On the Identity Problem for the Special Linear Group and the Heisenberg Group

Sang-Ki Ko
Korea Electronics Technology Institute, South Korea
sangkiko@keti.re.kr

Reino Niskanen
Department of Computer Science, University of Liverpool, UK
r.niskanen@liverpool.ac.uk

Igor Potapov
Department of Computer Science, University of Liverpool, UK
potapov@liverpool.ac.uk

Abstract

We study the identity problem for matrices, i.e., whether the identity matrix is in a semigroup generated by a given set of generators. In particular we consider the identity problem for the *special linear group* following recent NP-completeness result for $\mathrm{SL}(2,\mathbb{Z})$ and the undecidability for $\mathrm{SL}(4,\mathbb{Z})$ generated by 48 matrices. First we show that there is no embedding from pairs of words into 3×3 integer matrices with determinant one, i.e., into $\mathrm{SL}(3,\mathbb{Z})$ extending previously known result that there is no embedding into $\mathbb{C}^{2\times 2}$. Apart from theoretical importance of the result it can be seen as a strong evidence that the computational problems in $\mathrm{SL}(3,\mathbb{Z})$ are decidable. The result excludes the most natural possibility of encoding the Post correspondence problem into $\mathrm{SL}(3,\mathbb{Z})$, where the matrix products extended by the right multiplication correspond to the Turing machine simulation. Then we show that the *identity problem* is decidable in polynomial time for an important subgroup of $\mathrm{SL}(3,\mathbb{Z})$, the Heisenberg group $\mathrm{H}(3,\mathbb{Z})$. Furthermore, we extend the decidability result for $\mathrm{H}(n,\mathbb{Q})$ in any dimension n. Finally we are tightening the gap on decidability question for this long standing open problem by improving the undecidability result for the identity problem in $\mathrm{SL}(4,\mathbb{Z})$ substantially reducing the bound on the size of the generator set from 48 to 8 by developing a novel reduction technique.

2012 ACM Subject Classification Theory of computation → Models of computation,Computing methodologies → Symbolic and algebraic algorithms,Theory of computation → Program verification

Keywords and phrases matrix semigroup, identity problem, special linear group, Heisenberg group, decidability

Digital Object Identifier 10.4230/LIPIcs.ICALP.2018.132

Related Version [27], `https://arxiv.org/abs/1706.04166`

Funding This work was supported by EPSRC grant "Reachability problems for words, matrices and maps" (EP/M00077X/1).

1 Introduction

The dynamics of many systems can be represented by matrices and matrix products. The analysis of such systems lead to solving reachability questions in matrix semigroups which is essential part in verification procedures, control theory questions, biological systems'

45th International Colloquium on Automata, Languages, and Programming (ICALP 2018).
Editors: Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella;
Article No. 132; pp. 132:1–132:15

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

predictability, security etc. [9, 10, 16, 17, 20, 21, 28, 33, 35, 36, 37]. Many nontrivial algorithms for decision problems on matrix semigroups have been developed for matrices under different constraints on the dimension, the size of a generating set or for specific subclasses of matrices: e.g., commutative matrices [2], row-monomial matrices [30] or 2×2 matrix semigroups generated by non-singular integer matrices [41], upper-triangular integer matrices [25], matrices from the special linear group [4, 15], etc.

Despite visible interest in this research domain, we still see a significant lack of algorithms and complexity results for answering decision problems in matrix semigroups. Many computational problems for matrix (semi)groups are computationally hard starting from dimension two and very often become undecidable from dimensions three or four even in the case of integer matrices. The central decision problem in matrix semigroups is the membership problem, which was originally considered by A. Markov in 1947 [32]. Let $S = \langle G \rangle$ be a matrix semigroup finitely generated by a generating set of square matrices G. The *membership problem* is to decide whether or not a given matrix M belongs to the matrix semigroup S. By restricting M to be the identity matrix we call the problem the *identity problem*.

▶ **Problem 1** (Identity problem). *Let $S = \langle G \rangle$, where G is a finite set of n-dimensional matrices over $\mathbb{K} = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \ldots$. Is the identity matrix in the semigroup, i.e., does $\boldsymbol{I} \in S$ hold?*

The identity problem is computationally equivalent to another fundamental problem – the *subgroup problem* (i.e., to decide whether a semigroup contains a subgroup) as any subset of matrices, which can form a product leading to the identity also generate a group [15][1].

The decidability status of the identity problem was unknown for a long time for matrix semigroups of any dimension, see Problem 10.3 in "Unsolved Problems in Mathematical Systems and Control Theory" [10], but it was shown in [6] to be undecidable for 48 matrices from $\mathbb{Z}^{4\times4}$ by proving that the identity correspondence problem (a variant of the Post correspondence problem over a group alphabet) is undecidable, and embedding pairs of words over free group alphabet into $\mathrm{SL}(4,\mathbb{Z})$ as two blocks on the main diagonal and by a morphism f as follows $f(a) = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$, $f(a^{-1}) = \left(\begin{smallmatrix} 1 & -2 \\ 0 & 1 \end{smallmatrix}\right)$, $f(b) = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ and $f(b^{-1}) = \left(\begin{smallmatrix} 1 & 0 \\ -2 & 1 \end{smallmatrix}\right)$. In the seminal paper of Paterson in 1970, see [39], an injective morphism from pairs of words in alphabet $\Sigma = \{a, b\}$ into 3×3 integral matrices, $g(u, v) = \begin{pmatrix} n^{|u|} & 0 & 0 \\ 0 & n^{|v|} & 0 \\ \sigma(u) & \sigma(v) & 1 \end{pmatrix}$ (where σ represents each word as an n-adic number), was used to prove undecidability of the mortality problem (i.e., the membership problem of the zero matrix) and which later led to many undecidability results of matrix problems in dimension three, e.g., [12, 24]. Finding new injective morphisms is hard, but having them gives an opportunity to prove new undecidability results.

In 1999, Cassaigne, Harju and Karhumäki significantly boosted the research on finding algorithmic solutions for 2×2 matrix semigroups by showing that there is no injective semigroup morphism from pairs of words over any finite alphabet (with at least two elements) into complex 2×2 matrices [12]. This result led to substantial interest in finding algorithmic solutions for such problems as the identity problem, mortality, membership, vector reachability, freeness etc. for 2×2 matrices.

For example, in 2007 Gurevich and Schupp [23] showed that the membership problem is decidable in polynomial time for the finitely generated subgroups of the modular group and later in 2017 Bell, Hirvensalo and Potapov proved that the identity problem for a semigroup

[1] The product of matrices which is equal to the identity is still the identity element after a cyclic shift, so every element from this product has the inverse.

generated by matrices from $\mathrm{SL}(2,\mathbb{Z})$ is NP-complete by developing a new effective technique to operate with compressed word representations of matrices and closing the gap on complexity improving the original EXPSPACE solution proposed in 2005 [15]. The first algorithm for the membership problem which covers the cases beyond $\mathrm{SL}(2,\mathbb{Z})$ and $\mathrm{GL}(2,\mathbb{Z})$ has been proposed in [41] and provides the solution for a semigroup generated by non-singular 2×2 integer matrices. Later, these techniques have been applied to build another algorithm to solve the membership problem in $\mathrm{GL}(2,\mathbb{Z})$ extended by singular matrices [42]. The current limit of decidability is standing for 2×2 matrices which are defined over hypercomplex numbers (quaternions) for which most of the problems have been shown to be undecidable in [5] and correspond to reachability problems for 3-sphere rotation.

In our paper, we show that there is no embedding from pairs of words into 3×3 integer matrices with determinant one (i.e., into $\mathrm{SL}(3,\mathbb{Z})$), which is a strong evidence that computational problems in $\mathrm{SL}(3,\mathbb{Z})$ are decidable as all known undecidability techniques for low-dimensional matrices are based on encoding of Turing machine computations via the Post correspondence problem (PCP) which cannot be applied in $\mathrm{SL}(3,\mathbb{Z})$ following our result. In case of the PCP encoding the matrix products extended by the right multiplication correspond to the Turing machine simulation and the only known proof alternatives are recursively enumerable sets and Hilbert's tenth problem that provide undecidability for matrix equations, but of very high dimensions [3, 13, 26].

So in analogy to 1999 result from [12] on non-existence of embedding into 2×2 matrix semigroups over complex numbers, we expand a horizon of decidability area for matrix semigroups and show that there is no embedding from a set of pairs of words over a semigroup alphabet to any matrix semigroup in $\mathrm{SL}(3,\mathbb{Z})$. It follows almost immediately that there is no embedding from a set of pairs of group words into $\mathbb{Z}^{3\times 3}$.[2] The matrix semigroup in $\mathrm{SL}(3,\mathbb{Z})$ has attracted a lot of attention recently as it can be represented by a set of generators and relations [18, 19] similar to $\mathrm{SL}(2,\mathbb{Z})$ where it was possible to convert numerical problems into symbolic problems and solve them with novel computational techniques; see [4, 15, 41, 42]. Comparing to the relatively simple representation of $\mathrm{SL}(2,\mathbb{Z}) = \langle S,T \mid S^4 = \boldsymbol{I}_2, (ST)^6 = \boldsymbol{I}_2\rangle$, where $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ the case of $\mathrm{SL}(3,\mathbb{Z}) = \langle X,Y,Z \mid X^3 = Y^3 = Z^2 = (XZ)^3 = (YZ)^3 = (X^{-1}ZXY)^2 = (Y^{-1}ZYX)^2 = (XY)^6 = \boldsymbol{I}_3\rangle$ looks more challenging containing both non-commutative and partially commutative elements.

As the decidability status of the *identity problem* in dimension three is still a long standing open problem, we look for an important subgroup of $\mathrm{SL}(3,\mathbb{Z})$, the Heisenberg group $\mathrm{H}(3,\mathbb{Z})$, for which the *identity problem* could be decidable following our result on non-existence of embedding. The Heisenberg group is an important subgroup of $\mathrm{SL}(3,\mathbb{Z})$ which is useful in the description of one-dimensional quantum mechanical systems [11, 22, 29]. We show that the *identity problem* for a matrix semigroup generated by matrices from $\mathrm{H}(3,\mathbb{Z})$ and even $\mathrm{H}(3,\mathbb{Q})$ is decidable in polynomial time. Furthermore, we extend the decidability result for $\mathrm{H}(n,\mathbb{Q})$ in any dimension n. Moreover we tighten the gap between (un)decidability results for the *identity problem* substantially reducing the bound on the size of the generator set from 48 (see [6]) to 8 in $\mathrm{SL}(4,\mathbb{Z})$ by developing a novel reduction technique.

[2] The idea that such result may hold was motivated by analogy from combinatorial topology, where the identity problem is decidable for the braid group B_3 which is the universal central extension of the modular group $\mathrm{PSL}(2,\mathbb{Z})$ [40], an embedding for a set of pairs of words into the braid group B_5 exists, see [7], and non-existence of embeddings were proved for B_4 in [1]. So $\mathrm{SL}(3,\mathbb{Z})$ was somewhere in the goldilocks zone between B_3 and B_5.

2 Preliminaries

We say that a semigroup S is *generated* by a subset G of S if each element of S can be expressed as a composition of elements of G. In this case, we call G the *generating set* of S. Given an *alphabet* $\Sigma = \{a_1, \ldots, a_m\}$, a finite *word* u is an element of semigroup Σ^*. The *empty word* is denoted by ε. Let $\Gamma = \{a_1, \ldots, a_\ell, a_1^{-1}, \ldots, a_\ell^{-1}\}$ be a generating set of a free group $\mathrm{FG}(\Gamma)$. The elements of $\mathrm{FG}(\Gamma)$ are all *reduced* words over Γ, i.e., words not containing $a_i a_i^{-1}$ or $a_i^{-1} a_i$ as a subword. In this context, we call Γ a finite *group alphabet*, i.e., an alphabet with an involution. The multiplication of two elements (reduced words) $u, v \in \mathrm{FG}(\Gamma)$ corresponds to the unique reduced word of the concatenation uv. This multiplication is called *concatenation* throughout the paper. Later in the encoding of words over a group alphabet we denote a^{-1} by $\overline{a}$ and the alphabet of inverse letters is denoted as $\Sigma^{-1} = \{a^{-1} \mid a \in \Sigma\}$.

In the next lemma, we present an encoding from an arbitrary group alphabet to a binary group alphabet used in Section 5. The result is crucial as it allows us to present the results of the above section over the smallest domain.

▶ **Lemma 2** (Birget, Margolis [8]). *Let $\Gamma = \{z_1, \ldots, z_\ell, \overline{z_1}, \ldots, \overline{z_\ell}\}$ be a group alphabet and $\Gamma_2 = \{c, d, \bar{c}, \bar{d}\}$ be a binary group alphabet. Define the mapping $\alpha : \Gamma \to \mathrm{FG}(\Gamma_2)$ by $\alpha(z_i) = c^i d \bar{c}^i$, and $\alpha(\overline{z_i}) = c^i \bar{d} \bar{c}^i$, where $1 \leq i \leq \ell$. Then α is a monomorphism, that is, an injective morphism. Note that α can be extended to domain $\mathrm{FG}(\Gamma)$ in the usual way.*

The special linear group is $\mathrm{SL}(n, \mathbb{K}) = \{M \in \mathbb{K}^{n \times n} \mid \det(M) = 1\}$, where $\mathbb{K} = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \ldots$. The *identity matrix* is denoted by $\boldsymbol{I}_n$ and the *zero matrix* is denoted by $\mathbf{0}_n$. The *Heisenberg group* $\mathrm{H}(3, \mathbb{K})$ is formed by the 3×3 matrices of the form $M = \left(\begin{smallmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{smallmatrix}\right)$, where $a, b, c \in \mathbb{K}$. It is easy to see that the Heisenberg group is a non-commutative subgroup of $\mathrm{SL}(3, \mathbb{K})$. We can consider the Heisenberg group as a set of all triples with the following group law: $(a_1, b_1, c_1) \otimes (a_2, b_2, c_2) = (a_1 + a_2, b_1 + b_2, c_1 + c_2 + a_1 b_2)$. By $\psi(M)$ we denote the triple $(a, b, c) \in \mathbb{K}^3$ which corresponds to the upper-triangular coordinates of M. Let M be a matrix in $\mathrm{H}(3, \mathbb{K})$ such that $\psi(M) = (a, b, c)$. We define the *superdiagonal vector* of M to be $\vec{v}(M) = (a, b)$. Given two vectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$, the *cross product* of $\mathbf{u}$ and $\mathbf{v}$ is defined as $\mathbf{u} \times \mathbf{v} = u_1 v_2 - u_2 v_1$. Two vectors are *parallel* if the cross product is zero.

The Heisenberg group can be also defined in higher dimensions. The Heisenberg group of dimension n over $\mathbb{K}$ is denoted by $\mathrm{H}(n, \mathbb{K})$ and is the group of square matrices in $\mathbb{K}^{n \times n}$ of the form $\left(\begin{smallmatrix} 1 & \mathbf{a}^\mathsf{T} & c \\ 0 & \boldsymbol{I}_{n-2} & \mathbf{b} \\ 0 & 0 & 1 \end{smallmatrix}\right)$, where $\mathbf{a}, \mathbf{b} \in \mathbb{K}^{n-2}, c \in \mathbb{K}$. Similar to the Heisenberg group in dimension three, we can also consider the Heisenberg group in dimension n for any integer $n \geq 3$ as a set of all triples with the following group law: $(\mathbf{a}_1, \mathbf{b}_1, c_1) \otimes (\mathbf{a}_2, \mathbf{b}_2, c_2) = (\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}_1 + \mathbf{b}_2, c_1 + c_2 + \mathbf{a}_1 \cdot \mathbf{b}_2)$, where $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}_1, \mathbf{b}_2 \in \mathbb{K}^{n-2}$ and $\mathbf{a}_1 \cdot \mathbf{b}_2$ is the dot product of vectors $\mathbf{a}_1$ and $\mathbf{b}_2$.

We extend the function ψ to n-dimensional Heisenberg group: For a matrix M, $\psi(M)$ is the triple $(\mathbf{a}, \mathbf{b}, c) \in (\mathbb{K}^{n-2})^2 \times \mathbb{K}$ which corresponds to the upper-triangular coordinates of M. The product $M_1 M_2$ has $c_1 + c_2 + \mathbf{a}_1 \cdot \mathbf{b}_2$ in the upper-right corner whereas $M_2 M_1$ has $c_1 + c_2 + \mathbf{a}_2 \cdot \mathbf{b}_1$. The other coordinates are identical as we add numbers in the same coordinate. Clearly, the two products are equivalent if and only if $\mathbf{a}_1 \cdot \mathbf{b}_2 = \mathbf{a}_2 \cdot \mathbf{b}_1$ holds.

▶ **Lemma 3.** *Let M_1 and M_2 be two matrices from the Heisenberg group $\mathrm{H}(n, \mathbb{K})$ and $\psi(M_i) = (\mathbf{a}_i, \mathbf{b}_i, c_i)$ for $i = 1, 2$. Then $M_1 M_2 = M_2 M_1$ holds if and only if $\mathbf{a}_1 \cdot \mathbf{b}_2 = \mathbf{a}_2 \cdot \mathbf{b}_1$.*[3]

[3] Note that, in dimension three, the condition can be stated as superdiagonal vectors of M_1 and M_2 being parallel.

3 On embedding from pairs of words into $\mathrm{SL}(3,\mathbb{K})$

In this section, we show that there is no embedding from a set of pairs of words over a semigroup alphabet to the special linear group $\mathrm{SL}(3,\mathbb{Z})$, which can be seen as a strong evidence about decidability of computational problems for this class. If the injective morphism would exist, then we could encode Turing machine computations via the PCP which would provide undecidability proofs for various membership problems in $\mathrm{SL}(3,\mathbb{Z})$.

Let $\Sigma = \{0,1\}$. The monoid $\Sigma^* \times \Sigma^*$ has a generating set $S = \{(0,\varepsilon),(1,\varepsilon),(\varepsilon,0),(\varepsilon,1)\}$, where ε is the empty word. We simplify the notation by setting $a = (0,\varepsilon)$, $b = (1,\varepsilon)$, $c = (\varepsilon,0)$ and $d = (\varepsilon,1)$. It is easy to see that we have the following relations:

$$ac = ca, \qquad bc = cb, \qquad ad = da, \qquad bd = db. \tag{1}$$

In other words, a and b commute with c and d. Furthermore, these are the only relations. That is, a and b do not commute with each other, and neither do c and d. The monoid $\Sigma^* \times \Sigma^*$ is a partially commutative monoid or a trace monoid. A necessary and sufficient conditions for existence of an embedding of trace monoids into $\mathbb{N}^{2\times 2}$ was given in [14] but, to the authors' best knowledge, there are no similar results even for $\mathbb{N}^{3\times 3}$. Let $\varphi : \Sigma^* \times \Sigma^* \to \mathrm{SL}(3,\mathbb{K})$ be an injective morphism and denote $A = \varphi(a)$, $B = \varphi(b)$, $C = \varphi(c)$ and $D = \varphi(d)$. Our goal is to show that φ does not exist for $\mathbb{K} = \mathbb{Z}$. Additionally, we provide an embedding for $\mathbb{K} = \mathbb{Q}$. Unfortunately, the technique developed in [12], where the contradiction was derived from simple relations, resulting from matrix multiplication, cannot be used for a case of $\mathrm{SL}(3,\mathbb{Z})$ as it creates a large number of equations which do not directly limit the existence of φ. We found new techniques to show non-existence of φ by analysis of eigenvalues and the Jordan normal forms.

In the next theorem, we show that if we embed the generators of $\Sigma^* \times \Sigma^*$ into $\mathrm{SL}(3,\mathbb{Z})$, then, for each Jordan normal form, the matrices should satisfy additional equations.

▶ **Theorem 4.** *There is no injective morphism $\varphi : \Sigma^* \times \Sigma^* \to \mathrm{SL}(3,\mathbb{Z})$ for any $|\Sigma| \geq 2$.*

Proof (Sketch). Due to the obvious symmetries, it is enough to show that the claim holds for A. We conjugate the generators to transform A into Jordan normal form. Note that Jordan normal form J of an integer matrix A does not have to be integer or even real, but the contradictions we derive apply also to the original matrices. Also note that matrices J and A have integer trace and determinants are one. There are six possible Jordan normal forms for 3×3 matrices: $\left(\begin{smallmatrix}\lambda&0&0\\0&\mu&0\\0&0&\nu\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}\lambda&0&0\\0&\mu&0\\0&0&\mu\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}\lambda&0&0\\0&\mu&1\\0&0&\mu\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}\lambda&0&0\\0&\lambda&0\\0&0&\lambda\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}\lambda&1&0\\0&\lambda&0\\0&0&\lambda\end{smallmatrix}\right)$, or $\left(\begin{smallmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{smallmatrix}\right)$. The first and the fourth normal forms can be ruled out as the matrices commute with diagonal matrices, and then we prove that C and D commute with each other. In the second form, it follows from the fact that A has an integer trace and the determinant is one, that the eigenvalues are $\lambda = 2$ and $\mu = \frac{1}{2}$. Then, we can rule this form out as also the trace of A^2 should be an integer which does not hold for these eigenvalues. The third form is ruled out as from the relations $AC = CA$ and $AD = DA$, we can solve C and D, and see that necessary also $CD = DC$ holds. The final form is similar to the third Jordan normal form and is ruled out in a similar manner. The fifth form requires additional considerations. We solve most of the elements of $C = \left(\begin{smallmatrix}a&b&c\\d&e&f\\g&h&\ell\end{smallmatrix}\right)$ and $D = \left(\begin{smallmatrix}a'&b'&c'\\d'&e'&f'\\g'&h'&\ell'\end{smallmatrix}\right)$ from equations $AC = CA$ and $AD = DA$. To solve the remaining elements, we need to do further case analysis and we prove that matrices C and D do not commute if and only if $ch' \neq c'h$. We further solve B from $BC = CB$ and $BD = DB$, and show that necessarily $CD = DC$, which is not a valid relation. This is a brief sketch of high-level steps of showing that the injective morphism does not exist; see the complete proof in [27]. ◀

Note that some of the Jordan normal forms of the previous theorem can be ruled out even without assuming that the original matrices were in $\mathrm{SL}(3,\mathbb{Z})$. Using these additional constraints on matrices, we are able to find an embedding into $\mathrm{SL}(3,\mathbb{Q})$.

▶ **Theorem 5.** *Let* $\Sigma = \{0,1\}$. *The morphism* $\varphi : \Sigma^* \times \Sigma^* \to \mathrm{SL}(3,\mathbb{Q})$, *defined by* $\varphi((0,\varepsilon)) = \begin{pmatrix} 4 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$, $\varphi((1,\varepsilon)) = \begin{pmatrix} 9 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$, $\varphi((\varepsilon,0)) = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 4 \end{pmatrix}$ *and* $\varphi((\varepsilon,1)) = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 9 \end{pmatrix}$ *is an embedding.*

We can further extend the non-existence result to words over group alphabets. The only known results for undecidability of the identity problem rely on embedding of pairs of group words into matrices (Theorem 13 and [6]) suggesting that the problem might be decidable in dimension three over integers.

▶ **Corollary 6.** *There is no injective morphism* $\varphi : \mathrm{FG}(\Gamma) \times \mathrm{FG}(\Gamma) \to \mathbb{Z}^{3\times 3}$ *for any binary group alphabet* Γ.

Proof. We proceed by contradiction. Assume that there exists such an injective morphism φ from the set of pairs of words over a group alphabet to the set of matrices in $\mathbb{Z}^{3\times 3}$. Suppose that $A = \varphi((a,\varepsilon))$, where $a \in \Gamma$. Then the inverse matrix A^{-1} corresponding to $(\bar{a},\varepsilon)$ must be in $\mathbb{Z}^{3\times 3}$. This implies that the determinant of A is ± 1 because otherwise the determinant of A^{-1} becomes a non-integer. Consider then a morphism ψ such that $\psi(x) = \varphi(x)\varphi(x)$ for each $x \in \mathrm{FG}(\Gamma) \times \mathrm{FG}(\Gamma)$. It is clear that also ψ is injective and that the determinant of the image is 1. By Theorem 4, such injective morphism ψ does not exist even from semigroup alphabets and hence neither does φ. ◀

4 Decidability of the identity problem in the Heisenberg group

In this section, we prove that the identity problem is decidable for the Heisenberg group which is an important subgroup of the special linear group. First, we provide more intuitive solution for dimension three, i.e., $\mathrm{H}(3,\mathbb{Q})$, which still requires a number of techniques to estimate possible values of elements under permutations in matrix products. In the end of the section, we generalize the result for $\mathrm{H}(n,\mathbb{Q})$ using analogies in the solution for $\mathrm{H}(3,\mathbb{Q})$.

We prove that the identity problem for the Heisenberg group over rationals is decidable by analysing the behaviour of multiplications especially in the upper-right coordinate of matrices. From Lemma 3, it follows that the matrix multiplication is commutative in the Heisenberg group if and only if matrices have pairwise parallel superdiagonal vectors. So we analyse two cases of products for matrices with pairwise parallel and none pairwise parallel superdiagonal vectors and then provide algorithms that solve the problem in polynomial time. The most difficult part is showing that only limited number of conditions must be checked to guarantee the existence of a product that results in the identity.

▶ **Lemma 7.** *Let* $G = \{M_1,\ldots,M_r\} \subseteq \mathrm{H}(3,\mathbb{Q})$ *be a set of matrices from the Heisenberg group such that superdiagonal vectors of matrices are pairwise parallel. If there exists a sequence of matrices* $M = M_{i_1}\cdots M_{i_k}$, *where* $i_j \in [1,r]$ *for all* $1 \le j \le k$, *such that* $\psi(M) = (0,0,c)$ *for some* $c \in \mathbb{Q}$, *then,* $c = \sum_{j=1}^{k}(c_{i_j} - \frac{q}{2}a_{i_j}^2)$ *for some* $q \in \mathbb{Q}$ *dependent only on* G.

Proof. Consider the sequence $M_{i_1}\cdots M_{i_k}$ and let $M_i = \begin{pmatrix} 1 & a_i & c_i \\ 0 & 1 & b_i \\ 0 & 0 & 1 \end{pmatrix}$ for each $i \in [1,r]$. Since the superdiagonal vectors are parallel, i.e., $a_i b_j = b_i a_j$ for any $i,j \in [1,r]$, we have $q = \frac{b_i}{a_i} \in \mathbb{Q}$

and thus $a_i q = b_i$ for all $i \in [1, r]$. Let us consider the product of the matrices. Then the value c is equal to

$$c = \sum_{j=1}^{k} c_{i_j} + \sum_{\ell=1}^{k-1} \left(\sum_{j=1}^{\ell} a_{i_j} \right) a_{i_{\ell+1}} q = \sum_{j=1}^{k} c_{i_j} + \frac{q}{2} \left(\sum_{\ell=1}^{k} \sum_{j=1}^{k} a_{i_\ell} a_{i_j} - \sum_{j=1}^{k} a_{i_j}^2 \right) = \sum_{j=1}^{k} (c_{i_j} - \frac{q}{2} a_{i_j}^2).$$

Note that if $a_{i_j} = 0$ for some $i_j \in [1, r]$, then due to superdiagonal vectors being parallel, $a_{i_j} = 0$ for all i_j and the value c is equal to $\sum_{j=1}^{k} c_{i_j}$. ◀

Note that the previous lemma also holds for $\mathrm{H}(3, \mathbb{R})$. From the previous lemma we further see that the value c is preserved if the matrices are reordered due to their commutativity.

It is worth mentioning that the identity problem in the Heisenberg group is decidable if any two matrices have pairwise parallel superdiagonal vectors since now the problem reduces to solving a system of two linear homogeneous Diophantine equations. Hence, it remains to consider the case when there exist two matrices with non-parallel superdiagonal vectors in the sequence generating the identity matrix. In the following, we prove that the identity matrix is always constructible if we can construct any matrix with the zero superdiagonal vector by using matrices with non-parallel superdiagonal vectors.

▶ **Lemma 8.** *Let $S = \langle M_1, \ldots, M_r \rangle \subseteq \mathrm{H}(3, \mathbb{Q})$ be a finitely generated matrix semigroup. Then the identity matrix exists in S if there exists a sequence of matrices $M_{i_1} \cdots M_{i_k}$, where $i_j \in [1, r]$ for all $1 \leq j \leq k$, satisfying the following properties:*

(i) *$\psi(M_{i_1} \cdots M_{i_k}) = (0, 0, c)$ for some $c \in \mathbb{Q}$, and*

(ii) *$\vec{v}(M_{i_{j_1}})$ and $\vec{v}(M_{i_{j_2}})$ are not parallel for some $j_1, j_2 \in [1, k]$.*

To prove Lemma 8, we show that from a matrix $M = M_{i_1} \cdots M_{i_k}$, such that $\psi(M) = (0, 0, c)$, satisfying the conditions of the lemma, we can construct a matrix M' such that $\psi(M') = (0, 0, c')$ and $cc' < 0$. Given that M_i is the ith generator and $\psi(M_i) = (a_i, b_i, c_i)$, we have $\sum_{j=1}^{k} a_{i_j} = 0$ and $\sum_{j=1}^{k} b_{i_j} = 0$. Without loss of generality, $c > 0$, and the following also holds:

$$c = \sum_{\ell=1}^{k-1} \sum_{j=1}^{\ell} a_{i_j} b_{i_{\ell+1}} + \sum_{j=1}^{k} c_{i_j} > 0. \tag{2}$$

If the matrix semigroup $S \subseteq \mathrm{H}(3, \mathbb{Q})$ has two different matrices N_1 and N_2 such that $\psi(N_1) = (0, 0, c_1)$ and $\psi(N_2) = (0, 0, c_2)$ and $c_1 c_2 < 0$, then the identity matrix exists in S. Indeed, let $\psi(N_1) = (0, 0, \frac{p_1}{q_1})$ and $\psi(N_2) = (0, 0, \frac{p_2}{q_2})$, where $p_1, q_1, q_2 \in \mathbb{Z}$ are positive and $p_2 \in \mathbb{Z}$ is negative. Then it is easy to see that the matrix $N_1^{-q_1 p_2} N_2^{q_2 p_1}$ exists in S and that $\psi(N_1^{-q_1 p_2} N_2^{q_2 p_1}) = (0, 0, 0)$.

To construct the matrix M', we first classify the matrices into four types as follows. A matrix with a superdiagonal vector (a, b) is classified as

1) the $(+,+)$-type if $a, b > 0$,
2) the $(+,-)$-type if $a \geq 0$ and $b \leq 0$,
3) the $(-,-)$-type if $a, b < 0$, and
4) the $(-,+)$-type if $a < 0$ and $b > 0$.

Let $G = \{M_1, \ldots, M_r\}$ be the generating set of the matrix semigroup S. Then $G = G_{(+,+)} \sqcup G_{(+,-)} \sqcup G_{(-,-)} \sqcup G_{(-,+)}$ such that $G_{(\xi_1, \xi_2)}$ is the set of matrices of the (ξ_1, ξ_2)-type, where $\xi_1, \xi_2 \in \{+, -\}$.

Recall that we assume $M = M_{i_1} \cdots M_{i_k}$ and $\psi(M) = (0, 0, c)$ for some $c > 0$. The main idea of the proof is to generate a matrix M' such that $\psi(M') = (0, 0, c')$ for some $c' < 0$ by duplicating the matrices in the sequence $M = M_{i_1} \cdots M_{i_k}$ multiple times and

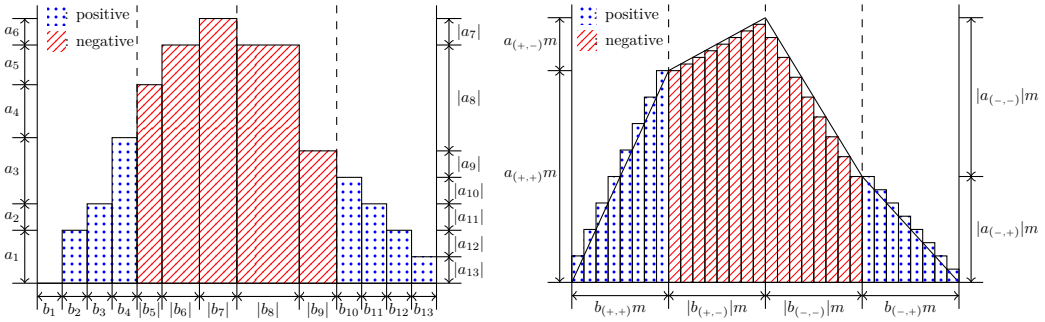

Figure 1 The histogram on the left describes how the upper-right corner of $M_1 \cdots M_{13}$ is computed by multiplications. The blue dotted (red lined) area implies the value which will be added to (subtracted from) the upper-right corner of the final matrix after multiplications of matrices in the sequence. The histogram on the right describes how the upper-right corner of $M^m_{(+,+)}M^m_{(+,-)}M^m_{(-,-)}M^m_{(-,+)}$ is computed by multiplications. Here $m = 8$.

reshuffling. Note that any permutation of the sequence generating the matrix M such that $\psi(M) = (0, 0, c)$ still generates matrices M' such that $\psi(M') = (0, 0, c')$ since the multiplication of matrices exchanges the first two coordinates in a commutative way. Also note that there exists a permutation such that $c \neq c'$ as we assumed that at least two matrices in the sequence do not commute. Moreover, we can still obtain matrices M'' such that $\psi(M'') = (0, 0, c'')$ for some $c'' \in \mathbb{Q}$ if we shuffle two different permutations of the sequence $M_{i_1} \cdots M_{i_k}$ by the same reason.

Let us illustrate the idea with the following example. See Figure 1 for pictorial descriptions of the idea. Let $\{M_i \mid 1 \leq i \leq 4\} \subseteq G_{(+,+)}$, $\{M_i \mid 5 \leq i \leq 7\} \subseteq G_{(+,-)}$, $\{M_i \mid 8 \leq i \leq 9\} \subseteq G_{(-,-)}$, and $\{M_i \mid 10 \leq i \leq 13\} \subseteq G_{(-,+)}$. Then assume that $M_1 \cdots M_{13} = \left(\begin{smallmatrix} 1 & 0 & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, where x is computed by (2). As we mentioned above, x changes if we change the order of multiplicands. In this example, we first multiply $(+,+)$-type matrices and accumulate the values in the superdiagonal coordinates since these matrices have positive values in the coordinates. Indeed, the blue dotted area implies the value we add to the upper-right corner by multiplying such matrices. Then we multiply $(+,-)$-type matrices and still increase the 'a'-value. The 'b'-values in $(+,-)$-type matrices are negative thus, the red lined area is subtracted from the upper-right corner. We still subtract by multiplying $(-,-)$-type matrices since the accumulated 'a'-value is still positive and 'b'-values are negative. We finish the multiplication by adding exactly the last blue dotted area to the upper-right corner. It is easy to see that the total subtracted value is larger than the total added value.

However, we cannot guarantee that x is negative since $\sum_{i=1}^{13} c_i$ could be larger than the contribution from the superdiagonal coordinates. This is why we need to copy the sequence of matrices generating the matrix corresponding to the triple $(0, 0, c)$ for some $c \in \mathbb{Q}$. In Figure 1, we describe an example where we duplicate the sequence eight times and shuffle and permute it in order to minimize the value in the upper-right corner. Now the lengths of both axes are m ($m = 8$ in this example) times larger than before and it follows that the area also grows quadratically in m. Since the summation $m \cdot \sum_{i=1}^{13} c_i$ grows linearly in m, we have $x < 0$ for large enough m. In [27], we formally prove this by bounding contributions of each matrix type and showing that the coefficient of the highest power of m is negative.

It should be noted that there are some subcases where some matrix types are not present in the product. In each case we need to show that the same idea can be used to construct a matrix M' such that $\psi(M') = (0, 0, c')$, where $c' < 0$. The full analysis of all cases can be found in [27].

▶ **Theorem 9.** *The identity problem for a semigroup generated by matrices from* $\mathrm{H}(3,\mathbb{Q})$ *is in polynomial time.*

Proof. Let S be the matrix semigroup in $\mathrm{H}(3,\mathbb{Q})$ generated by the set $G = \{M_1, \ldots, M_r\}$. There are two possible cases of having the identity matrix in the matrix semigroup in $\mathrm{H}(3,\mathbb{Q})$. Either the identity matrix is generated by a product of matrices where all superdiagonal vectors are parallel or there are at least two matrices with non-parallel superdiagonal vectors.

Consider the first case. Lemma 7 provides a formula to compute the value in the top corner regardless of the order of the multiplications. That is, we need to solve a system of linear homogeneous Diophantine equations with solutions over non-negative integers. We partition the set G into several disjoint subsets $G_1, \ldots, G_s$, where s is at most r, and each subset contains matrices with parallel superdiagonal vectors. Since superdiagonal vectors being parallel is a transitive and symmetric property, each matrix needs to be compared to a representative of each subset. If there are no matrices with parallel superdiagonal vectors, then there are r subsets G_i containing exactly one matrix and $O(r^2)$ tests were done. Let us consider $G_i = \{M_{k_1}, \ldots, M_{k_{s_i}}\}$, i.e., one of the subsets containing s_i matrices and $\psi(M_{k_j}) = (a_{k_j}, b_{k_j}, c_{k_j})$. By Lemma 7, the value $c_{k_j} - \frac{q_i}{2}a_{k_j}^2$, for a fixed $q_i \in \mathbb{Q}$, is added to the top corner when matrix M_{k_j} is multiplied.

We solve the system of two linear homogeneous Diophantine equations $A\mathbf{y} = \mathbf{0}$, where

$$A = \begin{pmatrix} a_{k_1} & a_{k_2} & \cdots & a_{k_{s_i}} \\ c_{k_1} - \frac{q_i}{2}a_{k_1}^2 & c_{k_2} - \frac{q_i}{2}a_{k_2}^2 & \cdots & c_{k_{s_i}} - \frac{q_i}{2}a_{k_{s_i}}^2 \end{pmatrix}$$

and $\mathbf{y}^\mathsf{T} \in \mathbb{N}^{s_i}$. The first row is the constraint that guarantees that the first component of the superdiagonal is zero in the matrix product constructed from a solution. Since the superdiagonal vectors are parallel, it also implies that the whole vector is zero. The second row guarantees that the upper corner is zero.

It is obvious that the identity matrix is in the semigroup if we have a solution in the system of two linear homogeneous Diophantine equations for any subset G_i. That is, we need to solve at most r systems of two linear homogeneous Diophantine equations.

Next, we consider the second case, where by Lemma 8, it is enough to check whether there exists a sequence of matrices generating a matrix with zero superdiagonal vector and containing two matrices with non-parallel superdiagonal vectors. Let us say that $M_{i_1}, M_{i_2} \in G$, where $1 \leq i_1, i_2 \leq r$ are the two matrices. Recall that $G = \{M_1, \ldots, M_r\}$ is a generating set of the matrix semigroup and let $\psi(M_i) = (a_i, b_i, c_i)$ for all $1 \leq i \leq r$. We can see that there exists such a product containing the two matrices by solving a system of two linear homogeneous Diophantine equations of the form $B\mathbf{y} = \mathbf{0}$, where $B = \left(\begin{smallmatrix} a_1 & a_2 & \cdots & a_r \\ b_1 & b_2 & \cdots & b_r \end{smallmatrix}\right)$, with an additional constraint that the numbers in the solution $\mathbf{y}$ that correspond to M_{i_1} and M_{i_2} are non-zero since we must use these two matrices in the product. We repeat this process at most $r(r-1)$ times until we find a solution. Therefore, the problem reduces again to solving at most $O(r^2)$ systems of two linear homogeneous Diophantine equations.

Finally, we conclude the proof by mentioning that the identity problem for matrix semigroups in the Heisenberg group over rationals $\mathrm{H}(3,\mathbb{Q})$ can be decided in polynomial time as the problem of existence of a positive integer solution to a system of linear homogeneous Diophantine equations is in polynomial time. Note that if the system is non-homogeneous, then solvability of a system of linear Diophantine equations with solutions over positive integers is an NP-complete problem; see for example [38]. Indeed, a system of linear homogeneous Diophantine equations with solutions over non-negative integers can be converted to a linear programming problem with a solution over rationals which is known to be solvable in polynomial time; see e.g., [43]. It is easy to add additional constraints to the linear programming

that ensure that solutions are positive and non-zero. As the system is homogeneous, any solution can be converted to an integer solution by multiplying by the denominators. ◀

Next, we generalize the above algorithm for the identity problem in the Heisenberg group $\mathrm{H}(3,\mathbb{Q})$ to the domain of the Heisenberg groups for any dimension over the rational numbers. Similarly to the case of dimension three, we establish the following result for the case of matrices where multiplication is commutative.

▶ **Lemma 10.** *Let $G = \{M_1, \ldots, M_r\} \subseteq \mathrm{H}(n,\mathbb{Q})$ be a set of matrices from the Heisenberg group such that $\psi(M_i) = (\mathbf{a}_i, \mathbf{b}_i, c_i)$ and $\psi(M_j) = (\mathbf{a}_j, \mathbf{b}_j, c_j)$ and $\mathbf{a}_i \cdot \mathbf{b}_j = \mathbf{a}_j \cdot \mathbf{b}_i$ for any $1 \leq i \neq j \leq r$. If there exists a sequence of matrices $M = M_{i_1} \cdots M_{i_k}$, where $i_j \in [1,r]$ for all $1 \leq j \leq k$, such that $\psi(M) = (\mathbf{0}, \mathbf{0}, c)$ for some $c \in \mathbb{Q}$, then $c = \sum_{j=1}^{k}(c_{i_j} - \frac{1}{2}\mathbf{a}_{i_j} \cdot \mathbf{b}_{i_j})$.*

Lemma 8 does not generalize to $\mathrm{H}(n,\mathbb{Q})$ in the same way as we cannot classify matrices according to types to control the value in upper-right corner, so we use a different technique to prove that the value in the upper corner will be diverging to both positive and negative infinity quadratically as we repeat the same sequence generating any matrix M such that $\psi(M) = (\mathbf{0}, \mathbf{0}, c)$.

▶ **Lemma 11.** *Let $S = \langle M_1, \ldots, M_r\rangle \subseteq \mathrm{H}(n,\mathbb{Q})$ be a finitely generated matrix semigroup. Then the identity matrix exists in S if there exists a sequence of matrices $M_{i_1} \cdots M_{i_k}$, where $i_j \in [1,r]$ for all $1 \leq j \leq k$, satisfying the following properties:*

(i) *$\psi(M_{i_1} \cdots M_{i_k}) = (\mathbf{0}, \mathbf{0}, c)$ for some $c \in \mathbb{Q}$, and*

(ii) *$\mathbf{a}_{i_{j_1}} \cdot \mathbf{b}_{i_{j_2}} \neq \mathbf{a}_{i_{j_2}} \cdot \mathbf{b}_{i_{j_1}}$ for some $j_1, j_2 \in [1,k]$, where $\psi(M_i) = (\mathbf{a}_i, \mathbf{b}_i, c_i)$ for $1 \leq i \leq r$.*

Proof. From the first property claimed in the lemma, we know that any permutation of the sequence of matrix multiplications of $M_{i_1} \cdots M_{i_k}$ results in matrices M' such that $\psi(M') = (\mathbf{0}, \mathbf{0}, y)$ for some $y \in \mathbb{Q}$ since the multiplication of matrices in the $\mathrm{H}(n,\mathbb{Q})$ performs additions of vectors which is commutative in the top row and the rightmost column excluding the upper-right corner. From the commutative behaviour in the horizontal and vertical vectors of matrices in the Heisenberg group, we also know that if we duplicate the matrices in the sequence $M_{i_1} \cdots M_{i_k}$ and multiply the matrices in any order, then the resulting matrix has a non-zero coordinate in the upper triangular coordinates only in the upper right corner.

Now let $j_1, j_2 \in [1,k]$ be two indices such that $\mathbf{a}_{i_{j_1}} \cdot \mathbf{b}_{i_{j_2}} \neq \mathbf{a}_{i_{j_2}} \cdot \mathbf{b}_{i_{j_1}}$ as claimed in the lemma. Then consider the following matrix M_d that can be obtained by duplicating the sequence $M_{i_1} \cdots M_{i_k}$ of matrices into ℓ copies and shuffling the order as follows: $M_d = M_{i_{j_1}}^{\ell} M_{i_{j_2}}^{\ell} M_x^{\ell}$, where M_x is a matrix that is obtained by multiplying the matrices in $M_{i_1} \cdots M_{i_k}$ except the two matrices M_{j_1} and M_{j_2}. Then it is clear that $\psi(M_d) = (\mathbf{0}, \mathbf{0}, d)$ for some d. Let $\psi(M_x) = (\mathbf{a}_x, \mathbf{b}_x, c_x)$. Then it is easy to see that $\mathbf{a}_{i_{j_1}} + \mathbf{a}_{i_{j_2}} + \mathbf{a}_x = \mathbf{0}$ and $\mathbf{b}_{i_{j_1}} + \mathbf{b}_{i_{j_2}} + \mathbf{b}_x = \mathbf{0}$. Consider then a product, where order of M_{j_1} and M_{j_2} is swapped. That is, $M_e = M_{i_{j_2}}^{\ell} M_{i_{j_1}}^{\ell} M_x^{\ell}$ and let us denote $\psi(M_e) = (\mathbf{0}, \mathbf{0}, e)$ for some $e \in \mathbb{Q}$. By solving values d and e, we notice that the coefficient of ℓ^2 in d is $\mathbf{a}_{i_{j_1}} \cdot \mathbf{b}_{i_{j_2}} - \mathbf{a}_{i_{j_2}} \cdot \mathbf{b}_{i_{j_1}}$ and in e is $\mathbf{a}_{i_{j_2}} \cdot \mathbf{b}_{i_{j_1}} - \mathbf{a}_{i_{j_1}} \cdot \mathbf{b}_{i_{j_2}}$. As we assumed that $\mathbf{a}_{i_{j_1}} \cdot \mathbf{b}_{i_{j_2}} \neq \mathbf{a}_{i_{j_2}} \cdot \mathbf{b}_{i_{j_1}}$, the coefficients are of different sign. Hence, for sufficiently large ℓ, the values d and e have opposite signs. Then, as in the proof Lemma 8, the identity matrix always exists in the semigroup as we can multiply these two matrices correct number of times to have zero in the upper right coordinate as well. ◀

Next, we prove that the identity problem is decidable for n-dimensional Heisenberg matrices. In contrast to Theorem 9, we do not claim that the problem is decidable in polynomial time since one of the steps of the proof is to partition matrices according to dot products which cannot be extended to higher dimensions than three. For higher dimensions,

partitioning matrices according to dot products takes an exponential time in the number of matrices in the generating set. Note that if the size of the generating set is fixed, i.e., only the matrices are part of the input, then the problem remains in P.

▶ **Theorem 12.** *The identity problem for a semigroup generated by matrices from* $\mathrm{H}(n,\mathbb{Q})$ *is decidable.*

Proof. Similarly to the proof of Theorem 9, there are two ways the identity matrix can be generated. Either all the matrices commute or there are at least two matrices that do not commute. Let S be the matrix semigroup in $\mathrm{H}(n,\mathbb{Q})$ generated by the set $G = \{M_1,\ldots,M_r\}$. Consider matrices N_1, N_2 and N_3, such that $\psi(N_1) = (\mathbf{a}_1,\mathbf{b}_1,c_1)$, $\psi(N_2) = (\mathbf{a}_2,\mathbf{b}_2,c_2)$ and $\psi(N_3) = (\mathbf{a}_3,\mathbf{b}_3,c_3)$. If $\mathbf{a}_1\cdot\mathbf{b}_2 = \mathbf{a}_2\cdot\mathbf{b}_1$ and $\mathbf{a}_2\cdot\mathbf{b}_3 = \mathbf{a}_3\cdot\mathbf{b}_2$, it does not imply that $\mathbf{a}_1\cdot\mathbf{b}_3 = \mathbf{a}_3\cdot\mathbf{b}_1$. Therefore, the number of subsets of G, where each subset contains matrices that commute with other matrices in the same subset, is exponential in r as two subsets are not necessarily disjoint. Now we examine whether it is possible to generate the identity matrix by multiplying matrices in each subset by Lemma 10. If it is not possible, we need to consider the case of having two matrices that do not commute with each other in the product with zero values in the upper-triangular coordinates except the corner. Let us say that $M_{i_1}, M_{i_2} \in G$, where $1 \le i_1, i_2 \le r$ are the two matrices. Recall that $G = \{M_1,\ldots,M_r\}$ is a generating set of the matrix semigroup and let $\psi(M_i) = (\mathbf{a}_i,\mathbf{b}_i,c_i)$ for all $1 \le i \le r$.

Then we can see that there exists such a product by solving a system of $2(n-2)$ linear homogeneous Diophantine equations of the form $B\mathbf{y} = \mathbf{0}$, where $B = \begin{pmatrix} \mathbf{a}_1^{\mathsf{T}}\cdots\mathbf{a}_r^{\mathsf{T}} \\ \mathbf{b}_1^{\mathsf{T}}\cdots\mathbf{b}_r^{\mathsf{T}} \end{pmatrix}$, with an additional constraint that the values in the solution $\mathbf{y}$ that correspond to M_{i_1} and M_{i_2} are non-zero since we must use these two matrices in the product. We repeat this process at most $r(r-1)$ times until we find a solution. Hence, we can view the identity problem in $\mathrm{H}(n,\mathbb{Q})$ as the problem of solving systems of $2(n-2)$ linear homogeneous Diophantine equations with some constraints on the solution. As we can solve systems of linear homogeneous Diophantine equations, we conclude that the identity problem in $\mathrm{H}(n,\mathbb{Q})$ is also decidable. ◀

5 The identity problem in matrix semigroups in dimension four

Now we tighten the decidability gap proving undecidability for 4×4 matrices, when the generating set has eight matrices (reducing from 48), with a new technique exploiting the anti-diagonal entries.

▶ **Theorem 13.** *Given a semigroup S generated by eight 4×4 integer matrices with determinant one, determining whether the identity matrix belongs to S is undecidable.*

Proof. We prove the claim by reducing from the PCP. We shall use an encoding to embed an instance of the PCP into a set of 4×4 integer matrices. An *instance* of the PCP consists of two morphisms $g,h:\Sigma^*\to B^*$, where Σ and B are alphabets. A nonempty word $u\in\Sigma^*$ is a *solution* of an instance (g,h) if it satisfies $g(u) = h(u)$.

Let α be the mapping of Lemma 2. We also define a monomorphism $f:\mathrm{FG}(\Gamma_2)\to\mathbb{Z}^{2\times 2}$ as $f(a) = \left(\begin{smallmatrix}1&2\\0&1\end{smallmatrix}\right)$, $f(\bar{a}) = \left(\begin{smallmatrix}1&-2\\0&1\end{smallmatrix}\right)$, $f(b) = \left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)$ and $f(\bar{b}) = \left(\begin{smallmatrix}1&0\\-2&1\end{smallmatrix}\right)$. Recall that the matrices $\left(\begin{smallmatrix}1&2\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)$ generate a free subgroup of $\mathrm{SL}(2,\mathbb{Z})$ [31]. The composition of two monomorphisms α and f gives us the embedding from an arbitrary group alphabet into $\mathrm{SL}(2,\mathbb{Z})$. We use the composition of two monomorphisms α and f to encode a set of pairs of words over an arbitrary group alphabet into a set of 4×4 integer matrices in $\mathrm{SL}(4,\mathbb{Z})$ and denote it by β.

Let (g,h) be an instance of the PCP, where $g,h:\{a_1,\ldots,a_n\}^*\to\Sigma_2^*$, where $\Sigma_2 = \{a,b\}$. Without loss of generality, we can assume that the solution starts with the letter

a_1. Moreover, we assume that this is the only occurence of a_1. We define the alphabet $\Gamma = \Sigma_2 \cup \Sigma_2^{-1} \cup \Sigma_B \cup \Sigma_B^{-1}$, where $\Sigma_B = \{q_0, q_1, p_0, p_1\}$ is the alphabet for the border letters that enforce the form of a solution.

Let us define the following sets of words $W_1 \cup W_2 \subseteq \mathrm{FG}(\Gamma) \times \mathrm{FG}(\Gamma)$, where

$$W_1 = \{(q_0 a \overline{q_0}, p_0 a \overline{p_0}),\ (q_0 b \overline{q_0}, p_0 b \overline{p_0}) \mid a, b \in \Sigma_2,\ \ q_0, p_0 \in \Sigma_B\} \text{ and}$$

$$W_2 = \left\{(q_0 \overline{g(a_1)} \overline{q_1}, p_0 \overline{h(a_1)} \overline{p_1}),\ (q_1 \overline{g(a_i)} \overline{q_1}, p_1 \overline{h(a_i)} \overline{p_1}) \mid 1 < i \leq n,\ \ q_0, q_1, p_0, p_1 \in \Sigma_B\right\}.$$

Intuitively, the words from set W_1 are used to construct words over Σ_2 and the words from set W_2 to cancel them according to the instance of the PCP.

Let us prove that $(q_0\overline{q_1}, p_0\overline{p_1}) \in \mathrm{FG}(W_1 \cup W_2)$ if and only if the PCP has a solution. It is easy to see that any pair of non-empty words in $\mathrm{FG}(W_1)$ is of the form $(q_0 w \overline{q_0}, p_0 w \overline{p_0})$ for $w \in \Sigma_2^+$. Then there exists a pair of words in $\mathrm{FG}(W_2)$ of the form $(q_0 \overline{w} \overline{q_1}, p_0 \overline{w} \overline{p_1})$ for some word $w \in \Sigma_2^+$ if and only if the PCP has a solution. Thus, $(q_0\overline{q_1}, p_0\overline{p_1})$ can be constructed by concatenating pairs of words in W_1 and W_2 if and only if the PCP has a solution.

For each pair of words $(u, v) \in \mathrm{FG}(W_1 \cup W_2)$, we define a matrix $A_{u,v}$ to be $\begin{pmatrix} \beta(u) & \mathbf{0}_2 \\ \mathbf{0}_2 & \beta(v) \end{pmatrix} \in \mathrm{SL}(4, \mathbb{Z})$, where $\mathbf{0}_2$ is the zero matrix in $\mathbb{Z}^{2\times 2}$. Moreover, we define the following matrix

$$B_{q_1\overline{q_0}, p_1\overline{p_0}} = \begin{pmatrix} \mathbf{0}_2 & \beta(q_1\overline{q_0}) \\ \beta(p_1\overline{p_0}) & \mathbf{0}_2 \end{pmatrix} \in \mathrm{SL}(4, \mathbb{Z}).$$

Let S be a matrix semigroup generated by the set $\{A_{u,v}, B_{q_1\overline{q_0}, p_1\overline{p_0}} \mid (u, v) \in W_1 \cup W_2\}$. We already know that the pair $(q_0\overline{q_1}, p_0\overline{p_1})$ of words can be generated by concatenating words in W_1 and W_2 if and only if the PCP has a solution. The matrix semigroup S has the corresponding matrix $A_{q_0\overline{q_1}, p_0\overline{p_1}}$ and thus, $\begin{pmatrix} \beta(q_0\overline{q_1}) & \mathbf{0}_2 \\ \mathbf{0}_2 & \beta(p_0\overline{p_1}) \end{pmatrix}\begin{pmatrix} \mathbf{0}_2 & \beta(q_1\overline{q_0}) \\ \beta(p_1\overline{p_0}) & \mathbf{0}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0}_2 & \beta(\varepsilon) \\ \beta(\varepsilon) & \mathbf{0}_2 \end{pmatrix} \in S$. Now, the identity matrix $\boldsymbol{I}_4$ exists in the semigroup S by repeating this product twice.

Now we prove that the identity matrix does not exist in S if the PCP has no solution. It is easy to see that we cannot obtain the identity matrix only by multiplying 'A' matrices since there is no possibility of cancelling every border letter. We need to multiply the matrix $B_{q_1\overline{q_0}, p_1\overline{p_0}}$ with a product of 'A' matrices at some point to reach the identity matrix. Note that the matrix $B_{q_1\overline{q_0}, p_1\overline{p_0}}$ cannot be the first matrix of the product, followed by the 'A' matrices, because the upper right block of $B_{q_1\overline{q_0}, p_1\overline{p_0}}$, which corresponds to the first word of the pair, should be multiplied with the lower right block of 'A' matrix, which corresponds to the second word of the pair.

Suppose that the 'A' matrix is of form $\begin{pmatrix} \beta(q_0 u \overline{q_1}) & \mathbf{0}_2 \\ \mathbf{0}_2 & \beta(p_0 v \overline{p_1}) \end{pmatrix}$. Since the PCP instance has no solution, either u or v is not the empty word. We multiply $B_{q_1\overline{q_0}, p_1\overline{p_0}}$ to the matrix and then obtain the following matrix $\begin{pmatrix} \beta(q_0 u \overline{q_1}) & \mathbf{0}_2 \\ \mathbf{0}_2 & \beta(p_0 v \overline{p_1}) \end{pmatrix}\begin{pmatrix} \mathbf{0}_2 & \beta(q_1\overline{q_0}) \\ \beta(p_1\overline{p_0}) & \mathbf{0}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0}_2 & \beta(q_0 u \overline{q_0}) \\ \beta(p_0 v \overline{p_0}) & \mathbf{0}_2 \end{pmatrix}$. We can see that either the upper right part or the lower left part cannot be $\beta(\varepsilon)$, which actually corresponds to the identity matrix in $\mathbb{Z}^{2\times 2}$. Now the only possibility of reaching the identity matrix is to multiply matrices which have $\mathrm{SL}(2, \mathbb{Z})$ matrices in the anti-diagonal coordinates like $B_{q_1\overline{q_0}, p_1\overline{p_0}}$. However, we cannot cancel the parts because the upper right block (the lower left block) of the left matrix is multiplied with the lower left block (the upper right block) of the right matrix as follows $\left(\begin{smallmatrix} \mathbf{0}_2 & A \\ B & \mathbf{0}_2 \end{smallmatrix}\right)\left(\begin{smallmatrix} \mathbf{0}_2 & C \\ D & \mathbf{0}_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} AD & \mathbf{0}_2 \\ \mathbf{0}_2 & BC \end{smallmatrix}\right)$, where A, B, C and D are matrices in $\mathbb{Z}^{2\times 2}$. As the first word of the pair is encoded in the upper right block of the matrix and the second word is encoded in the lower left block, it is not difficult to see that we cannot cancel the remaining blocks.

Currently, the undecidability bound for the PCP is five [34] and thus the semigroup S is generated by eight matrices. Recall that in the beginning of the proof, we assumed that letter a_1 of the PCP is used exacly once and is the first letter of a solution. This property is in fact present in [34]. ◀

Theorem 13 implies smaller undecidability bounds for the *special diagonal membership problem* from 14 [24] to eight and for the identity problem in $\mathbb{H}(\mathbb{Q})^{2\times 2}$ from 48 to eight [6].

References

1. Andrei M. Akimenkov. Subgroups of the braid group B_4. *Mathematical notes of the Academy of Sciences of the USSR*, 50(6):1211–1218, 1991. doi:10.1007/BF01158260.
2. László Babai, Robert Beals, Jin-yi Cai, Gábor Ivanyos, and Eugene M. Luks. Multiplicative equations over commuting matrices. In *Proceedings of SODA 1996*, pages 498–507. SIAM, 1996. URL: http://dl.acm.org/citation.cfm?id=313852.314109.
3. Paul Bell, Vesa Halava, Tero Harju, Juhani Karhumäki, and Igor Potapov. Matrix equations and Hilbert's Tenth Problem. *International Journal of Algebra and Computation*, 18(08):1231–1241, dec 2008. doi:10.1142/s0218196708004925.
4. Paul C. Bell, Mika Hirvensalo, and Igor Potapov. The identity problem for matrix semigroups in SL$(2,\mathbb{Z})$ is NP-complete. In *Proceedings of SODA 2017*, pages 187–206. SIAM, 2017. doi:10.1137/1.9781611974782.13.
5. Paul C. Bell and Igor Potapov. Reachability problems in quaternion matrix and rotation semigroups. *Information and Computation*, 206(11):1353–1361, 2008. doi:10.1016/j.ic.2008.06.004.
6. 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(6):963–978, 2010. doi:10.1142/S0129054110007660.
7. Vladimir N. Bezverkhnii and Irina V. Dobrynina. Undecidability of the conjugacy problem for subgroups in the colored braid group R_5. *Matematicheskie Zametki*, 65(1):15–22, 1999. doi:10.1007/BF02675004.
8. Jean-Camille Birget and Stuart W. Margolis. Two-letter group codes that preserve aperiodicity of inverse finite automata. *Semigroup Forum*, 76:159–168, 2008. doi:10.1007/s00233-007-9024-6.
9. Kenneth R. Blaney and Andrey Nikolaev. A PTIME solution to the restricted conjugacy problem in generalized Heisenberg groups. *Groups Complexity Cryptology*, 8(1):69–74, 2016. doi:10.1515/gcc-2016-0003.
10. Vincent D. Blondel and Alexandre Megretski, editors. *Unsolved problems in mathematical systems and control theory*. Princeton University Press, 2004.
11. Jean-Luc Brylinski. *Loop spaces, characteristic classes, and geometric quantization*. Birkhäuser, 1993.
12. Julien Cassaigne, Tero Harju, and Juhani Karhumäki. On the undecidability of freeness of matrix semigroups. *International Journal of Algebra and Computation*, 9(03n04):295–305, 1999. doi:10.1142/S0218196799000199.
13. Émilie Charlier and Juha Honkala. The freeness problem over matrix semigroups and bounded languages. *Information and Computation*, 237:243–256, 2014. doi:10.1016/j.ic.2014.03.001.
14. Christian Choffrut. A remark on the representation of trace monoids. *Semigroup Forum*, 40(1):143–152, 1990. doi:10.1007/bf02573262.
15. Christian Choffrut and Juhani Karhumäki. Some decision problems on integer matrices. *RAIRO - Theoretical Informatics and Applications*, 39(1):125–131, 2005. doi:10.1051/ita:2005007.
16. Ventsislav Chonev, Joël Ouaknine, and James Worrell. The orbit problem in higher dimensions. In *Proceedings of STOC 2013*, pages 941–950. ACM, 2013. doi:10.1145/2488608.2488728.
17. Ventsislav Chonev, Joël Ouaknine, and James Worrell. On the complexity of the orbit problem. *Journal of the ACM*, 63(3):23:1–23:18, 2016. doi:10.1145/2857050.

18 Marston Conder, Edmund Robertson, and Peter Williams. Presentations for 3-dimensional special linear groups over integer rings. *Proceedings of the American Mathematical Society*, 115(1):19–26, 1992. doi:10.2307/2159559.

19 Marston D. E. Conder. Some unexpected consequences of symmetry computations. In *SIGMAP 2014*, volume 159 of *PROMS*, pages 71–79. Springer, 2016. doi:10.1007/978-3-319-30451-9_3.

20 Jintai Ding, Alexei Miasnikov, and Alexander Ushakov. A linear attack on a key exchange protocol using extensions of matrix semigroups. *IACR Cryptology ePrint Archive*, 2015:18, 2015.

21 Esther Galby, Joël Ouaknine, and James Worrell. On matrix powering in low dimensions. In *Proceedings of STACS 2015*, volume 30 of *LIPIcs*, pages 329–340, 2015. doi:10.4230/LIPIcs.STACS.2015.329.

22 Razvan Gelca and Alejandro Uribe. From classical theta functions to topological quantum field theory. In *The influence of Solomon Lefschetz in geometry and topology*, volume 621 of *Contemprorary Mathematics*, pages 35–68. American Mathematical Society, 2014. doi:10.1090/conm/621.

23 Yuri Gurevich and Paul Schupp. Membership problem for the modular group. *SIAM Journal of Computing*, 37(2):425–459, 2007. doi:10.1137/050643295.

24 Vesa Halava, Tero Harju, and Mika Hirvensalo. Undecidability bounds for integer matrices using Claus instances. *International Journal of Foundations of Computer Science*, 18(5):931–948, 2007. doi:10.1142/S0129054107005066.

25 Juha Honkala. A Kraft-McMillan inequality for free semigroups of upper-triangular matrices. *Information and Computation*, 239:216–221, 2014. doi:10.1016/j.ic.2014.09.002.

26 Juha Honkala. Products of matrices and recursively enumerable sets. *Journal of Computer and System Sciences*, 81(2):468–472, 2015. doi:10.1016/j.jcss.2014.10.004.

27 Sang-Ki Ko, Reino Niskanen, and Igor Potapov. On the identity problem for the special linear group and the Heisenberg group. *CoRR*, abs/1706.04166, 2017. URL: https://arxiv.org/abs/1706.04166, arXiv:1706.04166.

28 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. doi:10.1090/conm/677/13625.

29 Bertram Kostant. Quantization and unitary representations. In *Lectures in Modern Analysis and Applications III*, pages 87–208. Springer, 1970. doi:10.1007/BFb0079068.

30 Alexei Lisitsa and Igor Potapov. Membership and reachability problems for row-monomial transformations. In *Proceedings of MFCS 2004*, volume 3153 of *LNCS*, pages 623–634. Springer, 2004. doi:10.1007/978-3-540-28629-5_48.

31 Roger C. Lyndon and Paul E. Schupp. *Combinatorial group theory*. Springer, 1977. doi:10.1007/978-3-642-61896-3.

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

33 Alexei Mishchenko and Alexander Treier. Knapsack problem for nilpotent groups. *Groups Complexity Cryptology*, 9(1):87–98, 2017. doi:10.1515/gcc-2017-0006.

34 Turlough Neary. Undecidability in binary tag systems and the Post correspondence problem for five pairs of words. In *Proceedings of STACS 2015*, volume 30 of *LIPIcs*, pages 649–661. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015. doi:10.4230/LIPIcs.STACS.2015.649.

35 Joël Ouaknine, João Sousa Pinto, and James Worrell. On termination of integer linear loops. In *Proceedings of SODA 2015*, pages 957–969. SIAM, 2015. doi:10.1137/1.9781611973730.65.

36 Joël Ouaknine and James Worrell. On the positivity problem for simple linear recurrence sequences. In *Proceedings of ICALP 2014*, volume 8573 of *LNCS*, pages 318–329. Springer, 2014. doi:10.1007/978-3-662-43951-7_27.

37 Joël Ouaknine and James Worrell. Ultimate positivity is decidable for simple linear recurrence sequences. In *Proceedings of ICALP 2014*, volume 8573 of *LNCS*, pages 330–341. Springer, 2014. doi:10.1007/978-3-662-43951-7_28.

38 Christos H. Papadimitriou. On the complexity of integer programming. *Journal of the ACM*, 28(4):765–768, 1981. doi:10.1145/322276.322287.

39 Michael S. Paterson. Unsolvability in 3×3 matrices. *Studies in Applied Mathematics*, 49(1):105, 1970. doi:10.1002/sapm1970491105.

40 Igor Potapov. Composition problems for braids. In *Proceedings of FSTTCS 2013*, volume 24 of *LIPIcs*, pages 175–187. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013. doi:10.4230/LIPIcs.FSTTCS.2013.175.

41 Igor Potapov and Pavel Semukhin. Decidability of the membership problem for 2×2 integer matrices. In *Proceedings of SODA 2017*, pages 170–186. SIAM, 2017. doi:10.1137/1.9781611974782.12.

42 Igor Potapov and Pavel Semukhin. Membership problem in $\mathrm{GL}(2,\mathbb{Z})$ extended by singular matrices. In *Proceedings of MFCS 2017*, LIPIcs, pages 44:1–44:13. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017. doi:10.4230/LIPIcs.MFCS.2017.44.

43 Alexander Schrijver. *Theory of Linear and Integer Programming*. Wiley, 1998.