum_{i=1}^k \ell_i \Delta_i = 0$, $\ell_i \geq 0, i = 1, \ldots, k$, has a non-zero *integer* solution $(\ell_1, \ldots, \ell_k)$. However, because all the equations and inequalities in the LP instance are *homogeneous*, the LP instance has a non-zero *integer* solution if and only if it has a non-zero *rational* solution. Furthermore, the total bit length of $\rho_i, \Delta_i, \mathcal{E}_i$ is linear with respect to the encoding size of $\mathcal{G}$. Therefore, the existence of a non-zero rational

solution is decidable in polynomial time. In particular, for $i = 1, \ldots, k$, one can decide whether this LP instance has a rational solution $(\ell_1, \ldots, \ell_k)$ with $\ell_i = 1$. Then, the LP instance has a non-zero rational solution if and only if it has a rational solution with $\ell_i = 1$ for some i. The decision procedure for U_2-Reachability is similar.

Next, we consider the Identity Problem. Define the set

$$\Lambda \coloneqq \left\{ (\ell_1, \ldots, \ell_k) \in \mathbb{Z}^k_{\geq 0} \middle| \sum_{i=1}^{k} \ell_i \rho_i = \sum_{i=1}^{k} \ell_i \Delta_i = \sum_{i=1}^{k} \ell_i \mathcal{E}_i = 0 \right\}.$$

By Proposition 14, the product of a string is in U_2 if and only if its Parikh vector is in Λ. It is easy to see that Λ is additively closed, meaning $\Lambda + \Lambda \subseteq \Lambda$. Define the *support* of a Parikh vector $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_k)$ to be $\mathrm{supp}(\boldsymbol{\ell}) = \{i \mid \ell_i \neq 0\}$, and the support of the set Λ to be

$$\mathrm{supp}(\Lambda) \coloneqq \bigcup_{\boldsymbol{\ell} \in \Lambda} \mathrm{supp}(\boldsymbol{\ell}) = \{i \mid \exists (\ell_1, \ldots, \ell_k) \in \Lambda, \ell_i \neq 0\}.$$

For $i = 1, \ldots, k$, we have $i \in \mathrm{supp}(\Lambda)$ if and only if the LP instance $\sum_{i=1}^{k} \ell_i \rho_i = \sum_{i=1}^{k} \ell_i \Delta_i = \sum_{i=1}^{k} \ell_i \mathcal{E}_i = 0$, $\ell_i > 0$ and $\ell_j \geq 0, j \neq i$ has an *integer* solution. Again, by homogeneity, this is decidable in polynomial time by deciding the existence of a *rational* solution. Hence, $\mathrm{supp}(\Lambda)$ is computable in polynomial time by deciding whether $i \in \mathrm{supp}(\Lambda)$ for all $i = 1, \ldots, k$.

If $\mathrm{supp}(\Lambda) \neq \{1, \ldots, k\}$, we can discard the elements $A_i \in \mathcal{G}$ with $i \notin \mathrm{supp}(\Lambda)$, then $\mathrm{card}(\mathcal{G})$ decreases and we are done by the induction hypothesis. Hence, we only need to consider the case where $\mathrm{supp}(\Lambda) = \{1, \ldots, k\}$. The following proposition answers the Identity Problem in this case. Again, the homogeneity yields a polynomial time deciding procedure.

▶ **Proposition 15.** *Suppose* $\varphi_0(A_i) = (\alpha, \beta, \kappa) \cdot \rho_i$ *for* $\rho_i \in \mathbb{Z}, i = 1, \ldots, k$, *and* $\mathrm{supp}(\Lambda) = \{1, \ldots, k\}$. *Define the values* $\Gamma_i = \alpha\epsilon_i - \kappa\delta_i, i = 1, \ldots, k$. *Then*

- **(i)** *When* $\rho_i \Gamma_j = \rho_j \Gamma_i$ *for all* $i, j \in \{1, \ldots, k\}$, *the identity matrix is reachable if and only if the set* $\{(\ell_1, \ldots, \ell_k) \in \Lambda \mid \sum_{i=1}^{k} \ell_i \Phi_i = 0\}$ *is not equal to* $\{\mathbf{0}\}$.
- **(ii)** *When* $\rho_i \Gamma_j \neq \rho_j \Gamma_i$ *for some* $i, j \in \{1, \ldots, k\}$, *the identity matrix is reachable.*

3.4 $\mathcal{C}$ has dimension 0

In this case, $\varphi_0(A_i) = \mathbf{0}$ for all i, so $\mathcal{G} \subset \mathsf{U}_1$. Since $\mathsf{U}_1 \overset{\tau}{\cong} \mathbb{Z}^3$, the Identity Problem and U_2, U_{10}-Reachability are decidable using linear programming. For example, deciding the Identity Problem amounts to deciding whether the LP instance $\sum_{i=1}^{k} \ell_i \cdot \tau(A_i) = \mathbf{0}$, $\ell_i \geq 0, i = 1, \ldots, k$ has a non-zero integer solution. As before, by the homogeneity of the LP instance, this is decidable in polynomial time by considering solutions in $\mathbb{Q}$.

3.5 $\mathcal{C}$ has dimension 2

Suppose now that there exist $p, q, r \in \mathbb{Z}$, not all zero, such that $p\alpha_i + q\beta_i + r\kappa_i = 0, i = 1, \ldots, k$. Consider the following cases on the values of p, q, r.

3.5.1 Case 1: there is at most one zero among p, q, r

The main difficulty of this case is as follows. By Lemma 7, (D_σ, E_σ) is constrained to the one dimensional subspace $\{(d, e) \mid pd - re = 0\} \subset \mathbb{R}^2$. Therefore, in order to decide whether the vectors $\tau(B(\sigma, t))$ can generate the neutral element, one needs to take into account their linear terms, i.e. $(\sum_{i=1}^{m} D_i, \sum_{i=1}^{m} E_i)$ as well. Define the additively closed set:

$$L \coloneqq \left\{ (\ell_1, \ldots, \ell_k) \in \mathbb{Z}^k_{\geq 0} \middle| \sum_{i=1}^{k} \ell_i \varphi_0(A_i) = 0 \right\}.$$

The product P of a string $B_1 \cdots B_m$ is in U_1 if and only if its Parikh vector is in L.

▶ **Lemma 16.** *When $\mathcal{C} = \mathcal{C}^{lin}$, we have* $\operatorname{supp}(L) = \{1, \ldots, k\}$.

We continue to adopt the notations from (8) for $\Delta_i, \mathcal{E}_i$. Consider the subset of L:

$$L_0 \coloneqq \left\{ (\ell_1, \ldots, \ell_k) \in L \middle| p \sum_{i=1}^{k} \ell_i \Delta_i - r \sum_{i=1}^{k} \ell_i \mathcal{E}_i = 0 \right\}.$$

L_0 can be described as the set of Parikh vectors whose corresponding strings have linear terms falling on the line $pd - re = 0$. Again, L_0 is additively closed. The main idea is that the quadratic term of $\varphi_1(B(\sigma, t))$ falls on the line $pd - re = 0$, therefore, if $P \in \mathsf{U}_2, \varphi_1(B(\sigma, t)) = 0$, then its linear term must also fall on the line $pd - re = 0$. This leads to the following lemma.

▶ **Lemma 17.** *Suppose* $\dim \mathcal{C} = 2$. *If the product P of a string $B_1 \cdots B_m$ is in U_2, then its Parikh vector $\boldsymbol{\ell}$ is in L_0.*

The following proposition gives a solution to the U_{10}-Reachability problem.

▶ **Proposition 18.** *Suppose* $\dim \mathcal{C} = 2$ *and at most one of p, q, r is zero.*

(i) *When $r \neq 0$, U_{10} is reachable.*

(ii) *When $r = 0, p \neq 0$, U_{10} is reachable if and only if L_0 is not equal to $\{\mathbf{0}\}$.*

In particular, whether L_0 equals $\{\mathbf{0}\}$ is decidable by linear programming, (again, by homogeneity, one can solve the linear programming instance in $\mathbb{Q}$). Hence, U_{10}-Reachability is decidable. We then treat the Identity Problem and U_2-Reachability. Consider the support of L_0. As before, $\operatorname{supp}(L_0) = \{i \mid \exists (\ell_1, \ldots, \ell_k) \in L_0, \ell_i \neq 0\}$ is computable using linear programming. By Lemma 17, in order to reach U_2 (or the identity matrix), we can only use matrices with index in $\operatorname{supp}(L_0)$. By discarding matrices and using the induction hypothesis on $\operatorname{card}(\mathcal{G})$, we only need to consider the case where $\operatorname{supp}(L_0) = \{1, \ldots, k\}$. The following proposition gives a positive answer to the Identity Problem and U_2-Reachability in this case.

▶ **Proposition 19.** *Suppose* $\dim \mathcal{C} = 2$ *and at most one of p, q, r is zero. If* $\operatorname{supp}(L_0) = \{1, \ldots, k\}$, *then the identity matrix is reachable. (In particular, U_2 is reachable.)*

Sketch of proof. Similarly to Proposition 12, we construct four elements in $\mathsf{U}_1 \cap \langle \mathcal{G} \rangle$ whose images under τ generate the two-dimensional linear subspace $\{(d, e, f) \in \mathbb{R}^3 \mid pd - re = 0\}$ as an $\mathbb{R}_{\geq 0}$-cone (see Figure 2 for an illustration). Consequently, the $\mathbb{Z}_{\geq 0}$-cone that they generate is two-dimensional lattice in $\{(d, e, f) \in \mathbb{Z}^3 \mid pd - re = 0\}$, which contains the neutral element. ◀

3.5.2 Case 2: $p = r = 0$

In this case, $\mathcal{G} \subset \mathrm{H}_5$, so the Identity Problem is decidable by Lemma 3. U_2 and U_{10}-Reachability reduce to the Identity Problem in $\mathbb{Z}^4$ and $\mathbb{Z}^3$, respectively, which are decidable in polynomial time using linear programming. Here, we claim an additional complexity result that strengthens Lemma 3, which is crucial for a polynomial complexity algorithm for $\mathsf{UT}(4, \mathbb{Z})$.

▶ **Proposition 20.** *For a fixed n, the Identity Problem in H_{2n+1} is decidable in polynomial time.*

3.5.3 Case 3: $p = q = 0$, $r \neq 0$ or $r = q = 0$, $p \neq 0$

The main technique in this case is a reduction from the Identity Problem to U_2-Reachability, from U_2-Reachability to U_{10}-Reachability, and from U_{10}-Reachability to linear programming or to the Identity Problem in H_3. If $p = q = 0$, $r \neq 0$, then $\kappa_i = 0, i = 1, \ldots, k$. If $r = q = 0$, $p \neq 0$, then $\alpha_i = 0, i = 1, \ldots, k$. Define the following matrices in H_3:

$$H_i := \begin{pmatrix} 1 & \alpha_i & \delta_i \\ 0 & 1 & \beta_i \\ 0 & 0 & 1 \end{pmatrix}, \quad i = 1, \ldots, k.$$

The following proposition along with Proposition 20 provides a solution to U_{10}-Reachability.

▶ **Proposition 21.**

(i) *When $\kappa_i = 0, i = 1, \ldots, k$, U_{10}-Reachability for $A_1, \ldots, A_k$ is equivalent to the Identity Problem for $H_1, \ldots, H_k$.*

(ii) *When $\alpha_i = 0, i = 1, \ldots, k$, U_{10} is reachable for $A_1, \ldots, A_k$ if and only if $\sum_{i=1}^{k} \ell_i \delta_i = \sum_{i=1}^{k} \ell_i \beta_i = \sum_{i=1}^{k} \ell_i \kappa_i = 0$ has a non-zero integer solution $(\ell_1, \ldots, \ell_k) \in \mathbb{Z}^k_{\geq 0}$.*

Next, consider the Identity Problem and U_2-Reachability. By symmetry, we can suppose $p = q = 0$, $r \neq 0$, so $\kappa_i = 0, i = 1, \ldots, k$. Define

$$A'_i := UT(\beta_i, \alpha_i, \epsilon_i; \delta_i, \phi_i, 0), i = 1, \ldots, k, \tag{9}$$

the following proposition reduces the Identity Problem and U_2-Reachability for $A_1, \ldots, A_k$ to reachability problems for $A'_1, \ldots, A'_k$:

▶ **Proposition 22.** *Suppose $\kappa_i = 0, i = 1, \ldots, k$.*

(i) *The Identity Problem for $A_1, \ldots, A_k$ is equivalent to U_2-Reachability for $A'_1, \ldots, A'_k$.*

(ii) *U_2-Reachability for $A_1, \ldots, A_k$ is equivalent to U_{10}-Reachability for $A'_1, \ldots, A'_k$.*

Together with the previous Subsections 3.2 - 3.5.2, we have completely reduced the Identity Problem for $\mathcal{G}$ to either the problem for a set of smaller cardinality, or to U_2-reachability of another set. We have also reduced U_2-reachability for $\mathcal{G}$ to either a problem for a set of smaller cardinality, or to U_{10}-reachability of another set. By Proposition 21 and the previous subsections, U_{10}-reachability is decidable. Hence, we have now exhausted all the possible cases for the dimension of $\mathcal{C}$, and we conclude that the Identity Problem, U_2-Reachability and U_{10}-Reachability in $\mathsf{UT}(4, \mathbb{Z})$ are decidable.

4 Complexity analysis and concluding remarks

In this paper, we have shown that the Identity Problem for $\mathsf{UT}(4, \mathbb{Z})$ is decidable. A brief analysis of our algorithm shows that it terminates in polynomial time. In fact, we can first show that the algorithm for U_{10}-Reachability terminates in polynomial time. Starting with $k = \mathrm{card}(\mathcal{G})$ matrices, we need to solve at most $O(k)$ linear equations, $O(k)$ homogeneous linear programming instances and one Identity Problem in H_3 before either $\mathrm{card}(\mathcal{G})$ decreases or a conclusion on U_{10}-Reachability is reached. All these problems have $O(k)$ inputs which are of polynomial size with respect to the coefficients of the matrices in $\mathcal{G}$, and are known to have polynomial complexity. Furthermore, the number $\mathrm{card}(\mathcal{G})$ decreases at most k times. Hence, the total complexity of our algorithm for U_{10}-reachability is polynomial with respect to the input $\mathcal{G}$. Then, using the same method, we can show that the algorithm for U_2-Reachability terminates in polynomial time: since after polynomial time, either $\mathrm{card}(\mathcal{G})$

decreases, or the problem is reduced to U_{10}-Reachability, or a conclusion on U_2-Reachability is reached. At last, we can show that the algorithm for the Identity Problem terminates in polynomial time: after polynomial time, either $\mathrm{card}(\mathcal{G})$ decreases, or the problem is reduced to U_2-Reachability or the Identity Problem in H_5, or a conclusion on the Identity Problem is reached. (In particular, the polynomial complexity of the Identity Problem in H_5 is a new result of our paper, see Proposition 20.)

It is likely that our method can be adapted to study the Identity Problem for other metabelian matrix groups, for instance the direct product H_3^n. There is also evidence that the arguments in this paper can be strengthened to tackle the Identity Problem for $\mathsf{UT}(n, \mathbb{Z})$ with $n \geq 5$, even though $\mathsf{UT}(5, \mathbb{Z})$ ceases to be metabelian. In fact, one can push the convex geometry arguments down the derived series of $\mathsf{UT}(n, \mathbb{Z})$, even when the series has length greater than two. Another natural follow-up question is the Membership Problem for $\mathsf{UT}(4, \mathbb{Z})$. An interesting idea would be to adapt the Register Automata method introduced in [5] for passing from the Identity Problem to the Membership Problem.

References

1 László Babai, Robert Beals, Jin-yi Cai, Gábor Ivanyos, and Eugene M. Luks. Multiplicative equations over commuting matrices. In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 498–507, 1996.

2 Paul C. Bell and Igor Potapov. On the undecidability of the identity correspondence problem and its applications for word and matrix semigroups. *International Journal of Foundations of Computer Science*, 21(06):963–978, 2010.

3 Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

4 Christian Choffrut and Juhani Karhumäki. Some decision problems on integer matrices. *RAIRO-Theoretical Informatics and Applications-Informatique Théorique et Applications*, 39(1):125–131, 2005.

5 Thomas Colcombet, Joël Ouaknine, Pavel Semukhin, and James Worrell. On reachability problems for low-dimensional matrix semigroups. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, volume 132 of *LIPIcs*, pages 44:1–44:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.ICALP.2019.44.

6 Sang-Ki Ko, Reino Niskanen, and Igor Potapov. On the identity problem for the special linear group and the heisenberg group. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 132:1–132:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ICALP.2018.132.

7 Daniel König, Markus Lohrey, and Georg Zetzsche. Knapsack and subset sum problems in nilpotent, polycyclic, and co-context-free groups. *Algebra and Computer Science*, 677:138–153, 2016.

8 V. M. Kopytov. Solvability of the problem of occurrence in finitely generated soluble groups of matrices over the field of algebraic numbers. *Algebra and Logic*, 7(6):388–393, 1968.

9 Engel Lefaucheux. Private Communication, 2022.

10 A. Markov. On certain insoluble problems concerning matrices. *Doklady Akad. Nauk SSSR*, 57(6):539–542, 1947.

11 K. A. Mikhailova. The occurrence problem for direct products of groups. *Matematicheskii Sbornik*, 112(2):241–251, 1966.

12 Michael S. Paterson. Unsolvability in 3 × 3 matrices. *Studies in Applied Mathematics*, 49(1):105–107, 1970.

13 Igor Potapov and Pavel Semukhin. Decidability of the membership problem for 2 × 2 integer matrices. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 170–186. SIAM, 2017.

14 Igor Potapov and Pavel Semukhin. Membership problem in GL(2, Z) extended by singular matrices. In Kim G. Larsen, Hans L. Bodlaender, and Jean-François Raskin, editors, *42nd International Symposium on Mathematical Foundations of Computer Science, MFCS 2017, August 21-25, 2017 – Aalborg, Denmark*, volume 83 of *LIPIcs*, pages 44:1–44:13. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.MFCS.2017.44.

15 Joseph J. Rotman. *An introduction to the theory of groups*, volume 148. Springer Science & Business Media, 2012.

16 Alexander Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, 1998.