

MEMBERSHIP PROBLEMS IN NILPOTENT GROUPS

CORENTIN BODART

Mathematical Institute, University of Oxford, UK

e-mail address: cobodart123@gmail.com

ABSTRACT. We study both the Submonoid Membership problem (SMM) and the Rational Subset Membership problem (RatM) in finitely generated nilpotent groups. We give two reductions with important applications:

- The SMM in any nilpotent group can be reduced to the RatM in *smaller* groups. As a corollary, we prove the existence of a group with decidable SMM and undecidable RatM, answering a question of Lohrey and Steinberg.
- The Rational Subset Membership problem in $H_3(\mathbb{Z})$ can be reduced to the Knapsack problem in the *same* group, and is therefore decidable.

We deduce that the filiform 3-step nilpotent group has decidable Submonoid Membership.

Decision problems are central motivating questions in combinatorial group theory. The first example is the Word Problem, introduced by Dehn in 1910 [6]. For matrix groups, the first hard problem is the Subgroup Membership problem. With it came one of the first undecidability results in $SL_4(\mathbb{Z})$ [24], and also some positive results in nilpotent and polycyclic groups [22]. This paper focuses on the next two decision problems in line, namely the (uniform) Submonoid and Rational Subset Membership problems. For a group G (given as a matrix group, or endowed with a finite generating set S) we attempt to produce algorithms with the following specifications:

Submonoid Membership problem (SMM(G))

Input: Elements g and $g_1, g_2, \dots, g_n \in G$ (defined as matrices or words over S).

Output: Decide whether g belongs to the submonoid $\{g_1, g_2, \dots, g_n\}^*$.

Rational Subset Membership problem (RatM(G))

Input: An element $g \in G$ (given either as a matrix or as a word over S) and a rational subset $R \subseteq G$ (defined by a finite state automaton, labeled by elements in G)

Output: Decide whether $g \in R$.

Key words and phrases: Submonoid Membership, Rational Subsets, Heisenberg group, nilpotent groups.
The author was partially supported by the Swiss NSF grant 200020-200400.
2020 *Mathematics Subject Classification.* 20F10, 20F18, 68Q45.

Both problems are known to be decidable for a few classes of groups, including free groups [1] and virtually abelian groups [14]. We suggest to have a look at the surveys [16, 7, 18] for a more exhaustive picture.

Finally, we recall the Knapsack Problem which plays a key role in our paper. This problem is a special case of Rational Subset Membership (see Figure 1), introduced in [27].

Knapsack problem (KS(G))

Input: Elements g and $g_1, g_2, \dots, g_n \in G$ (defined as matrices or words over S).

Output: Decide whether $g \in \{g_1\}^* \{g_2\}^* \dots \{g_n\}^*$.

This sub-problem is known to be decidable in a much larger class of group. This includes hyperbolic groups [27] and co-context-free groups [15].

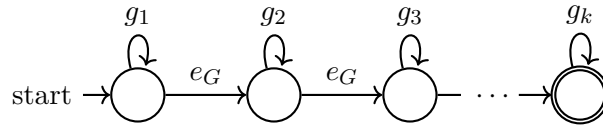


FIGURE 1. Finite state automaton relative to the Knapsack problem

Finitely generated submonoids are rational, hence the decidability of $\text{RatM}(G)$ implies the decidability of $\text{SMM}(G)$. A natural question is whether the reciprocal holds:

Question (Lohrey-Steinberg). Does there exist a finitely generated group with decidable Submonoid Membership and undecidable Rational Subset Membership?

Lohrey and Steinberg proved that both problems are recursively equivalent in RAAGs [19] and infinitely-ended groups [20]. However, they conjecture a positive answer for more general groups. They also note that the existence of such a group is equivalent to the property “ $\text{SMM}(G)$ is decidable” not being closed under free products [20, §4].

For (non virtually abelian) nilpotent groups, the full picture is not clear yet.

- The Knapsack problem (hence Rational Subset Membership) is undecidable in large nilpotent groups, most notably $H_3(\mathbb{Z})^k$ and $N_{2,k}$ (the free 2-step nilpotent group of rank k) for $k \gg 1$. [17, 15, 25]
- Submonoid Membership is undecidable in $H_3(\mathbb{Z})^k$ for $k \gg 1$. [30]

Both results rely on the negative solution to Hilbert’s 10th problem: there exists no algorithm deciding whether a Diophantine equation (or system of equations) admits an integer solution [23]. On the positive side, the list of results is even shorter:

- The Knapsack problem is decidable in $H_{2m+1}(\mathbb{Z})$ for all $m \geq 1$ [15], where

$$H_{2m+1}(\mathbb{Z}) = \left\{ \left(\begin{array}{ccc} 1 & \mathbf{a} & c \\ & I_m & \mathbf{b}^t \\ & & 1 \end{array} \right) \mid \mathbf{a}, \mathbf{b} \in \mathbb{Z}^m, c \in \mathbb{Z} \right\} \leq SL_{2m+1}(\mathbb{Z}).$$

- Colcombet, Ouaknine, Semukhin and Worrell proved that Submonoid Membership is decidable in $H_{2m+1}(\mathbb{Z})$ for all $m \geq 1$. [5]

Note that the former relies on deep results on quadratic Diophantine equations [13], whereas the latter is elementary. This points to $\text{RatM}(G)$ being harder than $\text{SMM}(G)$, hence the hope to separate both problems within the class of nilpotent groups.

In the first part of the paper, we re-interpret and extend Colcombet *et al.*'s result:

Theorem 2.6. *There exists an algorithm with the following specifications*

Input: *A finitely presented nilpotent group G (given by a finite presentation), a finite set $S \subset G$ and an element $g \in G$ (given as words).*

Output: *Finitely many instances $g_i \in R_i$ of Rational Subset Membership in a subgroup $H \leq G$ such that $g \in S^*$ if and only if $g_i \in R_i$ for some i .*

Moreover, it solves these instances if $h([H, H]) = h([G, G])$, with h the Hirsch length.

Note that, if $h([G, G]) \leq 1$, e.g. for $H_{2m+1}(\mathbb{Z})$, then either $h([H, H]) = h([G, G])$ or H is virtually abelian. In both cases the instances of $\text{RatM}(H)$ are decidable. In a more conceptual direction, Theorem 2.6 confirms the conjecture of Lohrey and Steinberg:

Corollary 2.8. *There exists a nilpotent group of class 2 with decidable Submonoid Membership and undecidable Rational Subset Membership.*

Remark. It is crucial to consider the *uniform* version of the Rational Subset Membership in the previous result. Indeed the instances of $\text{RatM}(H)$ we need to solve *will depend* on $g \in G$, even for a fixed submonoid $S^* \subseteq G$.

The second part of the paper focuses on the discrete Heisenberg group

$$H_3(\mathbb{Z}) \simeq \langle x, y \mid [x, [x, y]] = [y, [x, y]] = 1 \rangle.$$

The Heisenberg group is the smallest (non virtually-abelian) f.g. nilpotent group, in the sense that it embeds in all such groups. Our main result is the following:

Theorem 3.9. *$H_3(\mathbb{Z})$ has decidable Rational Subset Membership.*

Combining this result with Theorem 2.6, we get the following:

Corollary 0.1. *The filiform 3-step nilpotent group (also called “the Engel group”)*

$$E = \langle x, y_1, y_2, y_3 \mid [x, y_i] = y_{i+1} \text{ for } i = 1, 2; [x, y_3] = [y_i, y_j] = 1 \text{ for } i, j = 1, 2, 3 \rangle$$

has decidable Submonoid Membership.

Indeed, this nilpotent group has Hirsch length $h(E) = 4$, hence any infinite-index subgroup admits a finite-index subgroup isomorphic to 1 , \mathbb{Z} , \mathbb{Z}^2 , \mathbb{Z}^3 or $H_3(\mathbb{Z})$. It can be seen as a subgroup of the unitriangular matrices $UT_4(\mathbb{Z})$:

$$E \simeq \mathbb{Z}^3 \rtimes_X \mathbb{Z} = \left\{ \begin{pmatrix} X^n & \mathbf{y} \\ 0 & 1 \end{pmatrix} \mid n \in \mathbb{Z}, \mathbf{y} \in \mathbb{Z}^3 \right\} \quad \text{where } X = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

We expect that some extension of Theorem 3.9 holds:

Conjecture A. $H_{2m+1}(\mathbb{Z})$ has decidable Rational Subset Membership for all $m \geq 1$.

In turn, this would imply that Submonoid Membership is decidable in f.g. nilpotent groups G satisfying $h([G, G]) \leq 2$, or commensurable to $N_{2,3} \times \mathbb{Z}^n$ or $N_{3,2} \times \mathbb{Z}^n$. (Those groups satisfy $h([G, G]) = 3$, and contain no subgroup with $h([H, H]) = 2$.)

The proof of Theorem 3.9 relies on a technical proposition. We reduce the problem to the class of bounded regular languages, defined by the following equivalent conditions:

Theorem (See eg. [36]). *Fix $\mathcal{L} \subseteq \Sigma^*$ a regular language. The following are equivalent:*

- (a) \mathcal{L} has polynomial growth: $\beta_{\mathcal{L}}(n) := \#\mathcal{L} \cap \Sigma^{\leq n} \preceq n^d$ for some $d \geq 0$.
- (b) \mathcal{L} is bounded, i.e., $\mathcal{L} \subseteq \{w_1\}^* \{w_2\}^* \dots \{w_r\}^*$ for some $w_i \in \Sigma^*$.
- (c) \mathcal{L} is a finite union of languages $t_0 \{u_1\}^* t_1 \{u_2\}^* t_2 \dots \{u_s\}^* t_s$ with $t_i, u_i \in \Sigma^*$.

More precisely, we prove the following result:

Proposition 3.3. *Let $R \subseteq H_3(\mathbb{Z})$ be a rational subset, i.e., $R = \text{ev}(\mathcal{L})$ for some regular language $\mathcal{L} \subseteq G^*$ (see §1.1 for definitions). There exists a bounded regular language $\mathcal{L}' \subset G^*$ such that $R = \text{ev}(\mathcal{L}')$. Moreover, \mathcal{L}' can be effectively computed from \mathcal{L} .*

This essentially reduces $\text{RatM}(H_3(\mathbb{Z}))$ to $\text{KS}(H_3(\mathbb{Z}))$, which is decidable by [15].

Added in proof. A few weeks after this article was first made public, Markus Lohrey told me about unpublished work of Doron Shafrir (2018) proving that $\text{SMM}(G)$ is decidable and $\text{RatM}(G)$ is undecidable for $G = A \wr \mathbb{Z}^2$, with $A \neq 1$ finite abelian. This work has now been written down in Potthast's bachelor thesis [28], and generalised by Ruiwen Dong in [8].

Moreover, building on work of Shafrir [33], Dong proved that the decidability of SMM is not preserved by finite-index overgroups [8]. This leaves open the following basic questions:

Question B. Does there exist a finitely generated group G such that

- $\text{SMM}(G)$ is decidable and $\text{KS}(G)$ is undecidable?
- $\text{RatM}(G)$ (resp. SMM , KS) is decidable and $\text{RatM}(G \times \mathbb{Z})$ (resp. SMM , KS) is undecidable?

Finally Shafrir proved the above conjecture for rational subsets of a specific shape [32]. This is sufficient to conclude that $\text{SMM}(G)$ is decidable for G nilpotent satisfying $h([G, G]) \leq 2$, as well as $G = N_{2,3} \times \mathbb{Z}^n$ and $N_{3,2} \times \mathbb{Z}^n$. Extending this result further would require new ideas. However, we advertise the following problem, which still seems tractable (as the algorithms proposed in [5] or here do not rely on solving quadratic Diophantine equations).

Problem C. Give an upper bound on the time complexity of $\text{SMM}(H_3(\mathbb{Z}))$.

1. BACKGROUND ON RATIONAL SUBSETS

1.1. **Definition.** Fix a group G . An *automaton* over G is a tuple $M = (V, \delta, v_0, \text{accept})$ with

- V is a finite set of vertices/states.
- $\delta \subset V \times G \times V$ is a finite set of edges/transitions. Each element $(u, g, v) \in \delta$ should be thought as an oriented edge $u \rightarrow v$ with label g .
- $v_0 \in V$ is the initial/start vertex.
- $\text{accept} \subseteq V$ is the set of terminal/accept vertices.

★ Each automaton M recognizes a language $\mathcal{L}(M) \subseteq G^*$, namely the set of words $w \in G^*$ we can read along (oriented) paths from v_0 to some accept vertex.

★ There is a natural map $\text{ev}: G^* \rightarrow G: w \mapsto \bar{w}$, interpreting each word as a product in G .

★ A subset $R \subseteq G$ is *rational* if there exists an automaton M over G such that $\text{ev}(\mathcal{L}(M)) = R$. The subset R is *unambiguously rational* if we can moreover ensure that $\text{ev}: \mathcal{L}(M) \rightarrow R$ is bijective. In that case, $\mathcal{L}(M)$ is a *regular normal form* for R .

★ An automaton is *trim* if every vertex $p \in V$ lies on a path from v_0 to an accept vertex.

★ An automaton is *deterministic* if, for every pair $(u, g) \in V \times G$, there exists at most one $v \in V$ such that $(u, g, v) \in \delta$. This implies that each $w \in \mathcal{L}$ is accepted by a unique path.

★ Given a word $w \in G^*$, we denote its length by $\ell(w)$. Moreover, given $S \subseteq G$ and $g \in \langle S \rangle$, we define its *word length* as $\|g\|_S := \min\{\ell(w) \mid w \in S^* \text{ such that } \bar{w} = g\}$.

1.2. **Going to subgroups.** We state a lemma due to Gilman, himself inspired by Stallings. We recall the proof since the sets X and $\tilde{\mathcal{L}}$ defined in the proof will turn out useful.

Proposition 1.1 ([11, Lemma 5]). *Let G be a group and $R \subseteq G$ a rational subset recognized by an automaton M (over G). Suppose that R sits inside a subgroup $H \leq G$, then*

- *There exists an H -automaton \tilde{M} recognizing R (that is, R is rational in H). Moreover \tilde{M} is effectively computable from M .*
- *If \mathcal{L} is a normal form for R , then so is $\tilde{\mathcal{L}} = \mathcal{L}(\tilde{M})$.*
- *If \mathcal{L} is bounded, then so is $\tilde{\mathcal{L}}$.*

In particular, from the first point, we may drop the “over G ”.

Proof. We start with an automaton $M = (V, \delta, v_0, \text{accept})$ recognizing R . Using the powerset construction [29], we may suppose that this automaton is deterministic and trim. Let

$$X = \left\{ \bar{t}\bar{u}\bar{t}^{-1} \mid \begin{array}{l} t \in G^* \text{ labels a simple path } v_0 \rightarrow p \\ \exists p \in V \text{ s.t. } u \in G^* \text{ labels a simple cycle } p \rightarrow p \\ \text{both paths only intersect at } p \end{array} \right\}$$

$$Y = \{ \bar{y} \mid \exists q \in \text{accept s.t. } y \in G^* \text{ labels a simple path } v_0 \rightarrow q \}$$

We consider the regular language

$$\tilde{\mathcal{L}} = \left\{ x_1 x_2 \dots x_\ell y \in X^* Y \mid \begin{array}{l} t_i \text{ is a prefix of } t_{i+1} u_{i+1} \text{ for } 1 \leq i \leq \ell - 1 \\ t_\ell \text{ is a prefix of } y, \text{ where } \bar{t}_i \bar{u}_i \bar{t}_i^{-1} = x_i \end{array} \right\}$$

which is recognized by the automaton $\tilde{M} = (\tilde{V}, \tilde{\delta}, \varepsilon, \{\odot\})$ with

- $\tilde{V} = \{t \in G^* \mid \exists p \in V \text{ s.t. } t \text{ labels a simple path } v_0 \rightarrow p\} \sqcup \{\odot\}$,
- $\tilde{\delta}$ consists of all edges $s \rightarrow t$ labeled by $x \in X$ if $x = \bar{t} \bar{u} \bar{t}^{-1}$ and s is a prefix of tu , and all an edge $t \rightarrow \odot$ labeled by $\bar{y} \in Y$ if t is a prefix of y ,
- the empty string ε (which labels the simple path $v_0 \rightarrow v_0$) is the starting vertex,
- and the vertex \odot is the unique accept vertex.

► We first prove that $\text{ev}(\mathcal{L}) = \text{ev}(\tilde{\mathcal{L}})$.

Any word $w \in \mathcal{L}$ is accepted by a unique path $v_0 \rightarrow v_0$ in the automaton M . This path can be decomposed as a product of conjugates $t_i u_i t_i^{-1}$, using a loop-erasure algorithm:

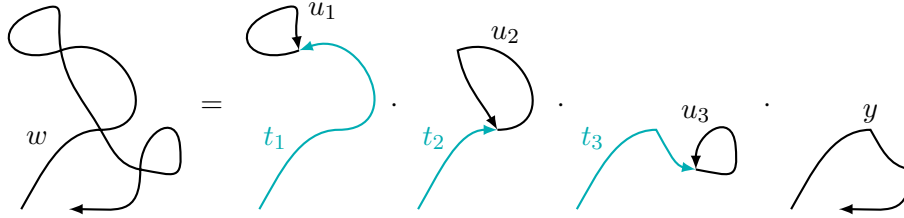


FIGURE 2. A path decomposed as a product of “freeze frames” of the loop-erasure algorithm.

The resulting word $\tilde{w} = (\bar{t}_1 \bar{u}_1 \bar{t}_1^{-1}) \dots (\bar{t}_\ell \bar{u}_\ell \bar{t}_\ell^{-1}) \cdot \bar{y} \in \tilde{\mathcal{L}}$ is the *decomposition* of w . Both w and its decomposition \tilde{w} evaluate to the same element in G . Moreover, words in $\tilde{\mathcal{L}}$ are exactly decompositions of words in \mathcal{L} . This proves that $\text{ev}(\mathcal{L}) = \text{ev}(\tilde{\mathcal{L}})$.

► We prove that $X, Y \subseteq H$.

Take $x = \bar{t} \bar{u} \bar{t}^{-1} \in X$, where t labels a simple path $v_0 \rightarrow p$ and u labels a simple cycle $p \rightarrow p$ with t and u only intersecting at p . As the automaton M is trim, there exists a path $p \rightarrow q \in \text{accept}$ labeled by some word $v \in G^*$. Consider $tuv \in \mathcal{L}$ and $xx_2 \dots x_\ell y \in \tilde{\mathcal{L}}$ its decomposition. Observe that $x_2 \dots x_\ell y \in \tilde{\mathcal{L}}$ too, so

$$x = (xx_2 \dots x_\ell y)(x_2 \dots x_\ell y)^{-1} \in R \cdot R^{-1} \subseteq H.$$

On the other side, we have $Y \subseteq R \subseteq H$.

► If \mathcal{L} is a normal form, then each $g \in R$ is represented by a unique word w in \mathcal{L} , hence has a unique representative \tilde{w} in $\tilde{\mathcal{L}}$.

► Observe that $\ell(w) \leq |V| \ell(\tilde{w})$ for all $w \in \mathcal{L}$ (as each letter in \tilde{w} corresponds to a simple cycle or a simple path in w), hence $\beta_{\mathcal{L}}(n) \geq \beta_{\tilde{\mathcal{L}}}(n)$. If \mathcal{L} is bounded, then \mathcal{L} and $\tilde{\mathcal{L}}$ have polynomial growth, hence $\tilde{\mathcal{L}}$ is bounded [36]. \square

Another lemma which complements nicely is the following:

Lemma 1.2. *Let G be a group and $H \leq G$ is a finite-index subgroup.*

(a) *If $R \subseteq G$ is rational, then $Hg \cap R$ is rational for each $g \in G$.*

If moreover we are given an automaton M such that $R = \text{ev}(\mathcal{L}(M))$ with labels inside a finite set $S \subseteq G$, and the Schreier graph $\text{Sch}(H \backslash G, S)$, then

(b) *we can effectively compute an automaton for $Hg \cap R$, and*

(c) *we can effectively compute $\{Hg \in H \backslash G \mid Hg \cap R \neq \emptyset\}$ and compute a subset $\mathcal{K} \subset \mathcal{L}(M)$ consisting of a unique representative for each of these cosets.*

Proof. (a-b) Given an automaton for $M = (V, \delta, v_0, \text{accept})$ for R , we construct the following automaton recognizing $Hg \cap R$:

- $V' = H \backslash G \times V$
- $v'_0 = (H, v_0)$
- $\text{accept}' = \{Hg\} \times \text{accept}$
- $(Hg_1, v_1) \xrightarrow{s} (Hg_2, v_2)$ if and only if $Hg_1s = Hg_2$ and $v_1 \xrightarrow{s} v_2$.

(c) For each coset, one can compute an automaton for $Hg \cap R$ and then decide if the intersection is non-empty: does there exist a path $v'_0 \rightarrow \text{accept}'$? If yes, a representative is given by $w \in \mathcal{L}(M)$ labeling the first path found. \square

Remark 1.3. Whenever $G = \langle A \mid B \rangle$ is finitely presented, the Todd-Coxeter algorithm allows to compute $\text{Sch}(H \backslash G, S)$ taking as input S and generators for H (given as words over A), so $\text{Sch}(H \backslash G, S)$ doesn't need to be part of the input of our algorithm.

1.3. Virtually abelian groups. We recall a result of Eilenberg and Schützenberger.

Theorem 1.4 ([10]). *Let G be an abelian group and R be a rational subset, then R is unambiguously rational. Moreover, given a language \mathcal{L} such that $\text{ev}(\mathcal{L}) = R$, we can effectively compute a regular normal form \mathcal{L}' for R .*

This result can easily be extended to virtually abelian groups using Lemma 1.2(b).

Corollary 1.5. *Any rational subset of a virtually abelian group can be represented by a bounded regular language \mathcal{L}' . Moreover, \mathcal{L}' can be computed effectively.*

Proof. Take \mathcal{L}' the normal form from Theorem 1.4, and let S be the (finite) set of elements appearing as labels on an automaton recognizing \mathcal{L}' . Note that $\ell(w) \geq \|\bar{w}\|_S$ for all $w \in S^*$ hence we can compare the volume growth of \mathcal{L}' and $(\langle S \rangle, S)$: we have $\beta_{\mathcal{L}'}(n) \leq \beta_{(\langle S \rangle, S)}(n)$, but $\langle S \rangle$ is a finitely generated virtually abelian group, hence has polynomial growth, so the rational language \mathcal{L}' has polynomial growth hence is bounded (see [36]). \square

2. DIMENSION GAIN FOR THE SUBMONOID MEMBERSHIP

2.1. Background on nilpotent groups. We first recall the classical notion of Hirsch length, and some related properties.

Definition 2.1. Let G be a group. The *lower central series* is the sequence of subgroups

$$\gamma_1(G) = G, \quad \gamma_{i+1}(G) = [\gamma_i(G), G] \text{ for all } i \geq 1$$

We say that G is *nilpotent of class c* (or c -step nilpotent) if $\gamma_{c+1}(G) = \{e\}$.

Definition 2.2. The *Hirsch length* of a c -step nilpotent group G is defined as

$$h(G) = \sum_{i=1}^c \text{rk}_{\mathbb{Q}}(\gamma_i(G)/\gamma_{i+1}(G)).$$

Proposition 2.3 (Folklore, [31, Exercise 8]). *Let G be a finitely generated nilpotent group.*

- (a) *For any exact sequence $1 \rightarrow N \rightarrow G \rightarrow Q \rightarrow 1$, we have $h(G) = h(N) + h(Q)$.*
- (b) *For any group $H \leq G$, we have $h(H) \leq h(G)$ with equality if and only if $[G : H] < \infty$.*
- (c) *Let $N_{2,m}$ be the free 2-step nilpotent group on m generators. Then $h(N_{2,m}) = m + \binom{m}{2}$.*

Next we quote a useful result, which can be understood as a discrete version of the Chow-Rashevskii theorem in sub-Riemannian geometry. This requires a definition/lemma:

Lemma 2.4 (Torsionfree part of the abelianization). *Let G be a group, then*

$$\overline{[G, G]} = \{g \in G \mid \exists n \neq 0, g^n \in [G, G]\}$$

is a characteristic subgroup of G . Moreover, if G is finitely generated, then $[G, G]$ has finite index inside $\overline{[G, G]}$, and $G/\overline{[G, G]} \simeq \mathbb{Z}^r$ for some r .

Proof. Let $\tau: G \rightarrow G/[G, G]$ be the abelianization map, then torsion elements of $G/[G, G]$ form a characteristic subgroup T and $\overline{[G, G]} = \tau^{-1}(T)$. If G is finitely generated, then $G/[G, G] \simeq \mathbb{Z}^r \times T$ hence $|\overline{[G, G]} : [G, G]| = |T|$ and $G/\overline{[G, G]} \simeq \mathbb{Z}^r$. \square

Proposition 2.5 ([3], see also [34]). *Consider an infinite, finitely generated nilpotent group G , and $\pi: G \twoheadrightarrow G/\overline{[G, G]} \simeq \mathbb{Z}^r$. For any $S \subseteq G$, the following are equivalent*

- (a) *$\text{ConvHull}(\pi(S)) \subseteq \mathbb{R}^r$ contains a ball $B(\mathbf{0}, \varepsilon)$ for some $\varepsilon > 0$.*
- (b) *For every non-zero homomorphism $f: G \rightarrow \mathbb{R}$, there exists $s \in S$ s.t. $f(s) < 0$.*
- (c) *The submonoid S^* is a finite-index subgroup of G .*

If S is finite, we can restrict to non-zero homomorphism $f: G \rightarrow \mathbb{Z}$.

2.2. Proof of Theorem 2.6. We prove the following algorithmic reduction:

Theorem 2.6. *There exists an algorithm with the following specifications:*

Input: *A finitely presented nilpotent group G , a finite set $S \subset G$, and $g \in G$.*

Output: *Finitely many instances $\{g_i \stackrel{?}{\in} R_i\}$ of the Rational Subset Membership in a subgroup $H \leq G$ such that $g \in S^*$ if and only if $g_i \in R_i$ for some i .*

Moreover, the algorithm solves these instances if $h([H, H]) = h([G, G])$.

Proof. First, we may assume $e_G \in S$. The proof splits into three steps

- (1) We compute the image of S through a map $\pi: G \rightarrow \mathbb{Z}^r$ as in Proposition 2.5. Using this quotient, we define a subgroup $H \leq G$ (depending only on S) and a partition $S = S_0 \sqcup S_+$, looking whether $\pi(s)$ is invertible or not in $\pi(S)^* \leq \mathbb{Z}^r$ for each $s \in S$.
- (2) We reduce the problem $g \in S^*$ to finitely many instances of $\text{RatM}(H)$.
- (3) In the case when $h([H, H]) = h([G, G])$, we solve the previous instances of $\text{RatM}(H)$.

The first step follows [5, Theorem 7] closely, while the last step generalizes ‘‘Case II’’ from the same proof. The observation of the second step seems new.

► First we compute a map $\pi: G \rightarrow \mathbb{Z}^r$ from a presentation for G . We consider the polytope $\text{ConvHull}(\pi(S)) \ni \mathbf{0}$ given by a V-representation (namely, $\pi(S)$). We can compute a H-representation (H for half-space), that is, a finite set of inequalities $\{f_i(\mathbf{v}) \geq a_i \mid i \in I\}$ with $f_i: \mathbb{Z}^r \rightarrow \mathbb{Z}$ non-zero linear forms and $a_i \in \mathbb{Z}$ such that

$$\text{ConvHull}(\pi(S)) = \{\mathbf{v} \in \mathbb{R}^r \mid \forall i \in I, f_i(\mathbf{v}) \geq a_i\}.$$

This is the classical *facet enumeration problem*. (See [37, Section 1.2] and references therein.) We compute a maximal linearly independent set $\{f_1, f_2, \dots, f_s\}$ inside $\{f_i \mid i \in I, a_i = 0\}$, and define a homomorphism $f: G \rightarrow \mathbb{Z}^s$ via

$$f(h) = (f_1(\pi(h)), f_2(\pi(h)), \dots, f_s(\pi(h))).$$

The image of $f: G \rightarrow \mathbb{Z}^s$ has finite index (by linear independence), and $f(S) \subset \mathbb{Z}_{\geq 0}^s$. We define $H = \ker f$ and partition S into $S_0 = S \cap H$ and $S_+ = S \setminus H$.

Geometric intermezzo: We are looking at the minimal face F of $\text{ConvHull}(\pi(S))$ containing $\mathbf{0}$ (i.e., no proper sub-face of F contains $\mathbf{0}$). This face is given by

$$F = \text{ConvHull}(\pi(S)) \cap K = \text{ConvHull}(\pi(S_0))$$

where $K = \{\mathbf{v} \in \mathbb{R}^r \mid f_i(\mathbf{v}) = 0 \text{ for } i = 1, 2, \dots, s\}$. The minimality of F can be stated as ‘‘there exists $\varepsilon > 0$ such that $B(\mathbf{0}, \varepsilon) \cap K \subset F$ ’’.

► There exists only finitely many words $w = u_1 u_2 \dots u_k \in S_+^*$ such that

$$f(\bar{w}) = \sum_{i=1}^k f(u_i) = f(g)$$

(Indeed, k is bounded by the sum of the components of $f(g)$.) For each word w , we need to decide whether $g \in S_0^* u_1 S_0^* u_2 S_0^* \dots S_0^* u_k S_0^*$. Observe that

$$S_0^* \cdot u_1 S_0^* \cdot u_2 S_0^* \cdot \dots \cdot u_k S_0^* = S_0^* \cdot (v_1 S_0 v_1^{-1})^* \cdot (v_2 S_0 v_2^{-1})^* \cdot \dots \cdot (v_k S_0 v_k^{-1})^* \cdot v_k$$

where $v_i = u_1 u_2 \dots u_i$. Hence the problem can be restated as

$$g v_k^{-1} \stackrel{?}{\in} R := S_0^* \cdot (v_1 S_0 v_1^{-1})^* \cdot (v_2 S_0 v_2^{-1})^* \cdot \dots \cdot (v_k S_0 v_k^{-1})^* \subseteq H$$

which is an instance of the Rational Subset Membership Problem in H .

We note that algorithms presented in [21] allow to compute a presentation for H , and then rewrite all elements $g v_k^{-1} \in H$ and $v_i S_0 v_i^{-1} \in H$ as words over the corresponding generating set. (Problems (III), (IV) and (II) in the article.)

► We solve the Membership problem under the hypothesis $h([G, G]) = h([H, H])$.

Observe that $\overline{[G, G]} = \overline{[H, H]}$. Indeed, we have $\overline{[G, G]} \leq H$ as H is the kernel of $f: G \rightarrow \mathbb{Z}^s$ with \mathbb{Z}^s torsionfree abelian. Moreover Proposition 2.3(b) and Lemma 2.4 imply that

$$m = \left[\overline{[G, G]} : [H, H] \right] = \left[\overline{[G, G]} : [G, G] \right] \left[[G, G] : [H, H] \right] < \infty.$$

For every $g \in \overline{[G, G]}$, we have $g^{m!} \in [H, H]$ hence $g \in \overline{[H, H]}$, proving the observation. In particular, we can identify the quotient $H/\overline{[H, H]} = H/\overline{[G, G]} = \pi(H)$.

Condition (a) of Proposition 2.5 (for H and S_0) now reads “ $F = \text{ConvHull}(\pi(S_0))$ contains a ball $B(\mathbf{0}, \varepsilon) \cap K$ ”, which holds by minimality of F . We conclude that S_0^* is a finite-index subgroup of H . Finally, the problem can be restated as

$$\langle S_0 \rangle g v_k^{-1} \cap (v_1 S_0 v_1^{-1})^* \cdot (v_2 S_0 v_2^{-1})^* \cdot \dots \cdot (v_k S_0 v_k^{-1})^* \stackrel{?}{=} \emptyset$$

which is easily decided using Lemma 1.2(c). □

2.3. Proof of Theorem 2.8. We give a lemma on abstract commensurability classes of subgroups of $N_{2,m} \times \mathbb{Z}^n$. Some ideas can already be found in [12, Theorem 7].

Lemma 2.7. *Any subgroup $H \leq N_{2,m} \times \mathbb{Z}^n$ admits a finite-index subgroup $H' \trianglelefteq H$ isomorphic to $N_{2,k} \times \mathbb{Z}^\ell$ for some $k \leq m$ and $\ell \leq \binom{m}{2} - \binom{k}{2} + n$.*

Proof. Let $G = N_{2,m} \times \mathbb{Z}^n$ and fix a subgroup $H \leq G$. Consider the composition $\alpha = \alpha_2 \circ \alpha_1$ where α_1 is the projection on the first factor, and α_2 is the abelianization map:

$$N_{2,m} \times \mathbb{Z}^n \xrightarrow{\alpha_1} N_{2,m} \xrightarrow{\alpha_2} \mathbb{Z}^m$$

The image $\alpha(H) \leq \mathbb{Z}^m$ is isomorphic to \mathbb{Z}^k for some $k \leq m$. Fix $g_1, \dots, g_k \in H$ such that $\{\alpha(g_1), \dots, \alpha(g_k)\}$ is a basis for $\alpha(H)$.

Claim 1. $\langle g_1, \dots, g_k \rangle \simeq N_{2,k}$ and $\{g_1, \dots, g_k\}$ is a basis.

Observe that the projections of $\alpha_1(g_1), \dots, \alpha_1(g_k) \in N_{2,m}$ in \mathbb{Z}^m are linearly independent, hence they generate a subgroup isomorphic to $N_{2,k}$ as a basis (see [26, Theorem 1.3]). Moreover, since $N_{2,k}$ is relatively free in the variety of 2-step nilpotent groups and $N_{2,m} \times \mathbb{Z}^n$ is 2-step nilpotent, we can define a morphism $\alpha_1(g_i) \mapsto g_i$, proving that the restriction $\alpha_1|_{\langle g_1, \dots, g_k \rangle}$ is invertible, hence the claim.

Observe that $K = \ker \alpha$ is the center of G , in particular is abelian. We have

$$K \twoheadrightarrow K / (\langle g_1, \dots, g_k \rangle \cap K) \xrightarrow{\sim} T \times \mathbb{Z}^{\binom{m}{2} - \binom{k}{2} + n} \twoheadrightarrow \mathbb{Z}^{\binom{m}{2} - \binom{k}{2} + n}$$

where T is some finite abelian group. Indeed $K \simeq \mathbb{Z}^{\binom{m}{2} + n}$ and $\langle g_1, \dots, g_k \rangle \cap K \simeq \mathbb{Z}^{\binom{k}{2}}$. We denote the whole composition by β , and the composition of the first two arrows by γ .

The image $\beta(H \cap K) \leq \mathbb{Z}^{\binom{m}{2} - \binom{k}{2} + n}$ is isomorphic to \mathbb{Z}^ℓ for some $\ell \leq \binom{m}{2} - \binom{k}{2} + n$. Fix $h_1, \dots, h_\ell \in H \cap K$ such that $\{\beta(h_1), \dots, \beta(h_\ell)\}$ is a basis of $\beta(H \cap K)$.

Claim 2. Since K is abelian, $\beta: \langle h_1, \dots, h_\ell \rangle \rightarrow \beta(H \cap K) \simeq \mathbb{Z}^\ell$ is an isomorphism.

Claim 3. We have $H' := \langle g_1, \dots, g_k, h_1, \dots, h_\ell \rangle \simeq N_{2,k} \times \mathbb{Z}^\ell$. It suffices to check that

- $\langle g_1, \dots, g_k \rangle$ and $\langle h_1, \dots, h_\ell \rangle$ commutes, which is true as $\langle h_1, \dots, h_\ell \rangle \leq K = \mathcal{Z}(G)$.
- $\langle g_1, \dots, g_k \rangle \cap \langle h_1, \dots, h_\ell \rangle = \{e\}$, which is true as restriction $\beta|_{\langle g_1, \dots, g_k \rangle \cap K}$ is zero while the restriction $\beta|_{\langle h_1, \dots, h_\ell \rangle}$ is injective.

Claim 4. $H' \trianglelefteq H$ and H/H' is finite.

Recall that $K = \ker \alpha$ is central hence

$$\forall g, g', h, h' \in G, (\alpha(g) = \alpha(g') \text{ and } \alpha(h) = \alpha(h')) \implies [g, h] = [g', h'].$$

Therefore $\alpha(H) = \alpha(H')$ implies $[H, H] = [H', H']$. It follows that $H' \trianglelefteq H$ (as any subgroup $H' \leq H$ containing $[H, H]$ is normal in H). Moreover,

$$H/H' \simeq (H \cap K)/(H' \cap K) \simeq \gamma(H \cap K)/\gamma(H' \cap K) \simeq \gamma(H \cap K) \cap T$$

using the Nine lemma, the third isomorphism theorem, and the Nine lemma again.

$$\begin{array}{ccccccc} & \downarrow & & \downarrow & & \downarrow & \\ 1 & \longrightarrow & H' \cap K & \longrightarrow & H' & \longrightarrow & \alpha(H') \longrightarrow 1 \\ & & \downarrow & & \downarrow & & \downarrow \\ 1 & \longrightarrow & H \cap K & \longrightarrow & H & \longrightarrow & \alpha(H) \longrightarrow 1 \\ & & \downarrow & & \downarrow & & \downarrow \\ 1 & \longrightarrow & (H \cap K)/(H' \cap K) & \xrightarrow{\sim} & H/H' & \longrightarrow & 1 \\ & & \downarrow & & \downarrow & & \\ & & 1 & & 1 & & \end{array} \quad \begin{array}{ccccccc} & \downarrow & & \downarrow & & \downarrow & \\ 1 & \longrightarrow & \gamma(H' \cap K) & \xrightarrow{\sim} & \beta(H' \cap K) & \longrightarrow & 1 \\ & & \downarrow & & \downarrow & & \downarrow \\ 1 & \longrightarrow & \gamma(H \cap K) \cap T & \longrightarrow & \gamma(H \cap K) & \longrightarrow & \beta(H \cap K) \longrightarrow 1 \\ & & \downarrow & & \downarrow & & \downarrow \\ 1 & \longrightarrow & \gamma(H \cap K) \cap T & \xrightarrow{\sim} & \gamma(H \cap K)/\gamma(H' \cap K) & \longrightarrow & 1 \\ & & \downarrow & & \downarrow & & \\ & & 1 & & 1 & & \end{array}$$

It should be noted that everything is effectively computable from generators for H . \square

Finally, we deduce the existence of a group separating $\text{SMM}(G)$ and $\text{RatM}(G)$:

Corollary 2.8. *There exist $m, n \geq 0$ such that $N_{2,m} \times \mathbb{Z}^n$ has decidable Submonoid Membership Problem and undecidable Rational Subset Membership Problem.*

Proof. Consider a group $G = N_{2,m} \times \mathbb{Z}^n$ with undecidable $\text{RatM}(G)$ and m minimal. Such a group does exist by [15], and [25] even proves that $m \leq 26$.

We prove that $\text{SMM}(G)$ is decidable. Using Theorem 2.6, any instance of $\text{SMM}(G)$ reduces to finitely many instances of $\text{RatM}(H)$ for a subgroup $H \leq G$ with $h([H, H]) < h([G, G])$. However Lemma 2.7 tells us such subgroups admit finite-index subgroups $H' \simeq N_{2,k} \times \mathbb{Z}^\ell$ with $k < m$ and $\ell \leq \binom{m}{2} - \binom{k}{2} + n$. In turn, $\text{RatM}(H)$ reduces to $\text{RatM}(N_{2,k} \times \mathbb{Z}^\ell)$ using Lemma 1.2(a-b) and Proposition 1.1 (see also [14]), and the latter is decidable as $k < m$. \square

Remark 2.9. It is important that subgroups of $N_{2,m} \times \mathbb{Z}^n$ fall into finitely many abstract commensurability classes, otherwise $\text{RatM}(H)$ could be decidable for each group H without any uniform algorithm working for all H . Here, we can compute the values of k and ℓ , and decide which of the finitely many algorithms to apply.

3. REDUCTION TO BOUNDED REGULAR LANGUAGES

3.1. The Heisenberg group. We recall the exponential coordinates on $H_3(\mathbb{Z})$, and a geometrical interpretation for them. The group $H_3(\mathbb{Z})$ is generated by the matrices

$$x = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Another notable element is $z := [x, y] = xyx^{-1}y^{-1}$. For a general element

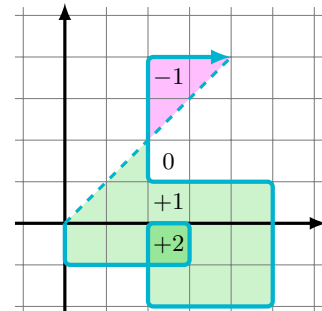
$$g = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \in H_3(\mathbb{Z}),$$

we define $\hat{g} = (a, b) \in \mathbb{Z}^2$ and $A(g) = c - \frac{1}{2}ab$. Note that $\pi: g \mapsto \hat{g}$ is the abelianization map. For instance $\hat{z} = \mathbf{0}$ (since $z \in [G, G]$) and $A(z) = 1$.

Given a product $g = s_1 s_2 \dots s_n$ with $s_i \in \{x^\pm, y^\pm\}$, the coordinates \hat{g} and $A(g)$ can be interpreted as follows: we consider the path γ_g in \mathbb{Z}^2 , starting from $\mathbf{0}$, where the i -th step goes in left, right, up or down direction depending if $s_i = x, x^{-1}, y$ or y^{-1} . Then

- \hat{g} is the endpoint of this path, and
- $A(g)$ is the algebraic area (or *balayage* area) of the curve obtained by closing the path by a straight segment from \hat{g} to $\mathbf{0}$. Each point of the plane is weighted according to the winding number of the path around it.

The adjacent figure is the path representing the product $g = y^{-1}x^3yx^{-1}y^{-2}x^3y^3x^{-3}y^3x^2$, with exponential coordinates given by $\hat{g} = (4, 4)$ and $A(g) = 12$.



Lemma 3.1 (See [2, Proposition 4.1]). *The exponential coordinates satisfy*

$$\widehat{gh} = \widehat{g} + \widehat{h} \quad \text{and} \quad A(gh) = A(g) + A(h) + \frac{1}{2} \det(\widehat{g}; \widehat{h}).$$

As a corollary, we also have

$$A(g_1 g_2 \dots g_n) = \frac{1}{2} \sum_{i < j} \det(\widehat{g}_i; \widehat{g}_j) + \sum_{i=1}^n A(g_i) \quad \text{and} \quad [g, h] = z^{\det(\widehat{g}; \widehat{h})}.$$

Lemma 3.2. *Let $N \trianglelefteq G = H_3(\mathbb{Z})$ be a finite-index normal subgroup. Fix $a, b \in N$ such that $\pi(N) = \langle \widehat{a}, \widehat{b} \rangle$ and $d \in \mathbb{Z}_{>0}$ such that $N \cap [G, G] = \langle z^d \rangle$. Two elements $g, h \in H_3(\mathbb{Z})$ lie in the same N -coset if and only if the following conditions hold:*

- (1) $\widehat{g} - \widehat{h} \in \langle \widehat{a}, \widehat{b} \rangle$, say $\widehat{g} - \widehat{h} = m\widehat{a} + n\widehat{b}$ with $m, n \in \mathbb{Z}$.
- (2) $A(g) \equiv A(h \cdot a^m b^n) = A(h) + A(a^m b^n) + \frac{1}{2} \det(\widehat{h}; m\widehat{a} + n\widehat{b}) \pmod{d}$.

In particular, if $g = h$ in $G/(N \cap [G, G])$ (i.e., $m = n = 0$), then $A(g) - A(h) \in d\mathbb{Z}$. For $N = G$, this means that $A(g) - A(h) \in \mathbb{Z}$ as soon as $\widehat{g} = \widehat{h}$.

Proof. Condition (1) is clearly necessary. Supposing (1), we have $a^m b^n \in N$ and $g[G, G] = ha^m b^n [G, G]$, hence $gN = hN$ is equivalent to $g\langle z^d \rangle = ha^m b^n \langle z^d \rangle$, which is condition (2). \square

3.2. Reduction to rational submonoids. The goal of §3.2, §3.3 is to prove the following:

Proposition 3.3. *Let $R \subseteq H_3(\mathbb{Z})$ be a rational subset, i.e., $R = \text{ev}(\mathcal{L})$ for some regular language $\mathcal{L} \subseteq G^*$. There exists a bounded regular language $\mathcal{L}' \subset G^*$ such that $R = \text{ev}(\mathcal{L}')$. Moreover, an automaton for \mathcal{L}' can be effectively computed from an automaton for \mathcal{L} .*

We begin by reducing to the case $\text{start} = \text{accept} = \{v\}$ (i.e., to the case of rational submonoids). For this step, G can still be an arbitrary group. We consider a general rational subset $R = \text{ev}(\mathcal{L})$, where \mathcal{L} is accepted by a automaton $M = (V, \delta, v_0, \text{accept})$ over G .

► Using a loop-erasure algorithm, we decompose each word $w \in \mathcal{L}$ as

$$w = w_0 s_1 w_1 s_2 \dots s_\ell w_\ell$$

where $s_1, s_2, \dots, s_\ell \in G$ label a simple path $v_0 \xrightarrow{s_1} v_1 \xrightarrow{s_2} \dots \xrightarrow{s_\ell} v_\ell$ with $v_\ell \in \text{accept}$, and each $w_i \in G^*$ labels a cycle $v_i \rightarrow v_i$. We proceed as follows: start at the vertex v_0 and skip directly to the last visit of v_0 , hence bypassing a (possibly empty) cycle w_0 from v_0 to v_0 , then go to the next vertex. Each time you enter a new vertex v_i , skip directly to the last visit of v_i (bypassing another cycle), then keep going. (See Figure 3.)

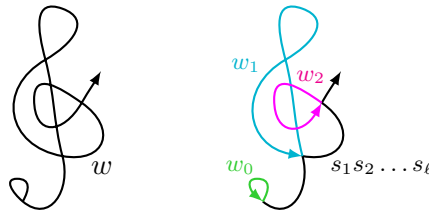


FIGURE 3. A path in the automaton and its decomposition

► It follows that

$$\mathcal{L} = \bigcup \mathcal{L}_{v_0 \rightarrow v_0} \cdot s_1 \cdot \mathcal{L}_{v_1 \rightarrow v_1} \cdot s_2 \cdot \dots \cdot s_\ell \cdot \mathcal{L}_{v_\ell \rightarrow v_\ell}$$

where the union is taken over simple paths $v_0 \xrightarrow{s_1} v_1 \xrightarrow{s_2} \dots \xrightarrow{s_\ell} v_\ell$ with $v_\ell \in \text{accept}$, and $\mathcal{L}_{v \rightarrow v}$ is the language of words labeling cycles $v \rightarrow v$. The union is finite, and each language $\mathcal{L}_{v \rightarrow v}$ is regular (accepted by the automaton $M_{v \rightarrow v} = (V, \delta, v, \{v\})$).

If we managed to find *bounded* regular languages $\mathcal{L}'_{v \rightarrow v}$ such that $\text{ev}(\mathcal{L}'_{v \rightarrow v}) = \text{ev}(\mathcal{L}_{v \rightarrow v})$, we would be done with the language

$$\mathcal{L}' = \bigcup_{s_1 s_2 \dots s_\ell} \mathcal{L}'_{v_0 \rightarrow v_0} \cdot s_1 \cdot \mathcal{L}'_{v_1 \rightarrow v_1} \cdot s_2 \cdot \dots \cdot s_\ell \cdot \mathcal{L}'_{v_\ell \rightarrow v_\ell}$$

which is bounded, regular, and evaluates to R . This is the subject of the next subsection.

3.3. Main discussion. We now fix $G = H_3(\mathbb{Z})$, and consider a trim automaton $M = (V, \delta, v, \{v\})$ recognizing a language \mathcal{L} and a rational subset R . In particular \mathcal{L} is a submonoid of G^* , and R is a submonoid of $H_3(\mathbb{Z})$. As in Proposition 1.1, let

$$X = \left\{ \bar{t} \bar{u} \bar{t}^{-1} \left| \begin{array}{l} t \in G^* \text{ labels a simple path } v \rightarrow p \\ \exists p \in V \text{ s.t. } u \in G^* \text{ labels a simple cycle } p \rightarrow p \\ \text{both paths only intersect at } p \end{array} \right. \right\}.$$

We will work with both \mathcal{L} and $\tilde{\mathcal{L}}$ defined in Proposition 1.1. (Note that $Y = \{\varepsilon\}$.)

Let $\pi: H_3(\mathbb{Z}) \rightarrow \mathbb{Z}^2: g \mapsto \hat{g}$ be the abelianization map. We discuss depending on the subset positively spanned by $\pi(X)$, i.e., depending on $\{\lambda_1 y_1 + \dots + \lambda_r y_r \mid \lambda_i \in \mathbb{R}_{\geq 0}, y_i \in \pi(X)\}$.

- (1) If $\pi(X)$ is included in a line.
- (2) If $\pi(X)$ spans the whole plane (i.e., $\mathbf{0}$ belong to the interior of $\text{ConvHull}(\pi(X))$).
- (3) If $\pi(X)$ spans a half-plane.
- (4) If $\pi(X)$ spans a cone.

In each case, we provide an (effectively computable) bounded regular language \mathcal{L}' such that $\text{ev}(\mathcal{L}') = \text{ev}(\mathcal{L})$. Case (4) will take most of our time.

3.3.1. $\pi(X)$ spans $\{\mathbf{0}\}$, a ray or a line. The subgroup $\langle X \rangle$ is abelian (isomorphic to $\{e\}$, \mathbb{Z} or \mathbb{Z}^2 , but one might as well work in \mathbb{Z}^X). As R is rational in $\langle X \rangle$ (Proposition 1.1), we can compute a bounded regular language $\mathcal{L}' \subset \langle X \rangle^*$ representing R (Corollary 1.5).

3.3.2. $\pi(X)$ spans the whole plane. For each $x = \bar{t} \bar{u} \bar{t}^{-1} \in X$ and $m \in \mathbb{Z}_{\geq 0}$, the set R contains $x^m \text{ev}(tw_{p \rightarrow v}) = \text{ev}(tu^m w_{p \rightarrow v})$, where $w_{p \rightarrow v}$ labels a path $p \rightarrow v$. Note that

$$\pi(x^m \text{ev}(tw_{p \rightarrow v})) = m\pi(x) + \pi(\text{ev}(tw_{p \rightarrow v})) = m\pi(x) + O(1).$$

It follows that $\text{ConvHull}(\pi(R)) = \mathbb{R}^2$, hence R is a subgroup by Proposition 2.5. More precisely $R = \langle X \rangle$ since $X \subseteq R \cdot R^{-1}$ (proof of Proposition 1.1). We compute

- $a, b \in \langle X \rangle$ such that $\pi(a), \pi(b)$ form a basis of $\pi(\langle X \rangle) \leq \mathbb{Z}^2$. These can be found using Gaussian elimination on the matrix with vectors $\pi(x)$ ($x \in X$) as rows.
- $z^d \in \langle X \rangle$ such z^d is a basis of $\langle X \rangle \cap [G, G]$. (d is the smallest positive integer such that $z^d \in \langle X \rangle$). Check if z, z^2, \dots belong to $\langle X \rangle$, for instance using the solution to the Subgroup Membership [22]. The algorithm terminates since $z^{\det(\pi(a); \pi(b))} = [a, b] \in \langle X \rangle$.)

These two conditions mean that $\{a, b, z^d\}$ is a Mal'cev basis of $\langle X \rangle$. We deduce a regular normal form $\mathcal{L}' = \{a^p b^q (z^d)^r \mid p, q, r \in \mathbb{Z}\} \subset G^*$ for R .

3.3.3. $\pi(X)$ spans a half-plane. By hypothesis, we can find

- $\bar{s} \bar{a} \bar{s}^{-1}$ and $\bar{t} \bar{c} \bar{t}^{-1} \in X$ such that $\hat{a}, \mathbf{0}$ and $\hat{c} \in \mathbb{Z}^2$ are on a line, in that order. We also fix $\tilde{s} \in G^*$ labeling a path from the endpoint of s to v . Define $\tilde{t} \in G^*$ similarly.
- $b \in \mathcal{L}$ such that \hat{b} doesn't lie on the same line. (Take $\bar{r} \bar{u} \bar{r}^{-1} \in X$ such that \hat{u} doesn't lie on the line. Fix $\tilde{r} \in G^*$ from the endpoint of r to v , and take $b = ru\tilde{r}$.)

For each $\mathbf{v} \in \mathbb{Z}^2$, there exists unique $\alpha(\mathbf{v}), \beta(\mathbf{v}) \in \mathbb{Q}$ such that $\mathbf{v} = \alpha(\mathbf{v}) \cdot \hat{a} + \beta(\mathbf{v}) \cdot \hat{b}$.

Lemma 3.4. *There exists a computable $K \geq 0$ such that, if $g \in R$ satisfies $\beta(\hat{g}) \geq K$, then g can be written as*

$$g = \text{ev}(sa^{n_1} \tilde{s} \cdot b^m \cdot tc^{n_2} \tilde{t} \cdot b^m \cdot sa^{n_3} \tilde{s} \cdot b^{n_4} \cdot w)$$

where m is fixed, $n_1, n_2, n_3, n_4 \in \mathbb{Z}_{\geq 0}$, and w varies in a fixed finite subset of \mathcal{L} .

Proof. Let $h(n_1, n_2, n_3, n_4) = \text{ev}(sa^{n_1} \tilde{s} \cdot b^m \cdot tc^{n_2} \tilde{t} \cdot b^m \cdot sa^{n_3} \tilde{s} \cdot b^{n_4})$. We fix $p, q \in \mathbb{Z}_{>0}$ such that $p\hat{a} + q\hat{c} = \mathbf{0}$. In particular, we have

$$\hat{h}(n_1 + p, n_2 + q, n_3, n_4) = \hat{h}(n_1, n_2, n_3, n_4) = \hat{h}(n_1, n_2 + q, n_3 + p, n_4). \quad (3.1)$$

Let us see how the area changes under the same transformations:

$$A(h(n_1 + p, n_2 + q, n_3, n_4)) - A(h(n_1, n_2, n_3, n_4)) = \underbrace{pA(a) + qA(c) + \det(p\hat{a}; m\hat{b} + \pi(\tilde{s}t))}_{d_+(m)}$$

$$A(h(n_1, n_2 + q, n_3 + p, n_4)) - A(h(n_1, n_2, n_3, n_4)) = \underbrace{pA(a) + qA(c) - \det(p\hat{a}; m\hat{b} + \pi(\tilde{t}s))}_{d_-(m)}$$

Equation (3.1) and Lemma 3.2 imply that $d_+(m)$ and $d_-(m)$ are integers. We fix m large enough so that $d_+(m) \cdot d_-(m) < 0$ and fix $d = \gcd(d_+(m), d_-(m))$.

The subgroup $N := \langle \bar{a}^d, \bar{b}^d, \bar{c}^d, z^d \rangle$ has finite index (Proposition 2.5). Moreover

$$\forall g \in N, g' \in G, \quad [g, g'] = z^{\det(\hat{g}; \hat{g}')} \in \langle z^d \rangle$$

as $\hat{g} \in d\langle \hat{a}, \hat{b}, \hat{c} \rangle$. It follows that $[N, G] \leq N$: the subgroup N is normal. The quotient Cayley graph can be constructed using Lemma 3.2. Using Lemma 1.2, we can construct a finite subset $\mathcal{K} \subset \mathcal{L}$ of representatives for each coset of N intersecting R .

We take $K = \beta(\pi(h(0, 0, 0, 0))) + \max_{w \in \mathcal{K}} \beta(\hat{w})$. Recall that R is a submonoid in G , hence its image in the finite quotient G/N is a subgroup. It follows that, for any $g \in R$, we can find $w \in \mathcal{K}$ such that $\bar{w} = h(0, 0, 0, 0)^{-1}g$ in G/N (as $g, h(0, 0, 0, 0) \in R$).

If furthermore $\beta(\hat{g}) \geq K$, there exists $m_1, m_2, m_4 \in \mathbb{Z}_{\geq 0}$ such that

$$\begin{aligned} \pi(g) &= \pi(h(0, 0, 0, 0)) + d(m_1\pi(\bar{a}) + m_2\pi(\bar{c}) + m_4\pi(\bar{b})) + \pi(\bar{w}) \\ &= \pi(h(dm_1, dm_2, 0, dm_4)\bar{w}) \end{aligned}$$

(We can ensure $m_1, m_2 \geq 0$ as $\pi(\bar{a}), \pi(\bar{c})$ are colinear, in opposite direction. Moreover, $dm_4 = \beta(\hat{g}) - \beta(\pi(h(0, 0, 0, 0))) - \beta(\hat{w}) \geq \beta(\hat{g}) - K \geq 0$.) Since $g, h(0, 0, 0, 0)\bar{w}$ and $h(dm_1, dm_2, 0, dm_4)\bar{w}$ lie in the same coset of N and $\hat{g} = \hat{h}(dm_1, dm_2, 0, dm_4)$, we have

$$A(g) \equiv A(h(dm_1, dm_2, 0, dm_4)\bar{w}) \pmod{d}$$

(Lemma 3.2). Using the two transformations described above, we can find $n_1, n_2, n_3, n_4 \in \mathbb{Z}_{\geq 0}$ such that $g = h(n_1, n_2, n_3, n_4)\bar{w}$. \square

Finally, we can decompose into two (effectively computable) regular languages

$$\mathcal{L} = \{w \in \mathcal{L} \mid \beta(\hat{w}) \geq K\} \sqcup \{w \in \mathcal{L} \mid \beta(\hat{w}) < K\} =: \mathcal{L}_{\text{reg}} \sqcup \mathcal{L}_{\text{abn}}.$$

- The first term can be replaced by a bounded regular language

$$\mathcal{L}'_{\text{reg}} = \{sa^{n_1}\tilde{s} \cdot b^m \cdot tc^{n_2}\tilde{t} \cdot b^m \cdot sa^{n_3}\tilde{s} \cdot b^{n_4} \cdot w \mid w \in \mathcal{K}, n_1, n_2, n_3, n_4 \in \mathbb{Z}_{\geq 0}\} \subseteq \mathcal{L}.$$

Using Lemma 3.4, we have $\text{ev}(\mathcal{L}_{\text{reg}}) \subseteq \text{ev}(\mathcal{L}'_{\text{reg}}) \subseteq \text{ev}(\mathcal{L}) = R$.

- For the second term, we compute a trim automaton for \mathcal{L}_{abn} and compute the set X associated to each strongly connected component. For each component, $\pi(X)$ is contained in the line through $\mathbf{0}$ and $\pi(\bar{a})$. (Otherwise we would be able to pump along a cycle u for which $\beta(\hat{u}) > 0$ hence produce words $w \in \mathcal{L}_{\text{abn}}$ with $\beta(\hat{w})$ arbitrarily large.)

We can therefore apply the arguments of §3.2 and §3.3.1 to get a bounded regular language $\mathcal{L}'_{\text{abn}}$ such that $\text{ev}(\mathcal{L}_{\text{abn}}) = \text{ev}(\mathcal{L}'_{\text{abn}})$.

The language we are looking for is $\mathcal{L}' = \mathcal{L}'_{\text{reg}} \cup \mathcal{L}'_{\text{abn}}$.

3.3.4. $\pi(X)$ spans a cone. We construct a bounded regular language \mathcal{L}'_+ such that

$$\{g \in \text{ev}(\mathcal{L}) \mid A(g) \geq 0\} \subseteq \text{ev}(\mathcal{L}'_+) \subseteq \text{ev}(\mathcal{L}).$$

Obviously, we can do the same thing for elements of negative area, hence taking the union of both languages gives the desired bounded regular language \mathcal{L}' .

Consider

- tat^{-1} with t a simple path from v to some p , and a a simple loop from p to p such that $\pi(\bar{a})$ belongs to the lower side of the cone, and $\frac{A(tat^{-1})}{\|\pi(\bar{a})\|}$ is maximized, where $\|\cdot\|$ is your favorite norm on \mathbb{R}^2 . We also fix $\tilde{t} \in G^*$ labeling a path back from p to v .
- $s^{-1}bs$ with s a simple path from some q to v , and b a simple loop from q to q such that $\pi(\bar{b})$ belongs to the upper side of the cone, and $\frac{A(s^{-1}bs)}{\|\pi(\bar{b})\|}$ is maximized. We also fix $\tilde{s} \in G^*$ labeling a path from v to q .

For each $\mathbf{v} \in \mathbb{Z}^2$, there exists unique $\alpha(\mathbf{v}), \beta(\mathbf{v}) \in \mathbb{Q}$ such that $\mathbf{v} = \alpha(\mathbf{v}) \cdot \hat{a} + \beta(\mathbf{v}) \cdot \hat{b}$. Moreover, \mathbf{v} belongs to the cone if and only if $\alpha(\mathbf{v}) \geq 0$ and $\beta(\mathbf{v}) \geq 0$.

3.3.4.i. *There does not exist $x \in X$ such that $\pi(x) = \mathbf{0}$ and $A(x) > 0$.*

This is the difficult case. If \hat{g} is far from the border of the cone, and $A(g)$ is far from its maximum (for a fixed value of \hat{g}), then there is enough leeway to find a word representing g in a bounded sub-language of \mathcal{L} which we are going to specify.

Lemma 3.5. *There exist computable $K, m \geq 0$ such that, if $g \in R$ satisfies*

$$0 \leq A(g) \leq \frac{1}{2} \det(\hat{a}; \hat{b}) \cdot \alpha(\hat{g})\beta(\hat{g}) + A(tat^{-1}) \cdot \alpha(\hat{g}) + A(s^{-1}bs) \cdot \beta(\hat{g}) - K,$$

$$\alpha(\hat{g}) \geq K \quad \text{and} \quad \beta(\hat{g}) \geq K,$$

then g can be written as

$$g = \text{ev}(ta^{n_1}\tilde{t} \cdot \tilde{s}b^{n_2}s \cdot ta^{n_3}\tilde{t} \cdot w \cdot \tilde{s}b^{p_1}s \cdot ta\tilde{t} \cdot \tilde{s}b^{p_2}s \cdot \dots \cdot ta\tilde{t} \cdot \tilde{s}b^{p_m}s)$$

where $n_1, n_2, n_3, p_1, p_2, \dots, p_m$ vary in $\mathbb{Z}_{\geq 0}$ and w varies in a fixed finite subset of \mathcal{L} .

Proof. Let

$$h(n_1, n_2, n_3; w; p_1, \dots, p_m) = \text{ev}(ta^{n_1}\tilde{t} \cdot \tilde{s}b^{n_2}s \cdot ta^{n_3}\tilde{t} \cdot w \cdot \tilde{s}b^{p_1}s \cdot ta\tilde{t} \cdot \tilde{s}b^{p_2}s \cdot \dots \cdot ta\tilde{t} \cdot \tilde{s}b^{p_m}s).$$

The operation $(p_i, p_{i+1}) \rightarrow (p_i + 1, p_{i+1} - 1)$ preserves \hat{h} , and decreases the area by

$$A(h(\dots, p_i, p_{i+1}, \dots)) - A(h(\dots, p_i + 1, p_{i+1} - 1, \dots)) = \det(\pi(s \cdot ta\tilde{t} \cdot \tilde{s}); \hat{b}).$$

Note that $st\tilde{t}\tilde{s}$ labels a cycle in the automaton, so $\alpha(\pi(st\tilde{t}\tilde{s})) \geq 0$. Let

$$d = \det(\pi(s \cdot ta\tilde{t} \cdot \tilde{s}); \hat{b}) = (1 + \alpha(\pi(st\tilde{t}\tilde{s}))) \cdot \det(\hat{a}; \hat{b}) \geq \det(\hat{a}; \hat{b}) > 0.$$

We also fix $m = 2d + 1$. We consider $N = \langle \bar{a}^d, \bar{b}^d, z^d \rangle$. Using Lemma 1.2(c), we compute a finite set $\mathcal{K} \subset \mathcal{L}$ containing a representative for each coset of N intersecting R .

For each $g \in R$ satisfying the condition from Lemma 3.5, we find parameters so that $h(n_1, n_2, n_3; w; p_1, \dots, p_m) = g$. We split this into four steps:

(1) We find $w \in \mathcal{K}$ such that $g = h(0, 0, 0; w; 0, \dots, 0)$ in G/N .

(2) For K computably large enough, we find $N_1, P_m \geq 0$ such that

$$g = h(N_1, 0, 0; w; 0, \dots, 0, P_m)$$

in $G/(N \cap [G, G])$. Moreover $A(h(N_1, 0, 0; w; 0, \dots, 0, P_m)) - A(g) \geq 0$.

(3) For K large enough, we find $n_1, n_2, n_3, p_m \geq 0$ with $n_2 \leq \frac{3}{5}\beta(\hat{g})$ and

$$g = h(n_1, n_2, n_3; w; 0, \dots, 0, p_m)$$

in $G/(N \cap [G, G])$. Moreover

$$0 \leq A(h(n_1, n_2, n_3; w; 0, \dots, 0, p_m)) - A(g) \leq d \left(\frac{3}{5}\beta(\hat{g}) + \beta(\pi(\tilde{t}\tilde{s}st)) \right) \det(\hat{a}; \hat{b}).$$

(4) For K large enough, we find $p_1, \dots, p_m \geq 0$ such that $g = h(n_1, n_2, n_3; w; p_1, \dots, p_m)$.

(1) Since the image of the monoid R in G/N is a subgroup, we can find $w \in \mathcal{K}$ such that

$$w = (t\tilde{t} \cdot \tilde{s}s \cdot t\tilde{t})^{-1} \cdot g \cdot ((\tilde{s}s \cdot ta\tilde{t})^{m-1} \cdot \tilde{s}s)^{-1} \quad \text{in } G/N.$$

It follows that $g = h(0, 0, 0; w; 0, \dots, 0)$ in G/N .

(2) In particular, we have $\hat{g} - \hat{h}(0, 0, 0; w; 0, \dots, 0) \in d(\hat{a}, \hat{b})$ (Lemma 3.2). It follows that

$$N_1 = \alpha(\hat{g}) - \alpha(\hat{h}(0, 0, 0; w; 0, \dots, 0)),$$

$$P_m = \beta(\hat{g}) - \beta(\hat{h}(0, 0, 0; w; 0, \dots, 0))$$

are integer multiples of d . Moreover $N_1, P_m \geq 0$ for K computably large enough. We have

$$g = h(N_1, 0, 0; w; 0, \dots, 0, P_m) \quad \text{in } G/(N \cap [G, G]).$$

The area of $h(N_1, 0, 0; w; 0, \dots, 0, P_m)$ is given by

$$A(ta^{N_1} \cdot v \cdot b^{P_m}s) = \frac{1}{2} \det(\hat{a}; \hat{b}) \cdot \alpha(\hat{g})\beta(\hat{g}) + A(tat^{-1}) \cdot \alpha(\hat{g}) + A(s^{-1}bs) \cdot \beta(\hat{g}) - C(w)$$

with $v = \tilde{t}\tilde{s}st\tilde{t}w(\tilde{s}sta\tilde{t})^{m-1}\tilde{s}$, and $C(w)$ is a computable constant that only depends on w . For K large enough, we may therefore assume $A(g) \leq A(h(N_1, 0, 0; w; 0, \dots, 0, P_m))$.

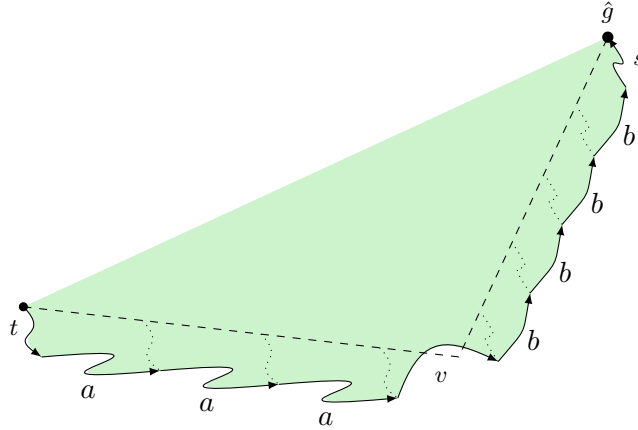


FIGURE 4. The path corresponding to $h(N_1, 0, 0; w; 0, \dots, 0, P_m)$.

Indeed, the large triangle (Figure 4) has area $\frac{1}{2} \det(\hat{a}; \hat{b}) \cdot \alpha(\hat{g})\beta(\hat{g})$. The N_1 (resp. P_m) smaller regions bordered by a (resp. b) have area $A(tat^{-1})$ (resp. $A(s^{-1}bs)$).

(3) Next, we apply two transformations in order:

- $(n_2, p_m) \rightarrow (n_2 + d, p_m - d)$. This doesn't change \hat{h} , and decreases the area by

$$A(h(N_1, n_2, 0; w; 0, \dots, 0, p_m)) - A(h(N_1, n_2 + d, 0; w; 0, \dots, 0, p_m - d))$$

$$= \det\left(\pi(s \cdot t\tilde{t} \cdot w \cdot (\tilde{s}s \cdot ta\tilde{t})^{m-1} \cdot \tilde{s}); d \cdot \pi(b)\right) = d\left((m-1) + \alpha(\pi(w(\tilde{s}st\tilde{t})^m))\right) \det(\hat{a}; \hat{b}) > 0.$$

(Note that $\alpha(\pi(\tilde{s}s)), \alpha(\pi(t\tilde{t})), \alpha(\hat{w}) \geq 0$ since $\tilde{s}s, t\tilde{t}, w \in \mathcal{L}$.) Repeat as long as possible while keeping $n_2 \leq \frac{3}{5}\beta(\hat{g})$ and $A(h) \geq A(g)$. (The first condition ensures that $p_m \geq 0$ since $p_m = P_m - n_2 \geq \frac{2}{5}\beta(\hat{g}) - \beta(\hat{h}(0, 0, 0; w; 0, \dots, 0)) > 0$ for K large enough.)

- $(n_1, n_3) \rightarrow (n_1 - d, n_3 + d)$. This doesn't change \hat{h} , and decreases the area by

$$\begin{aligned} & A(h(n_1, n_2, n_3; w; 0, \dots, 0, p_m)) - A(h(n_1 - d, n_2, n_3 + d; w; 0, \dots, 0, p_m)) \\ &= \det(d\pi(a); \pi(\tilde{t} \cdot \tilde{s}b^{n_2}s \cdot t)) = d(n_2 + \beta(\pi(\tilde{t}\tilde{s}st))) \det(\hat{a}; \hat{b}) \geq 0. \end{aligned}$$

Repeat as long as possible while keeping $n_1 \geq 0$ and $A(h) \geq A(g)$.

At the end of this process, we have found an element $h(n_1, n_2, n_3; w; 0, \dots, 0, p_m)$ with $n_1, n_2, n_3, p_m \geq 0$ and $n_2 \leq \frac{3}{5}\beta(\hat{g})$, which coincides with g in $G/(N \cap [G, G])$. In particular, $A(h) - A(g) \in d\mathbb{Z}$. Moreover, since we cannot apply the second step anymore, we have

$$0 \leq A(h) - A(g) \leq d(n_2 + \beta(\pi(\tilde{t}\tilde{s}st))) \det(\hat{a}; \hat{b}) \leq d \left(\frac{3}{5}\beta(\hat{g}) + \beta(\pi(\tilde{t}\tilde{s}st)) \right) \det(\hat{a}; \hat{b}).$$

Here we use that $A(g) \geq 0$. If we had continued until the conditions $n_2 \leq \frac{3}{5}\beta(\hat{g})$ and $n_1 \geq 0$ were the limiting conditions, we would have reached an area of

$$A(h) \sim A \left(h \left(0, \frac{3}{5}P_m, N_1; w; 0, \dots, 0, \frac{2}{5}P_m \right) \right) \sim -\frac{1}{10} \det(\hat{a}; \hat{b}) \cdot \alpha(\hat{g})\beta(\hat{g}) < 0 \leq A(g)$$

where \sim means that the quotient converges to 1 when $\min\{\alpha(\hat{g}), \beta(\hat{g})\} \rightarrow \infty$. For K computably large enough, this implies that the condition $A(h) \geq A(g)$ stopped us.

(4) Finally, we use the operation $(p_i, p_{i+1}) \rightarrow (p_i + 1, p_{i+1} - 1)$. We can use this operation up to $(m-1)(P_m - n_2) = 2d(P_m - n_2)$ times, decreasing the area by exactly d each time. We can therefore reduce the area by any multiple of d up to

$$2d \cdot \left(\frac{2}{5}\beta(\hat{g}) - \beta(\hat{h}(0, 0, 0; w; 0, \dots, 0)) \right) \cdot d \geq d \left(\frac{3}{5}\beta(\hat{g}) + \beta(\pi(\tilde{t}\tilde{s}st)) \right) \det(\hat{a}; \hat{b})$$

as $d \geq \det(\hat{a}; \hat{b})$ and $\beta(\hat{g}) \geq K$ with K large enough. Since $A(h) - A(g)$ was an integer multiple of d in the correct range after step (3), we get $A(h) = A(g)$ at some point. \square

As in the Case 3.3.3, we decompose \mathcal{L} into two regular languages, with all elements in the first set treated by Lemma 3.5, and all elements in the second case following the border of the cone (hence being easier to treat). We first need a lemma to take care of the ‘‘area condition’’ with a finite state automaton.

Lemma 3.6. *Consider a word $w \in \mathcal{L}$ and $\tilde{w} = x_1x_2 \dots x_\ell \in X^*$ its decomposition. Let*

$$\begin{aligned} M(w) &= \# \{i \mid 1 \leq i \leq \ell \text{ and } \alpha(x_i), \beta(x_i) > 0\}, \\ N(w) &= \# \{(i, j) \mid 1 \leq i < j \leq \ell \text{ and } \alpha(x_j), \beta(x_i) > 0\}. \end{aligned}$$

There exists a computable constant L such that, if

$$A(\tilde{w}) \geq \frac{1}{2} \det(\hat{a}; \hat{b}) \cdot \alpha(\hat{w})\beta(\hat{w}) + A(tat^{-1}) \cdot \alpha(\hat{w}) + A(s^{-1}bs) \cdot \beta(\hat{w}) - K,$$

then $M(w), N(w) \leq L$.

Proof. Let

$$\begin{aligned} \Delta &= \max \left(\{A(x) \mid x \in X\} \cup \{0\} \right), \\ \delta &= \min \left(\{\alpha(x), \beta(x) \mid x \in X\} \setminus \{0\} \right). \end{aligned}$$

We can bound the area of \bar{w} by

$$\begin{aligned}
A(\bar{w}) &= \sum_{i < j} \frac{1}{2} \det(\hat{x}_i; \hat{x}_j) + \sum_i A(x_i) \\
&= \frac{1}{2} \det(\hat{a}; \hat{b}) \left(\sum_{i < j} \alpha(\hat{x}_i) \beta(\hat{x}_j) - \alpha(\hat{x}_j) \beta(\hat{x}_i) \right) + \sum_i A(x_i) \\
&\leq \frac{1}{2} \det(\hat{a}; \hat{b}) \left(\sum_i \alpha(\hat{x}_i) \cdot \sum_j \beta(\hat{x}_j) - N(w) \cdot \delta^2 \right) + \\
&\quad + \sum_{i: \beta(\hat{x}_i)=0} \alpha(\hat{x}_i) A(tat^{-1}) + \sum_{i: \alpha(\hat{x}_i)=0} \beta(\hat{x}_i) A(s^{-1}bs) + M(w) \cdot \Delta \\
&\leq \frac{1}{2} \det(\hat{a}; \hat{b}) \cdot \alpha(\hat{w}) \beta(\hat{w}) + A(tat^{-1}) \cdot \alpha(\hat{w}) + A(s^{-1}bs) \cdot \beta(\hat{w}) + \\
&\quad + M(w) \cdot \Delta - \frac{1}{2} \det(\hat{a}; \hat{b}) \cdot N(w) \cdot \delta^2.
\end{aligned}$$

The first inequality is the only part using that $A(x) \leq 0$ for all $x \in X$ such that $\pi(x) = \mathbf{0}$, and that $A(tat^{-1})/\|\pi(a)\|$ and $A(s^{-1}bs)/\|\pi(b)\|$ are maximal. This ensures that

- $A(x_i) \leq 0$ if $\alpha(\hat{x}_i) = \beta(\hat{x}_i) = 0$,
- $A(x_i) \leq \alpha(x_i)A(tat^{-1})$ if $\alpha(\hat{x}_i) \neq 0$ and $\beta(\hat{x}_i) = 0$,
- $A(x_i) \leq \beta(x_i)A(s^{-1}bs)$ if $\alpha(\hat{x}_i) = 0$ and $\beta(\hat{x}_i) \neq 0$.

It follows that $M(w) \cdot \Delta - \frac{1}{2} \det(\hat{a}; \hat{b}) \delta^2 \cdot N(w) \geq -K$. Since $\binom{M(w)}{2} \leq N(w)$ we get

$$K \geq \frac{\delta^2 \det(\hat{a}; \hat{b})}{4} \cdot M(w)(M(w) - 1) - \Delta \cdot M(w)$$

hence $M(w) \leq L$ for some computable L . Finally $N(w) \leq \frac{2}{\delta^2 \det(\hat{a}; \hat{b})} (M(w) \cdot \Delta + K)$. \square

We decompose the language $\tilde{\mathcal{L}}$ given by Proposition 1.1 into two regular subsets:

$$\mathcal{L}_{\text{reg}} := \left\{ \tilde{w} \in \tilde{\mathcal{L}} \mid \alpha(\hat{w}), \beta(\hat{w}) \geq K \text{ and } N(w) \geq L \right\},$$

and $\mathcal{L}_{\text{abn}} := \tilde{\mathcal{L}} \setminus \mathcal{L}_{\text{reg}}$.

- We replace \mathcal{L}_{reg} by the bounded regular language

$$\mathcal{L}'_{\text{reg}} = \bigcup_{w \in \mathcal{K}} ta^* \tilde{t} \cdot \tilde{s} b^* s \cdot ta^* \tilde{t} \cdot w \cdot (\tilde{s} b^* s \cdot ta \tilde{t})^{m-1} \cdot \tilde{s} b^* s,$$

so that $\{g \in \text{ev}(\mathcal{L}_{\text{reg}}) \mid A(g) \geq 0\} \subseteq \text{ev}(\mathcal{L}'_{\text{reg}}) \subseteq R$.

- For the second term, we compute a trim automaton for \mathcal{L}_{abn} and compute the set X associated to each strongly connected component. For each component, $\pi(X)$ is contained in the line through $\pi(a)$ or through $\pi(b)$. We can apply §3.2 and §3.3.1 to get a bounded regular language $\mathcal{L}'_{\text{abn}}$ such that $\text{ev}(\mathcal{L}'_{\text{abn}}) = \text{ev}(\mathcal{L}_{\text{abn}})$.

The language we are looking for is $\mathcal{L}'_+ = \mathcal{L}'_{\text{reg}} \cup \mathcal{L}'_{\text{abn}}$.

3.3.4.ii. *There does exist $x \in X$ such that $\pi(x) = \mathbf{0}$ and $A(x) > 0$.*

This case is similar to §3.3.3. We fix $r, c, \tilde{r} \in G^*$ labeling a path $v \rightarrow p$, a cycle $p \rightarrow p$ and a path $p \rightarrow v$ respectively, such that $\hat{c} = \mathbf{0}$ and $A(\bar{c}) > 0$.

Lemma 3.7. *There exists a computable $K \geq 0$ such that, if $g \in R$ satisfies*

$$A(g) \geq 0, \quad \alpha(\hat{g}) \geq K \quad \text{and} \quad \beta(\hat{g}) \geq K,$$

then g can be written as $g = \text{ev}(\tilde{s}b^{n_1}s \cdot ta^{n_2}\tilde{t} \cdot rc^{n_3}\tilde{r} \cdot w)$ where n_1, n_2, n_3 varies in $\mathbb{Z}_{\geq 0}$ and w varies in a fixed finite subset of \mathcal{L} .

Proof. Let $d = A(c) > 0$. We introduce the notation

$$h(n_1, n_2, n_3) = \text{ev}(\tilde{s}b^{n_1}s \cdot ta^{n_2}\tilde{t} \cdot rc^{n_3}\tilde{r}).$$

We consider the finite-index normal subgroup $N = \langle \bar{a}^d, \bar{b}^d, \bar{c} = z^d \rangle$. Using Lemma 1.2(c), we construct a finite subset $\mathcal{K} \subset \mathcal{L}$ of representatives for each coset of N intersecting R . For $g \in R$ satisfying the condition from the Lemma 3.7, we take

- $w \in \mathcal{K}$ such that $\bar{w} = h(0, 0, 0)^{-1}g$ in G/N .
- $n_1 = \alpha(\pi(g)) - \alpha(\pi(h(0, 0, 0)\bar{w}))$ and $n_2 = \beta(\hat{g}) - \beta(\pi(h(0, 0, 0)\bar{w}))$. These are integer multiples of d (Lemma 3.2). Moreover, they are non-negative for K large enough.
- $n_3 = \frac{1}{d}(A(g) - A(h(n_1, n_2, 0) \cdot \bar{w}))$. Note that n_3 is an integer since

$$g = h(n_1, n_2, 0)\bar{w} \quad \text{in } G/(N \cap [G, G]).$$

Moreover, $n_3 \geq 0$ for K large enough. Indeed, we have $n_1, n_2 \geq K - C_1(w)$ hence

$$A(h(n_1, n_2, 0) \cdot \bar{w}) = -\frac{1}{2} \det(\hat{a}; \hat{b}) \cdot n_1 n_2 + C_2(w) \cdot n_1 + C_3(w) \cdot n_2 + A(w) \leq 0 \leq A(g)$$

for K large enough, where $C_1(w), C_2(w), C_3(w)$ are constant depending on w .

We conclude with $\pi(h(n_1, n_2, n_3)\bar{w}) = \pi(h(n_1, n_2, 0)\bar{w}) = \pi(g)$ and $A(h(n_1, n_2, n_3)\bar{w}) = A(h(n_1, n_2, 0)\bar{w}) + n_3 d = A(g)$, hence $h(n_1, n_2, n_3)\bar{w} = g$. \square

Finally, we decompose \mathcal{L} into two regular languages

$$\mathcal{L} = \{w \in \mathcal{L} \mid \alpha(\hat{w}), \beta(\hat{w}) \geq K\} \sqcup \{w \in \mathcal{L} \mid \alpha(\hat{w}) < K \text{ or } \beta(\hat{w}) < K\} =: \mathcal{L}_{\text{reg}} \sqcup \mathcal{L}_{\text{abn}}.$$

- The first term can be replaced by a bounded regular language

$$\mathcal{L}'_{\text{reg}} = \{\tilde{s}b^{n_1}s \cdot ta^{n_2}\tilde{t} \cdot rc^{n_3}\tilde{r} \cdot w \mid n_1, n_2, n_3 \geq 0, w \in \mathcal{K}\} \subseteq \mathcal{L}.$$

Using Lemma 3.7, we have $\{g \in \text{ev}(\mathcal{L}_{\text{reg}}) \mid A(g) \geq 0\} \subseteq \text{ev}(\mathcal{L}'_{\text{reg}}) \subseteq \text{ev}(\mathcal{L})$.

- For the second term, we compute a trim automaton for \mathcal{L}_{abn} and compute the set X associated to each strongly connected component. For each component, $\pi(X)$ is contained in the line through $\pi(a)$ or through $\pi(b)$. We can apply §3.2 and §3.3.1 to get a bounded regular language $\mathcal{L}'_{\text{abn}}$ such that $\text{ev}(\mathcal{L}'_{\text{abn}}) = \text{ev}(\mathcal{L}_{\text{abn}})$.

The language we are looking for is $\mathcal{L}'_+ = \mathcal{L}'_{\text{reg}} \cup \mathcal{L}'_{\text{abn}}$.

3.4. Equations under rational constraints. Finally, we prove that $\text{RatM}(H_3(\mathbb{Z}))$ is decidable, using Proposition 3.3. We prove a stronger result about solving equations under rational constraints.

Definition 3.8. An equation with rational constraints in G consists of

- an element $w \in F(x_1, \dots, x_n) * G$ and
- n rational subsets $R_1, \dots, R_n \subseteq G$.

This equation admits a solution if there exists $(g_1, \dots, g_n) \in R_1 \times \dots \times R_n$ such that $w(g_1, \dots, g_n) := f(w) = e_G$, where f is the homomorphism

$$f: \begin{pmatrix} F(x_1, \dots, x_n) * G & \longrightarrow & G \\ g & \longmapsto & g \\ x_i & \longmapsto & g_i \end{pmatrix}.$$

This generalizes the Rational Subset Membership $g \stackrel{?}{\in} R$, as we may consider the equation $x_1 = g$ under the rational constraint $x_1 \in R$.

Theorem 3.9. *There exists an algorithm which takes as input an equation with rational constraints in $H_3(\mathbb{Z})$, and decides whether it admits a solution.*

This extends the analogous result without rational constraints, due to Duchin, Liang and Shapira. The proof is a straightforward adaptation of [9, Theorem 3] and [15, Theorem 6.8], with Proposition 3.3 as a starting point.

Proof. We first prove the statement under the extra assumption that each rational constraint is given as $R_i = h_{i,0} \{k_{i,1}\}^* h_{i,1} \{k_{i,2}\}^* h_{i,2} \dots h_{i,\ell_i-1} \{k_{i,\ell_i}\}^* h_{i,\ell_i}$. (Such a set will be called “Knapsack-like”.) In particular, a generic element of R_i is given as

$$g_i = g_i(n_1, \dots, n_{\ell_i}) = h_{i,0} k_{i,1}^{n_1} h_{i,1} k_{i,2}^{n_2} h_{i,2} \dots h_{i,\ell_i-1} k_{i,\ell_i}^{n_{\ell_i}} h_{i,\ell_i}$$

with $n_j \in \mathbb{N}$. In coordinates, the existence of a solution reduces to the system

$$\begin{cases} \hat{w}(g_1, \dots, g_n) = \mathbf{0} \\ A(w(g_1, \dots, g_n)) = 0 \end{cases}$$

This system consists of two linear and one quadratic equations, with coefficients in $\frac{1}{2}\mathbb{Z}$ and unknowns $n_{i,j} \in \mathbb{N}$, hence we can decide if it admits a solution using [13].

In general, we can write each rational constraint as $R_i = \bigcup_{j=1}^{m_i} R_{ij}$ where each R_{ij} is Knapsack-like using Proposition 3.3. Therefore, we only have to check whether one of $m_1 \dots m_n$ systems with Knapsack-like constraints admits a solution. The final answer is “Yes” if any of these answers is yes, “No” otherwise. \square

Remark 3.10. It should be noted that the result about Rational Subset Membership can be extended to $H_3(\mathbb{Q})$ as every finitely generated subgroup of $H_3(\mathbb{Q})$ is (effectively) isomorphic to a subgroup of $H_3(\mathbb{Z})$. Indeed, given a finite set of matrices

$$M_i = \begin{pmatrix} 1 & a_i & c_i \\ 0 & 1 & b_i \\ 0 & 0 & 1 \end{pmatrix} \in H_3(\mathbb{Q}) \quad (i = 1, 2, \dots, r),$$

find $N \neq 0$ such that $Na_i, Nb_i, N^2c_i \in \mathbb{Z}$. The dilation $\delta_N: (a, b, c) \mapsto (Na, Nb, N^2c)$ is an automorphism $H_3(\mathbb{Q}) \rightarrow H_3(\mathbb{Q})$ such that $\delta_N(\langle M_i \rangle) \leq H_3(\mathbb{Z})$.

4. FURTHER QUESTIONS AND REMARKS

One may wonder for which class of groups can the Rational Subset Membership be reduced to the Knapsack problem (as in Proposition 3.3):

Question D. Characterize groups G such that every rational subset $R \subseteq G$ can be effectively represented by a *bounded* regular language $\mathcal{L}' \subset G^*$. In particular, what if

- (1) G is 2-step nilpotent with cyclic derived subgroup?
- (2) G is 2-step nilpotent?
- (3) $G = E$ is the following “filiform” 3-step nilpotent group

$$E = \langle x, y_1, y_2 \mid [x, y_i] = y_{i+1} \text{ for } i = 1, 2, [x, y_3] = [y_i, y_j] = 1 \rangle ?$$

- (4) Does any group of super-polynomial growth have this property?

► Using Proposition 1.1 and Lemma 1.2, we can prove that this property passes to quotients, subgroups and finite-index overgroups.

► Recall that 2-step nilpotent groups with infinite cyclic derived subgroup contain copies of $H_{2m+1}(\mathbb{Z}) \times \mathbb{Z}^n$ as finite-index subgroups, see [35, Lemma 7.1]. In particular, we only have to consider $G = H_{2m+1}(\mathbb{Z}) \times \mathbb{Z}^n$ to answer Question D.1.

A positive answer to Question D.1 would be particularly interesting as it would imply decidability for the Rational Subset Membership for those groups. Indeed, the Knapsack problem is also decidable for those groups (adapting the proof of [15, Theorem 6.8] from $H_3(\mathbb{Z}) \times \mathbb{Z}^n$ to $H_{2m+1}(\mathbb{Z}) \times \mathbb{Z}^n$, then using [15, Theorem 7.3]).

► An instance of Question D.2 of special interest is

$$G = N_{2,r} = \langle x_1, x_2, \dots, x_r \mid \text{2-step nilpotent} \rangle$$

with $R = \{x_1, x_2, \dots, x_r\}^*$. A positive answer (algorithmic in r) would allow to reduce the Submonoid Membership Problem to the Knapsack Problem in any 2-step nilpotent groups. In turn, this would provide an example of a group with decidable Submonoid Membership and undecidable Knapsack problem (adapting the proof of Corollary 2.8). Note that the answer is positive for $r = 2$, as we have the equality

$$\{x, y\}^* = \{x\}^* \{y\}^* x \{y\}^* \{x\}^* \sqcup \{y\}^*$$

in $N_{2,2} = H_3(\mathbb{Z})$. This equality is one of the key motivations behind §3.3.4.

► The property fails for the free 3-step nilpotent group of rank 2

$$G = N_{3,2} = \langle x, y \mid \text{3-step nilpotent} \rangle.$$

This group admits a path model similar to the model for elements of $H_3(\mathbb{Z})$ defined in §3.1, where we keep track of an additional parameter $B(g) \in \mathbb{R}^2$ (see [4, Section 3.1] for all necessary definitions). We argue that $R = \{x, y\}^*$ cannot be represented by a bounded regular language. Let $\mathcal{L}' \subset G^*$ be a bounded regular language such that $\text{ev}(\mathcal{L}') \subseteq R$, say

$$\mathcal{L}' = \bigcup_{i=1}^I v_{i,0} \{w_{i,1}\}^* v_{i,1} \{w_{i,2}\}^* v_{i,2} \dots v_{i,\ell_i-1} \{w_{i,\ell_i}\}^* v_{i,\ell_i}.$$

Observe that $\hat{w}_{i,j} \in \mathbb{Z}_{\geq 0}^2$, otherwise the word $w = v_{i,0} v_{i,1} \dots v_{i,j-1} w_{i,j}^N v_{i,j} \dots v_{i,\ell_i} \in \mathcal{L}'$ will project to $\hat{w} = N \cdot \hat{w}_{i,j} + O(1) \notin \mathbb{Z}_{\geq 0}^2 = \pi(R)$ for N large enough.

Take a direction $u \in \mathbb{R}_{\geq 0}^2$ which is not proportional to any of the $\hat{w}_{i,j}$ and take an element $\gamma_{u,n} \in \text{ev}\{x, y\}^*$ of length n which best follows the ray \mathbb{R}^+u , as defined in [4, Section 3.2]. Adapting computations from [2, Section 4], we have

$$\langle B(\gamma_{u,n}) - B(\bar{w}); u^\perp \rangle = -O_u(n) + \Theta_u(n^3) = \Theta_u(n^3)$$

for all elements $w \in \mathcal{L}'$ such that $\hat{w} = \hat{\gamma}_{u,n}$. In particular, $\gamma_{u,n} \notin \text{ev}(\mathcal{L}')$ for n large.

ACKNOWLEDGEMENT

I would like to thank Laura Ciobanu for suggesting to look at the Rational Subset Membership in $H_3(\mathbb{Z})$, Alex Levine for asking about constrained equations, and Tatiana Nagnibeda for her constant support. This article has also greatly benefited from the anonymous referee's exhaustive suggestions and comments.

REFERENCES

- [1] Michèle Benois and Jacques Sakarovitch. On the complexity of some extended word problems defined by cancellation rules. *Information Processing Letters*, 23(5):281–287, 1986.
- [2] Coentín Bodart. Intermediate geodesic growth in virtually nilpotent groups. *Groups, Geometry, and Dynamics*, 2025.
- [3] Coentín Bodart, Laura Ciobanu, and George Metcalfe. Ordering groups and the Identity Problem, 2024. <https://arxiv.org/abs/2411.15639>.
- [4] Coentín Bodart and Kenshiro Tashiro. Horofunctions on the Heisenberg and Cartan groups, 2024. <https://arxiv.org/abs/2407.11943>.
- [5] Thomas Colcombet, Joël Ouaknine, Pavel Semukhin, and James Worrell. On reachability problems for low-dimensional matrix semigroups. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132, pages 44:1–44:15, 2019.
- [6] Max Dehn. Über die Topologie des dreidimensionalen Raumes. *Mathematische Annalen*, 69:137–168, 1910.
- [7] Ruiwen Dong. Recent advances in algorithmic problems for semigroups. *ACM SIGLOG News*, 10(4):3–23, 2023.
- [8] Ruiwen Dong. Submonoid membership in n -dimensional lamplighter groups and S-unit equations. *52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025)*, 2025.
- [9] Moon Duchin, Hao Liang, and Michael Shapiro. Equations in nilpotent groups. *Proceedings of the American Mathematical Society*, 143(11):4723–4731, 2015.
- [10] Samuel Eilenberg and Marcel-Paul Schützenberger. Rational sets in commutative monoids. *Journal of Algebra*, 13(2):173–191, 1969.

- [11] Robert H. Gilman. Groups with a rational cross-section. In S. M. Gersten and John R. Stallings, editors, *Combinatorial Group Theory and Topology*, pages 175–184. Princeton University Press, 1987.
- [12] Oleg N. Golovin and N.P. Gol'dina. Subgroups of free metabelian groups. *Matematicheskii Sbornik*, 79(2):323–336, 1955.
- [13] Fritz Grunewald and Dan Segal. On the integer solutions of quadratic equations. *Journal für die reine und angewandte Mathematik*, 2004(569):13–45, 2004.
- [14] Zephian Grunschlag. *Algorithms in Geometric Group Theory*. PhD thesis, University of California, Berkeley, 1999.
- [15] Daniel König, Markus Lohrey, and Georg Zetsche. *Knapsack and subset sum problems in nilpotent, polycyclic, and co-context-free groups*, pages 129–144. American Mathematical Society Providence, RI, 01 2016.
- [16] Markus Lohrey. The rational subset membership problem for groups: a survey. In *Groups St Andrews 2013*, page 368–389. Cambridge University Press, 2015.
- [17] Markus Lohrey. Rational subsets of unitriangular groups. *International Journal of Algebra and Computation*, 25(01n02):113–121, 2015.
- [18] Markus Lohrey. Membership problems in infinite groups. In *Twenty Years of Theoretical and Practical Synergies*, pages 44–59. Springer Nature Switzerland, 2024.
- [19] Markus Lohrey and Benjamin Steinberg. The submonoid and rational subset membership problems for graph groups. *Journal of Algebra*, 320(2):728–755, 2008. Computational Algebra.
- [20] Markus Lohrey and Benjamin Steinberg. Submonoids and rational subsets of groups with infinitely many ends. *Journal of Algebra*, 324(5):970–983, 2010. Computational Algebra.
- [21] Jeremy Macdonald, Alexei Myasnikov, Andrey Nikolaev, and Svetla Vassileva. Logspace and compressed-word computations in nilpotent groups. *Transactions of the American Math. Soc.*, 375:5425–5459, 2022.
- [22] Anatolii Malcev. On homomorphisms onto finite groups. *Transl., Ser. 2, Am. Math. Soc.*, 119:67–79, 1983.
- [23] Yuri V. Matiyasevich. *Hilbert's tenth problem*. MIT Press, Cambridge, MA, USA, 1993.
- [24] K. A. Mihailova. The occurrence problem for direct products of groups. *Doklady Akad. Nauk SSSR*, 119:1103–1105, 1958.
- [25] Alexei Mishchenko and Alexander Treier. Knapsack problem for nilpotent groups. *Groups Complexity Cryptology*, 9(1):87–98, 2017.
- [26] S. Moran. A subgroup theorem for free nilpotent groups. *Transactions of the American Mathematical Society*, 103(3):495–515, 1962.
- [27] Alexei G. Myasnikov, Andrey Nikolaev, and Alexander Ushakov. Knapsack problems in groups. *Mathematics of Computation*, 84:987–1016, 2013.
- [28] Felix Potthast. Submonoid membership for wreath products, 2020. Bachelor thesis, University of Siegen.
- [29] Michael O. Rabin and Dana Scott. Finite automata and their decision problems. *IBM Journal of Research and Development*, 3(2):114–125, 1959.
- [30] Vitaly Roman'kov. Undecidability of the submonoid membership problem for a sufficiently large finite direct power of the Heisenberg group. *Siberian Electronic Mathematical Reports*, 20(1):293–305, 2023.
- [31] Daniel Segal. *Polycyclic Groups*. Cambridge Tracts in Mathematics. Cambridge University Press, 1983.
- [32] Doron Shafir. Bounded generation of submonoids of Heisenberg groups, 2024. <https://arxiv.org/abs/2405.05939>.
- [33] Doron Shafir. Is decidability of the submonoid membership problem closed under finite extensions?, 2024. <https://arxiv.org/abs/2405.12921>.
- [34] Doron Shafir. A saturation theorem for submonoids of nilpotent groups and the Identity Problem, 2024. <https://arxiv.org/abs/2402.07337>.
- [35] Michael Stoll. Rational and transcendental growth series for the higher Heisenberg groups. *Inventiones mathematicae*, 126:85–109, 1996.
- [36] Vladimir I. Trofimov. Growth functions of some classes of languages. *Cybernetics*, 17:727–731, 1981.
- [37] Günter M. Ziegler. *Lectures on Polytopes*. Graduate Texts in Mathematics. Springer New York, 1994.