# THE LENGTHS OF CONJUGATORS IN THE MODEL FILIFORM GROUPS

M. R. BRIDSON AND T. R. RILEY

ABSTRACT. The conjugator length function of a finitely generated group $\Gamma$ gives the optimal upper bound on the length of a shortest conjugator for any pair of conjugate elements in the ball of radius $n$ in the Cayley graph of $\Gamma$. We prove that polynomials of arbitrary degree arise as conjugator length functions of finitely presented groups. To establish this, we analyse the geometry of conjugation in the discrete model filiform groups $\Gamma_d = \mathbb{Z}^d \rtimes_\phi \mathbb{Z}$ where is $\phi$ is the automorphism of $\mathbb{Z}^d$ that fixes the last element of a basis $a_1, \ldots, a_d$ and sends $a_i$ to $a_i a_{i+1}$ for $i < d$. The conjugator length function of $\Gamma_d$ is polynomial of degree $d$. **2020 Mathematics Subject Classification: 20F65, 20F10, 20F18**

*Key words and phrases:* nilpotent groups, model filiform groups, conjugator length

## 1. INTRODUCTION

Whereas Dehn (isoperimetric) functions are the most natural measure of the complexity of a direct approach to the word problem in a finitely presented group, *conjugator length functions* are the most obvious measure of the complexity of a direct approach to the conjugacy problem in a finitely generated group. By definition, $\mathrm{CL}_G(n)$ is the least integer $N$ such that whenever a pair of words $u$ and $v$ on the generators represent conjugate elements of $G$ and the sum of their lengths is at most $n$, there is a word $w$ of length at most $N$ such that $uw = wv$ in $G$. The study of conjugator length functions is far less developed than the study of Dehn functions. In particular, whereas it is essentially known which functions arise as Dehn functions [BORS02], the class of functions that arise as conjugator length functions is poorly understood. Indeed, until recently, very few sharp estimates on conjugator length functions had been established. In part, this reflects the greater delicacy of the conjugacy problem: for example, there exist pairs of finitely generated groups $H < G$ with $|G/H| = 2$ such that $H$ has a solvable conjugacy problem and $G$ does not [CM77]. Such examples warn us that, in contrast to the study of Dehn functions, we cannot hope to

*Date*: 29 May 2025.

The first author thanks the Mathematics Department of Stanford University for its hospitality. The second author gratefully acknowledges the financial support of the National Science Foundation (NSF OIA-2428489).)

understand conjugacy length functions using the techniques of coarse geometry. On the other hand, the conjugacy problem is fundamentally geometric in nature: understanding conjugacy in a group $\Gamma$ is closely tied to an understanding of the geometry of annuli with prescribed boundaries in any space with fundamental group $\Gamma$. We refer to our article with Andrew Sale [BRSb] for a more detailed introduction to these ideas and a survey of what is known about conjugator length functions.

In this article we will advance the state of the art by showing that polynomials of arbitrary degree arise as conjugator length functions of finitely presented groups. We shall do so by investigating the geometry of the conjugacy problem in the discrete model filiform groups $\Gamma_d$. By definition, $\Gamma_d = \mathbb{Z}^d \rtimes_\phi \mathbb{Z}$, where, with respect to a fixed basis $a_1, \ldots, a_d$ for $\mathbb{Z}^d$, the automorphism $\phi$ fixes $a_d$ and maps $a_i \mapsto a_i a_{i+1}$ for $i = 1, \ldots, d-1$.

$$\Gamma_d = \langle a_1, \ldots, a_d, t \mid \forall 1 \leq i < j \leq d,\ [a_i, a_j] = 1 = [t, a_d],\ t^{-1} a_i t = a_i a_{i+1} \rangle.$$

The corresponding nilpotent Lie groups $G_d = \mathbb{R}^d \rtimes_\phi \mathbb{R}$ have been studied extensively in the context of Carnot geometry, while the lattices $\Gamma_d < G_d$ have served as key examples for nilpotent and polynomial phenomena in geometric group theory. In particular, $(\Gamma_d)_{d \in \mathbb{N}}$ was among the first families of groups used to prove Dehn functions can be polynomial of arbitrary degree; see [BMS93, BP94]. Here, we prove that the groups $(\Gamma_d)_{d \in \mathbb{N}}$ illustrate the same diversity of behaviour among conjugator length functions. In a companion article [BR], we will show that 2-step nilpotent groups can also have conjugator length functions of arbitrary polynomial degree.

**Theorem 1.1.** *The conjugator length function of $\Gamma_d$ is polynomial of degree $d$.*

Theorem 1.1 provides a counterpoint to our work with Andrew Sale on the conjugator length functions of free-by-cyclic groups: in [BRSa] we proved that $\mathrm{CL}(n) \simeq n$ for the group $F_d \rtimes \mathbb{Z}$ obtained by removing the relations $[a_i, a_j] = 1$ from the presentation of $\Gamma_d$. It also contrasts with Sale's result [Sal16, Theorem 4.1] that if $\theta$ is diagonalisable over $\mathbb{R}$ and all its eigenvalues have absolute value greater than 1, then $\mathbb{Z}^d \rtimes_\theta \mathbb{Z}$ has $\mathrm{CL}(n) \simeq n$.

**An outline of the proof.** Whenever one tries to understand a conjugacy problem in a group $\Gamma$, one is inevitably led to study the structure of centralisers in $\Gamma$, because the set of solutions $x$ to an equation $x^{-1}\gamma_1 x = \gamma_2$ is a coset of $C(\gamma_1)$. Thus a feature of our study is that we will need a close understanding of the structure of centralisers in $\Gamma_d$; this is described in Section 4. It will emerge from this analysis of centralisers that we also need tight control on the geometry of roots in $\Gamma_d$, by which we mean solutions to equations $x^p = \gamma$; this is the subject of Section 5. With these tools in hand, we will prove Theorem 1.1 by means of an induction on $d$ that exploits the observation that the centre of $\Gamma_d$ is infinite cyclic, generated by $a_d$, and the quotient of $\Gamma_d$ by its centre is isomorphic to $\Gamma_{d-1}$, with the images of $t$ and $a_i$ $(i < d)$ satisfying the defining the relations of $\Gamma_{d-1}$. When working with this inductive structure, it is important to keep in mind that if two words $u, v$ in the letters $t, a_1, \ldots, a_{d-1}$ define the same element in $\Gamma_{d-1}$, they will in general not be equal in $\Gamma_d$; rather, an equality of the form $u = va_d^r$ will hold in $\Gamma_d$. The first lesson to be taken from this observation is that we must be careful to specify in which group equalities between words are taking place. The second point to absorb is that we will have reason to control $|r|$ in expressions of the form $u = va_d^r$.

With these preparatory thoughts in mind, we can outline our strategy of proof. The actual proof requires us to keep track of various constants, but for the purposes of this outline it makes sense to absord these into symbols such as $\preceq$.

The lower bound $n^d \preceq \mathrm{CL}_{\Gamma_d}(n)$ is established by arguing that $t^r$ is the unique shortest word conjugating $a_{d-1}$ to $a_{d-1}a_d^r$ in $\Gamma_d$ and that $d(1,\, a_{d-1}a_d^{n^d}) \simeq n$ (Section 3). The argument to establish that $\mathrm{CL}_{\Gamma_d}(n) \preceq n^d$ is more involved. We must show that if $u$ and $g$ are conjugate in $\Gamma_d$ and both lie in the ball of radius $n$ about 1, then there exists $x$ such that $x^{-1}ux = g$ and $|x| \preceq n^d$. We will see that the difficult case to deal with is when $g$ does not lie in the normal subgroup $\mathbb{Z}^d \rtimes 1$, so let us assume here that this is the case. Assuming that the theorem has been proved for $\Gamma_{d-1}$, we get a word $\theta$ with length $|\theta| \preceq n^{d-1}$ in the letters $t, a_1, \ldots, a_{d-1}$ conjugating the images of $u$ and $g$ in $\Gamma_{d-1}$. This tells us that $\theta^{-1}u\theta = ga_d^r$ in $\Gamma_d$ for some integer $r$. We will use our knowledge of the Dehn function in $\Gamma_{d-1}$ to ensure that $|r| \preceq n^{d(d-1)}$. To complete the induction, we must argue that if $ga_d^r$ is conjugate to $g$ and $|r| \preceq n^{d(d-1)}$, then the conjugacy can be achieved by a conjugator $h$ with $d(1, h) \preceq n^d$. (Note that $ga_d^r$ may lie outside the ball of radius $n$ in $\Gamma_d$, which complicates the structure of the argument.)

It transpires that it is best to realise the conjugacy of $g$ to $ga_d^r$ in two steps. In the first step, we conjugate by a suitable power of $a_{d-1}$: the aim here is to reduce $|r|$ to something of significantly smaller order by conjugating by $a_{d-1}^M$ with $M \approx n^{d(d-1)}$, and since the distortion of $\langle a_{d-1} \rangle < \Gamma_d$ is polynomial of degree $d-1$, the length of such a conjugator is roughly $n^d$, which is what we are aiming for. We will see that, under the constraint $M \approx n^{d(d-1)}$, we can reduce $|r|$ to something less than $pqe < ne$ where $p$ is the maximal power such that the image of $g$ in $\Gamma_{d-1}$ can be written in the form $\overline{g} = g_0^p a_{d-1}^*$ and $q$ is the exponent sum of $t$ in $g_0$, and $e = \gcd(p, q)$. A surprisingly large amount of work is required for the final step in the proof, wherein we must bound the length of a conjugator taking

$g$ to $ga_d^r$ for small values of $r$: this involves understanding the action on the coset $g\langle a_d\rangle$ of the centraliser of $\overline{g} \in \Gamma_{d-1}$ through a homomorphism $\zeta_{\overline{g}} : C(\overline{g}) \to \mathbb{Z}$ defined in Definition 7.3. Our understanding of this action relies on an estimate $d(1, g_0) \preceq d(1, g)$ that comes from the analysis of the lengths of roots in $\Gamma_{d-1}$ in Section 5.

## 2. Preliminaries

Throughout, $\Gamma_d$ will denote the group $\mathbb{Z}^d \rtimes_\phi \mathbb{Z}$ where, with respect to a fixed basis, $\phi$ and $\phi^{-1}$ are represented by the following matrices, respectively:

$$\begin{pmatrix} 1 & & & & \\ 1 & 1 & & & \\ & 1 & \ddots & & \\ & & \ddots & 1 & \\ & & & 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ 1 & -1 & \ddots & & \\ -1 & 1 & \ddots & 1 & \\ \vdots & \vdots & \ddots & -1 & 1 \end{pmatrix}.$$

We work with the standard presentation

$$\Gamma_d = \langle a_1, \ldots, a_d, t \mid \forall 1 \leq i < j \leq d,\ [a_i, a_j] = 1 = [t, a_d],\ t^{-1}a_i t = a_i a_{i+1}\rangle. \tag{1}$$

All distances in $\Gamma_d$ will be taken with respect to the generating set of this presentation and the area of any null-homotopic word will be calculated with respect to this set of defining relations. It is worth noting, however, that $\Gamma_d$ is generated by just two elements, $a_d$ and $t$.

We are interested in the *conjugator length function* $\mathrm{CL} : \mathbb{N} \to \mathbb{N}$ of these groups. By definition, $\mathrm{CL}(n)$ is the minimal integer $N$ such that whenever a pair of words $u$ and $v$ of length at most $n$ in the generators represent conjugate elements in $\Gamma_d$, there is a word $w$ of length at most $N$ such that $uw = wv$ in $\Gamma_d$. If we worked with a different finite generating set, the resulting conjugator length function would be equivalent in the sense of the following standard definition.

*For functions $f$ and $g$ mapping $\mathbb{N} \to \mathbb{N}$ write $f \preceq g$ when there exists $C > 0$ such that $f(n) \leq Cg(Cn + C) + C$ for all $n \in \mathbb{N}$, and write $f \simeq g$ when $f \preceq g$ and $g \preceq f$.*

The following standard lemma will be needed to control the size of the discrepancy $r$ that arises when we lift an equality $u = v$ in $\Gamma_{d-1}$ to an equality of the form $u = va_d^r$ in $\Gamma_d$.

**Lemma 2.1.** *If two words $u, v$ in the letters $t, a_1, \ldots, a_{d-1}$ define the same element of $\Gamma_{d-1}$, then $u = va_d^r$ in $\Gamma_d$, where $|r| \leq \mathrm{Area}_{\Gamma_{d-1}}(uv^{-1})$.*

*Proof.* By definition, $\mathrm{Area}_{\Gamma_{d-1}}(uv^{-1})$ is the least integer $A$ for which there is an equality

$$uv^{-1} = \prod_{i=1}^{A} \theta_i^{-1} \rho_i \theta_i$$

in the free group $F = F(t, a_1, \ldots, a_{d-1})$ with each $\rho_i^{\pm 1}$ one of the defining relations of $\Gamma_{d-1}$. We modify the righthand side of this expression by replacing each instance of $[a_{d-1}, t]$ with $[a_{d-1}, t]a_d$, leaving the other relations untouched. Let $\rho_i'$ be the resulting relations. Then, in $F \times \langle a_d \rangle$ we have the equality

$$uv^{-1} = a_d^r \prod_{i=1}^{A} \theta_i^{-1} \rho_i' \theta_i,$$

where $-r$ is the number of letters $a_d$ added during the editing of the $\rho_i$. The righthand side has image $a_d^r$ under the surjection $F \times \langle a_d \rangle \to \Gamma_d$ that is implicit in the notation. Thus $u = va^r$ in $\Gamma_d$, with $|r| \leq A$. □

The following well-known result will be needed in connection with the above lemma.

**Theorem 2.2** ([BP94])**.** *The Dehn function of $\Gamma_d$ is $\simeq n^{d+1}$.*

Positive powers of the matrices for $\phi$ and $\phi^{-1}$ can be calculated using the binomial theorem, writing each as $I + N$ and expanding $(I + N)^n$. For $\phi$, the entry in column $i$ on the bottom row for $\phi^n$ is $\epsilon(i, n) = \binom{n}{d-i}$, while the corresponding entry $\epsilon(i, -n)$ for $\phi^{-n}$ is a signed sum on binomial coefficients. The following crude estimate will suffice for our purposes.

**Lemma 2.3.** *There is a constant $\varepsilon_d$ such that $|\epsilon(i, m)| \leq \varepsilon_d \, |m|^{d-i}$ for all $m \in \mathbb{Z}$ and $i = 1, \ldots, d-1$.*

The significance for us of the integers $\epsilon(i, n)$ is that for $i = 1, \ldots, d-1$, if $\phi_{d-1}^n(a_i) \in \Gamma_{d-1}$ and $\phi_d^n(a_i) \in \Gamma_d$ are written as words in the letters $a_1, \ldots, a_d$, then in $\Gamma_d$ we have

$$\phi_d^n(a_i) = \phi_{d-1}^n(a_i) \, a_d^{\epsilon(i,n)}. \tag{2}$$

## 3. Normal forms and distance

We will push the letters $t^{\pm 1}$ to the left when writing elements of $\Gamma_d$ in the standard normal form for polycyclic groups.

**Lemma 3.1.** *Every element $g \in \Gamma_d$ can be expressed uniquely as a word in normal form $t^r a_1^{p_1} a_2^{p_2} \cdots a_d^{p_d}$, and if $d(1, g) \leq n$ then $|r| \leq n$ and $|p_i| \leq n^i$ for $i = 1, \ldots, d$.*

*Proof.* The only non-trivial assertion is the one about the size of the exponents. We proceed by induction on $n$. Given a word $w = b_1 \cdots b_n \beta$ of length $n+1$, we apply the inductive hypothesis to write $w' = b_1 \cdots b_n$ in normal form, then shuffle $\beta$ past the letters of this normal form to write $w$ in normal form. If $\beta = a_j^{\pm 1}$ then this shuffling leaves $r$ and $p_i$ unchanged for $i \neq j$ and alters $p_j$ by $\pm 1$. If $\beta = t$ then this shuffling leaves $p_1$ unchanged and for $i \geq 1$ increases $p_i$ by $p_{i-1}$, so the observation $n^i + n^{i-1} < (n+1)^i = n^i + in^{i-1} + \cdots$ finishes the induction. If $\beta = t^{-1}$ then the shuffling replaces $p_i$ by

$$(-1)^{i+1}(p_1 - p_2 + \cdots + (-1)^{i+1} p_i),$$

whose absolute value is, by induction, less than $n + n^2 + \cdots + n^i$, which is less than $(n+1)^i$. □

Our next result provides a converse to the preceding lemma.

**Proposition 3.2.** *There is a constant $C_d \geq 1$ such that for every $g \in \Gamma_d$ with normal form $t^r a_1^{p_1} a_2^{p_2} \ldots a_d^{p_d}$, where $|r| \leq n$ and $|p_i| \leq n^i$, we have $d(1,g) \leq C_d\, n$.*

The following special case of the proposition contains the main idea in the proof; it also facilitates a simple induction to prove the general case.

**Lemma 3.3.** *For every $d \geq 1$ there is a constant $D_d$ such that for all positive integers $0 < p \leq n^d$ we have $d(1, a_d^p) \leq D_d\, n$ in $\Gamma_d$.*

*Proof.* We proceed by induction on $d$. The base case is $\Gamma_1 = \mathbb{Z}^2$, where the assertion is trivial.

The Hilbert-Waring Theorem provides us with an integer $M_d$ such that every positive integer $p$ can be written as a sum $p = k_1^d + \cdots + k_m^d$ with $k_i \in \mathbb{N}$ and $m \leq M_d$. Then $\sum_{i=1}^m k_i \leq M_d n$, because $p \leq n^d$ implies that $k_i \leq n$ for all $i$. In $\Gamma_d$, for any $k \in \mathbb{N}$ we have the equation

$$a_{d-1}^{-k^{d-1}} t^{-k} a_{d-1}^{k^{d-1}} t^k = a_d^{k^d}.$$

By induction, for $i = 1, \ldots, m$, there is a word $w_i$ in the letters $t, a_1, \ldots, a_{d-1}$ of length at most $D_{d-1}\, k_i$ such that $w_i = a_{d-1}^{k_i^{d-1}}$ in $\Gamma_{d-1}$. Recalling that $\Gamma_{d-1}$ is obtained from $\Gamma_d$ by killing the central subgroup $\langle a_d \rangle$, we have $w_i = a_{d-1}^{k_i^{d-1}} a_d^{l_i}$ in $\Gamma_d$ for some $l_i \in \mathbb{Z}$. As $a_d$ is central, in $\Gamma_d$ we have

$$w_i^{-1} t^{-k_i} w_i t^{k_i} = a_{d-1}^{-k_i^{d-1}} t^{-k_i} a_{d-1}^{k_i^{d-1}} t^{k_i} = a_d^{k_i^d}.$$

Hence, in $\Gamma_d$,

$$a_d^p = \prod_{i=1}^m w_i^{-1} t^{-k_i} w_i t^{k_i},$$

and the righthand side is a word of length at most $2(D_{d-1}+1)\sum_{i=1}^m k_i \leq 2(D_{d-1}+1)M_d\,n$. Taking $D_d = 2(D_{d-1}+1)M_d$ completes the induction. □

*Proof of Proposition 3.2.* We proceed by induction on $d$. Again, the base case $d = 1$ is trivial. Assume that the proposition is true in $\Gamma_{d-1}$. Given $g = t^r a_1^{p_1} a_2^{p_2} \dots a_d^{p_d} \in \Gamma_d$, with $|r| \leq n$ and each $|p_i| \leq n^i$, we know by induction that there is a word $W = W(t, a_1, \dots, a_{d-1})$ of length at most $C_{d-1}n$ such that in $\Gamma_{d-1}$ we have

$$W = t^r a_1^{p_1} a_2^{p_2} \cdots a_{d-1}^{p_{d-1}}.$$

In $\Gamma_d$, then, for a unique $q \in \mathbb{Z}$ we have

$$W = t^r a_1^{p_1} a_2^{p_2} \cdots a_{d-1}^{p_{d-1}} a_d^q,$$

and by Lemma 3.1 we know $|q| \leq |W|^d \leq C_{d-1}^d n^d$. So, $|p_d - q| \leq |p_d| + |q| \leq (1+C_{d-1}^d)n^d$, and then Lemma 3.3 leads to $d(1, a_d^{p_d-q}) \leq (1+C_{d-1}^d)D_d n$. Therefore, by the triangle inequality in $\Gamma_d$, noting that $g = W a_d^{p_d-q}$,

$$d(1, g) \leq d(1, W) + d(1, a_d^{p_d-q}) \leq C_{d-1}n + D_d(1 + C_{d-1}^d)n,$$

and defining $C_d = C_{d-1} + D_d(1 + C_{d-1}^d)$ completes the induction. □

Lemma 3.1 and Proposition 3.2 provide the complementary upper and lower bounds for the following statement.

**Corollary 3.4.** *For $i = 1, \dots, d$, the function $n \mapsto d_{\Gamma_d}(1, a_i^n)$ is Lipschitz equivalent to $n^{1/i}$.*

*Remark* 3.5. This corollary played an important role in [Bri99] where it was proved with an argument that is less elementary but which is more concise and provides greater geometric intuition. This argument, which is due to Gromov [Gro93], takes place in the Malcev completion of $\Gamma_d$, which is the nilpotent Lie group $G_d = \mathbb{R}^d \rtimes_\phi \mathbb{R}$, where $\Gamma_d$ is a lattice. The Lie algebra of $G_d$, which is graded, has presentation

$$(X_1, ..., X_d, T : [X_i, X_j] = 0 = [X_d, T],\ [T, X_i] = X_{i-1},\ i = 1, ..., d-1).$$

This Lie algebra is generated under brackets by $\{X_1, T\}$, so there is a Carnot-Caratheodory metric on $G_d$: this is defined by fixing an inner product on the plane $\langle X_1, T\rangle$ so that $\{X_1, T\}$ is orthonormal, transporting this plane around $G_d$ by left translation, then defining the distance $d_{cc}(x, y)$ between two points $x, y \in G_d$ to be the infimum of the lengths of piecewise smooth curves from $x$ to $y$ that are everywhere tangent to this field of 2-planes, with the length of tangent vectors measured using the transported inner product.

The Lie algebra of $G_d$ admits the 1-parameter family of automorphisms $m_s : X_i \mapsto s^i.X_i$, $T \mapsto s.T$ for $s \neq 0$. These automorphisms induce (via the exponential map) homotheties of the Carnot-Caratheodory metric on $G_d$. For all $s \in \mathbb{R}$ we have $d_{cc}(1, \exp(sX_i)) = s^{1/i}$. The restriction of $d_{cc}$ to the lattice $\Gamma_d$ is Lipschitz equivalent to the word metric on $\Gamma_d$ and $a_i^n = \exp(nX_i)$.

**Lower Bound for Theorem 1.1.** The key observation here is that an arbitrary element $t^r x \in \Gamma_d = \mathbb{Z}^d \rtimes_\phi \mathbb{Z}$ with $x \in \mathbb{Z}^d \times 1$ conjugates $a_{d-1}$ to $a_{d-1}a_d^r$, with $x$ playing no role. It follows that for each positive integer $r > 0$, the unique shortest element of $\Gamma_d$ conjugating $a_{d-1}$ to $a_{d-1}a_d^r$ is $t^r$.

In Lemma 3.3 we established the existence of a constant $D_d$ such that $a^{n^d}$ lies in the ball of radius $D_d n$ about $1 \in \Gamma_d$, hence $d(1, a_{d-1}a_d^{n^d}) \leq D_d n + 1$. As $t^{n^d}$ is the unique shortest element conjugating $a_{d-1}$ to $a_{d-1}a_d^{n^d}$, we deduce

$$\mathrm{CL}_{\Gamma_d}(D_d n + 1) \geq n^d,$$

hence $n^d \preceq \mathrm{CL}_{\Gamma_d}(n)$.

## 4. Centralisers in $\Gamma_d$

Picking up a theme from the last argument of the previous section, we note that $g \in \Gamma_d$ will conjugate $\gamma$ to an element of the form $\gamma a_d^*$ if and only if the image of $g$ in $\Gamma_{d-1}$ centralises the image of $\gamma$. This highlights our need to understand centralisers in the groups $\Gamma_d$.

**Notation.** We will need to discuss the subgroup $\langle a_1, \ldots, a_d \rangle = \mathbb{Z}^d \rtimes 1 < \Gamma_d$ repeatedly in what follows, and it will be important to distinguish it from an abstract copy of $\mathbb{Z}^d$. For this reason, we shall henceforth denote it by $\mathbb{A}_d$.

**Proposition 4.1.** *Given $\gamma \in \Gamma_d$, if $\gamma = \gamma_0^p a_d^q$ with $p > 0$ maximal, then*

(1) *$C(\gamma) = \langle \gamma_0, a_d \rangle \cong \mathbb{Z}^2$ if $\gamma_0 \notin \mathbb{A}_d$,*
(2) *$C(\gamma) = \mathbb{A}_d$ if $\gamma \in \mathbb{A}_d \smallsetminus \langle a_d \rangle$,*
(3) *$C(\gamma) = \Gamma_d$ if $\gamma \in \langle a_d \rangle$.*

*Proof.* We again proceed by induction on $d$, viewing $\Gamma_{d-1}$ as the quotient of $\Gamma_d$ by $\langle a_d \rangle$ and writing $\overline{\gamma}$ for the image of $\gamma$ in $\Gamma_{d-1}$. The base step $\Gamma_1 \cong \mathbb{Z}^2$ is trivial (with case (2) vacuous).

Cases (2) and (3) are easily verified, so in the inductive step we concentrate on the case $\gamma = \gamma_0^p a_d^q$ with $\gamma_0 \notin \mathbb{A}_d$.

Some preliminary observations are in order: if $\overline{\gamma} = g^m$ in $\Gamma_{d-1}$, then for any preimage $\tilde{g}$ in $\Gamma_d$ of $g$ we have $\tilde{g}^m = \gamma a_d^r$ for some $r$, so $m \leq p$, by the maximality of $p$, and if $m = p$ then $g = \overline{\gamma_0}$ by the uniqueness of roots in torsion-free nilpotent groups. Thus $\overline{\gamma_0}$ is the unique maximal root of $\overline{\gamma}$ in $\Gamma_{d-1}$ (even though $\gamma_0$ itself is only uniquely defined up to multiplication by a power of $a_d$).

By induction, $C(\overline{\gamma}) < \Gamma_{d-1}$ is a free abelian group of rank 2 that contains $\langle a_{d-1} \rangle$. The image of $C(\gamma)$ in $C(\overline{\gamma})$ intersects $\langle a_{d-1} \rangle$ trivially, because either $\gamma$ or $\gamma^{-1}$ is $xt^{-j}$ with $x \in \mathbb{A}_d$ and $j > 0$, and in $\Gamma_d$ (n.b. $\Gamma_d$ rather than $\Gamma_{d-1}$)

$$\tilde{a}_{d-1}^{-k} x t^{-j} \tilde{a}_{d-1}^{k} = a_d^{kj} x t^{-j} \neq x t^{-j}$$

for any preimage $\tilde{a}_{d-1} = a_{d-1} a_d^{*}$ in $\Gamma_d$ of $a_{d-1} \in \Gamma_{d-1}$. Thus the image of $C(\gamma)$ in $C(\overline{\gamma})$ is a cyclic group. This cyclic group contains $\overline{\gamma}$, so by the considerations of the previous paragraph, it is contained in $\langle \overline{\gamma_0} \rangle$. As $\gamma_0$ commutes with $\gamma$, the image is exactly $\langle \overline{\gamma_0} \rangle$.

The preimage in $\Gamma_d$ of $\langle \overline{\gamma_0} \rangle$ is $\langle \gamma_0, a_d \rangle$, which centralises $\gamma$. Thus $C(\gamma) = \langle \gamma_0, a_d \rangle$ and the induction is complete. □

## 5. The lengths of roots in $\Gamma_d$

The emergence of $\gamma_0$ in Proposition 4.1 throws up a need to understand roots in the groups $\Gamma_d$. In particular, when it comes to estimating the lengths of minimal conjugating elements, we will need to bound the lengths of roots of elements in $\Gamma_d$.

The bounds in the following proposition are not optimal, but they will suffice for our purposes and a sharpening would involve an unwelcome increase in notation.

**Proposition 5.1.** *There is a constant $k_d > 0$ so that for all $p, n > 0$ and all $\gamma \in \Gamma_d$ with $d(1, \gamma) \leq n$, if $h^p = \gamma$ then*

$$h = t^{m_0} a_1^{m_1} \cdots a_d^{m_d}$$

*with $|m_0| \leq n/p$ and $|m_i| \leq k_d n^i / p$ for $i = 1, \ldots, d$.*

For the most part, we shall only need the following easy consequence of Propositions 3.2 and 5.1.

**Corollary 5.2.** *There is a constant $K_d$ so that if $h^p = \gamma$ in $\Gamma_d$ and $d(1, \gamma) \leq n$, then $d(1, h) \leq K_d\, n$.*

Before proving Proposition 5.1, we consider the case $d = 2$ in order to elucidate the geometry of what is happening.

*Example* 5.3. In the 3-dimensional Heisenberg group $\Gamma_2 = \langle a_1, a_2, t\rangle$ we have

$$(a_1 t)^n = t^n a_1^n a_2^{\sigma(n)},$$

where

$$\sigma(n) = 1 + \cdots + n = n(n-1)/2 = \mathrm{Area}_{\langle a_1, t\rangle}(u^n v^{-1})$$

with $u \equiv a_1 t$ and $v \equiv t^n a_1^n$. One can see that $\sigma(n)$ really is $\mathrm{Area}_{\langle a_1,t\rangle}(u^n v^{-1})$ either by considering the van Kampen enclosed by the loop $u^n v^{-1}$ in the $(a_1, t)$-plane tessellated by squares, or else by arguing algebraically: we reduce $u^n v^{-1} \equiv (a_1 t)^n a_1^{-n} t^{-n}$ to the empty word by starting in the middle and replacing the subword $a_1 t a_1^{-1}$ by $t$ at a cost of applying 1 commutator relation, then $a_1 t^2 a_1^{-1}$ by $t^2$ at an additional cost of 2, etc., until finally $a_1 t^n a_1^{-1}$ is replaced by $t^n$ at a cost of $n$, after which we can freely reduce $t^n t^{-n}$ to the empty word.

If $n$ is odd, then

$$(a_1 t a_2^{(1-n)/2})^n = t^n a_1^n.$$

Thus $h_n := a_1 t a_2^{(1-n)/2}$ is the unique $n$-th root of $g := t^n a_1^n$ in $\Gamma_2$. Projecting to $\Gamma_1 = \mathbb{Z}^2$, it is easy to see that $d(1, t^n a_1^n) = 2n$ in $\Gamma_2$. On the other hand, $d(1, h_n)$ is approximately $\sqrt{n}$, even though $h_n^n = g$.

We now consider the more general situation where $\gamma \in \Gamma_2 \smallsetminus \langle a_1, a_2\rangle$ is a $p$-th power, say $\gamma = h^p$. Counting occurrences of $t$, we see that $|p| \leq d(1, \gamma)$. Our goal is to estimate the length of $h$. Suppose $n = d(1,\gamma)$ and fix a geodesic word $w$ representing $\gamma$. Let $\overline{w}$ be the word obtained from $w$ by deleting all occurrences of $a_2$ and note that this word equals the image of $\gamma$ in $\Gamma_1$. We write the image of $h$ in $\Gamma_1 = \mathbb{Z}^2$ as a geodesic word $v$ in the free abelian group on $\{a_1, t\}$. Replacing $\gamma$ and $h$ by their inverses if necessary, we may assume that only positive powers of $t$ arise in $v$. In $\Gamma_1 = \mathbb{Z}^2$ we know that $v^p = \overline{w}$ implies $|v| = d(1, \overline{w})/p$, so $|v| \leq |w|/p = n/p$. As in Lemma 2.1, in $\Gamma_2$ we have

$$v^p = w a_2^r,$$

where $|r| \leq \mathrm{Area}_{\Gamma_1}(v^{-p}\overline{w})$, which is less than $n^2$, by the quadratic isoperimetric inequality in $\mathbb{Z}^2$. (In more detail, $v^{-p}\overline{w}$ is a non-empty word $\tau$ in $a_1$ and $t$ of length at most $2n$ that equals the identity in $\Gamma_1$ and any such $\tau$ has a subword which freely reduces to $a_1^{\pm 1} t^l a_1^{\mp 1}$ or $t^{\pm 1} a_1^l t^{\mp 1}$ for some $|l| \leq n/2$; this subword can be replace by $t^l$ or $a_1^l$, respectively, at a cost of applying at most $n/2$ commutator relations, from which it follows that $\mathrm{Area}(\tau) < n^2$, by an induction on length.) But $v = h a_2^j$ in $\Gamma_2$ for some $j \in \mathbb{Z}$, so

$$w a_2^r = v^p = h^p a_2^{jp} = \gamma a_2^{jp} = w a_2^{jp}.$$

Therefore $r = jp$.

We saw earlier that $|v| \leq n/p$ and $|r|<n^2$, and from Proposition 3.2 we know $d(1, a_2^{-j}) \leq C_2 \sqrt{j} = C_2\sqrt{r/p}$, which is less than $C_2\, n/\sqrt{p}$. Thus, writing $v \in \langle a_1, t\rangle$ in normal form $v = t^{m_0} a_1^{m_1}$ with $|m_0|+|m_1| \leq n/p$, we have $h = v a_2^{-j} = t^{m_0} a_1^{m_1} a_2^{-j}$ in the form required by Proposition 5.1, since $|j| \leq n^2/p$. Moreover,

$$d(1, h) \leq |v| + d(1, a_2^{-j}) \leq \frac{1}{p}(n + C_2\, np^{1/2}) \leq C_2\ \frac{1}{p}(n + np^{1/2}) \leq 2C_2\ np^{-1/2}.$$

*Proof of Proposition 5.1.* We proceed by induction on $d$. The case $d = 1$ is trivial and the case $d = 2$ was covered in the preceding example. Assume true up to $d-1$. We have $\gamma, h \in \Gamma_d$ with $h^p = \gamma$ and $d(1, \gamma) \leq n$. Let $\overline{h}$ be the image of $h$ in $\Gamma_{d-1}$. Then, by induction, $\overline{h} = t^{m_0} a_1^{m_1} \ldots a_{d-1}^{m_{d-1}}$, in $\Gamma_{d-1}$, with the $|m_i|$ bounded as stated in the proposition. Let

$$h_0 := t^{m_0} a_1^{m_1} \cdots a_{d-1}^{m_{d-1}} \in \Gamma_d.$$

Then $h_0 = h a_d^m$ for some $m \in \mathbb{Z}$, and if the normal form for $h_0^p \in \Gamma_d$ is

$$h_0^p = t^{pm_0} a_1^{s_1} \cdots a_d^{s_d},$$

then the normal form for $\gamma = h^p = (h_0 a_d^{-m})^p$ is

$$\gamma = t^{pm_0} a_1^{s_1} \cdots a_{d-1}^{s_{d-1}} a_d^{s_d - mp},$$

and from Lemma 3.1 we have $|s_d - mp| \leq n^d$. So, if we can prove that there is a constant $\alpha_d$ such that

$$|s_d| \leq \alpha_d n^d, \tag{3}$$

then we will have $|mp| \leq (\alpha_d + 1)n$ and defining $k_d := \max\{\alpha_d + 1, k_{d-1}\}$ will complete the induction.

To obtain the required bound on $s_d$, we transform $h_0^p$ into normal form carefully, pushing all occurrences of $t^{\pm 1}$ to the left and then shuffling the resulting element of $\mathbb{A}_d \cong \mathbb{Z}^d$ into normal form. The pushing of the $t$ letters amounts to using free identities $a_i t^{jm_0} = t^{jm_0}(t^{-jm_0} a_i t^{jm_0})$ and then evaluating $t^{-jm_0} a_i t^{jm_0}$ as $\phi_d^{jm_0}(a_i)$. At this point, it is important to recall from equation (2) that for each $i \leq d-1$

$$\phi_d^{jm_0}(a_i) = \phi_{d-1}^{jm_0}(a_i) a_d^{\epsilon(i, jm_0)}.$$

(This is equally valid for the negative powers of $\phi$ arising from pushing $t^{-1}$ if $m_0 < 0$.)

The result of shuffling $h_0^p$ into normal form in $\Gamma_{d-1}$ (producing the word $t^{pm_0} a_1^{s_1} \cdots a_{d-1}^{s_{d-1}}$) differs from the result of doing it in $\Gamma_d$ by a power of $a_d$ that can be calculated by considering how many copies of each letter $a_i$ in $h_0^p$ we have to move each $t^{m_0}$ past, adding up the number of $a_d$ letters produced by each push. The total (ignoring cancellation due to signs if $m_0 < 0$) is

$$M = \sum_{j=1}^{p-1} \sum_{i=1}^{d-1} |m_i \, \epsilon(i, jm_0)|.$$

We estimate this (crudely) by using the inequality $|\epsilon(i,m)| \leq \varepsilon_d \, m^{d-i}$ from Lemma 2.3, noting that $|jm_0|$ is less than $|pm_0|$, which is the absolute value of the exponent sum of $t$ in $\gamma$, which is less than $n$. Thus

$$M \leq \sum_{j=1}^{p-1} \sum_{i=1}^{d-1} |m_i \, \varepsilon_d \, (jm_0)^{d-i}| \leq \varepsilon_d \, (p-1) \sum_{i=1}^{d-1} |m_i| \, n^{d-i},$$

and our induction assures us that $|pm_i| \leq k_{d-1} n^i$, so

$$|s_d| \leq M \leq \varepsilon_d \, k_{d-1} \sum_{i=1}^{d-1} n^i n^{d-i} = \varepsilon_d \, k_{d-1} \sum_{i=1}^{d-1} n^d.$$

Thus, to obtain the required bound (3), it suffices to set $\alpha_d := d\varepsilon_d k_{d-1}$. $\square$

## 6. Quatifying Bézout's Lemma

We will need an elementary observation concerning Bézout's Lemma.

**Lemma 6.1.** *For positive integers $A$ and $B$, neither of which divides the other, if $e = \gcd(A, B)$ then there exist positive integers $\mu < A$ and $\lambda \leq B$ such that $\lambda A - \mu B = e$.*

*Proof.* By Bézout's Lemma, there exist non-zero integers $a$ and $b$ such that $aA + bB = e$. As $A$ does not divide $e$, it does not divide $b$, so $b = kA - r$ with $0 < r < A$. Then $(kB + a)A - rB = e$ and

$$(kB + a) = (r/A)B + (e/A)$$

is a positive integer less than or equal to $B$, because $r/A$ and $e/A$ are both positive numbers less than 1. Let $\lambda = (kB + a)$ and let $\mu = r$. $\square$

## 7. The proof of Theorem 1.1

At the end of Section 3 we proved that $\mathrm{CL}_{\Gamma_d}(n) \succeq n^d$, so the following proposition completes the proof of Theorem 1.1.

**Proposition 7.1.** *There is an integer $B_d$ such that if $u, v \in \Gamma_d$ are conjugate and* $\max\{d(1,u),\, d(1,v)\} \leq n$*, then there exists $g \in \Gamma_d$ with $g^{-1}ug = v$ and $d(1,g) \leq B_d n^d$.*

We will again proceed by induction on $d$. If $d = 1$ then $\Gamma_d = \mathbb{Z}^2$ and the proposition is trivial. If $d = 2$ then $\Gamma_d$ is the 3-dimensional integral Heisenberg group and since $u$ and $v$ are conjugate in $\Gamma_2$, their images in $\Gamma_1 \cong \mathbb{Z}^2$ are equal. So $u = t^r a_1^\alpha a_2^\eta$ and $v = t^r a_1^\alpha a_2^\xi$ for some $r, \alpha, \eta, \xi \in \mathbb{Z}$, and there exist $s, \omega \in \mathbb{Z}$ such that $g = t^s a_1^\omega$ satisfies $ug = gv$ in $\Gamma_d$. Now in $\Gamma_2$,

$$ug = (t^r a_1^\alpha a_2^\eta)(t^s a_1^\omega) = t^{r+s}\phi^s(a_1^\alpha a_2^\eta)a_1^\omega = t^{r+s}a_1^{\alpha+\omega}a_2^{s\alpha+\eta}$$

and

$$gv = (t^s a_1^\omega)(t^r a_1^\alpha a_2^\xi) = t^{r+s}\phi^r(a_1^\omega)a_1^\alpha a_2^\xi = t^{r+s}a_1^{\alpha+\omega}a_2^{r\omega+\xi},$$

and these are equal if and only if $s\alpha + \eta = r\omega + \xi$. This condition can be expressed as $s\alpha - r\omega = \xi - \eta$, and if we regard this as an equation with variables $s$ and $\omega$, Lemma 6.1 implies that this equation has an integer solution with $|s|$ and $|\omega|$ both at most $\max\{|\alpha|, |r|, |\xi - \eta|\}$. But, by Lemma 3.1, $|\alpha|$ and $|r|$ are at most $n$, and $|\xi|$ and $|\eta|$ are at most $n^2$. The existence of a suitable constant $B_2$ follows.

Assume now that $d \geq 3$ and that the existence of $B_{d-1}$ has been established. Suppose that $u, v \in \Gamma_d$ are conjugate and $\max\{d(1,u),\, d(1,v)\} \leq n$.

**Lemma 7.2.** *Proposition 7.1 is true for elements $u, v \in \mathbb{A}_d = \mathbb{Z}^d \rtimes 1 < \Gamma_d$.*

*Proof.* We first consider the case $v \in \langle a_{d-1}, a_d \rangle$. In this case, the conjugates of $v$ all have the form $va_d^{em}$ where $e$ is the exponent sum of $a_{d-1}$ in $v$, and the unique shortest element conjugating $v$ to $va_d^{em}$ is $t^m$. Also, $d(1, va_d^{em}) \leq n$ implies $d(1, a_d^{em}) \leq d(1,v) + d(1, va_d^{em}) \leq 2n$, hence $|m| \leq |em| \leq (2n)^d$ by Lemma 3.1.

Next we suppose $v \in \mathbb{A}_d \smallsetminus \langle a_{d-1}, a_d \rangle$. As $\langle a_{d-1}, a_d \rangle$ and $\mathbb{A}_d$ are normal, $u$ also lies in $\mathbb{A}_d \smallsetminus \langle a_{d-1}, a_d \rangle$. The unique shortest element conjugating $u$ to $v$ is again a power of $t$. Proposition 4.1(2) tells us that this same power of $t$ is the unique shortest element conjugating the image of $u$ to the image of $v$ in $\Gamma_{d-1}$, and by induction this has length less than $B_{d-1}n^{d-1}$. □

**The generic case in Proposition 7.1.** Continuing our proof of Proposition 7.1, we now consider the case where the normal forms of $u$ and $v$ contain a non-zero power of $t$, in other words $u, v \notin \mathbb{A}_d$. Let $\overline{u}$ and $\overline{v}$ be geodesic words representing the images of $u$ and $v$ in $\Gamma_{d-1}$. Note that $|\overline{u}| \leq d(1,u)$ and $|\overline{v}| \leq d(1,v)$. Therefore, by induction, there is a word $w$ in the free group on $\{t, a_1, \ldots, a_{d-1}\}$ of length at most $B_{d-1}n^{d-1}$ such that $w^{-1}\overline{u}w = \overline{v}$ in $\Gamma_{d-1}$.

Theorem 2.2 provides an integer $A_{d-1}$ and such that

$$\mathrm{Area}_{\Gamma_{d-1}}(w^{-1}\overline{u}w\overline{v}^{-1}) \leq A_{d-1}|w^{-1}\overline{u}w\overline{v}^{-1}|^d \leq A_{d-1}(2n+2B_{n-1}n^{d-1})^d \leq A'_{d-1}n^{d(d-1)},$$

where for convenience we define $A'_{d-1} := 2^d A_{d-1}(1+B_{n-1})^d$. It follows from Lemma 2.1 that $w^{-1}uw = va_d^{\ell}$ in $\Gamma_d$ where $|\ell| \leq A'_{d-1}n^{d(d-1)}$.

Our inductive proof will be complete if we can prove the following lemma. For then, in the case $v = \gamma$ with $A = A'_{d-1}$, we have $(wz)^{-1}u(wz) = v$ with

$$d(1, wz) \leq |w| + d(1,z) \leq B_{d-1}n^{d-1} + A'_{d-1}E_d n^d$$

and we can take $B_d = B_{d-1} + A'_{d-1}E_d$.

The proof of the lemma involves the homomorphisms $\zeta_g : C_{\Gamma_{d-1}}(g) \to \mathbb{Z}$ defined as follows.

**Definition 7.3.** [The homomorphisms $\zeta_g$] Let $g \in \Gamma_{d-1}$. For each element of the centralizer $x \in C_{\Gamma_{d-1}}(g)$ and all preimages $\tilde{x}, \tilde{g} \in \Gamma_d$ we have $\tilde{x}^{-1}\tilde{g}\tilde{x} = \tilde{g}a_d^m$ in $\Gamma_d$, where $m \in \mathbb{Z}$ is independent of the choices of $\tilde{x}$ and $\tilde{g}$ because different choices differ by a power of $a_d$, which is central. Define $\zeta_g(x) := m$ and note that $\zeta_g : C_{\Gamma_{d-1}}(g) \to \mathbb{Z}$ is a homomorphism that vanishes on $\langle g \rangle$.

**Lemma 7.4.** *For $d \geq 3$ and $A$ a positive integer, there is a constant $E_d$ such that whenever $\gamma \in \Gamma_d \smallsetminus \mathbb{A}_d$ is conjugate to $\gamma a_d^{\ell}$ with $d(1,\gamma) \leq n$ and $|\ell| \leq A\,n^{d(d-1)}$, then there exists $z \in \Gamma_d$ such that $z^{-1}\gamma z = \gamma a_d^{\ell}$ and $d(1,z) \leq AE_d\,n^d$.*

*Proof.* The elements $c \in \Gamma_d$ that conjugate $\gamma$ to an element of the form $\gamma a_d^{\ell}$ are precisely those whose image $\overline{c} \in \Gamma_{d-1}$ centralise the image $\overline{\gamma}$ of $\gamma$ and have $\zeta_{\overline{\gamma}}(\overline{c}) = \ell$. In the light of these comments, the lemma is a consequence of the following claim.

<u>*Claim 1:*</u> *If $g \in \Gamma_{d-1} \smallsetminus \mathbb{A}_{d-1}$ with $d(1,g) \leq n$ and $\ell \in \operatorname{im}\zeta_g$ with $|\ell| \leq A\,n^{d(d-1)}$, then there exists $y \in C_{\Gamma_{d-1}}(g)$ with $d(1,y) < AE_d\,n$ and $\zeta_g(y) = \ell$.*

Before proving Claim 1, we examine $\operatorname{im}\zeta_g$. From Proposition 4.1(1) we know that $C_{\Gamma_{d-1}}(g) = \langle g_0, a_{d-1}\rangle \cong \mathbb{Z}^2$ where

$$g = g_0^p a_{d-1}^r$$

with $p > 0$ maximal. A count of the letters $t$ gives $p \leq n$. Note that by moving the $a_{d-1}$ letters into $g_0$ if necessary, we may assume $0 < r < p$.

<u>*Claim 2:*</u> *If $g = g_0^p a_{d-1}^r = t^{pq}x$ with $x \in \mathbb{A}_{d-1}$ and $0 < r < p$, then* $\operatorname{im} \zeta_g < \mathbb{Z}$ *is generated by $\zeta_g(a_{d-1}^{-1}) = pq$ and $\zeta_g(g_0) = rq$. Hence* $\operatorname{im} \zeta_g = qe\mathbb{Z}$*, where* $e = \gcd(p, r)$.

To see the truth of Claim 2, recall that $\zeta_g(h)$ is independent of the choices of preimages that we take in $\Gamma_d$, so writing $x$ as a geodesic word in the letters $a_i$, we can take $t^{pq}x$ as the preimage of $g$ in $\Gamma_d$ and calculate

$$-\zeta_g(a_{d-1}) = \zeta_g(a_{d-1}^{-1}) = pq \tag{4}$$

as follows

$$a_{d-1}t^{pq}xa_{d-1}^{-1} = t^{pq}\phi_d^{pq}(a_{d-1})xa_{d-1}^{-1} = t^{pq}a_{d-1}a_d^{pq}xa_{d-1}^{-1} = t^{pq}xa_d^{pq},$$

and we can verify that $\zeta_g(g_0) = qr$ by fixing a preimage $\tilde{g_0} = t^q y$ with $y \in \mathbb{A}_d$ then defining $\tilde{g} := \tilde{g_0}^p a_{d-1}^r$ and calculating

$$\tilde{g_0}^{-1}(\tilde{g_0}^p a_{d-1}^r)\tilde{g_0} = \tilde{g_0}^p \tilde{g_0}^{-1} a_{d-1}^r \tilde{g_0} = \tilde{g_0}^p y^{-1}\phi_d^q(a_{d-1}^r)y = \tilde{g_0}^p y^{-1} a_{d-1}^r a_d^{qr} y = \tilde{g} a_d^{qr}.$$

This proves Claim 2.

With Claim 2 in hand, our next goal is to find a small element of $C_{\Gamma_{d-1}}(g)$ whose image generates $\operatorname{im} \zeta_g$. Lemma 6.1 provides us with positive integers $\lambda < p$ and $\mu \leq r < p$ such that $\lambda r - \mu p = e$. Then, for $g = g_0^p a_{d-1}^r$ we have

$$\zeta_g(g_0^\lambda a_{d-1}^\mu) = \lambda\zeta_g(g_0) - \mu\zeta_g(a_{d-1}^{-1}) = \lambda rq - \mu pq = qe. \tag{5}$$

Corollary 5.2, applied to the equality $ga_{d-1}^{-r} = g_0^p$ in $\Gamma_{d-1}$, gives us the first of the following inequalities and $r < p$ gives the second:

$$d(1, g_0) \leq K_{d-1} d(1, ga_{d-1}^{-r}) \leq K_{d-1}(d(1, g) + p). \tag{6}$$

So we estimate that in $\Gamma_{d-1}$ we have

$$d(1, g_0^\lambda a_{d-1}^\mu) \leq \lambda\, d(1, g_0) + \mu \leq \lambda K_{d-1}\,(d(1, g) + p) + \mu \leq p(K_{d-1}(n+p)+1). \tag{7}$$

**The Final Argument.** We are now ready to prove Claim 1. We have $q \neq 0$ because $g \notin \mathbb{A}_{d-1}$. So, given $\ell \in \operatorname{im} \zeta_g$ with $|\ell| \leq A\, n^{d(d-1)}$, because $pq \in \operatorname{im} \zeta_g = qe\mathbb{Z}$, we can write

$$\ell = Mpq + \rho qe$$

for integers $\rho$ and $M$ with $0 \leq \rho < p$ and $|M| \leq A\, n^{d(d-1)}$. Thus, using (4) and (5), we have

$$\ell = M\, \zeta_g(a_{d-1}^{-1}) + \rho\, \zeta_g(g_0^\lambda a_{d-1}^\mu) = \zeta_g(a_{d-1}^{-M}(g_0^\lambda a_{d-1}^\mu)^\rho).$$

We are going to argue that the term

$$y := a_{d-1}^{-M}(g_0^{\lambda}a_{d-1}^{\mu})^{\rho}$$

from the last bracket satisfies Claim 1.

From the triangle inequality in $\Gamma_{d-1}$ we have, invoking Lemma 3.3 and estimate (7) in the second line, and using $|M| \leq A\, n^{d(d-1)}$ and $p \leq n$ in the third,

$$\begin{aligned} d(1,y) = d(1, a_{d-1}^{-M}(g_0^{\lambda}a_{d-1}^{\mu})^{\rho}) &\leq d(1, a_{d-1}^{M}) + \rho\, d(1, g_0^{\lambda}a_{d-1}^{\mu}) \\ &\leq D_{d-1}|M|^{1/(d-1)} + p^2(K_{d-1}(n+p)+1) \\ &\leq D_{d-1}An^d + n^2(2K_{d-1}n+1) \\ &\leq AE_d n^d \end{aligned}$$

provided $d \geq 3$ and we choose $E_d$ to be sufficiently larger than $D_{d-1}$ and $K_{d-1}$. This completes the proof of Claim 1, hence Lemma 7.4, Proposition 7.1 and Theorem 1.1. □

## References

[BMS93] G. Baumslag, C. F. Miller, III, and H. Short. Isoperimetric inequalities and the homology of groups. *Invent. Math.*, 113(3):531–560, 1993.

[BORS02] J.-C. Birget, A. Yu. Ol′shanskii, E. Rips, and M. V. Sapir. Isoperimetric functions of groups and computational complexity of the word problem. *Ann. of Math. (2)*, 156(2):467–518, 2002.

[BP94] M. R. Bridson and Ch. Pittet. Isoperimetric inequalities for the fundamental groups of torus bundles over the circle. *Geom. Dedicata*, 49(2):203–219, 1994.

[BR] M. R. Bridson and T. R. Riley. Linear Diophantine equations and conjugator length in 2-step nilpotent groups. preprint, `arXiv:2506.01239`.

[Bri99] M. R. Bridson. Fractional isoperimetric inequalities and subgroup distortion. *J. of the A. Math. Soc.*, 12(4):1103–1118, 1999.

[BRSa] M. R. Bridson, T. R. Riley, and A. Sale. Conjugacy in a family of free-by-cyclic groups. preprint, `arXiv:2506.01248`.

[BRSb] M. R. Bridson, T. R. Riley, and A. Sale. Conjugator length in finitely presented groups. In preparation.

[CM77] D. J. Collins and C. F. Miller, III. The conjugacy problem and subgroups of finite index. *Proc. London Math. Soc. (3)*, 34(3):535–556, 1977.

[Gro93] M. Gromov. Asymptotic invariants of infinite groups. In G. Niblo and M. Roller, editors, *Geometric group theory II*, number 182 in LMS lecture notes. Camb. Univ. Press, 1993.

[Sal16] A. W. Sale. Conjugacy length in group extensions. *Comm. Algebra*, 44(2):873–897, 2016.

Martin R. Bridson
Mathematical Institute, Andrew Wiles Building, Oxford OX2 6GG, United Kingdom
bridson@maths.ox.ac.uk, https://people.maths.ox.ac.uk/bridson/

Timothy R. Riley
Department of Mathematics, 310 Malott Hall, Cornell University, Ithaca, NY 14853, USA
tim.riley@math.cornell.edu, http://www.math.cornell.edu/∼riley/