On Word Representations and Embeddings in Complex Matrices

Paul C. Bell[1], George Kenison[2], Reino Niskanen[1],
Igor Potapov[3], and Pavel Semukhin[1]

[1] Liverpool John Moores University, UK
{p.c.bell,r.niskanen,p.semukhin}@ljmu.ac.uk
[2] KU Leuven, Belgium george.kenison@kuleuven.be
[3] University of Liverpool, UK potapov@liverpool.ac.uk

Abstract. Embeddings of word structures into matrix semigroups provide a natural bridge between combinatorics on words and linear algebra. However, low-dimensional matrix semigroups impose strong structural restrictions on possible embeddings. Certain finitely generated groups admit faithful representations in $\mathrm{SL}(2,\mathbb{C})$ and other similar matrix groups. On the other hand, it is known that the product of two free semigroups on two generators cannot be embedded into 2×2 complex matrices. In this paper we study embeddings of word structures into low-dimensional matrix semigroups over the complex numbers and develop new techniques for constructing word representations of the Euclidean Bianchi groups. These representations provide a symbolic framework and a natural first step towards analysing fundamental decision problems in 2×2 matrix semigroups.

1 Introduction

Matrix products are one of the most fundamental operations in mathematics. Originally introduced as a compact way to represent and solve systems of linear equations, matrix products later came to play an essential role in many fields, with applications ranging from engineering and control theory to modern data science and large-scale information systems such as search ranking algorithms.

The study of embeddings of word structures into matrices connects combinatorics on words with linear algebra. By representing words or generators as matrices, questions about concatenation of words and computational models operating on words can be translated into questions about matrix multiplication. Thus solutions to problems in combinatorics on words and formal language theory can employ algebraic and analytic tools from matrix theory. Moreover, embeddings into integer or complex matrices provide a concrete and well-understood algebraic framework in which structural properties of word semigroups and groups can be analysed, while also enabling the transfer of results between the theory of words and the theory of matrix semigroups.

Consider that it is not always possible to embed word structures within certain matrix classes or dimensions. This observation clarifies the expressive power of small matrix semigroups, showing that some algebraic structures generated by words cannot be faithfully represented in restricted matrix domains. Further, it helps determine which techniques from linear algebra can be applied to word problems, and which problems require different approaches.

The origins of the connection between combinatorial group theory and linear representations go back to the work of Nielsen in the 1920s on free groups [22]. These techniques provided an early structural understanding of free groups and later influenced approaches to representing group elements by algebraic objects. Later Magnus showed that free groups admit faithful representations into groups of matrices over certain rings of formal power series. Interest in such embeddings grew further with the development of algorithmic problems in algebra and theoretical computer science, when researchers began investigating whether problems concerning words and formal languages could be translated into algebraic problems involving matrices and vice versa [7,8,11,20,24,25].

Recent work has also highlighted the importance of decision problems for matrix semigroups over complex numbers. In particular, the first decidability results for the identity and group problems in the complex Heisenberg group $\mathrm{H}(3, \mathbb{C})$ demonstrate that matrix semigroups with non-trivial structural constraints require new number theoretical techniques [3] and new group theoretic techniques for algorithmic analysis [12]. At the same time, many fundamental questions for low-dimensional matrices over the complex numbers remain open. A central example concerns embeddings of word semigroups into 2×2 complex matrices. It is known that there exists no injective semigroup morphism from a pair of words over an alphabet with at least two symbols into $\mathbb{C}^{2\times 2}$ [7]. On the other hand, deciding whether a finitely generated torsion-free group embeds in $\mathrm{SL}(2, \mathbb{C})$ (see [5]) has deep connections to questions in geometric group theory. These results indicate that the expressive power of 2×2 complex matrices lies at a delicate boundary—while they allow rich algebraic representations, they also impose strong structural limitations on possible embeddings.

$\nexists$	$\mathrm{S}(\Sigma_1)$	$\mathrm{S}(\Sigma_2)$	$\mathrm{F}(\Sigma_1)$	$\mathrm{F}(\Sigma_2)$
$\{\varepsilon\}$	?	?	?	$\mathrm{U}(n, \mathbb{C}), n \geq 1$ (Thm. 9)
$\mathrm{S}(\Sigma_1)$	?	?	?	$\mathrm{SL}(2, \mathcal{O}_d)$ (Prop. 16)
$\mathrm{S}(\Sigma_2)$		$\mathbb{C}^{2\times 2}$[7], $\mathrm{SL}(3, \mathbb{Z})$[20]	$\mathcal{O}_d^{2\times 2}$ (Prop. 16)	$\mathbb{Z}^{3\times 3}$ (Prop. 16)
$\mathrm{F}(\Sigma_1)$			?	$\mathcal{O}_d^{2\times 2}$ (Prop. 16)
$\mathrm{F}(\Sigma_2)$				$\mathbb{Z}^{3\times 3}$[20]

Table 1. The state of the art for the non-existence of embeddings from pairs of words into different matrix semigroups. Entries in red are new to the present paper.

This paper explores the boundary between combinatorics on words and matrix semigroups, aiming to characterise which word structures admit faithful low-dimensional matrix representations over complex numbers and develops new techniques to find word representations for the Euclidean Bianchi groups (Theorem 4), an essential first step to study fundamental decision problems in matrix semigroups within their symbolic representations, like membership, freeness, and vector reachability in 2×2 matrix semigroups. Table 1 summarises non-existence results for embeddings in different settings, where $\mathrm{S}(\Sigma_1)$ and $\mathrm{S}(\Sigma_2)$ denote free semigroups over unary and binary alphabets, while $\mathrm{F}(\Sigma_1)$ and $\mathrm{F}(\Sigma_2)$ denote free groups over unary and binary group alphabets. The entries are read as "column" × "row". The grey background indicates symmetrical cases that do not require consideration, e.g., $\mathrm{S}(\Sigma_2) \times \mathrm{S}(\Sigma_1) = \mathrm{S}(\Sigma_1) \times \mathrm{S}(\Sigma_2)$. Table 3 details existing embeddings and can be found in the appendix.

2 Preliminaries

2.1 Words, semigroups, and groups

Given an *alphabet* $\Sigma = \{a_1, a_2, \ldots, a_m\}$, a finite *word* u is a finite sequence of letters, $u = u_1 u_2 \cdots u_n$, where $u_i \in \Sigma$. We denote the set of all finite words over Σ by $\mathrm{S}(\Sigma)$. Note, that we do not use the usual notation of Σ^* to simplify our notation. The *empty word* is denoted by ε. The length of a finite word u is denoted by $|u|$ and $|\varepsilon| = 0$. Consider Σ as a generating set of a free group $\mathrm{F}(\Sigma)$. The free group contains both the letters a_i as well as their inverses a_i^{-1}. Naturally, $a_i a_i^{-1} = a_i^{-1} a_i = \varepsilon$. The elements of $\mathrm{F}(\Sigma)$ are all *reduced* words over Σ, i.e., words not containing $a_i a_i^{-1}$ or $a_i^{-1} a_i$ as a subword. In this context, we call Σ a finite *group alphabet*, i.e., an alphabet with an involution. The multiplication of two elements (reduced words) $u, v \in \mathrm{F}(\Sigma)$ corresponds to the unique reduced word of the concatenation uv.

Let φ be a mapping from $\mathrm{S}(\Sigma)$ into $\mathbb{K}^{n \times n}$, where $\mathbb{K} = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{C}, \ldots$ and $n \geq 1$. That is, a mapping of semigroup words into n-by-n matrices with elements from $\mathbb{K}$. We say that φ is an *injective morphism* or an *embedding* if $\varphi(u)\varphi(v) = \varphi(uv)$ for every $u, v \in \mathrm{S}(\Sigma)$ and if $\varphi(u) = \varphi(v)$ implies that $u = v$. We define *group embeddings* over free group $\mathrm{F}(\Sigma)$ analogously.

The following proposition is folklore. It allows us to focus on small alphabets, namely unary or binary alphabets, in our subsequent results.

Proposition 1. *Let $\Sigma_k = \{a_1, a_2, \ldots, a_k\}$ be an alphabet, defined for any $k \geq 2$. If there exists an embedding $\sigma : \mathrm{S}(\Sigma_2) \to \mathbb{F}^{n \times n}$, then there exists an embedding $\sigma : \mathrm{S}(\Sigma_k) \to \mathbb{F}^{n \times n}$ for any $k \geq 2$. If there exists a group embedding $\sigma : \mathrm{F}(\Sigma_2) \to \mathbb{F}^{n \times n}$, then there exists a group embedding $\sigma : \mathrm{F}(\Sigma_k) \to \mathbb{F}^{n \times n}$ for any $k \geq 2$.*

2.2 Integer rings, Euclidean domains, and quadratic fields

We shall assume some familiarity with concepts from algebraic number theory (cf. [27]). We refer the reader to the background material on properties of certain imaginary quadratic fields to Appendix A.

Let $\overline{\mathbb{Q}}$ denote the field of algebraic numbers. Recall that a number is *algebraic* if it is the root of a non-zero polynomial $p \in \mathbb{Z}[x]$ with integer coefficients. Further, a number is an *algebraic integer* if it is the root of a non-zero monic polynomial $p \in \mathbb{Z}[x]$. We can effectively represent and compute algebraic numbers [9].

An *integral domain* is a non-zero commutative ring such that the product of any two non-zero elements is itself non-zero. An integral domain E is Euclidean (or a *Euclidean domain*) if there exists a non-negative integer valued function g defined on the non-zero elements of E such that for every $x, y \in E \setminus \{0\}$,

- $g(xy) \geq g(x)$;
- if $x, y \in E$ and $y \neq 0$. The there exist $q, r \in E$ with $a = qb + r$ and either $r = 0$ or $g(r) < g(b)$.

A function g with such properties is a *Euclidean function* [13,17].

The group of *units* $R^\times$ of a ring R is the subset of elements that possess a multiplicative inverse element in the ring. Two elements $a, b \in R$ of an integral domain are *associates* if there exists a unit $u \in R$ such that $a = bu$.

Recall that a number field $\mathbb{K}$ is *quadratic* (or a *quadratic field*) if there is a square-free integer $-d \in \mathbb{Z}$ such that $\mathbb{K} = \mathbb{Q}(\sqrt{-d})$. If, in addition, $d \in \mathbb{N}$, then the field $\mathbb{Q}(\sqrt{-d})$ is an *imaginary quadratic field.* In the work that follows, we denote by $\mathcal{O}_d$ the ring of algebraic integers in $\mathbb{Q}(\sqrt{-d})$. For $d \in \mathbb{Z}$ square-free, the quadratic field $\mathbb{Q}(\sqrt{-d})$ admits a *field norm* $N\colon \mathbb{Q}(\sqrt{d}) \to \mathbb{Q}$ [27]. This norm is multiplicative, so that $N(\alpha\beta) = N(\alpha)N(\beta)$, for algebraic integers α we have $N(\alpha) \in \mathbb{Z}$, and for imaginary quadratic fields we have $N(\alpha) = |\alpha|^2$. For the ring of algebraic integers in a number field, the units are precisely those integers u for which $N(u) = 1$.

The next result precisely characterises that imaginary quadratic fields $\mathbb{Q}(\sqrt{-d})$ whose integer rings $\mathcal{O}_d$ are Euclidean domains.

Lemma 2 ([18, Proposition 8.9]). *The ring of integers $\mathcal{O}_d$ for the imaginary quadratic field $\mathbb{Q}(\sqrt{-d})$ is a Euclidean domain if and only if $d \in \{1, 2, 3, 7, 11\}$. For each such d, the field norm on $\mathbb{Q}(\sqrt{-d})$ is Euclidean.*

2.3 Matrix groups

We shall assume some familiarity with (semi)groups of square matrices such as $\mathrm{GL}(2, R) = \{M \in R^{2\times 2} : \det(M) \in R^\times\}$, the *General Linear group* of 2×2 matrices with entries in the ring R whose inverses also have entries in the ring R, the *Special Linear group* $\mathrm{SL}(2, R) = \{M \in \mathrm{GL}(2, R) : \det(M) = 1\}$, and the quotient group $\mathrm{SL}(2, R)/\{\pm \mathrm{Id}_2\} =: \mathrm{PSL}(2, R)$ called the *Projective Special Linear group.* We are interested in the related Euclidean Bianchi groups $\mathrm{PSL}(2, \mathcal{O}_d)$ with $d = 1, 2, 3, 7, 11$.

Recall that among the imaginary quadratic number rings, only $\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3, \mathcal{O}_7$, and $\mathcal{O}_{11}$ are Euclidean domains (Lemma 2). In the literature $\mathcal{O}_1 = \mathbb{Z}[\mathrm{i}]$ and so $\mathrm{PSL}(2, \mathcal{O}_1)$ $(= \mathrm{PSL}(2, \mathbb{Z}[\mathrm{i}]))$ is the Picard group [14].

Theorem 3 ([15, Theorem 4.3.1]). *The Euclidean Bianchi groups admit the following presentations:*

$$\begin{aligned}
\mathrm{PSL}(2,\mathcal{O}_1) &= \langle a,\ell,t,u \mid a^2=\ell^2=(a\ell)^2=(t\ell)^2 \\
&\qquad\qquad\qquad = (u\ell)^2=(at)^3=(ua\ell)^3=[t,u]=\varepsilon\rangle,\\
\mathrm{PSL}(2,\mathcal{O}_2) &= \langle a,t,u \mid a^2=(at)^3=(u^{-1}aua)^2=[t,u]=\varepsilon\rangle,\\
\mathrm{PSL}(2,\mathcal{O}_3) &= \langle a,\ell,t,u \mid a^2=(at)^3=\ell^3=(a\ell)^2=(ua\ell)^3=[t,u]=\varepsilon\rangle,\\
\mathrm{PSL}(2,\mathcal{O}_7) &= \langle a,t,u \mid a^2=(at)^3=(u^{-1}auat)^2=[t,u]=\varepsilon\rangle, \quad \textit{and}\\
\mathrm{PSL}(2,\mathcal{O}_{11}) &= \langle a,t,u \mid a^2=(at)^3=(u^{-1}auat)^3=[t,u]=\varepsilon\rangle.
\end{aligned}$$

The notation $[t,u] \coloneqq t^{-1}u^{-1}tu$ *is the* commutator *element. The relation* $[t,u]=\varepsilon$ *indicates that* t *and* u *commute, i.e.,* $tu=ut$.

3 Word Representation Procedure

By definition, we associate to each element $m \in \mathrm{PSL}(2,R)$ a set $\{M,-M\}$ of two matrices in $\mathrm{SL}(2,R)$. Henceforth we employ a slight abuse of notation and write $m=\pm M$, or choose either matrix M or $-M$ to represent m. Intuitively, one can take $\mathrm{PSL}(2,R)$ as $\mathrm{SL}(2,R)$ by ignoring the sign. The matrices associated to each of the generators a, t, u, and ℓ in Theorem 3 are:

$$A=\begin{pmatrix}0&-1\\1&0\end{pmatrix}, \quad T=\begin{pmatrix}1&1\\0&1\end{pmatrix}, \quad U=\begin{pmatrix}1&\omega\\0&1\end{pmatrix},$$

and

$$L=\begin{cases}\begin{pmatrix}\mathrm{i}&0\\0&-\mathrm{i}\end{pmatrix} & \text{if } d=1, \text{ and}\\ \begin{pmatrix}\omega^2&0\\0&\omega\end{pmatrix} & \text{if } d=3 \text{ where } \omega=-\frac{1}{2}+\frac{\sqrt{3}\mathrm{i}}{2}.\end{cases}$$

In the above, the top-right entry of U depends on the specific matrix group $\mathrm{PSL}(2,\mathcal{O}_d)$. We give further details in Appendix A.

The following theorem generalises the procedure in [2, Lemma 3.1] that generates word representations for elements of $\mathrm{PSL}(2,\mathbb{Z})$.

Theorem 4. *There is a procedure that, given an element* M *in a Euclidean Bianchi group* $\mathrm{PSL}(2,\mathcal{O}_d)$*, outputs a word representation for* M *of the form*

$$\begin{gathered}(L^{\epsilon}T^{p_0}U^{q_0})AT^{p_k}U^{q_k}AT^{p_{k-1}}U^{q_{k-1}}\cdots AT^{p_1}U^{q_1} \ \textit{if } d\in\{1,3\}, \ \textit{and}\\ (T^{p_0}U^{q_0})AT^{p_k}U^{q_k}AT^{p_{k-1}}U^{q_{k-1}}\cdots AT^{p_1}U^{q_1} \ \textit{if } d\in\{2,7,11\}.\end{gathered}$$

Let $\|M\| \coloneqq \max_{1\le i,j\le 2}|M_{ij}|^2$*. Here* $\epsilon\in\{0,1,2\}$*, the exponent pairs* $p_\ell,q_\ell\in\mathbb{Z}$ *each satisfy* $|p_\ell+q_\ell\omega|^2\le\|M\|$*, and* $k<1-\log_{\kappa(d)}\|M\|$ *where* $\kappa(d)$ *is the Euclidean minimum of* $\mathcal{O}_d$ *(see Appendix A). Moreover, this procedure runs in time polynomial in* $-\log_{\kappa(d)}\|M\|$.

Proof. For the ease of presentation, we give the procedure for elements of the Picard group $\mathrm{PSL}(2,\mathbb{Z}[\mathrm{i}])$ here and relegate the analogous arguments for the remaining Euclidean Bianchi groups to the appendix (Appendix B). There are two parts to our proof: the construction of the word representation and the polynomial runtime.

Construction of the word representation. Suppose that $M = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right) \in \mathrm{PSL}(2,\mathbb{Z}[\mathrm{i}])$. Our first step is to construct an element $H \in \mathrm{PSL}(2,\mathbb{Z}[\mathrm{i}])$ such that MH is upper-triangular (thus reducing the representation problem to that of representing an upper-triangular element). The second step constructs the word representation of an upper-triangular element.

We begin with the first step. In the case that $\gamma = 0$, we choose $H = \mathrm{Id}_2$. We continue under the assumption that $\gamma \neq 0$. We claim that there is an $H_1 \in \mathrm{PSL}(2,\mathbb{Z}[\mathrm{i}])$ such that $MH_1 = \left(\begin{smallmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{smallmatrix}\right)$ where $N(\gamma_1) < N(\gamma)$. (Here N is the field norm on $\mathbb{Q}(\mathrm{i})$.) Since $\mathbb{Z}[\mathrm{i}]$ equipped with the field norm N is a Euclidean domain, we can write $\delta = -\theta_1\gamma + \gamma_1$ with $\theta_1, \gamma_1 \in \mathbb{Z}[\mathrm{i}]$ such that $N(\gamma_1) = N(\theta_1\gamma + \delta) < N(\gamma)$. Let us write $\theta_1 = -p_1 - q_1\mathrm{i} \in \mathbb{Z}[\mathrm{i}]$ in terms of the integral basis $\{1, \mathrm{i}\}$, then

$$\begin{aligned} MT^{-p_1}U^{-q_1}A &= \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^{-p_1}\begin{pmatrix}1 & \mathrm{i}\\ 0 & 1\end{pmatrix}^{-q_1}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\\ &= \begin{pmatrix}\alpha & \theta_1\alpha+\beta\\ \gamma & \theta_1\gamma+\delta\end{pmatrix}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\theta_1\alpha+\beta & -\alpha\\ \theta_1\gamma+\delta & -\gamma\end{pmatrix} =: \begin{pmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{pmatrix}.\end{aligned}$$

Since $N(\gamma_1) < N(\gamma)$, the element $H_1 := U^{-p_1}T^{-q_1}A$ has the claimed properties.

We loop the above construction in order to generate a sequence of matrices of the form $MH_1 \cdots H_\ell$. Let γ_ℓ be the bottom-left entry of matrix $MH_1 \cdots H_\ell$. We repeat this process until we obtain a matrix $MH_1 \cdots H_k$ that satisfies the condition $\gamma_k = 0$.

The following observations guarantee that the above process both terminates and does so correctly. First, $N(\gamma_\ell) = |\gamma_\ell|^2 \in \mathbb{Z}_{\geq 0}$ for each ℓ since $\gamma_\ell \in \mathbb{Z}[\mathrm{i}]$. Second, the sequence of these norms is strictly decreasing (and so $N(\gamma_{\ell+1}) \leq N(\gamma_\ell) - 1$ for $1 \leq \ell \leq k-1$). Hence there exists $k \in \mathbb{N}$ such that $N(\gamma_k) = 0$, from which we deduce that $\gamma_k = 0$. Thus we have constructed an element $H_1 \cdots H_k =: H \in \mathrm{PSL}(2,\mathbb{Z}[\mathrm{i}])$ such that MH is upper-triangular. For our second step, consider

$$MH = \begin{pmatrix}\alpha_k & \beta_k\\ 0 & \delta_k\end{pmatrix} =: \begin{pmatrix}\rho & \sigma\\ 0 & \tau\end{pmatrix}$$

with $\rho, \sigma, \tau \in \mathbb{Z}[\mathrm{i}]$. Since $1 = \det(MH) = \rho\tau$, we deduce that $\rho, \tau \in \mathbb{Z}[\mathrm{i}]^\times = \{\pm 1, \pm \mathrm{i}\}$ and, moreover, $\overline{\tau} = \tau^{-1} = \rho$. The possible pairs (ρ, τ) ensure that

$$MH = \begin{pmatrix}\mathrm{i} & 0\\ 0 & -\mathrm{i}\end{pmatrix}^\epsilon \begin{pmatrix}1 & \sigma'\\ 0 & 1\end{pmatrix} = \begin{pmatrix}\mathrm{i} & 0\\ 0 & -\mathrm{i}\end{pmatrix}^\epsilon \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^{p_0}\begin{pmatrix}1 & \mathrm{i}\\ 0 & 1\end{pmatrix}^{q_0} = L^\epsilon T^{p_0} U^{q_0}$$

for some $\epsilon \in \{0,1\}$ and $\sigma' = p_0 + q_0\mathrm{i} \in \mathbb{Z}[\mathrm{i}]$ (an associate of σ).

Taken together, these two procedural steps construct representations for elements $H \in \mathrm{PSL}(2, \mathbb{Z}[\mathrm{i}])$ and $L^{\epsilon}T^{p_0}U^{q_0}$ for which $M = L^{\epsilon}T^{p_0}U^{q_0}H^{-1}$. Using the above products, we can output a word representation of the desired form for a given M.

Polynomial runtime. It is useful to introduce the following notations. For $1 \le \ell \le k$, we shall write

$$M_\ell := MH_1 \cdots H_\ell = \begin{pmatrix} \theta_\ell\alpha_{\ell-1} + \beta_{\ell-1} & -\alpha_{\ell-1} \\ \theta_\ell\gamma_{\ell-1} + \delta_{\ell-1} & -\gamma_{\ell-1} \end{pmatrix} =: \begin{pmatrix} \alpha_\ell & \beta_\ell \\ \gamma_\ell & \delta_\ell \end{pmatrix}. \tag{1}$$

The update matrix M_ℓ is obtained by performing a Euclidean division in the Gaussian integers $\mathcal{O}_1 = \mathbb{Z}[\mathrm{i}]$. Recall that a division $a = qb + r$ in the Gaussian integers satisfies $N(r) < N(b)$ where N is the associated field norm. In fact, working in $\mathbb{Z}[\mathrm{i}]$ we have the tighter upper bound $N(r) < \frac{1}{2}N(b)$ since $\frac{1}{2}$ is the Euclidean minimum of $\mathbb{Z}[\mathrm{i}]$. We give further background on Euclidean minima in Appendix A.

For $\ell < k$, repeated application of the Euclidean minimum gives

$$1 \le N(\gamma_\ell) < \tfrac{1}{2}N(\gamma_{\ell-1}) < \cdots < \tfrac{1}{2^{\ell-1}}N(\gamma_1) < \tfrac{1}{2^\ell}N(\gamma) = \tfrac{1}{2^\ell}|\gamma|^2 \le \tfrac{1}{2^\ell}\|M\|.$$

Taking logarithms, we find the upper bound $\log_2 \|M\| > \ell$, from which the desired inequality $1 + \log_2 \|M\| > k$ follows.

We now exhibit bounds on the matrix norms $\langle \|M_\ell\| \rangle_{\ell=1}^{k}$. The determinant condition on M_ℓ tells us that

$$\begin{aligned} \alpha_\ell\delta_\ell - \beta_\ell\gamma_\ell = -(\theta_\ell\alpha_{\ell-1} + \beta_{\ell-1})\gamma_{\ell-1} + (\theta_\ell\gamma_{\ell-1} + \delta_{\ell-1})\alpha_{\ell-1} \\ = -(\theta_\ell\alpha_{\ell-1} + \beta_{\ell-1})\gamma_{\ell-1} + \gamma_\ell\alpha_{\ell-1} = 1. \end{aligned}$$

For $\ell - 1 < k$, we have that $\gamma_{\ell-1} \neq 0$ and so

$$\begin{aligned} |\alpha_\ell|^2 = N(\alpha_\ell) = N(\theta_\ell\alpha_{\ell-1} + \beta_\ell) = N\left(\frac{\gamma_\ell\alpha_{\ell-1} - 1}{\gamma_{\ell-1}}\right) \le \\ \frac{N(\gamma_\ell)N(\alpha_{\ell-1})}{N(\gamma_{\ell-1})} + \frac{1}{N(\gamma_{\ell-1})} \le \frac{1}{2}N(\alpha_{\ell-1}) + 1. \end{aligned} \tag{2}$$

Here the rightmost inequality follows from the division $-\delta_{\ell-1} = \theta_\ell\gamma_{\ell-1} + \gamma_\ell$.

We make the following claim.

Claim 5. *For $M_{\ell-1}$ and M_ℓ in $\mathrm{PSL}(2, \mathbb{Z}[\mathrm{i}])$ as above, we necessarily have that $\|M_\ell\| \le \|M_{\ell-1}\|$.*

All that remains is to bound the sizes of the integer exponents in the word representation. For $\ell \le k$, we bound the quotients $\theta_\ell := -p_\ell - q_\ell\mathrm{i}$ as follows,

$$\begin{aligned} |p_\ell|^2 + |q_\ell|^2 = N(\theta_\ell) = N\left(\frac{\delta_{\ell-1} + \gamma_\ell}{\gamma_{\ell-1}}\right) \\ \le N(\delta_{l-1}) + \frac{N(\gamma_\ell)}{N(\gamma_{\ell-1})} \le \|M_{\ell-1}\| + \tfrac{1}{2} \le \|M\| + \tfrac{1}{2}. \end{aligned}$$

Since $N(\theta_\ell), \|M\| \in \mathbb{Z}$, we have $N(\theta_\ell) \leq \|M\|$ and so $|p_\ell|^2, |q_\ell|^2 \leq \|M\|$. Observe that the rightmost inequality follows from Claim 5. We consider the matrix MH in order to bound p_0 and q_0. Once again, it is clear that

$$|p_0|^2, |q_0|^2 \leq |\sigma|^2 \leq \max_{1\leq i,j\leq 2} |(MH)_{ij}|^2 = \|M_k\| \leq \|M\|.$$

The above observations bound the number of iterations in the initial looping procedure as well as the sizes of the exponents in the word representation. Taken together with standard results on the division algorithm, we obtain the stated polynomial runtime. □

Proof (Proof of Claim 5). By (1), we have $|\beta_\ell|^2, |\delta_\ell|^2, |\gamma_\ell|^2 \leq \|M_{\ell-1}\|$. Thus $\|M_\ell\| \leq \|M_{\ell-1}\|$ unless $|\alpha_\ell|^2 > \|M_{\ell-1}\|$. Let us assume, for a contradiction, that $|\alpha_\ell|^2 > \|M_{\ell-1}\|$. By (2), it follows that

$$\|M_{\ell-1}\| < \|M_\ell\| = N(\alpha_\ell) \leq \frac{1}{2}N(\alpha_{\ell-1}) + 1 < \frac{1}{2}\|M_{\ell-1}\| + 1,$$

and so we deduce that $\|M_{\ell-1}\| < 2$. Thus $(M_{\ell-1})_{ij} \in \{0, \pm1, \pm\mathrm{i}\}$ for each (i,j).

We first argue that it is not possible that each entry of $M_{\ell-1}$ is non-zero; for otherwise, each entry is a unit and the determinant condition $\alpha_{\ell-1}\delta_{\ell-1} - \beta_{\ell-1}\gamma_{\ell-1} = 1$ is not satisfied by any tuple of units in $\mathbb{Z}[\mathrm{i}]^\times$. The determinant condition also ensures that $M_{\ell-1} \neq 0_{2\times2}$. Thus $M_{\ell-1}$ takes one of the following forms

$$\begin{pmatrix} 0 & \beta_{\ell-1} \\ \gamma_{\ell-1} & \delta_{\ell-1} \end{pmatrix}, \begin{pmatrix} 0 & \beta_{\ell-1} \\ \gamma_{\ell-1} & 0 \end{pmatrix}, \begin{pmatrix} \alpha_{\ell-1} & 0 \\ \gamma_{\ell-1} & \delta_{\ell-1} \end{pmatrix}, \text{ or } \begin{pmatrix} \alpha_{\ell-1} & \beta_{\ell-1} \\ \gamma_{\ell-1} & 0 \end{pmatrix}$$

where an entry not denoted by 0 is a unit (i.e., in $\{\pm1, \pm\mathrm{i}\} = \mathbb{Z}[\mathrm{i}]^\times$). In the first two forms, where $\alpha_{\ell-1} = 0$, it is clear that the update (1) ensures that $|\alpha_\ell|^2 = |\beta_{\ell-1}|^2 > 0$ and so the required inequality trivially holds. Likewise, the inequality holds for the third form since, by the update (1), $|\alpha_\ell|^2 = |\alpha_{\ell-1}|^2$. All that remains is to treat the fourth form. The update to the fourth form gives $\left(\begin{smallmatrix} \alpha_\ell & \beta_\ell \\ \gamma_\ell & \delta_\ell \end{smallmatrix}\right) = \left(\begin{smallmatrix} \beta_{\ell-1} & -\alpha_{\ell-1} \\ 0 & -\gamma_{\ell-1} \end{smallmatrix}\right)$ because $0 = \theta_\ell\gamma_{\ell-1} + \gamma_\ell = \theta_\ell\gamma_{\ell-1}$ implies that $\theta_\ell = 0$. Thus $|\alpha_\ell|^2 = |\beta_{\ell-1}|^2 = 1 = |\alpha_{\ell-1}|^2$. We conclude that the required inequality holds. □

As seen in [2], a procedure that generates a word representation of an element in $\mathrm{PSL}(2,\mathbb{Z})$ implies a procedures that generates the word representation of an element in $\mathrm{SL}(2,\mathbb{Z})$. In a similar vein, we have the following corollary.

Corollary 6. *For each of the groups* $\mathrm{SL}(2,\mathcal{O}_d)$ *with* $d \in \{1,2,3,7,11\}$*, there is a procedure that, given an element* $M \in \mathrm{SL}(2,\mathcal{O}_d)$ *outputs a word representation for* M *in terms of the generators of* $\mathrm{SL}(2,\mathcal{O}_d)$*.*

4 Word Embeddings into Triangular Matrix Groups

Let us consider two known group embeddings from the literature (cf. [4,7]).

Proposition 7. *Let $\Sigma_2 = \{a, b\}$. Then $\varphi : \mathrm{F}(\Sigma_2) \to \mathbb{Z}^{2\times 2}$ defined by*

$$\varphi(a) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad \varphi(b) = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

is an embedding.

Proposition 8. *Let $\Sigma_2 = \{a, b\}$. Then $\varphi : \mathrm{F}(\Sigma_2) \to \mathbb{C}^{2\times 2}$ defined by*

$$\varphi(a) = \begin{pmatrix} \frac{3}{5} + \frac{4}{5}\mathrm{i} & 0 \\ 0 & \frac{3}{5} - \frac{4}{5}\mathrm{i} \end{pmatrix}, \quad \varphi(b) = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

is an embedding.

Observe that the two embeddings presented in Propositions 7 and 8 use all four entries of a two-by-two matrix. We show that there is no such embedding into upper triangular matrices of any dimension.

Theorem 9. *There does not exist an embedding $\sigma : \mathrm{F}(\Sigma_2) \to \mathrm{U}(\mathbb{C}, n)$ for any n.*

Let G be a group. The *lower central series* $G = \gamma_1 G \trianglerighteq \gamma_2 G \trianglerighteq \cdots \trianglerighteq \gamma_k G \trianglerighteq \cdots$ is a sequence of subgroups of G whereby $\gamma_1 G \coloneqq G$ and $\gamma_{k+1} G \coloneqq [\gamma_k G, G]$. We recognise $\gamma_2 G = [G, G]$ as the *commutator* or *derived subgroup* of G. The structure of the lower central series of free groups is discussed in [16,22]. A group G is *nilpotent* if there is a $k \geq 1$ for which $\gamma_k(G) = \{1\}$.

Proposition 10. *There is no embedding of $\mathrm{F}(\Sigma_2)$ into a nilpotent group.*

Proof. Let G be a nilpotent group. Suppose, for a contradiction, there exists an embedding $\varphi \colon \mathrm{F}(\Sigma_2) \to G$.

Since $\mathrm{F}(\Sigma_2)$ is not nilpotent, for each $k \geq 1$ there is a non-trivial $x \in \gamma_k(\mathrm{F}(\Sigma_2))$. Additionally, as φ is a homomorphism it commutes with the commutator brackets (i.e., $\varphi([g, h]) = [\varphi(g), \varphi(h)]$ for all $g, h \in \mathrm{F}(\Sigma_2)$). Thus it is clear that $\varphi(x) \in \gamma_k(G)$. We have arrived at a contradiction. As G is nilpotent, there is a $k \geq 1$ and non-trivial element $x \in \gamma_k(\mathrm{F}(\Sigma_2))$ such that $\varphi(x) = 1$, which contradicts the injectivity of the embedding. □

For $n \geq 1$, the group of $n \times n$ upper unitriangular matrices $\mathrm{UT}(n, \mathbb{C})$ is nilpotent. Thus the following corollary is immediate.

Corollary 11. *For $n \geq 1$, There is no embedding of $\mathrm{F}(\Sigma_2)$ into $\mathrm{UT}(n, \mathbb{C})$.*

A group G is *nilpotent-by-abelian* if i) G possesses a normal subgroup N that is nilpotent and ii) the quotient group G/N is abelian. Lemma 12 is well known (cf. [26, Theorem 2.23]) and Corollary 13 follows immediately.

Lemma 12. *The commutator subgroup $\gamma_2 G$ of a group G is a normal subgroup. Suppose that $N \triangleleft G$ is a normal subgroup of G. Then G/N is abelian if and only if $\gamma_2 G \subseteq N$. Further, $G/\gamma_2 G$ is abelian.*

Corollary 13. *If G is nilpotent-by-abelian, then $\gamma_2 G = [G, G]$ is nilpotent.*

Proposition 14 generalises the result in Proposition 10.

Proposition 14. *There is no embedding of $\mathrm{F}(\Sigma_2)$ into a nilpotent-by-abelian group.*

Proof. Let G be a nilpotent-by-abelian group. We suppose, for a contradiction, that there exists an embedding $\varphi\colon\ \mathrm{F}(\Sigma_2) \to G$.

Recall that $[\mathrm{F}(\Sigma_2), \mathrm{F}(\Sigma_2)]$ is a free group of infinite rank (cf. [26, Theorem 11.48]). Since $[\mathrm{F}(\Sigma_2), \mathrm{F}(\Sigma_2)]$ is not nilpotent, for each $k \geq 1$ there is a non-trivial element $x \in \gamma_k([\mathrm{F}(\Sigma_2), \mathrm{F}(\Sigma_2)])$. Moreover, as φ is a homomorphism it commutes with the commutator brackets. Thus it is clear that $\varphi(x) \in \gamma_k([G, G])$. From our assumption that G is nilpotent-by-abelian and Corollary 13, $[G, G]$ is nilpotent. It follows that there is a $k \geq 1$ for which $\gamma_k([G, G]) = \{1\}$. Thus we arrive at a contradiction: there is a non-trivial element $x \in \mathrm{F}(\Sigma_2)$ for which $\varphi(x) = 1$. □

Theorem 9 follows straightforwardly as a corollary to Proposition 14. Indeed, recall that $[\mathrm{U}(n, \mathbb{C}), \mathrm{U}(n, \mathbb{C})] = \mathrm{UT}(n, \mathbb{C})$ and so, by Lemma 12, $\mathrm{U}(n, \mathbb{C})$ is nilpotent-by-abelian. Thus the desired result follows immediately from Proposition 14.

5 Word Embeddings Into Low-Dimensional Matrix Rings

In this section discuss word embeddings into matrix rings $R^{n\times n}$ where R is a ring and n is the order (sometimes *size*) of each matrix. Motivated by the non-existence of an embedding from a binary free group into upper-triangular matrices, we investigate other settings where embeddings exist or can be proven to not exist. Recall, that Paterson's classical embedding of pairs or binary semigroup words into $\mathbb{N}^{3\times 3}$ [24] is contrasted by non-existence of an embedding into $\mathbb{C}^{2\times 2}$ [7]. The next two propositions provide a summarise of the results. For the rest of the section, we assume that $\Sigma_2 = \{a, b\}$ and $\Sigma_1 = \{c\}$.

Proposition 15.

- *There exists an embedding from $\mathrm{F}(\Sigma_1) \times \mathrm{F}(\Sigma_1)$ into $\mathrm{U}(2, \mathbb{N})$.*
- *There exists an embedding from $\mathrm{F}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ into $\mathbb{Q}^{2\times 2}$.*
- *There exists an embedding from $\mathrm{S}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ into $\mathbb{Z}^{3\times 3}$.*
- *There exists an embedding from $\mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_1)$ into $\mathbb{Z}^{2\times 2}$.*

We then turn our attention to non-existence results in similar settings using linear algebraic properties of the matrices being mapped to.

Proposition 16. *Let $d \in \{1, 2, 3, 7, 11\}$.*

- *There is no embedding from $\varphi_1 : \mathrm{S}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ into $\mathcal{O}_d^{2\times 2}$.*
- *There is no embedding from $\varphi_2 : \mathrm{F}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ into $\mathcal{O}_d^{2\times 2}$.*
- *There is no embedding from $\varphi_3 : \mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_2)$ into $\mathcal{O}_d^{3\times 3}$.*
- *There is no embedding from $\varphi_4 : \mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_1)$ into $\mathrm{SL}(2, \mathcal{O}_d)$.*

Proof. Let us prove the first and second claims. Proofs of the other two claims use similar approaches that were used in [7,20]. Additional details can be found in Appendix C.

Let $\Sigma_2 = \{a, b\}$ and $\Sigma_1 = \{c\}$. Assume to the contrary that such an embedding exists. Now, $\varphi : \mathrm{S}(\Sigma_2) \times \mathrm{F}(\Sigma_1) \to \mathcal{O}_d^{2\times 2}$ maps the generators as follows:

$$(a,\varepsilon) \mapsto \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \quad (b,\varepsilon) \mapsto \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix},$$
$$(\varepsilon,c) \mapsto \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}, \quad (\varepsilon,c^{-1}) \mapsto \frac{1}{c_1c_4 - c_2c_3}\begin{pmatrix} c_4 & -c_2 \\ -c_3 & c_1 \end{pmatrix}.$$

The last mapping implies that $c_1c_4 - c_2c_3 \in \mathcal{O}_d^{\times}$ as those are the only elements whose inverse is in $\mathcal{O}_d$. Furthermore, we have the following relations:

$$\begin{aligned} \varphi((a,\varepsilon)(b,\varepsilon)) &\neq \varphi((b,\varepsilon)(a,\varepsilon)) \\ \varphi((a,\varepsilon)(\varepsilon,c)) &= \varphi((\varepsilon,c)(a,\varepsilon)) \\ \varphi((b,\varepsilon)(\varepsilon,c)) &= \varphi((\varepsilon,c)(b,\varepsilon)). \end{aligned} \tag{3}$$

Let us consider products $\varphi((a,\varepsilon)(\varepsilon,c))$ and $\varphi((\varepsilon,c)(a,\varepsilon))$:

$$\varphi((a,\varepsilon)(\varepsilon,c)) = \begin{pmatrix} a_1c_1 + a_2c_3 & a_1c_2 + a_2c_4 \\ a_3c_1 + a_4c_3 & a_3c_2 + a_4c_4 \end{pmatrix} \tag{4}$$

$$\varphi((\varepsilon,c)(a,\varepsilon)) = \begin{pmatrix} a_1c_1 + a_3c_2 & a_2c_1 + a_4c_2 \\ a_1c_3 + a_3c_4 & a_2c_3 + a_4c_4 \end{pmatrix}. \tag{5}$$

In order to derive a contradiction, let us assume first that $c_2 = 0$. Now, the top-right corner elements in $\varphi((a,\varepsilon)(\varepsilon,c))$ and $\varphi((\varepsilon,c)(a,\varepsilon))$ are a_2c_4 and a_2c_1, respectively. Because the two matrices are equal by (3), we observe that $c_1 = c_4$. It follows that the bottom-left corner elements are $a_3c_1 + a_4c_3$ and $a_1c_3 + a_3c_1$. Again, due to the relation in (3), $a_1 = a_4$. By the same logic $b_1 = b_4$ follows from the relation $\varphi((b,\varepsilon)(\varepsilon,c)) = \varphi((\varepsilon,c)(b,\varepsilon))$.

Now, the bottom-right corner elements in (4) and (5) are a_1c_1 and $a_2c_3 + a_1c_1$. Recall that these are equal, which implies that either $a_2 = 0$ (and analogously $b_2 = 0$) or $c_3 = 0$. In the latter case, $\varphi(\varepsilon,c) = \left(\begin{smallmatrix} c_1 & 0 \\ 0 & c_1 \end{smallmatrix}\right)$ and $\varphi(\varepsilon,c^{-1}) = \frac{1}{c_1^2}\left(\begin{smallmatrix} c_1 & 0 \\ 0 & c_1 \end{smallmatrix}\right)$. This implies that $c_1 \in \mathcal{O}_d^{\times}$. Recall that all elements of each $\mathcal{O}_d^{\times}$ are given in Lemma 18. It is straightforward to see that $\varphi(\varepsilon,c)^k = \mathrm{Id}_2$ for some k, which contradicts the fact that φ is an embedding. As an example, consider $d = 3$ and the unit $\omega = \frac{1}{2} + \frac{\sqrt{-3}}{2}$. In this case, $\varphi(\varepsilon,c)^6 = \left(\begin{smallmatrix} \omega & 0 \\ 0 & \omega \end{smallmatrix}\right)^6 = \mathrm{Id}_2$. That is, $c_3 \neq 0$ and we now assume that $a_2 = 0$ and $b_2 = 0$. However, if we consider $\varphi((a,\varepsilon)(b,\varepsilon))$ and $\varphi((b,\varepsilon)(a,\varepsilon))$, we observe that they commute. Indeed,

$$\varphi((a,\varepsilon)(b,\varepsilon)) = \begin{pmatrix} a_1b_1 & 0 \\ a_3b_1 + a_1b_3 & a_1b_1 \end{pmatrix}, \quad \varphi((b,\varepsilon)(a,\varepsilon)) = \begin{pmatrix} a_1b_1 & 0 \\ a_1b_3 + a_3b_1 & a_1b_1 \end{pmatrix}.$$

Analogously, we can show that matrices, $\varphi((a,\varepsilon)(b,\varepsilon))$ and $\varphi((b,\varepsilon)(a,\varepsilon))$ commute if we assume that $c_3 = 0$ instead of our assumption on c_2 in (4) and (5).

Observe that in equations (4) and (5), the top-left corner elements are equal if and only if $a_2c_3 = a_3c_2$. We have shown that both c_2 and c_3 are non-zero. We observe that a_2 and a_3 (and b_2 and b_3) are also non-zero. Indeed, if one is then both are, and furthermore

$$\varphi((a,\varepsilon)(b,\varepsilon)) = \begin{pmatrix} a_1b_1 & a_1b_2 \\ a_4b_3 & a_4b_4 \end{pmatrix}, \qquad \varphi((b,\varepsilon)(a,\varepsilon)) = \begin{pmatrix} a_1b_1 & a_4b_2 \\ a_1b_3 & a_4b_4 \end{pmatrix} \text{ and}$$

$$\varphi((a,\varepsilon)(\varepsilon,c)) = \begin{pmatrix} a_1c_1 & a_1c_2 \\ a_4c_3 & a_4c_4 \end{pmatrix}, \qquad \varphi((\varepsilon,c)(a,\varepsilon)) = \begin{pmatrix} a_1c_1 & a_4c_2 \\ a_1c_3 & a_4c_4 \end{pmatrix}.$$

By relations (3), the two matrices on the first line should be different while the two matrices on the second line should be the same. For the first two matrices to be unequal, $a_1 \neq a_4$ has to hold, but for the last two matrices to be equal $a_1 = a_4$ has to hold.

We finally observe that the top-left corners of the products imply that $a_2c_3 = a_3c_2$, $b_2c_3 = b_3c_2$ and $a_2b_3 \neq a_3b_2$. As a_2, a_3, b_2, b_3, c_2 and c_3 are all non-zero, we have $\frac{a_2}{a_3} = \frac{c_2}{c_3} = \frac{b_2}{b_3}$ which contradicts the inequality $a_2b_3 \neq a_3b_2$.

The second claim follows directly: if there is no embedding from a semigroup alphabet, then there is also no embedding from a group alphabet. □

Remark 1 *In the proof of Proposition 16, we use the fact that the unit group $\mathcal{O}_d^\times$ is finite. Mutatis Mutandis, the same non-existence results hold for any integer ring with finite unit group. By Dirichlet's unit theorem (cf. [23, Chapter 1, Theorem 7.4]), a number field $\mathbb{K}$ has only finitely many unit integers if and only if $\mathbb{K} = \mathbb{Q}$ or $\mathbb{K} = \mathbb{Q}(\sqrt{-d})$ for square-free $d \geq 1$. Thus the non-existence statements in Proposition 16 for embeddings into $\mathcal{O}_d^{2\times 2}$ with $d \in \{1,2,3,7,11\}$ also hold for the codomains $\mathbb{Z}^{2\times 2}$ and $\mathcal{O}_d^{2\times 2}$ for square-free $d \geq 1$.*

6 Conclusion and Future Directions

In Section 3, we presented the first step towards showing the decidability of the identity problem for matrices over $\mathrm{SL}(2,\mathcal{O}_d)$. Indeed, in [2], the word representation algorithm for $\mathrm{SL}(2,\mathbb{Z})$ was used to construct finite petal graphs where an existence of a short path can be checked in NP. The construction utilised a complete, confluent and monadic term rewriting system for the generators of $\mathrm{SL}(2,\mathbb{Z})$. No such system is known for Bianchi groups. One can obtain either a confluent or a monadic rewriting system, but this is not sufficient for the construction as the finiteness of the petal graph is no longer guaranteed. A pertinent question is whether there exists a confluent and monadic rewriting system for $\mathrm{SL}(2,\mathcal{O}_d)$ or whether one can ensure that the resulting petal graph is finite.

In Sections 4 and 5, we considered various matrix semigroups and showed that there is no way to embed pairs of words over different alphabets. It is interesting to ask, how narrow is the boundary between existence and non-existence? In particular, when considering $\mathbb{K}^{n\times n}$, what is the smallest n and the largest $\mathbb{K}$ such that the embedding no longer exists. It is also worth noting that we considered embeddings of pairs of words. One can extend the problem to the k-fold products, i.e., $\mathrm{S}(\Sigma)^k$ and $\mathrm{F}(\Sigma)^k$; see [6].

References

1. Bayer Fluckiger, E.: Upper bounds for Euclidean minima of algebraic number fields. Journal of Number Theory **121**(2), 305–323 (2006). https://doi.org/10.1016/j.jnt.2006.03.002
2. Bell, P.C., Hirvensalo, M., Potapov, I.: The Identity Problem for Matrix Semigroups in $\mathrm{SL}_2(\mathbb{Z})$ is NP-complete. In: Klein, P.N. (ed.) Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19. pp. 187–206. SIAM (2017). https://doi.org/10.1137/1.9781611974782.13
3. Bell, P.C., Niskanen, R., Potapov, I., Semukhin, P.: On the Identity and Group Problems for Complex Heisenberg Matrices. In: Bournez, O., Formenti, E., Potapov, I. (eds.) Reachability Problems - 17th International Conference, RP 2023, Nice, France, October 11-13, 2023, Proceedings. Lecture Notes in Computer Science, vol. 14235, pp. 42–55. Springer (2023). https://doi.org/10.1007/978-3-031-45286-4_4
4. Bell, P.C., Potapov, I.: Reachability problems in quaternion matrix and rotation semigroups. Information and Computation **206**(11), 1353–1361 (2008). https://doi.org/10.1016/j.ic.2008.06.004
5. Button, J.O.: Groups and Embeddings in $\mathrm{SL}(2, \mathbb{C})$. Communications in Algebra **44**(1), 265–278 (2016). https://doi.org/10.1080/00927872.2014.975347
6. Campagnolo, C., Kammeyer, H.: Products of free groups in Lie groups. Journal of Algebra **579**, 237–255 (2021). https://doi.org/10.1016/j.jalgebra.2021.03.023
7. Cassaigne, J., Harju, T., Karhumäki, J.: On the undecidability of freeness of matrix semigroups. International Journal of Algebra and Computation **9**(03n04), 295–305 (1999). https://doi.org/10.1142/S0218196799000199
8. Cassaigne, J., Nicolas, F.: On the decidability of semigroup freeness. RAIRO Theor. Informatics Appl. **46**(3), 355–399 (2012). https://doi.org/10.1051/ITA/2012010
9. Cohen, H.: A course in computational algebraic number theory, vol. 138. Springer Science & Business Media (2013)
10. Conway, J.H., Smith, D.A.: On Quaternions and Octonions. AK Peters/CRC Press (2003)
11. Dong, R.: Semigroup algorithmic problems in metabelian groups. In: Mohar, B., Shinkar, I., O'Donnell, R. (eds.) Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024. pp. 884–891. ACM (2024). https://doi.org/10.1145/3618260.3649609
12. Dong, R.: Semigroup Intersection Problems in the Heisenberg Groups. SIAM J. Discret. Math. **38**(4), 3176–3197 (2024). https://doi.org/10.1137/23M1581467
13. Dubois, D.W., Steger, A.: A note on division algorithms in imaginary quadratric number fields. Canadian J. Math. **10**, 285–286 (1958). https://doi.org/10.4153/CJM-1958-030-3
14. Fine, B.: The Picard group and the modular group, p. 150–163. London Mathematical Society Lecture Note Series, Cambridge University Press (1987)
15. Fine, B.: Algebraic theory of the Bianchi groups, Monographs and Textbooks in Pure and Applied Mathematics, vol. 129. Marcel Dekker, Inc., New York (1989)
16. Hall, Jr., M.: The theory of groups. The Macmillan Company, New York (1959)
17. Hardy, G.H., Wright, E.M.: An introduction to the theory of numbers. Oxford University Press, Oxford, sixth edn. (2008), revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles
18. Hatcher, A.: Topology of numbers. American Mathematical Society, Providence, RI (2022)

19. Kenison, G.: georgekenison/EuclidBianchiWordRep: EuclideanBianchiWordRepresentation v1.0 (Mar 2026). https://doi.org/10.5281/zenodo.18939459
20. Ko, S.K., Niskanen, R., Potapov, I.: On the identity problem for the special linear group and the Heisenberg group. In: Proceedings of ICALP 2018. LIPIcs, vol. 107, pp. 132:1–132:15 (2018). https://doi.org/10.4230/lipics.icalp.2018.132
21. Lemmermeyer, F.: The Euclidean algorithm in algebraic number fields. Expositiones Mathematicae **13**, 385–416 (1 1995)
22. Magnus, W., Karrass, A., Solitar, D.: Combinatorial group theory: Presentations of groups in terms of generators and relations. Interscience Publishers [John Wiley & Sons], New York-London-Sydney (1966)
23. Neukirch, J.: Algebraic number theory, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 322. Springer-Verlag, Berlin (1999). https://doi.org/10.1007/978-3-662-03983-0
24. Paterson, M.S.: Unsolvability in 3×3 matrices. Studies in Applied Mathematics **49**(1), 105 (1970). https://doi.org/10.1002/sapm1970491105
25. Potapov, I., Semukhin, P.: Vector and scalar reachability problems in $\mathrm{SL}(2,\mathbb{Z})$. Journal of Computer and System Sciences **100**, 30–43 (03 2019). https://doi.org/10.1016/j.jcss.2018.09.003
26. Rotman, J.J.: An Introduction to the Theory of Groups. Springer New York (1995). https://doi.org/10.1007/978-1-4612-4176-8
27. Stewart, I., Tall, D.: Algebraic number theory and Fermat's last theorem. CRC Press, Boca Raton, FL, fourth edn. (2016)
28. The Sage Developers: SageMath, the Sage Mathematics Software System (Version 10.7) (2026), `https://www.sagemath.org`

A Background Material on Imaginary Quadratic Fields

We have the following characterisation of the ring of algebraic integers in the imaginary quadratic field $\mathbb{Q}(\sqrt{-d})$ [27, Chapter 3].

Lemma 17. *Suppose that $d \in \mathbb{Z}$ is square-free. The subring of algebraic integers $\mathcal{O}_d$ in the quadratic field $\mathbb{Q}(\sqrt{-d})$ has the form $\mathcal{O}_d = \mathbb{Z}[\omega] \coloneqq \{x + y\omega \mid x, y \in \mathbb{Z}\}$ where*

$$\omega = \begin{cases} \sqrt{-d} & \text{if } -d \equiv 2,3 \pmod 4, \text{ and} \\ (1+\sqrt{-d})/2 & \text{if } -d \equiv 1 \pmod 4. \end{cases}$$

Let $x + y\omega \in \mathcal{O}_d$ be an integer in $\mathbb{Q}(\sqrt{d})$, then

$$N(x+y\omega) = \begin{cases} x^2 + y^2 d & \text{if } -d \equiv 2,3 \pmod 4, \\ x^2 + xy + y^2\left(\frac{1+d}{4}\right) & \text{if } -d \equiv 1 \pmod 4. \end{cases}$$

Lemma 18. *Suppose that $d \in \mathbb{N}$ is square-free. Then the group of units $\mathcal{O}_d^\times$ in the ring of imaginary quadratic integers $\mathcal{O}_d$ is finite. Further,*

$$\mathcal{O}_d^\times = \begin{cases} \{\pm 1, \pm \mathrm{i}\} & \text{if } \; d = 1, \\ \{\pm 1, \pm\frac{1}{2} \pm \frac{\sqrt{-3}}{2}, \pm\frac{1}{2} \mp \frac{\sqrt{-3}}{2}\} & \text{if } \; d = 3, \\ \{\pm 1\} & \text{if } \; d = 2, 7, 11. \end{cases}$$

Recall the classical Euclidean algorithm for elements $a, b \in \mathbb{Z}$ allows one to choose $q, r \in \mathbb{Z}$ such that $a = qb + r$ with $|r| < |b|/2$. The following definition abstracts the contraction factor of $1/2$ on the absolute value of the remainder to the setting of algebraic integers (although we limit ourselves to quadratic integer rings). Given a field $\mathbb{Q}(\sqrt{-d})$, the corresponding ring of integers $\mathcal{O}_d$, and field norm N as above, we define the *Euclidean minimum* [21] of $\mathbb{Q}(\sqrt{-d})$ by

$$\kappa(d) \coloneqq \sup_{\alpha \in \mathbb{Q}(\sqrt{-d})} \inf_{\beta \in \mathcal{O}_d} |N(\alpha - \beta)|.$$

The connection to the Euclidean algorithm is made clear by the equivalent definition

$$\kappa(d) = \inf\{\kappa > 0 \mid \forall \alpha, \beta \in \mathcal{O}_d \backslash \{0\},\ \exists q \in \mathcal{O}_d \text{ such that } N(\alpha - q\beta) < \kappa N(\beta)\}.$$

Surveys on Euclidean minima in the general algebraic setting are given in [1,21]. For $\mathcal{O}_d$ to be a Euclidean domain, it is sufficient for $\kappa(d) < 1$, cf. [10, §2.3–2.4] and [18, Chapter 8]). The minima for the Euclidean domains we discuss are given below.

Proposition 19 ([21, Proposition 4.2]). *For d and $\mathcal{O}_d$ as above,*

$$\kappa(d) = \begin{cases} \dfrac{d+1}{4} & \text{if } d = 1, 2, \text{ and} \\ \dfrac{(d+1)^2}{16d} & \text{if } d = 3, 7, 11. \end{cases}$$

The minima for the five Euclidean imaginary quadratic integer rings are given in Table 2.

d	basis element ω	$\kappa(d)$	$\frac{1}{1-\kappa(d)}$	$\{x \in \mathcal{O}_d : N(x) < \frac{1}{1-\kappa(d)}\}$
1	i	$\frac{1}{2}$	2	$\{0, \pm 1, \pm \mathrm{i}\}$
2	$\sqrt{\text{-}2}$	$\frac{3}{4}$	4	$\{0, \pm 1, \pm\sqrt{-2}, -1 \pm \sqrt{-2}, 1 \pm \sqrt{-2}\}$
3	$\frac{1}{2} + \frac{\sqrt{-3}}{2}$	$\frac{1}{3}$	$\frac{3}{2}$	$\{0, \pm 1, \pm\omega, \pm\overline{\omega}\}$
7	$\frac{1}{2} + \frac{\sqrt{-7}}{2}$	$\frac{4}{7}$	$\frac{7}{3}$	$\{0, \pm 1, \pm\omega, \pm\overline{\omega}\}$
11	$\frac{1}{2} + \frac{\sqrt{-11}}{2}$	$\frac{9}{11}$	$\frac{11}{2}$	$\{0, \pm 1, \pm 2, \pm\omega, -1 \pm \omega, 1 \pm \omega, \pm(\omega - 2)\}$

Table 2. Data for the integer rings $\mathcal{O}_d$ relevant to the Euclidean Bianchi groups $\mathrm{PSL}(2, \mathcal{O}_d)$.

B Word Representation Procedure for the Euclidean Bianchi Groups

Recall the main result in Section 3.

Theorem 4. *There is a procedure that, given an element M in a Euclidean Bianchi group* $\mathrm{PSL}(2, \mathcal{O}_d)$*, outputs a word representation for M of the form*

$$(L^{\epsilon}T^{p_0}U^{q_0})AT^{p_k}U^{q_k}AT^{p_{k-1}}U^{q_{k-1}}\cdots AT^{p_1}U^{q_1} \text{ if } d \in \{1,3\}, \text{ and}$$
$$(T^{p_0}U^{q_0})AT^{p_k}U^{q_k}AT^{p_{k-1}}U^{q_{k-1}}\cdots AT^{p_1}U^{q_1} \text{ if } d \in \{2,7,11\}.$$

Let $\|M\| := \max_{1\le i,j\le 2}|M_{ij}|^2$. Here $\epsilon \in \{0,1,2\}$, the exponent pairs $p_\ell, q_\ell \in \mathbb{Z}$ each satisfy $|p_\ell + q_\ell\omega|^2 \le \|M\|$, and $k < 1 - \log_{\kappa(d)}\|M\|$ where $\kappa(d)$ is the Euclidean minimum of $\mathcal{O}_d$ (see Appendix A). Moreover, this procedure runs in time polynomial in $-\log_{\kappa(d)}\|M\|$.

We previously showed that Theorem 4 holds for the Picard group $\mathrm{PSL}(2, \mathbb{Z}[\mathrm{i}])$. In this section, we will complete the proof of the proposition by establishing the result for each of the remaining Euclidean Bianchi groups.

Proof (Proof of Theorem 4 for $\mathrm{PSL}(2, \mathcal{O}_d)$ *with $d = 2, 3, 7, 11$).* We split the proof of Theorem 4 into two parts: the construction of the word representation and the polynomial runtime.

Construction of the word representation. Suppose that $M = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right) \in \mathrm{PSL}(2, \mathcal{O}_d)$. Our first step is to construct an element $H \in \mathrm{PSL}(2, \mathcal{O}_d)$ such that MH is upper-triangular. In the case that $\gamma = 0$, we choose $H = \mathrm{Id}_2$.

We continue under the assumption that $\gamma \neq 0$. We claim that there is an $H_1 \in \mathrm{PSL}(2, \mathcal{O}_d)$ such that $MH_1 = \left(\begin{smallmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{smallmatrix}\right)$ where $N(\gamma_1) < N(\gamma)$. (Here N is the field norm on $\mathbb{Q}(\sqrt{-d})$.) Since $\mathcal{O}_d$ equipped with the field norm N is a Euclidean domain, we can write $\delta = -\theta_1\gamma + \gamma_1$ with $\theta, \gamma_1 \in \mathcal{O}_d$ such that $N(\gamma_1) = N(\theta_1\gamma + \delta) < N(\gamma)$. Let us write $\theta_1 = -p_1 - q_1\omega \in \mathcal{O}_d$ in terms of the integral basis $\{1, \omega\}$, then

$$\begin{aligned} MT^{-p_1}U^{-q_1}A &= \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^{-p_1}\begin{pmatrix}1 & \omega\\ 0 & 1\end{pmatrix}^{-q_1}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\\ &= \begin{pmatrix}\alpha & \theta_1\alpha + \beta\\ \gamma & \theta_1\gamma + \delta\end{pmatrix}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}\theta_1\alpha + \beta & -\alpha\\ \theta_1\gamma + \delta & -\gamma\end{pmatrix} =: \begin{pmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{pmatrix}. \end{aligned}$$

The constructed element $H_1 := U^{-p_1}T^{-q_1}A$ has the required property that $N(\gamma_1) < N(\gamma)$.

We loop the above construction in order to generate a sequence of matrices of the form $MH_1 \cdots H_\ell$. Let γ_ℓ be the bottom-left entry of the matrix $MH_1 \cdots H_\ell$. We repeat this loop until we reach a matrix $MH_1 \cdots H_k$ that satisfies the condition $\gamma_k = 0$.

The following observations guarantee that the above loop both terminates and does so correctly. First, $N(\gamma_\ell) = |\gamma_\ell|^2 \in \mathbb{Z}_{\ge 0}$ for each ℓ since $\gamma_\ell \in \mathcal{O}_d$. Second, the sequence of these norms is strictly decreasing (and so $N(\gamma_{\ell+1}) \le N(\gamma_\ell) - 1$ for $1 \le \ell \le k - 1$). Hence there exists $k \in \mathbb{N}$ such that $N(\gamma_k) = 0$, from which we deduce that $\gamma_k = 0$. Thus we have constructed an element $H_1 \cdots H_k =: H \in \mathrm{PSL}(2, \mathcal{O}_d)$ such that MH is upper-triangular. This ends our first step, which has reduced the representation problem to that of representing an upper-triangular element.

For our second step, consider

$$MH = \begin{pmatrix} \alpha_k & \beta_k \\ 0 & \delta_k \end{pmatrix} =: \begin{pmatrix} \rho & \sigma \\ 0 & \tau \end{pmatrix}$$

with $\rho, \sigma, \tau \in \mathcal{O}_d$. Since $1 = \det(MH) = \rho\tau$, we deduce that $\rho, \tau \in \mathcal{O}_d^\times$ and, moreover, $\overline{\tau} = \tau^{-1} = \rho$. The remainder of this step is broken into two cases that depend on the group of units $\mathcal{O}_d^\times$ (see Lemma 18).

- Suppose that $d = 3$. Then for any $\rho \in \mathcal{O}_d^\times = \{\pm 1, \pm\omega, \pm\omega^2\}$ where $\omega = -\frac{1}{2} + \frac{\sqrt{3}\mathrm{i}}{2}$, we pick $\tau = \overline{\rho}$. From the possible pairings, we deduce that
$$MH = \begin{pmatrix} \omega^2 & 0 \\ 0 & \omega \end{pmatrix}^\epsilon \begin{pmatrix} 1 & \sigma' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \omega^2 & 0 \\ 0 & \omega \end{pmatrix}^\epsilon \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{p_0} \begin{pmatrix} 1 & \omega \\ 0 & 1 \end{pmatrix}^{q_0} = L^\epsilon T^{p_0} U^{q_0}$$
where $\epsilon \in \{0, 1, 2\}$ and $\sigma' = p_0 + q_0\omega \in \mathbb{Z}[\omega] = \mathcal{O}_3$ is an associate of σ.
- Suppose that $d \in \{2, 7, 11\}$. Then $\mathcal{O}_d^\times = \{\pm 1\}$. In this case, our only options are $\rho = \tau = \pm 1$. Thus
$$MH = \begin{pmatrix} 1 & \sigma' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{p_0} \begin{pmatrix} 1 & \omega \\ 0 & 1 \end{pmatrix}^{q_0} = T^{p_0} U^{q_0}$$
where $\pm\sigma = \sigma' = p_0 + q_0\omega \in \mathcal{O}_d$.

Thus ends the second step.

Taken together, the two preceding construct representations for elements H and $T^{p_0}U^{q_0}$ (or $L^\epsilon T^{p_0} U^{q_0}$) for which $M = T^{p_0}U^{q_0}H^{-1}$ (or $M = L^\epsilon T^{p_0} U^{q_0} H^{-1}$). Using the above products, we can output a word representation of the desired form for a given M.

Polynomial runtime. We shall employ the same notations $\|\cdot\|$ for the matrix norm and M_ℓ for matrix products as before. Recall that $\kappa(d)$ is the Euclidean minima on $\mathbb{Q}(\sqrt{-d})$ (see Proposition 19 and Table 2). For $\ell < k$, repeated application of the Euclidean minima gives

$$\begin{aligned} 1 \le |\gamma_\ell|^2 = N(\gamma_\ell) < \kappa(d) N(\gamma_{\ell-1}) < \kappa(d)^2 N(\gamma_{\ell-2}) < \cdots \\ < \kappa(d)^{\ell-1} N(\gamma_1) < \kappa(d)^\ell N(\gamma) = \kappa(d)^\ell |\gamma|^2 = \kappa(d)^\ell \|M\|. \end{aligned}$$

Taking logarithms, we find the upper bound $-\log_{\kappa(d)} \|M\| > \ell$. Arguing the contrapositive, we obtain the upper bound $-\log_{\kappa(d)} \|M\| + 1 > k$ on termination of the first step.

The determinant condition on M_ℓ tells us that

$$\begin{aligned} \alpha_\ell\delta_\ell - \beta_\ell\gamma_\ell = -(\theta_\ell\alpha_{\ell-1} + \beta_{\ell-1})\gamma_{\ell-1} + (\theta_\ell\gamma_{\ell-1} + \delta_{\ell-1})\alpha_{\ell-1} \\ = -(\theta_\ell\alpha_{\ell-1} + \beta_{\ell-1})\gamma_{\ell-1} + \gamma_\ell\alpha_{\ell-1} = 1. \end{aligned}$$

For $\ell - 1 < k$, we have that $\gamma_{\ell-1}^{-1}$ is well-defined and so

$$N(\alpha_\ell) = N(\theta_\ell \alpha_{\ell-1} + \beta_\ell) = N\left(\frac{\gamma_\ell \alpha_{\ell-1} - 1}{\gamma_{\ell-1}}\right) \\ \leq \frac{N(\gamma_\ell)N(\alpha_{\ell-1})}{N(\gamma_{\ell-1})} + \frac{1}{N(\gamma_{\ell-1})} \leq \kappa(d)N(\alpha_{\ell-1}) + 1. \quad (6)$$

Here the rightmost inequality follows from the Euclidean division $-\delta_{\ell-1} = \theta_\ell \gamma_{\ell-1} + \gamma_\ell$, which gives the bound $N(\gamma_\ell) < \kappa(d)N(\gamma_{\ell-1})$.

Claim 20. *For each $d \in \{2, 3, 7, 11\}$, there are no matrices $M_{\ell-1}$ and M_ℓ (as above) for which $\|M_\ell\| > \|M\|_{\ell-1}$.*

The proof of Claim 20 is given in Appendix B.1. All that remains is to bound the sizes of the integer exponents in the word representation. For $\ell \leq k$, we give the following bounds on the quotients $\theta_\ell := -p_\ell - q_\ell \omega$,

$$|p_\ell + q_\ell \omega|^2 = N(\theta_\ell) = N\left(\frac{\delta_{\ell-1} + \gamma_\ell}{\gamma_{\ell-1}}\right) \\ \leq N(\delta_{l-1}) + \frac{N(\gamma_\ell)}{N(\gamma_{\ell-1})} \leq \|M_{\ell-1}\| + \kappa(d) \leq \|M\| + \kappa(d).$$

Since $N(\theta_\ell), \|M\| \in \mathbb{Z}$, we have $N(\theta_\ell) \leq \|M\|$, as desired. Observe that the rightmost inequality follows from Claim 20. We consider the matrix MH in order to bound $|p_0 + q_0\omega|^2$. Once again, it is clear that

$$|p_0 + q_0\omega|^2 = |\sigma|^2 \leq \max_{1 \leq i,j \leq 2} |(MH)_{ij}|^2 = \|M_k\| \leq \|M\|$$

The above observations bound the number of iterations in the initial looping procedure as well as the sizes of the exponents in the word representation. Taken together, with standard results on the division algorithm leads us to the stated polynomial runtime. □

B.1 Proof of Claim 20

In this section, we shall prove the pending cases for Claim 20. We take this opportunity to sketch a proof outline. For each $d = 2, 3, 7, 11$, consider ordered pairs of matrices (M, M') in $\mathrm{PSL}(2, \mathcal{O}_d)$ given by $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ and $M' = \left(\begin{smallmatrix} \theta\alpha+\beta & -\alpha \\ \theta\gamma+\delta & -\gamma \end{smallmatrix}\right)$ (as in (1)) where $-\theta, r \in \mathcal{O}_d$ are the quotient and remainder in $\delta = -\theta \cdot \gamma + r$, respectively. We need to show that each such pair (M, M') satisfies $\|M\| < \|M'\|$.

The proof proceeds by contradiction, any such pair (M, M') requires that

$$\|M\| = \max_{1 \leq i,j \leq 2} |M_{ij}|^2 \leq \frac{1}{1 - \kappa(d)} \quad (7)$$

(where $\kappa(d)$ is the euclidean minima of $\mathcal{O}_d$). Thus, to prove the claim, it is sufficient to compute and compare the norms $\|M\|$ and $\|M'\|$ for all such pairs

(M, M'). The number of matrices M that satisfy (7) is too large to deal with by hand. Thus, with computer assistance[4], we perform an exhaustive search for said matrices and establish that $\|M\| > \|M'\|$ holds for each candidate. Thus ends our sketch proof.

Proof (Proof of Claim 20). As in the Gaussian integer case of Theorem 4, the matrix norm satisfies $\|M_\ell\| > \|M\|_{\ell-1}$ only if $|\alpha_\ell|^2 > |\alpha_{\ell-1}|^2$. Assume, for a contradiction, that $\|M_\ell\| > \|M_{\ell-1}\|$ then, by (6),

$$\|M_{\ell-1}\| < \|M_\ell\| = N(\alpha_\ell) \leq \kappa(d)N(\alpha_{\ell-1}) + 1 < \kappa(d)\|M_{\ell-1}\| + 1,$$

and so we deduce that $(1 - \kappa(d))\|M\|_{\ell-1} < 1$. The possible entries for a matrix in $\mathrm{PSL}(2, \mathcal{O}_d)$ are listed in Table 2. Our proof continues with an exhaustive search for pairs $(M_{\ell-1}, M_\ell)$ of matrices that falsify the claim (i.e., pairs with $\|M_{\ell-1}\| < \|M_\ell\|$). We have implemented this search in SageMath [28]; our code is available at [19]. A description of our search procedure is given below (Algorithm 1).

In Algorithm 1, we employ a function $\textsc{Quotient}\colon \mathcal{O}_d \times \mathcal{O}_d \to \mathcal{O}_d$ that takes an ordered pair of algebraic integers and returns their quotient (in the ring $\mathcal{O}_d$). SageMath does not automatically recognise that the integer rings $\mathcal{O}_d$ (with $d = 1, 2, 3, 7, 11$) are Euclidean domains. Thus for integers $a, b \in \mathcal{O}_d$, the SageMath command `a.quo_rem(b)[0]` will output the quotient a/b obtained by division in the field $\mathbb{Q}(\sqrt{-d})$. We circumvent this problem by implementing a straightforward search for the closest algebraic integer in $\mathcal{O}_d$ to the quotient a/b [19].

From our implementation and testing, we confirm that Claims 5 and 20 hold for each $d \in \{1, 2, 3, 7, 11\}$. □

C Omitted Embeddings Proofs

Proposition 15.

- *There exists an embedding from* $\mathrm{F}(\Sigma_1) \times \mathrm{F}(\Sigma_1)$ *into* $\mathrm{U}(2, \mathbb{N})$.
- *There exists an embedding from* $\mathrm{F}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ *into* $\mathbb{Q}^{2\times 2}$.
- *There exists an embedding from* $\mathrm{S}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ *into* $\mathbb{Z}^{3\times 3}$.
- *There exists an embedding from* $\mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_1)$ *into* $\mathbb{Z}^{2\times 2}$.

Proof. Consider mappings $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ defined as follows:

- Let $\varphi_1 : \mathrm{S}(\Sigma_1) \times \mathrm{S}(\Sigma_1) \to \mathrm{U}(2, \mathbb{N})$ be defined by

$$a \mapsto \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad c \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

- Let $\varphi_2 : \mathrm{F}(\Sigma_2) \times \mathrm{F}(\Sigma_1) \to \mathbb{Q}^{2\times 2}$ be defined by

$$a \mapsto \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \qquad b \mapsto \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \qquad c \mapsto \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

[4] Our computations were performed in SageMath [28].

```
Input: $d \in \{1, 2, 3, 7, 11\}$
  matrices $\leftarrow$ empty list
  $S \leftarrow \left\{x \in \mathcal{O}_d : N(x) < \frac{1}{1-\kappa(d)}\right\}$;    ▷ set of matrix entries (see Table 2)
  for ordered tuples $(\alpha, \beta, \gamma, \delta)$, with $\alpha, \beta, \gamma, \delta \in S$, $\gamma \neq 0$, and $\alpha\delta - \beta\gamma = 1$ do
    $\left(M, \|M\|\right) \leftarrow \left(\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right), \max\{|\alpha|^2, |\beta|^2, |\gamma|^2, |\delta|^2\}\right)$;    ▷ matrix and its norm
    $\theta \leftarrow$ Quotient$(-\delta, \gamma)$;    ▷ output quotient $-\theta \in \mathcal{O}_d$
    $(\alpha, \beta) \leftarrow (\theta \cdot \alpha + \beta, -\alpha)$;
    $(\gamma, \delta) \leftarrow (\theta \cdot \gamma + \delta, -\gamma)$;
    $\left(N, \|N\|\right) \leftarrow \left(\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right), \max\{|\alpha|^2, |\beta|^2, |\gamma|^2, |\delta|^2\}\right)$;    ▷ update matrix and its norm
    if $\|M\| < \|N\|$ then
      add $(M, N)$ to matrices
    end if
  end for
  if matrices is an empty list then
    return "Claim holds for matrices in $\mathrm{PSL}(2, \mathcal{O}_d)$"
  else
    return "Claim does not hold for the pairs matrices"
  end if
```

Algorithm 1: For $d = 1$, this procedure determines whether Claim 5 holds. For $d = 2, 3, 7, 11$, this procedure determines whether Claim 20 holds.

- Let $\varphi_3 : \mathrm{S}(\Sigma_2) \times \mathrm{F}(\Sigma_1) \to \mathbb{Z}^{3\times 3}$ be defined by

$$a \mapsto \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad b \mapsto \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad c \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

- Let $\varphi_4 : \mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_1) \to \mathbb{Z}^{2\times 2}$ be defined by

$$a \mapsto \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \qquad b \mapsto \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \qquad c \mapsto \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Consider first the mapping φ_1. The images are diagonal matrices that have two 1-by-1 blocks for counting the number of letters seen.

It is straightforward to see that mappings φ_2, φ_3 and φ_4 are injective morphisms. Indeed, they all use well-known embeddings from the literature for the first component. In the first and third mappings, the determinant is additionally used to count the number of unary letters. While in the second mapping, the block $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ is used to encode words over the unary alphabet. □

$\exists$	$\mathrm{S}(\Sigma_1)$	$\mathrm{S}(\Sigma_2)$	$\mathrm{F}(\Sigma_1)$	$\mathrm{F}(\Sigma_2)$
$\{\varepsilon\}$		$\mathrm{U}(2,\mathbb{N})$ (folklore)	$U(2,\mathbb{Z})$ (folklore), $\mathbb{Q}$ (folklore)	$\mathbb{C}^{2\times 2}$ (folklore), $\mathbb{Z}^{2\times 2}$ (folklore)
$\mathrm{S}(\Sigma_1)$	$U(2,\mathbb{N})$ (Prop. 15)	$\mathbb{N}^{2\times 2}$ (Prop. 15)	?	$\mathbb{Z}^{2\times 2}$ (Prop. 15)
$\mathrm{S}(\Sigma_2)$		$\mathrm{SL}(3,\mathbb{Q})$ [20], $\mathrm{U}(3,\mathbb{N})$ [24]	$\mathbb{Q}^{2\times 2}$, $\mathbb{Z}^{3\times 3}$ (Prop. 15)	?
$\mathrm{F}(\Sigma_1)$			$\mathbb{Q}$ (folklore)	$\mathbb{Q}^{2\times 2}$ (Prop. 15)
$\mathrm{F}(\Sigma_2)$				$\mathbb{Z}^{4\times 4}$, $\mathbb{H}(\mathbb{Q})^{2\times 2}$ [4]

Table 3. The state of the art for the existence of embeddings from pairs of words into different matrix semigroups. Entries in blue are straightforward extensions of results in the literature.

Proposition 16. *Let $d \in \{1, 2, 3, 7, 11\}$.*

- *There is no embedding from $\varphi_1 : \mathrm{S}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ into $\mathcal{O}_d^{2\times 2}$.*
- *There is no embedding from $\varphi_2 : \mathrm{F}(\Sigma_2) \times \mathrm{F}(\Sigma_1)$ into $\mathcal{O}_d^{2\times 2}$.*
- *There is no embedding from $\varphi_3 : \mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_2)$ into $\mathcal{O}_d^{3\times 3}$.*
- *There is no embedding from $\varphi_4 : \mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_1)$ into $\mathrm{SL}(2, \mathcal{O}_d)$.*

Proof. We prove the final two cases. Let us first consider the third case. The proof follows the proof of non-existence of an embedding from $\mathrm{S}(\Sigma_2) \times \mathrm{S}(\Sigma_2)$ into $\mathrm{SL}(3, \mathbb{Z})$ found in [20]. Intuitively, the main differences in the statements are extending the codomain with elements in $\mathbb{Z}$ to the codomain with elements in

$\mathcal{O}_d$, and moving the explicit condition on determinants of matrices to implicit condition on the domain. Indeed, while we do not restrict ourselves to images with determinant one, the same constraint applies to the first component since letters have inverses and for the image matrix to have inverse, the determinant has to be a unit. After this shift of a constraint, the proof of [20] is easily adapted to prove our result.

Let us consider the final case. Let $\Sigma_2 = \{a, b\}$ and $\Sigma_1 = \{c\}$. Assume to the contrary that such an embedding exists. Now, $\varphi : \mathrm{F}(\Sigma_2) \times \mathrm{S}(\Sigma_1) \to \mathrm{SL}(2, \mathcal{O}_d)$ maps the generators as follows:

$$(a, \varepsilon) \mapsto \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \ (b, \varepsilon) \mapsto \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}, \ (\varepsilon, c) \mapsto \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}.$$

Without loss of generality we can assume that $\left(\begin{smallmatrix} a_1 & a_2 \\ a_3 & a_4 \end{smallmatrix}\right)$ is in the Jordan normal form as conjugating by invertible matrices does not affect the injectivity of a mapping. There are three potential Jordan normal forms for 2-by-2 matrices. Namely,

$$(a, \varepsilon) \mapsto \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \text{ or } \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

where the first two forms are possible if the matrix has repeated eigenvalues and the final form is when there are two distinct eigenvalues. We will rule out each form, thus proving that the embedding does not exist. Recall that we have the following relations:

$$\begin{aligned} \varphi((a, \varepsilon)(b, \varepsilon)) &\neq \varphi((b, \varepsilon)(a, \varepsilon)) \\ \varphi((a, \varepsilon)(\varepsilon, c)) &= \varphi((\varepsilon, c)(a, \varepsilon)) \\ \varphi((b, \varepsilon)(\varepsilon, c)) &= \varphi((\varepsilon, c)(b, \varepsilon)). \end{aligned} \tag{8}$$

Consider the first form, $(a, \varepsilon) \mapsto \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)$. As this matrix is diagonal, it commutes with every matrix and in particular $\varphi((a, \varepsilon)(b, \varepsilon)) = \varphi((b, \varepsilon)(a, \varepsilon))$, which contradicts the first relation in (8). The second form, $(a, \varepsilon) \mapsto \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)$, is also straightforward. It is easy to see that matrices of this form commute only with matrices of form $\left(\begin{smallmatrix} x & 1 \\ 0 & x \end{smallmatrix}\right)$. Hence, $\varphi(\varepsilon, c)$ has to be of this form. But now, for $\varphi((b, \varepsilon)(\varepsilon, c)) = \varphi((\varepsilon, c)(b, \varepsilon))$ to hold, also $\varphi(b, \varepsilon)$ has to be of the same form. This means that $(b, \varepsilon) \mapsto \left(\begin{smallmatrix} y & 1 \\ 0 & y \end{smallmatrix}\right)$, for some y and now $\varphi((a, \varepsilon)(b, \varepsilon)) = \varphi((b, \varepsilon)(a, \varepsilon))$ contradicting the first relation of (8).

Let us consider the last Jordan normal form. Assume that $(a, \varepsilon) \mapsto \left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$. Consider the products $\varphi((a, \varepsilon)(\varepsilon, c))$ and $\varphi((\varepsilon, c)(a, \varepsilon))$:

$$\begin{aligned} \varphi((a, \varepsilon)(\varepsilon, c)) &= \begin{pmatrix} \lambda_1 c_1 & \lambda_1 c_2 \\ \lambda_2 c_3 & \lambda_2 c_4 \end{pmatrix} \\ \varphi((\varepsilon, c)(a, \varepsilon)) &= \begin{pmatrix} \lambda_1 c_1 & \lambda_2 c_2 \\ \lambda_1 c_3 & \lambda_2 c_4 \end{pmatrix}. \end{aligned}$$

For these two matrices to be equal, we observe that $c_2 = c_3 = 0$. Consider then the products $\varphi((b,\varepsilon)(\varepsilon,c))$ and $\varphi((\varepsilon,c)(b,\varepsilon))$:

$$\begin{aligned}\varphi((b,\varepsilon)(\varepsilon,c)) &= \begin{pmatrix} b_1c_1 & b_2c_4 \\ b_3c_1 & b_4c_4 \end{pmatrix} \\ \varphi((\varepsilon,c)(b,\varepsilon)) &= \begin{pmatrix} b_1c_1 & b_2c_1 \\ b_3c_4 & b_4c_4 \end{pmatrix}.\end{aligned}$$

Assume first that $c_1 \neq c_4$. Now, for the equality to hold, $b_2 = b_3 = 0$. However, we now violate the first relation of (8) as $\varphi((a,\varepsilon)(b,\varepsilon)) = \left(\begin{smallmatrix} \lambda_1 b_1 & 0 \\ 0 & \lambda_2 b_4 \end{smallmatrix}\right) = \varphi((b,\varepsilon)(a,\varepsilon))$. We conclude that $c_1 = c_4$ and $(\varepsilon, c) \mapsto \left(\begin{smallmatrix} c_1 & 1 \\ 0 & c_1 \end{smallmatrix}\right)$. Recall that this matrix is not necessarily in $\mathrm{SL}(2,\mathcal{O}_d)$ as it was conjugated by some invertible matrix in order to turn $\varphi(a,\varepsilon)$ into the Jordan normal form. However, the determinant remains 1 and thus $c_1 = \pm 1$. In either case, $\varphi(\varepsilon,c)^2 = \varphi(\varepsilon,c)$ which contradicts the assumption that φ was an embedding. □