Séminaire Lotharingien de Combinatoire **86B** (2022) *Proceedings of the 34th Conference on Formal Power*
Article #79, 12 pp. *Series and Algebraic Combinatorics (Bangalore)*

Growth of Unbounded Sets in Nilpotent Groups and Random Mapping Statistics

Be'eri Greenfeld*[1] and Hagai Lavner†[2]

[1] *Department of Mathematics, University of California-San Diego, La Jolla, California 92093, USA*
[2] *Einstein Institute of Mathematics, The Hebrew University of Jerusalem, Israel*

Abstract. For a group G, we asymptotically quantify the maximum number of length-n words over an arbitrary n-letter subset of G. If G is finitely generated and residually finite, then either this function is exponentially bounded, which happens if and only if G is virtually abelian, or else it is bounded from below by $\left(e^{-\frac{1}{4}} + o(1)\right) n^n$; the latter bound cannot be improved, namely, it is attained for the Heisenberg group. For higher-step free nilpotent groups, the asymptotic behavior becomes $(1 - o(1))n^n$. As a key ingredient in the proof, we calculate the number of pair histograms of functions $f\colon [n] \to [n]$ and the probability that a random function $f\colon [n] \to [n]$ can be uniquely determined by its pair histogram. A geometric interpretation of group laws of the Heisenberg group by means of closed paths attached to words in the free group and their projected oriented polygons is given.

Keywords: random functions, growth of groups, nilpotent group, group laws, probabilistic group theory

1 Introduction

Let G be a finitely generated infinite group. The most important large-scale geometric measurement associated with G is its growth rate with respect to a fixed finite (symmetric) generating subset S:

$$\gamma_{G,S}(n) = \# \bigcup_{i=1}^{n} S^i,$$

namely the volume of n-balls in the associated Cayley graph. Nilpotent groups have polynomial growth rate, and conversely, Gromov's celebrated theorem [5] ensures that a finitely generated group with polynomial growth rate is virtually nilpotent. De la Harpe [6] introduced the notion of *uniform* polynomial growth. A group G is said to have a uniform

*bgreenfeld@ucsd.edu. Partially supported by ISF Grant #1994/20.
†hagai.lavner@mail.huji.ac.il. Partially supported by the Kaye Einstein scholarship.

polynomial growth if there exist functions $A, B\colon \mathbb{N} \to \mathbb{N}$ such that for any subset $S \subset G$ of cardinality k we have

$$|S^n| \le A(k)n^{B(k)}.$$

This fits into the framework of the following natural measurement of a (not necessarily finitely generated) group G:

$$\gamma_G^{\max}(k,n) = \sup_{S\subset G,\ |S|\le k} |S^n|.$$

Bożejko [1] proved that nilpotent groups have uniform polynomial growth. By a result of Mann [7], if G has uniform polynomial growth then there exist functions f, g such that every m-generated subgroup of G admits a nilpotent subgroup of nilpotency class $\le f(m)$ and index $\le g(m)$. This can be thought of as a uniform version of Gromov's theorem.

For a group not of uniform polynomial growth, it turns out that (at least for residually finite groups) $\gamma_G^{\max}(k,n)$ tends to be close to the maximum possible value, k^n. Semple and Shalev [8, 9] called a group G for which $\gamma_G^{\max}(n) := \gamma_G^{\max}(n,n) < n^n$ (for some, and hence all sufficiently large n) *collapsing*. A group is collapsing if and only if $\gamma_G^{\max}(k,n) < k^n$ for some k, n.

Any virtually nilpotent group is collapsing, and Semple and Shalev proved the opposite for finitely generated residually finite groups [8]; this was further developed in [9]. In a slightly broader context, recall that a group law of a given group G is a non-trivial word in the free group (on an arbitrary number of generators) vanishing for any substitution from G, and a positive group law is a group law of the form $u = v$ where u, v are positive elements[1] in the free group. Finitely generated residually finite groups, as well as solvable groups which satisfy a positive law are virtually nilpotent [2]. Then [8] fits into the following implication diagram:

$$\begin{tikzcd}
\text{Virtually nilpotent} \arrow[r, Rightarrow] & \text{Positive law} \arrow[r, Rightarrow] & \text{Collapsing} \arrow[r, Rightarrow] \arrow[ll, bend right, "\text{+ f.g. residually finite ([8])}"'] & \text{Group law}
\end{tikzcd}$$

The question of whether every collapsing group (not necessarily residually finite) satisfies a positive group law is open [9, Question 1, page 61].

1.1 Quantifying growth of unbounded subsets in groups

Let G be a finitely generated group. Suppose that G is collapsing, namely, $\gamma_G^{\max}(n) < n^n$ for $n \gg 1$. Then how does $\gamma_G^{\max}(n)$ grow compared to n^n? By the Semple-Shalev theorem, this question reduces from residually finite groups to virtually nilpotent groups. Our main result is the following dichotomy for $\gamma_G^{\max}(n)$:

[1]That is, involve no inverses of the generators; for instance, $xy = yx$ is a positive law, but $[w,x][y,z] = [y,z][w,x]$ cannot be written as a positive law.

Theorem 1.1. *Let G be a finitely generated residually finite group. Then either:*

- $\gamma_G^{\max}(n)$ *is exponentially bounded, which happens if and only if G is virtually abelian; or*
- $\gamma_G^{\max}(n) \geq \left(e^{-\frac{1}{4}} + o(1)\right) n^n$.

The second assertion cannot be improved, namely, it becomes an equality for the integral Heisenberg group.

We denote by $\mathcal{N}_{d,n}$ the n-generated free nilpotent group of class d. Recall that the integral Heisenberg group is:

$$\mathcal{H}_3(\mathbb{Z}) = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \; : \; a, b, c \in \mathbb{Z} \right\} \cong \mathcal{N}_{2,2}$$

henceforth denoted $\mathcal{H}$.

It turns out that the 2-step nilpotent case is somewhat singular:

Theorem 1.2. *Let $\mathcal{N}$ be a finitely generated free nilpotent group of class $c > 2$. Then:*

$$\gamma_{\mathcal{N}}^{\max}(n) = (1 - o(1))n^n.$$

1.2 Statistics of random functions

An interesting interpretation of $\gamma_G^{\max}(n)$ is related to enumerative combinatorics. This interpretation is the key ingredient in our proof of Theorems 1.1 and 1.2.

Fix $f\colon [n] \to [n]$, thought of as a length-n word over the alphabet $[n] = \{1, \ldots, n\}$ (whose elements are thought of as 'letters'), that is,

$$f(1) \cdots \; f(n).$$

We call these *n-fold functions*. Fix $a, b \in [n]$. We denote:

$$m_a^f = \# f^{-1}(a),$$
$$m_{a,b}^f = \#\{1 \leq i < j \leq n \mid f(i) = a, f(j) = b\}.$$

Notice that for $a \neq b$, we have $m_{a,b}^f = m_a^f m_b^f - m_{b,a}^f$, and therefore the data $(m_a^f)_{a\in[n]}$ and $(m_{a,b}^f)_{a,b\in[n]}$ is equivalent to the data $(m_a^f)_{a\in[n]}$ and $(m_{a,b}^f)_{a,b\in[n]\colon\, a<b}$.

We call the data $(m_a^f)_{a\in[n]}$ the *letter histogram* of f, and the data consisting of both $(m_a^f)_{a\in[n]}$ and $(m_{a,b}^f)_{a,b\in[n]}$ the *pair histogram* of f. Denote the set of all pair histograms of n-fold functions by $\mathcal{P}_n$.

A useful observation is that for a free 2-step nilpotent group $\mathcal{N}$, the quantity $\gamma_{\mathcal{N}}^{\max}(n)$ is equal to the number of distinct pair histograms of functions $f\colon [n] \to [n]$; for nilpotent groups of higher class we encounter a modification of that count.

The following is well-known:

Proposition 1.3. *Let $\mathcal{H}$ be the integral Heisenberg group. Then for all $n \geq 1$:*

$$\gamma_{\mathcal{H}}^{\max}(n) = \#\mathcal{P}_n.$$

Moreover, the same equality holds for any free 2-step nilpotent group.

The following might be of independent interest from a purely combinatorial point of view:

Theorem 1.4. *The number of pair histograms of functions $f\colon [n] \to [n]$ is $\left(e^{-\frac{1}{4}} + o(1)\right) n^n$.*

To this end, we calculate the distribution of the number of functions $f\colon [n] \to [n]$ sharing the same pair histogram with a random function, and achieve the following result along the way:

Theorem 1.5. *The probability that a random function $f\colon [n] \to [n]$ is uniquely determined by its pair histogram converges to $e^{-\frac{1}{2}}$ as $n \to \infty$.*

(It is well-known that a permutation is uniquely determined by its pair histogram, which is equivalent to its inversion set; however, the set of all permutations on n letters constitutes of an exponentially small portion of the set of functions $f\colon [n] \to [n]$.)

2 Elementary properties of growth of unbounded subsets

2.1 Elementary properties

We begin with some elementary properties of $\gamma_G^{\max}(k,n)$ which will be freely utilized in the rest of the paper.

Remark 2.1. For any group G:

- $\gamma_G^{\max}(k,n)$ is monotone non-decreasing in each entry;
- $\gamma_G^{\max}(k,n) \leq k^n$;
- If either $H \leq G$ or H is a homomorphic image of G then $\gamma_H^{\max}(k,n) \leq \gamma_G^{\max}(k,n)$.

Proposition 2.2. *Let G be a group containing a non-commutative free subsemigroup. Then:*

$$\gamma_G^{\max}(k,n) = k^n$$

for every k, n. In particular, $\gamma_G^{\max}(n) = n^n$.

Remark 2.3. Having $\gamma_G^{\max}(k,n) = k^n$ for all k,n (equivalently, being non-collapsing) does not imply the existence of a noncommutative free subsemigroup. However, we have the following combinatorics-free characterization of collapsing groups. Recall that two groups are *elementarily equivalent* if they share exactly the same first-order sentences. We can show that a group G is non-collapsing if and only if it is elementarily equivalent to a group containing a noncommutative free subsemigroup.

It turns out that finite index extensions do not significantly affect the growth.

Proposition 2.4. *Let G be a finitely generated infinite group. If $H \leq G$ is a finite index subgroup then there exist some $C_1, C_2 \in \mathbb{N}$ such that:*

$$\gamma_G^{\max}(k,n) \leq C_1 \cdot \gamma_H^{\max}(C_2 k, n).$$

2.2 Abelian groups

Proposition 2.5. *Let G be a finitely generated free abelian group. Then:*

$$\gamma_G^{\max}(k,n) = \binom{n+k-1}{k-1} = \binom{n+k-1}{n}$$

for every k,n. In particular, $\gamma_G^{\max}(n)^{1/n} \xrightarrow{n\to\infty} 4$.

As a consequence:

Corollary 2.6. *Let G be an infinite, finitely generated, virtually abelian group. Then $\gamma_G^{\max}(n)$ grows exponentially.*

In this case, the quantity $\gamma_G(k,n)$ can be interpreted as the number of possible displacements of n balls in k bins. This is valid in any abelian group with infinite exponent.

For a positive integer c, define $b_c(k,n)$ to be the number of possible displacements of n balls in k bins, such that each bin contains less than c balls (bounded capacity).

Proposition 2.7. *Let G be an infinite abelian group of finite exponent and let e be the minimum number such that eG is a finite group (the 'essential' exponent). Then:*

$$\gamma_G^{\max}(k,n) = \Theta\left(\sum_{0\leq r\leq n,\ r\equiv n_{(mod\ e)}} b_e(k,r)\right)$$

It follows that the exponents $\limsup_{n\to\infty} \gamma_G^{\max}(n)^{1/n}$ are determined by the essential exponent e. They form an increasing sequence, starting from 3 and accumulating to 4.

Using the notion of *rewritable groups* (see [4]) it can be shown that an exponential (and in fact, even much higher) behavior of $\gamma_G^{\max}(n)$ is equivalent in the finitely generate case to being virtually abelian:

Proposition 2.8. *Let G be an infinite group. Then either $\gamma_G^{\max}(n) \geq n!$ or $\gamma_G^{\max}(n)$ grows exponentially. If G is finitely generated, the latter case is equivalent to G being virtually abelian.*

3 Partitions and word reconstruction

3.1 Partitions

In order to count distinct pair histograms of n-fold functions, we investigate how much information on a function can be recovered from its pair histogram.

Definition 3.1. For two n-fold functions $f, g\colon [n] \to [n]$ we write $f \sim g$ if they share the same pair histograms, that is: $m_a^f = m_a^g$, $m_{a,b}^f = m_{a,b}^g$ for all $a, b \in [n]$.

Notation 3.2. *If $I = [i,j] \subseteq [n]$ is an interval and $f\colon [n] \to [n]$, we denote by $m_a^{f|_I}$, $m_{a,b}^{f|_I}$ the letter and pair histograms (respectively) of $f|_I\colon I \to [n]$. That is, $m_a^{f|_I}$ counts the number of occurrences of a within $f(I) = f(i)\cdots f(j)$ and $m_{a,b}^{f|_I}$ counts the number of occurrences of the pattern $\star\ a\ \star\ b\ \star$ within $f(i)\cdots f(j)$.*

Suppose we are given an n-fold function f. If $g \sim f$ then f, g share the same *unique letters*, namely letters a for which since $m_a^f = 1$. In this case, we say that $f^{-1}(a)$ is a *unique index*. Unique letters even appear in the same displacement in both f, g:

Lemma 3.3. *If $f \sim g$ and a is a unique letter in f (and hence in g) then $f^{-1}(a) = g^{-1}(a)$.*

Proof. We have
$$f^{-1}(a) = \sum_{b \neq a} m_{b,a}^f + 1 = \sum_{b \neq a} m_{b,a}^g + 1 = g^{-1}(a).$$
□

Definition 3.4. A *partition* Δ of $[n]$ is a collection of disjoint intervals $I_1 = [1, n_1], \ldots, I_k = [n_{k-1}+1, n]$ such that:
$$[n] = \bigcup_{j=1}^{k} I_j.$$
A partition is called *reconstructing* for f if whenever $g \sim f$, we have that $f(I_j) = g(I_j)$ as multisets, for all j. In other words, a partition reconstructs f if the pair histogram of f determines the letter histogram of $g|_{I_j}$, for every j (given that $g \sim f$).

Note that the trivial partition $\Delta = \{[n]\}$ is reconstructing, as the pair histogram contains the letter histogram. On the other hand, if the partition $\Delta = \{\{1\}, \ldots, \{n\}\}$ is reconstructing for f, then $g = f$ for all $g \sim f$.

Lemma 3.5. *If $a_1 = f(i_1), \ldots, a_k = f(i_k)$ are the unique letters of f then:*
$$\Delta_u = \{[1, i_1 - 1], \{i_1\}, [i_1 + 1, i_2 - 1], ..., \{i_k\}, [i_k + 1, n]\}$$
is a reconstructing partition of f.

We can further dissect Δ_u into a finer partition, which is still reconstructing for f. Note that we only know the letter histogram of $g|_I$, rather than its pair histogram. Therefore a naive approach of recursive refinement of each interval would not result in a reconstructing partition. To circumvent that we need the following.

Definition 3.6. Let Δ be a reconstructing partition of f. We say that a letter $a \in [n]$ appearing in an interval I, that is $a \in f(I)$, has *unique neighbors* (in the interval I) if a is unique in I ($m_a^{f|_I} = 1$ and therefore $m_a^{g|_I} = 1$ for every $g \sim f$, since Δ is reconstructing), and for every interval $J \neq I$ such that $a \in f(J)$, we have $f(I) \cap f(J) = \{a\}$.

Lemma 3.7. *Given a partition $\Delta = \{I_1 = [1, n_1], \ldots, I_k = [n_{k-1} + \ 1, n]\}$, which is reconstructing for f, and a letter $a = f(i)$ with unique neighbors in an interval $i \in I_j \in \Delta$, the following refinement of Δ:*

$$\Delta' = \{I_1, \ldots, I_{j-1}, I_{j,L} = [n_{j-1} + 1, \ldots, i - 1], \{i\}, I_{j,R} = [i + 1, n_{j+1} - 1], I_{j+1}, \ldots, I_k\}$$

is reconstructing for f.

We refine Δ_u (from lemma 3.3) iteratively using lemma 3.7, until we remain with a reconstructing partition with respect to which no letter has unique neighbors in any interval. Call such a partition a *terminal partition*. Given f, fix a terminal partition Λ. To conclude:

Proposition 3.8. *For every interval $J \in \Lambda$ and every letter $a \in f(J)$ appearing in J only once (that is, $m_a^{f|_J} = 1$), there exists another letter $b \in f(J)$ with which it appears in another interval $J \neq J' \in \Lambda$ (that is $m_a^{f|_{J'}}, m_b^{f|_{J'}} > 0$).*

3.2 Bounding intervals in the terminal partition

For any n-fold function, let Λ be a terminal partition, as in Proposition 3.8. From now on we pick an n-fold function f chosen uniformly at random (each function $f : [n] \to [n]$ is chosen with probability $\frac{1}{n^n}$). Define two sequences of events:

$$A_n = \{\text{Every interval of } \Lambda \text{ is of length} \leq 10 \ln n\},$$
$$B_n = \{\text{Every interval of } \Lambda \text{ is of length} \leq 2\}.$$

Our goal is to show that $\Pr(B_n)$ converges to 1, but this needs to be done in two steps. First, we show that 'long' intervals occur in the terminal partition with negligible probability. This is achieved by a careful analysis of the correlations of the indicators of appearances of each letter within a restricted interval, utilizing a Chernoff-type bound (essentially a slightly weakened version of a theorem of Imapgliazzo and Kabanets).

Lemma 3.9. *We have*

$$\Pr(A_n) \xrightarrow{n \to \infty} 1.$$

On the other hand, logarithmically bounded intervals have almost full images:

Lemma 3.10. *With probability tending to* 1*, for any* $\ell \leq \lceil 10 \ln n \rceil$ *and for any length-*ℓ *interval* $I \subseteq [n]$ *the set* $f(I)$ *contains at least* $\ell - 1$ *letters.*

These observations are key ingredients in proving the following fundamental proposition, ensuring that only *very small* intervals occur in the terminal partition with non-negligible probability:

Proposition 3.11. *We have*

$$\Pr(B_n) \xrightarrow{n \to \infty} 1.$$

The proof of this result is very technical and combinatorially involved, but we now roughly sketch the general strategy. Recall that in the terminal partition, any pair of letters in each interval share at least one additional interval in which they both appear. By Lemma 3.10, the existence of long intervals in the terminal partition would yield 'too many' coinciding appearances of pairs within additional intervals. This is where Lemma 3.9 gets into the picture: it ensures that the probability that pairs of letters appear together in more than one interval is sufficiently low.

4 The number of pair histograms

4.1 Quantifying the number of pair histograms

We proved that with probability tending to 1 the maximal interval is of length at most 2, which we henceforth assume. We now clarify how this affects the probability that an n-fold function f can be fully reconstructed out of its pair histogram. Let us call such functions *recoverable* (equivalently, those are the functions for which $\{\{1\}, \ldots, \{n\}\}$ is a reconstructing partition).

Fix a length-2 interval $[i, i+1] \in \Lambda$, say, $f(i)f(i+1) = ab$. Then there exists another length-2 interval $f(j)f(j+1) = ab$ or $f(j)f(j+1) = ba$, by Proposition 3.8. We have $|i - j| \geq 2$ since intervals in Λ are disjoint. It can be shown that length-2 intervals split, with probability tending to 1, into *disjoint* pairs of length-2 intervals, which we call *bi-pairs*. Bi-pairs are either *compatible*, namely, the two intervals are the same, or else *incompatible*. In the following example, consider

$$12345612754.$$

The bi-pair **12**3456**12**754 is compatible, whereas 123**45**61275**4** is incompatible. A *switch* of a bi-pair is the function obtained by interchanging the length-2 intervals of the bi-pair, *e.g.*, switching the incompatible bi-pair above we get:

$$123\mathbf{45}6127\mathbf{54} \rightsquigarrow 123\mathbf{54}6127\mathbf{45}.$$

Lemma 4.1.

1. *Incompatible bi-pairs belong to any reconstructing partition, and thus functions with incompatible bi-pairs are not recoverable.*

2. *Suppose that a reconstructing partition Λ consisting only of singletons and disjoint bi-pairs has two intervals $[i, i+1]$, $[j, j+1]$ such that $f(i)f(i+1)$, $f(j)f(j+1)$ form a compatible bi-pair. Then Λ can be refined to a reconstructing partition Λ' containing $\{i\}, \{i+1\}, \{j\}, \{j+1\}$ as intervals.*

Thus, among functions whose terminal partition consists of intervals of length at most two, the underlying function is recoverable if and only if all bi-pairs are compatible. We have the following.

Lemma 4.2. *The distribution of the number of incompatible bi-pairs in a random function converges to Poisson distribution with mean $\lambda = \frac{1}{2}$.*

We are now able to compute the amount of n-fold functions which can be completely recovered out of their pair histograms:

Theorem 4.3 (Theorem 1.5)**.** *The probability that a random function $f\colon [n] \to [n]$ is uniquely determined by its pair histogram converges to $e^{-\frac{1}{2}}$.*

Moreover, we can calculate the number of distinct pair histograms:

Theorem 4.4 (Theorem 1.4)**.** *We have*

$$\#\mathcal{P}_n = \left(e^{-\frac{1}{4}} + o(1)\right) n^n.$$

The constant $e^{-\frac{1}{4}}$ arises as $\mathbb{E}\left(2^{-X}\right)$ where $X \sim Poi\left(\frac{1}{2}\right)$.

5 Growth of unbounded subsets of nilpotent groups

Building on the previous section, we deduce the following from Proposition 1.3 and Theorem 1.4:

Corollary 5.1. *We have*

$$\gamma_{\mathcal{H}}^{\max}(n) = \left(e^{-\frac{1}{4}} + o(1)\right) n^n.$$

Using an algebraic analysis of nilpotent groups, it can be shown that any finitely generated nilpotent group is either virtually abelian, or admits a chain of subgroups and homomorphic images which ends with the Heisenberg group. This gives a proof of Theorem 1.1.

A natural question now is how typical this behavior of $\gamma_G^{\max}(n)$ is for other finitely generated nilpotent groups. It turns out to be very atypical. To analyze growth of unbounded subsets in the 2-generated free 3-nilpotent group $\mathcal{N}_{3,2}$, we carefully refine the notion of pair histogram, with the additional data:

$$m^f_{a,\{b,c\}} = \#\{i < j < k : \ f(i) = a, \ f(j) = \blacklozenge, \ f(k) = \clubsuit, \ \{\blacklozenge, \clubsuit\} = \{b, c\}\}.$$

We show that the growth of subsets in $\mathcal{N}_{3,2}$ is controlled by this modification of pair histograms. We then apply the analysis of reconstructing partitions to quantify the number of such 'finer' histograms.

Proposition 5.2. *We have*

$$(1 - o(1))n^n \le \gamma^{\max}_{\mathcal{N}_{3,2}}(n) \lneq n^n.$$

Any free nilpotent groups of class > 2 homomorphically maps onto $\mathcal{N}_{3,2}$. Thus, any free nilpotent group of class > 2 has 'almost maximal' growth function, as Theorem 1.2 asserts.

6 Geometric interpretation of group laws

Consider an element in the free group $w = x_{i_1}^{\epsilon_1} \cdots x_{i_n}^{\epsilon_n} \in F_d$, where $i_1, \ldots, i_n \in [d]$ and $\epsilon_1, \ldots, \epsilon_n \in \{\pm 1\}$. Then w determines a directed path in $\mathbb{R}^d = \mathrm{Span}_{\mathbb{R}}\{e_1, \ldots, e_d\}$, given by a concatenation of unit vectors:

$$0 \to \epsilon_1 e_{i_1} \to \epsilon_1 e_{i_1} + \epsilon_2 e_{i_2} \to \cdots \to \sum_{j=1}^{n} \epsilon_j e_{i_j}.$$

Denote this path by $\gamma(w)$.

Remark 6.1. A word $w \in F_d$ is a group law for any abelian group if and only if it is a product of commutators, if and only if $\gamma(w)$ is a closed path.

Next, how can one geometrically characterize products of iterated commutators? Let $\pi_{i,j} \colon \mathbb{R}^d \to \mathbb{R}^2$ denote the standard projection onto the plane $\mathrm{Span}_{\mathbb{R}}\{e_i, e_j\}$. If w is indeed a product of iterated commutators then it is in particular a product of commutators, so $\gamma(w)$, and hence all of its projections $\pi_{i,j}(\gamma(w))$, are closed paths. Therefore each $\pi_{i,j}(\gamma(w))$ can be written as $\partial M_{i,j}$ for an oriented polygon $M_{i,j} \subset \mathrm{Span}_{\mathbb{R}}\{e_i, e_j\}$. For instance, the word $w = abccaaba^{-1}c^{-1}b^{-1}b^{-1}a^{-1}a^{-1}c^{-1}$ yields the following $\gamma(w)$ and its projections to the three principal planes:

We have

Theorem 6.2. *Let $w \in F_d$. Then w is a product of iterated commutators if and only if $\gamma(w)$ is a closed path and the signed area bounded by the projection $\pi_{a,b}(\gamma(w))$ is zero for every a, b.*

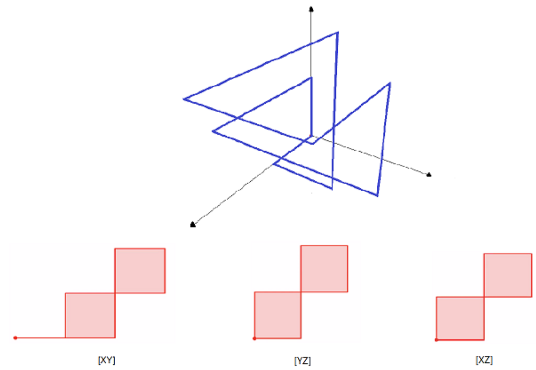

Figure 1: The path $\gamma(abccaaba^{-1}c^{-1}b^{-1}b^{-1}a^{-1}a^{-1}c^{-1})$ and its projections onto the three planes.

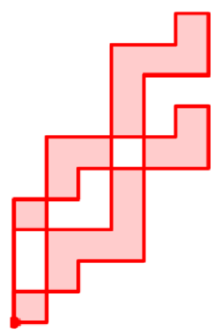

Figure 2: The area bounded by $\gamma(w)$ for $w = w = [b^3, [ba^2ba, ab^3a^2b^2]]$.

For instance, the iterated commutator $w = [b^3, [ba^2ba, ab^3a^2b^2]]$ yields the following $\gamma(w) \subset \mathbb{R}^2$, inscribing a region (that is, $M_{a,b}$) of zero area:

Thus, Theorem 1.1 can be restated as counting equivalence classes of paths of length n in the positive octant of $\mathbb{Z}^n$, where we identify two paths if walking along the first and then returning along the second (appropriately concatenated) gives a closed path whose projections to any plane $\mathrm{Span}\{e_i, e_j\} \subset \mathbb{R}^n$ have zero signed area. We note that this notion of area also appears naturally in the context of (Carnot–Carathéodory) geometry of the real Heisenberg group. For more on the relation between the word metric and the sub-Finsler metric on the Heisenberg group, see [3].

References

[1] M. Bożejko. “Uniformly amenable discrete groups”. *Math. Ann.* **251** (1980), pp. 1–6.

[2] R. G. Burns, O. Macedońska, and Y. Medvedev. “Groups satisfying semigroup laws, and nilpotent-by-Burnside varieties”. *J. Algebra* **2**.195 (1997), pp. 510–525.

[3] M. Duchin and M. Shapiro. “The Heisenberg group is pan-rational”. *Adv. Math.* **346** (2019), pp. 219–263.

[4] M. I. Elashiry and D. S. Passman. “Rewritable groups”. *J. Algebra* **345** (2011), pp. 190–201.

[5] M. Gromov. “Groups of polynomial growth and expanding maps”. *Inst. Hautes Études Sci. Publ. Math.* 53 (1981), pp. 53–73.

[6] P. de la Harpe. *Topics in Geometric Group Theory.* Chicago Lectures in Mathematics. University of Chicago Press, Chicago, 2000, p. 340.

[7] A. Mann. “Growth conditions in infinitely generated groups”. *Groups Geom. Dyn.* 1 (2007), pp. 613–622.

[8] J. F. Semple and A. Shalev. “Combinatorial conditions in residually finite groups I”. *J. Algebra* **157** (1993), pp. 43–50.

[9] A. Shalev. “Combinatorial conditions in residually finite groups II”. *J. Algebra* **157** (1993), pp. 51–62.