

# HYDRA GROUPS

W. DISON AND T.R. RILEY

**ABSTRACT.** We give examples of  $CAT(0)$ , biautomatic, free-by-cyclic, one-relator groups which have finite-rank free subgroups of huge (Ackermannian) distortion. This leads to elementary examples of groups whose Dehn functions are similarly extravagant. This behaviour originates in manifestations of *Hercules-versus-the-hydra battles* in string-rewriting.

**2010 Mathematics Subject Classification:** 20F65, 20F10, 20F67

*Key words and phrases:* free-by-cyclic group, subgroup distortion, Dehn function, hydra

## 1. INTRODUCTION

**1.1. Hercules versus the hydra.** Hercules' second labour was to fight the Lernaean hydra, a beast with multiple serpentine heads enjoying magical regenerative powers: whenever a head was severed, two grew in its place. Hercules succeeded with the help of his nephew, Iolaus, who stopped the regrowth by searing the stumps with a burning torch after each decapitation. The extraordinarily fast-growing functions we will encounter in this article stem from a re-imagining of this battle.

For us, a *hydra* will be a finite-length *positive* word on the alphabet  $a_1, a_2, a_3, \dots$  — that is, it includes no inverse letters  $a_1^{-1}, a_2^{-1}, a_3^{-1}, \dots$ . Hercules fights a hydra by striking off its first letter. The hydra then regenerates as follows: each remaining letter  $a_i$ , where  $i > 1$ , becomes  $a_i a_{i-1}$  and the  $a_1$  are unchanged. This process — removal of the first letter and then regeneration — repeats, with Hercules victorious *when* (not *if*!) the hydra is reduced to the empty word  $\varepsilon$ .

For example, Hercules defeats the hydra  $a_2 a_3 a_1$  in five strikes:

$$a_2 a_3 a_1 \rightarrow a_3 a_2 a_1 \rightarrow a_2 a_1 a_1 \rightarrow a_1 a_1 \rightarrow a_1 \rightarrow \varepsilon.$$

(Each arrow represents the removal of the first letter and then regeneration.)

**Proposition 1.1.** *Hercules defeats all hydra.*

*Proof.* When fighting a hydra in which the highest index present is  $k$ , no  $a_i$  with  $i > k$  will ever appear, and nor will any new  $a_k$ . The prefix before the first  $a_k$  is itself a hydra, which, by induction, we can assume Hercules defeats. Hercules will then remove that  $a_k$ , decreasing the total number of  $a_k$  present. It follows that Hercules eventually wins.  $\square$

However these battles are of extreme duration. Define  $\mathcal{H}(w)$  to be the number of strikes it takes Hercules to vanquish the hydra  $w$ , and for integers  $k \geq 1$ ,  $n \geq 0$ , define  $\mathcal{H}_k(n) :=$

$\mathcal{H}(a_k^n)$ . We call the  $\mathcal{H}_k$  *hydra functions*. Here are some values of  $\mathcal{H}_k(n)$ .

	1	2	3	4	...	$n$	...
1	1	2	3	4	...	$n$	...
2	1	3	7	15	...	$2^n - 1$	...
3	1	4	46	$3 \cdot 2^{46} - 2$	...	...	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$k$	1	$k + 1$	$\vdots$	$\vdots$			
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$			

To see that  $\mathcal{H}_2(n) = 2^n - 1$  for all  $n$ , note that

$$\mathcal{H}(a_2^{n+1}) = \mathcal{H}(a_2^n) + \mathcal{H}(a_2 a_1^{\mathcal{H}(a_2^n)}) = 2\mathcal{H}(a_2^n) + 1.$$

And  $\mathcal{H}_3(n)$  is essentially an  $n$ -fold iterated exponential function because, for all  $n > 0$ ,

$$\mathcal{H}_3(n+1) = 3 \cdot 2^{\mathcal{H}_3(n)} - 2,$$

by the calculations

$$\begin{aligned} \mathcal{H}(a_3^{n+1}) &= \mathcal{H}(a_3^n) + 1 + \mathcal{H}(a_2 a_1 a_2 a_1^2 \dots a_2 a_1^{\mathcal{H}(a_3^n)}), \\ \mathcal{H}(a_2 a_1 a_2 a_1^2 \dots a_2 a_1^m) &= 3 \cdot 2^m - m - 3. \end{aligned}$$

Extending this line of reasoning, we will derive relationships (15) and (19) in Section 3 from which it will follow that

$$\mathcal{H}_4(3) = 3 \cdot 2^{3 \cdot 2^{3 \cdot 2^{3 \cdot 2^{5-1-1-1}}} - 1} - 1.$$

So these functions are extremely wild. The reason behind the extreme fast growth is a nested recursion. What we have is a variation on Ackermann's functions  $A_k : \mathbb{N} \rightarrow \mathbb{N}$ , defined for integers  $k, n \geq 0$  by:

$$A_0(n) = n + 2 \text{ for } n \geq 0,$$

$$A_k(0) = \begin{cases} 0 & \text{for } k = 1 \\ 1 & \text{for } k \geq 2, \end{cases}$$

$$\text{and } A_{k+1}(n+1) = A_k(A_{k+1}(n)) \text{ for } k, n \geq 0.$$

So, in particular,  $A_1(n) = 2n$ ,  $A_2(n) = 2^n$  and  $A_3(n) = \exp_2^{(n)}(1)$ , the  $n$ -fold iterated power of 2. (Definitions of Ackermann's functions occur with minor variations in the literature.) Ackermann's functions represent the successive levels of the Grzegorzcyk hierarchy, which is a grading of all primitive recursive functions — see, for example, [31].

We will prove the following relationship in Section 3. Our notation in this proposition and henceforth in this article is that for  $f, g : \mathbb{N} \rightarrow \mathbb{N}$ , we write  $f \leq g$  when there exists  $C > 0$  such that for all  $n$  we have  $f(n) \leq Cg(Cn + C) + Cn + C$ . This gives an equivalence relation capturing qualitative agreement of the growth rate:  $f \simeq g$  if and only if  $f \leq g$  and  $g \leq f$ .

**Proposition 1.2.** *For all  $k \geq 1$ ,  $\mathcal{H}_k \simeq A_k$ .*

Other hydra dwell in the mathematical literature, particularly in the context of results concerning independence from Peano arithmetic and other logical systems. The hydra of Kirby and Paris [22], based on finite rooted trees, are particularly celebrated. Similar, but yet more extreme hydra were later constructed by Buchholz [10]. And creatures that, like ours, are finite strings that regenerate on *decapitation* were defined by Hamano and Okada [20] and then independently by Beklemishev [5]. They go by the name of *worms*, are descended from Buchholz's hydra, involve more complex regeneration rules, and withstand Hercules even longer.

**1.2. Wild subgroup distortion.** The distortion function  $\text{Dist}_H^G : \mathbb{N} \rightarrow \mathbb{N}$  for a subgroup  $H$  with finite generating set  $T$  inside a group  $G$  with finite generating set  $S$  compares the intrinsic word metric  $d_T$  on  $H$  with the extrinsic word metric  $d_S$ :

$$\text{Dist}_H^G(n) := \max \{ d_T(1, g) \mid g \in H \text{ with } d_S(1, g) \leq n \}.$$

Up to  $\simeq$  it does not depend on the particular finite generating sets used.

A manifestation of our Hercules–versus–the–hydra battle leads to the result that even for apparently benign  $G$  and  $H$ , distortion can be wild.

**Theorem 1.3.** *For each integer  $k \geq 1$ , there is a finitely generated group  $G_k$  that*

- *is free-by-cyclic,*
- *can be presented with only one defining relator,*
- *is CAT(0),*
- *is biautomatic,*
- *and enjoys the rapid decay property,*

*and yet has a rank- $k$  free subgroup  $H_k$  that is distorted like the  $k$ -th of Ackermann’s functions — that is,  $\text{Dist}_{H_k}^{G_k} \simeq A_k$ .*

This distortion of a free subgroup of a CAT(0) group stands in stark contrast to that of any abelian subgroup — they are always quasi-isometrically embedded (see Theorem 4.10 of Chapter III.Γ in [9], for example) and so no more than linearly distorted. The distortion we achieve exceeds that found in the hyperbolic groups of Mitra [26] and the subsequent 2-dimensional CAT(−1) groups of Barnard, Brady and Dani [2]: first, for all  $k$ , they give examples with a free subgroup of distortion  $\simeq \exp^{(k)}(n)$ , and then they give examples with free subgroups whose distortion functions grow faster than  $\exp^{(k)}(n)$  for every  $k$ . However, our examples contain  $\mathbb{Z}^2$  subgroups and so are not hyperbolic.

Explicitly, our examples are

$$(1) \quad G_k = \langle a_1, \dots, a_k, t \mid t^{-1}a_1t = a_1, t^{-1}a_it = a_ia_{i-1} \ (\forall i > 1) \rangle$$

and their subgroups

$$H_k := \langle a_1t, \dots, a_kt \rangle.$$

So  $G_k$  is the free-by-cyclic group  $F(a_1, \dots, a_k) \rtimes \mathbb{Z}$  where  $\mathbb{Z} = \langle t \rangle$  and  $t$  acts by the automorphism of  $F(a_1, \dots, a_k)$  that is the restriction of the automorphism  $\theta$  of  $F(a_1, a_2, \dots)$  defined by

$$(2) \quad \theta(a_i) = \begin{cases} a_1 & i = 1, \\ a_ia_{i-1} & i > 1. \end{cases}$$

For  $i \leq j$ , the canonical homomorphism  $G_i \rightarrow G_j$  is an inclusion as the free-by-cyclic normal forms of an element of  $G_i$  and its image in  $G_j$  are the same. So the direct limit of the  $G_i$  under these inclusions is

$$G = \langle t, a_1, a_2, \dots \mid t^{-1}a_1t = a_1, t^{-1}a_it = a_ia_{i-1} \ (\forall i > 1) \rangle.$$

Also, the subgroup  $H := \langle a_1t, a_2t, \dots \rangle$  of  $G$  is  $\varinjlim H_i$  and  $H_k = G_k \cap H$ .

Our convention is that  $[a, b] = a^{-1}b^{-1}ab$ . By re-expressing the original relations as  $[a_1, t] = 1$  and  $a_{i-1} = [a_i, t]$  for  $i > 1$  and then eliminating  $a_1, \dots, a_{k-1}$ , one can present  $G_k$  with one relation, a left-normed commutator:

$$G_k \cong \langle a_k, t \mid [a_k, \underbrace{t, \dots, t}_k] = 1 \rangle.$$

As we will explain in Section 4,  $G_k$  can also be presented as

$$\langle \alpha_1, \dots, \alpha_k, s \mid \alpha_1^{-1}s\alpha_1 = s, \alpha_i^{-1}s\alpha_i = \alpha_{i-1} \ (i > 1) \rangle.$$

By checking the link condition (see, for example, [9, II.5.24]) one finds that the Cayley 2-complex of this presentation (that is, the universal cover of the associated presentation 2-complex), metrized so that each 2-cell is a Euclidean square, is CAT(0). Gersten & Short [18] proved that all such groups are automatic, and later Niblo & Reeves [27] proved that a more general class of groups, those acting geometrically on CAT(0) cube complexes, are biautomatic.

The groups  $G_k$  enjoy the property of *rapid decay* as a consequence of [21, Corollary 2.1.10].

We remark that a corollary of our recursive upper bound on  $\text{Dist}_{H_k}^{G_k}$  is that the membership problem for  $H_k$  in  $G_k$  is decidable.

The family  $G_k$  have received attention before. The group  $G_2$  was shown by Gersten in [17] to be CAT(0) with quadratic divergence function. He gave the free-by-cyclic, the one-relator, and the CAT(0) presentations of  $G_2$  we have described. In Example 1.3.3 of [24] Macura identified  $G_3$  as being CAT(0) and used it to give a CAT(0) complex which has cubic divergence function and admits a discrete cocompact action by a group of isometries. Macura also mentions  $G_2$  and  $G_3$  in [23] as examples in the context of Kolchin maps and quadratic isoperimetric functions. It is stated in [4] (Example 4) that  $G_3$  is biautomatic. Bridson uses  $G_k$  in [8] as a starting point to construct free-by-free groups with Dehn functions that are polynomial of degree  $k+1$  and he shows them to be subgroups of  $\text{Out}(F_n)$  for suitable  $n$ . Additionally he shows his example are asynchronously automatic via normal forms which have length  $\simeq n^k$ , but by no shorter normal form. En route he shows (Section 4.1(3)) that free-by-cyclic  $F_k \rtimes \mathbb{Z}$  groups, such as  $G_k$ , embed in  $\text{Aut}(F_k)$ .

Examples of yet more extreme distortion are known, even for subgroups of hyperbolic groups. Arzhantseva & Osin [1, §3.4] and Pittet [29] explain an argument attributed to Sela in [19, §3, 3.K''\_3]: the Rips construction, applied to a finitely presentable group with unsolvable word problem yields a hyperbolic (indeed,  $C'(1/6)$  small-cancellation) group  $G$  with a finitely generated subgroup  $N$  such that  $\text{Dist}_N^G$  is not bounded above by any recursive function. The reason is that when  $N$  is a finitely generated normal subgroup of a finitely presented group  $G$ , there is an upper bound for the Dehn function of  $G/N$  in terms of the Dehn function of  $G$  and the distortion of  $N$  in  $G$  — see [14, Corollary 8.2], [29]. Ol'shanskii & Sapir in [28, Theorem 2] provide another source of extreme examples — using Mikhailova's construction as their starting point, they show that the set of distortion functions of finitely generated subgroups of  $F_2 \times F_2$  coincides (up to  $\simeq$ ) with the set of Dehn functions of finitely presented groups. However, it seems harder to gauge what distortion functions arise for finitely *presented* subgroups of CAT(0) or hyperbolic groups.

**1.3. Extreme Dehn functions.** The Dehn function  $\text{Area}(n)$  of a finitely presented group  $\langle A \mid R \rangle$  is related to the group's word problem in that  $\text{Area}(n)$  is the minimal  $N$  such that given any word  $w$  of length at most  $n$  that represents the identity,  $w$  freely equals some product  $\prod_{i=1}^N u_i^{-1} r_i u_i$  of  $N' \leq N$  conjugates of relators  $r_i \in R^{\pm 1}$ , or, equivalently, one can reduce  $w$  to the empty word by applying defining relations at most  $N$  times and removing or inserting inverse pairs of letters. At the same time, the Dehn function is a natural geometric invariant (in fact, a quasi-isometry invariant up to  $\simeq$ ) of a group:  $\text{Area}(n)$  is the minimal  $N$  such that any edge-loop of length at most  $n$  in the Cayley 2-complex of  $\langle A \mid R \rangle$  can be spanned by a combinatorial filling disc (a van Kampen diagram) with area (that is, number of 2-cells) at most  $N$ . This geometric perspective is related to the classical notion of an isoperimetric function in Riemannian geometry in that if  $\langle A \mid R \rangle$  is the fundamental group

of a closed Riemannian manifold  $M$ , then its Dehn function is  $\simeq$ -equivalent to the minimal isoperimetric function of the universal cover of  $M$ .

Theorem 1.3 leads to strikingly simple examples of finitely presented groups with huge Dehn functions, namely the HNN-extensions of  $G_k$  with stable letter commuting with all elements of the subgroup  $H_k$ .

**Theorem 1.4.** *For  $k \geq 2$ , the Dehn function of the group*

$$\Gamma_k := \langle a_1, \dots, a_k, t, p \mid t^{-1}a_1t = a_1, t^{-1}a_it = a_ia_{i-1} \ (i > 1), [p, a_it] = 1 \ (i > 0) \rangle.$$

*is  $\simeq$ -equivalent to  $A_k$ .*

So, together with  $\Gamma_1$ , which has Dehn function  $\simeq$ -equivalent to  $n \mapsto n^2$  (see Proposition 10.1), these groups have Dehn functions that are representative of each graduation of the Grzegorzcyk hierarchy of primitive recursive functions. Details of the proof are in Section 10.

These are not the only such examples (but we believe they are the first that are explicit and elementary): Cohen, Madlener and Otto [12, 13, 25] embedded algorithms (*modular Turing machines*, in fact) with running times like  $n \mapsto A_k(n)$  in groups so that the running of the algorithm is displayed in van Kampen diagrams so as to make the Dehn function reflect the time-complexity of the algorithms. They state that their techniques produce yet more extreme examples as they also apply to an algorithm with running time like  $n \mapsto A_n(n)$ , and so yield a group with Dehn function that is recursive but not primitive recursive. More extreme still, any finitely presentable group with undecidable word problem is not bounded above by any recursive function.

Elementary examples of groups with large Dehn function are described by Gromov in [19, §4], but their behaviour is not so extreme. There is the family

$$\langle x_0, \dots, x_k \mid x_{i+1}^{-1}x_ix_{i+1} = x_i^2 \ (i < k) \rangle,$$

which has Dehn function  $\simeq$ -equivalent to  $n \mapsto \exp_2^{(k)}(n)$ . [We write  $\exp_2(n)$  to denote  $2^n$ .] And Baumslag's group [3]

$$(3) \quad \langle a, b \mid (b^{-1}a^{-1}b)a(b^{-1}ab) = a^2 \rangle,$$

which contains  $\langle x_0, \dots, x_k \mid x_{i+1}^{-1}x_ix_{i+1} = x_i^2 \ (i \geq 0) \rangle$  as a normal subgroup, was shown by Platonov [30] to have Dehn function  $\simeq$ -equivalent to  $n \mapsto \exp^{(\log_2 n)}(1)$ . (Prior partial results towards this direction are in [6, 15, 16].)

**1.4. The organisation of the article.** We believe the most compelling assertion of Theorem 1.3 to be the existence of groups  $H_k$  and  $G_k$  with  $H_k$  free of rank  $k$ ,  $G_k$  enjoying the bulleted list of properties, and  $\text{Dist}_{H_k}^{G_k}$  bounded below by  $A_k$ . In particular, this shows that there is no uniform upper bound on the level in the Grzegorzcyk hierarchy at which the functions  $\text{Dist}_{H_k}^{G_k}$  appear. The reader who is primarily interested in these components of Theorem 1.3 need only read up to the end of Section 6. The contents of this half of the article are as follows. In Section 2 we derive a collection of elementary properties of the Ackermann functions that will be used elsewhere in the paper. Section 3 contains a proof of Proposition 1.2 comparing the hydra functions to Ackermann's functions. In Sections 4 and 5 we prove, respectively, that the groups  $G_k$  are CAT(0) and that the groups  $H_k$  are free. And in Section 6 we prove that each function  $\text{Dist}_{H_k}^{G_k}$  is bounded below by  $\mathcal{H}_k$  — combining this result with Proposition 1.2 gives the lower bound  $A_k$ .

Our proof that each function  $\text{Dist}_{H_k}^{G_k}$  lies in the same  $\simeq$ -equivalence class of functions as  $A_k$  — *i.e.* that  $A_k$  is an upper bound for  $\text{Dist}_{H_k}^{G_k}$  — is considerably more involved than that of the lower bound and occupies most of the second half of the article: Sections 7, 8 and 9.

In deriving the upper bound, a key notion will be that of passing a power of  $t$  through a word  $w$  on the letters  $a_i$ . We explain this idea in Section 7, where we also identify some recursive structure that will be crucial in facilitating an inductive analysis. In Section 8 we focus in detail on the situation where  $w$  is of the form  $\theta^n(a_k^{\pm 1})$  and derive preliminary result that will feed into the main proof, presented in Section 9, that  $\text{Dist}_{H_k}^{G_k} \leq A_k$ .

Finally, in Section 10, we prove Theorem 1.4, which gives the Dehn functions of the groups  $\Gamma_k$ .

We illustrate some of our arguments using van Kampen diagrams, particularly observing their *corridors* (also known as *bands*). For an introduction see, for example, I.8A.4 and the proof of Proposition 6.16 in III.Γ of [9].

We denote the length of a word  $w$  by  $\ell(w)$ . We write  $w = w(a_1, \dots, a_k)$  when  $w$  is a word on  $a_1^{\pm 1}, \dots, a_k^{\pm 1}$ .

**1.5. Acknowledgements.** We are grateful to Martin Bridson for a number of conversations on this work, to Indira Chatterji for pointing out that the groups  $G_k$  enjoy the rapid decay property, to Volker Diekert for a discussion on Ackermann's functions, and to John McCammond for help with some computer explorations of Hercules' battle with the hydra.

## 2. ACKERMANN'S FUNCTIONS

Throughout this article we will frequently compare functions to Ackermann's functions and will find the following relationships useful.

**Lemma 2.1.** *For integers  $k, l, m, n$ , the following relations hold within the given domains:*

$$\begin{aligned}
 (4) \quad & A_k(A_{k+1}(n)) = A_{k+1}(n+1), & k, n \geq 0, \\
 (5) \quad & A_k(1) = 2, & k \geq 1, \\
 (6) \quad & A_k(2) = 4, & k \geq 0, \\
 (7) \quad & A_k(n) \leq A_{k+1}(n), & k \geq 1; n \geq 0, \\
 (8) \quad & A_k(n) < A_k(n+1), & k, n \geq 0, \\
 (9) \quad & n \leq A_k(n), & k, n \geq 0,
 \end{aligned}$$

(with equality holding in (9) if and only if  $(k, n) = (1, 0)$ )

$$\begin{aligned}
 (10) \quad & mA_k(n) \leq A_k(nm), & k, n \geq 1; m \geq 0, \\
 (11) \quad & mA_k^{(l)}(n) \leq A_k^{(l+m)}(n), & k \geq 1; l, m, n \geq 0, \\
 (12) \quad & A_k(n) + A_k(m) \leq A_k(n+m), & k, n, m \geq 1, \\
 (13) \quad & A_k(n) + m \leq A_k(n+m), & k, n, m \geq 0, \\
 (14) \quad & (A_k(n))^m \leq A_k(nm), & k \geq 2; n, m \geq 0.
 \end{aligned}$$

*Proof.* Equation (4) follows immediately from the definition of the Ackermann functions. Equations (5) and (6) follow from (4) by an easy induction on  $k$ .

Before proving (7), (8) and (9), we first prove that non-strict versions of these inequalities hold. The proof is by induction on  $k$  and  $n$ . It is easy to check that (7) holds if  $k = 1$  or if  $n = 0$  and that (8) and (9) hold if  $k = 0$ , if  $k = 1$  or if  $n = 0$ . Now let  $k' > 1$  and  $n' > 0$  and suppose, as an inductive hypothesis, that (7), (8) and (9) hold (not necessarily strictly) if  $k < k'$  or if  $k = k'$  and  $n < n'$ . We prove that the inequalities hold if  $k = k'$  and  $n = n'$ . For (7), we calculate that  $A_{k'}(n') = A_{k'-1}(A_{k'}(n' - 1)) \leq A_{k'-1}(A_{k'+1}(n' - 1)) \leq$

$A_{k'}(A_{k'+1}(n' - 1)) = A_{k'+1}(n')$ , where we have applied (4) and the inductive hypothesis versions of (7) and (8). For (8), we calculate that  $A_{k'}(n') \leq A_{k'-1}(A_{k'}(n')) = A_{k'}(n' + 1)$ , where we have used (4) and the inductive hypothesis version of (9). For (9), we calculate that  $n' \leq 2n' = A_1(n') \leq A_{k'}(n')$ , where we have used the inductive hypothesis version of (7). This completes the proof that (7), (8) and (9) hold in non-strict form. Now observe that equality in (9) at  $(k, n) = (k', n')$  requires  $n' = 2n'$ , whence  $n' = 0$ . Since  $A_k(0) = 1$  for all  $k \geq 2$ , equality in (9) holds if and only if  $(k, n) = (1, 0)$ . It follows that equality in (8) at  $(k, n) = (k', n')$  would require that  $A_{k'}(n') = 0$  and  $k' - 1 = 1$ , whence  $A_2(n') = 0$ . But  $A_2(n) = 2^n > 0$  for all  $n$  and so the inequality (8) is strict.

We now prove inequality (10). This clearly holds if  $m = 0$ , so suppose that  $m \geq 1$ . The proof is by induction on  $k$  and  $n$ . It is clear that (10) holds if  $k = 1$ . The inequality also holds if  $n = 1$  since, applying (5) and (7), we calculate that  $mA_k(1) = 2m = A_1(m) \leq A_k(m)$ . Now let  $k', n' > 1$  and suppose, as an inductive hypothesis, that (10) holds if  $k < k'$  or if  $k = k'$  and  $n < n'$ . We calculate that  $mA_{k'}(n') = mA_{k'-1}(A_{k'}(n' - 1)) \leq A_{k'-1}(mA_{k'}(n' - 1)) \leq A_{k'-1}(A_{k'}(mn' - m)) \leq A_{k'-1}(A_{k'}(mn' - 1)) = A_{k'}(mn')$ , where we have used (4) and (8). Thus the inequality holds if  $(k, n) = (k', n')$ , completing the proof of (10).

For inequality (11) observe that, by (9),  $mA_k^{(l)}(n) \leq A_{k+1}(m)A_k^{(l)}(n) = A_k^{(m)}(1)A_k^{(l)}(n)$ . It also follows from (9) that  $A_k^{(i)}(1) \geq 1$  for all  $i \geq 0$ . We can thus apply (10), together with (8), to show that  $A_k^{(m)}(1)A_k^{(l)}(n) \leq A_k^{(m)}(A_k^{(l)}(n)) = A_k^{(l+m)}(n)$ .

We prove (12) by induction on  $k$ . It is clear that the inequality holds if  $k = 1$ , so suppose that  $k > 1$  and that the result is true for smaller values of  $k$ . Without loss of generality suppose that  $n \leq m$ . It follows from (9) that  $A_{k-1}^{(i)} \geq 1$  for all  $i \geq 0$ , and so we can apply the induction hypothesis to calculate that  $A_k(n) + A_k(m) = A_{k-1}^{(m)}(1) + A_{k-1}^{(m)}(1) \leq A_{k-1}^{(m)}(1 + A_{k-1}^{(m-n)}(1)) = A_{k-1}^{(m)}(1 + A_k(m - n))$ . Applying (8) gives that this quantity is at most  $A_{k-1}^{(m)}(A_k(m - n + 1)) = A_k(m + 1) \leq A_k(m + n)$ .

We now prove inequality (13). This clearly holds if  $k = 0$ ,  $k = 1$  or  $m = 0$ . If  $k \geq 2$  and  $n = 0$ , then  $A_k(n) + m = m + 1 \leq A_k(m) = A_k(n + m)$  by (9). It remains to prove (13) if  $k, n, m \geq 1$ . But in this case  $A_k(n) + m \leq A_k(n) + A_k(m) \leq A_k(n + m)$  by (9) and (12).

Finally, we prove (14) by induction on  $k$ . It is clear that the inequality holds if  $k = 2$ , so suppose that  $k \geq 3$  and that the result holds for smaller values of  $k$ . It is also clear that the inequality holds if  $n = 0$  or if  $m = 0$ ; suppose that  $n, m \geq 1$ . Applying the induction hypothesis, together with (4), we calculate that  $A_k(n)^m = A_{k-1}(A_k(n - 1))^m \leq A_{k-1}(mA_k(n - 1))$ . Applying (4), (8) and (10), we see that this quantity is at most  $A_{k-1}(A_k(nm - m)) \leq A_{k-1}(A_k(nm - 1)) = A_k(nm)$ .  $\square$

### 3. COMPARING THE HYDRA FUNCTIONS TO ACKERMANN'S FUNCTIONS

In this section we prove Proposition 1.2 comparing Ackermann's functions to the hydra functions. The proof will proceed via a third family of functions  $\phi_k$ . In this section  $\phi_k(n)$  will be defined for  $n \geq 0$ ; subsequently we will give a more general definition with an expanded domain.

For integers  $k \geq 1$  and  $n \geq 0$ , define  $\phi_k(n) := \mathcal{H}(\theta^n(a_k))$ . The functions  $\mathcal{H}_k$  satisfy the recursion relation

$$(15) \quad \mathcal{H}_k(n + 1) = \mathcal{H}_k(n) + \phi_k(\mathcal{H}_k(n))$$

since after  $\mathcal{H}_k(n)$  strikes the word  $a_k^{n+1}$  has become  $\theta^{\mathcal{H}_k(n)}(a_k)$ . We will need the following elementary properties of the functions  $\phi_k$ .

**Lemma 3.1.** For integers  $k \geq 1$  and  $n \geq 0$ ,

$$(16) \quad \phi_k(0) = 1,$$

$$(17) \quad \phi_2(n) = n + 1,$$

$$(18) \quad \phi_k(n) \geq 1,$$

$$(19) \quad \phi_{k+1}(n+1) = \phi_{k+1}(n) + \phi_k(\phi_{k+1}(n) + n).$$

For integers  $k \geq 2$  and  $n \geq 0$ ,

$$(20) \quad \phi_k(n) < \phi_k(n+1),$$

$$(21) \quad \phi_k(n) \geq n.$$

*Proof.* Assertions (16), (17) and (18) are straightforward. For (19), note that, by induction on  $n$ ,  $\theta^{n+1}(a_{k+1}) = a_{k+1}a_k\theta(a_k) \dots \theta^n(a_k)$  and hence  $\theta^{n+1}(a_{k+1}) = \theta^n(a_{k+1})\theta^n(a_k)$ . Thus, after  $\phi_{k+1}(n)$  strikes,  $\theta^{n+1}(a_{k+1})$  has become  $\theta^{\phi_{k+1}(n)}(\theta^n(a_k)) = \theta^{\phi_{k+1}(n)+n}(a_k)$ . Inequality (20) follows immediately from (18) and (19) and inequality (21) follows from (18) and (20).  $\square$

It is easy to check that  $\phi_1 \simeq A_0$  and  $\phi_2 \simeq A_1$ . As such, the next result is sufficient to establish that  $\phi_k \simeq A_{k-1}$  for  $k \geq 1$ .

**Lemma 3.2.**

- (i) For integers  $k \geq 3$  and  $n \geq 0$ ,  $\phi_k(n) \geq A_{k-1}(n)$ .
- (ii) For integers  $k \geq 2$  and  $n \geq 0$ ,  $\phi_k(n) \leq A_{k-1}(n+k) - n - k$ .

*Proof.* We prove (i) by simultaneous induction on  $k$  and  $n$ . It is immediate from (16) that the inequality holds if  $n = 0$ . Solving the recursion relation (19) with the initial condition given by (16), one checks that  $\phi_3(n) = 3 \cdot 2^n - n - 2$ . Since  $A_2(n) = 2^n$ , it is easy to check that (i) holds if  $k = 3$ . Now let  $k' > 3$  and  $n' > 0$  and suppose, as an inductive hypothesis, that the result is true if  $k < k'$  or if  $k = k'$  and  $n < n'$ . Applying (4), (18) and (20), we calculate that  $\phi_{k'}(n') = \phi_{k'}(n'-1) + \phi_{k'-1}(\phi_{k'}(n'-1) + n' - 1) \geq \phi_{k'-1}(\phi_{k'}(n'-1)) \geq \phi_{k'-1}(A_{k'-1}(n'-1)) \geq A_{k'-2}(A_{k'-1}(n'-1)) = A_{k'-1}(n')$ . Thus the result holds at  $(k, n) = (k', n')$ , completing the proof of (i).

We now make the following claim: for all  $k \geq 2$ ,  $n \geq 0$  and  $c \geq k$ ,

$$(22) \quad \phi_k(n) \leq A_{k-1}(n+c) - n + k - 2c.$$

Assertion (ii) will follow by setting  $c = k$ . The proof of this inequality is by simultaneous induction on  $k$  and  $n$ . Since  $A_1(n) = 2n$  and, by (17),  $\phi_2(n) = n + 1$ , it is straightforward to check that (22) holds if  $k = 2$ . The inequality also holds for  $n = 0$  since, by (7) and (16),  $\phi_k(0) = 1 \leq k = A_1(c) + k - 2c \leq A_{k-1}(c) + k - 2c$ . Now let  $c \geq k' > 2$  and  $n' > 0$  and suppose, as an induction hypothesis, that (22) holds if  $k < k'$  or if  $k = k'$  and  $n < n'$ . We calculate that

$$\begin{aligned} \phi_{k'}(n') &= \phi_{k'}(n'-1) + \phi_{k'-1}(\phi_{k'}(n'-1) + n' - 1) && \text{by (19)} \\ &\leq \phi_{k'}(n'-1) + A_{k'-2}(\phi_{k'}(n'-1) + n' + c - 1) \\ &\quad - \phi_{k'}(n'-1) - n' + k' - 2c \\ &= A_{k'-2}(\phi_{k'}(n'-1) + n' + c - 1) - n' + k' - 2c \\ &\leq A_{k'-2}(A_{k'-1}(n' + c - 1) + k' - c) - n' + k' - 2c && \text{by (8)} \\ &\leq A_{k'-2}(A_{k'-1}(n' + c - 1)) - n' + k' - 2c && \text{by (8)} \\ &= A_{k'-1}(n' + c) - n' + k' - 2c. && \text{by (4)} \end{aligned}$$

Thus the inequality holds if  $(k, n) = (k', n')$ , completing the proof of (22).  $\square$

Since  $A_1(n) = 2n$ ,  $\mathcal{H}_1(n) = n$ ,  $A_2(n) = 2^n$  and  $H_2(n) = 2^n - 1$ , the next result is sufficient to establish Proposition 1.2.

**Proposition 3.3.**

- (i) For integers  $k \geq 3$  and  $n \geq 2$ ,  $\mathcal{H}_k(n) \geq A_k(n)$ .
- (ii) For integers  $k \geq 1$  and  $n \geq 0$ ,  $\mathcal{H}_k(n) \leq A_k(n + k)$ .

*Proof of Proposition 3.3.* We prove (i) by induction on  $n$ . The inequality certainly holds for  $n = 2$  since, by (6),  $\mathcal{H}_k(2) = \mathcal{H}(a_k a_{k-1} a_{k-1} a_{k-2}) \geq 4 = A_k(2)$ . Now let  $n' > 2$  and suppose that (i) holds for  $n < n'$ . Applying (4), (15) and (20), together with Lemma 3.2 (i), we calculate that  $\mathcal{H}_k(n') = \mathcal{H}_k(n' - 1) + \phi_k(\mathcal{H}_k(n' - 1)) \geq \phi_k(\mathcal{H}_k(n' - 1)) \geq \phi_k(A_k(n' - 1)) \geq A_{k-1}(A_k(n' - 1)) = A_k(n')$ . Thus the inequality holds for  $n = n'$ , completing the proof of (i).

For (ii), we prove the stronger claim that, for all  $k \geq 1$ ,  $n \geq 0$ ,

$$(23) \quad \mathcal{H}_k(n) \leq A_k(n + k) - k.$$

The proof is by simultaneous induction on  $k$  and  $n$ . Since  $A_1(n) = 2n$  and  $\mathcal{H}_1(n) = n$ , it is straightforward to check that (23) holds if  $k = 1$ . The inequality holds if  $n = 0$  since, by (7),  $\mathcal{H}_k(0) = 0 \leq k = A_k(k) - k \leq A_k(k) - k$ . Now let  $k' > 1$  and  $n' > 0$  and suppose, as an inductive hypothesis, that (23) holds if  $k < k'$  or if  $k = k'$  and  $n < n'$ . We calculate that

$$\begin{aligned} \mathcal{H}_{k'}(n') &= \mathcal{H}_{k'}(n' - 1) + \phi_{k'}(\mathcal{H}_{k'}(n' - 1)) && \text{by (15)} \\ &\leq \mathcal{H}_{k'}(n' - 1) + A_{k'-1}(\mathcal{H}_{k'}(n' - 1) + k') \\ &\quad - \mathcal{H}_{k'}(n' - 1) - k' && \text{by Lemma 3.2 (ii)} \\ &= A_{k'-1}(\mathcal{H}_{k'}(n' - 1) + k') - k' \\ &\leq A_{k'-1}(A_{k'}(n' + k' - 1)) - k' && \text{by (8)} \\ &= A_{k'}(n' + k') - k' && \text{by (4)}. \end{aligned}$$

Thus the inequality holds if  $(k, n) = (k', n')$ , completing the proof of (23).  $\square$

#### 4. A PRESENTATION THAT REVEALS $\tilde{G}_k$ TO BE CAT(0)

In this section we will establish that  $G_k$  admits the presentation which, as we explained in Section 1.2, has a CAT(0) squared complex as its Cayley 2-complex.

One checks that the automorphism  $\theta$  of Section 1.2 satisfies

$$(24) \quad \theta^{-1}(a_i) = \begin{cases} a_i a_{i-2} \dots a_1 a_2^{-1} a_4^{-1} \dots a_{i-1}^{-1} & \text{for } i \text{ odd,} \\ a_i a_{i-2} \dots a_2 a_1^{-1} a_3^{-1} \dots a_{i-1}^{-1} & \text{for } i \text{ even,} \end{cases}$$

and that, for each  $r \geq 0$ ,  $i \geq 2$ ,

$$(25) \quad \theta^{-r}(a_i) = a_i \theta^{-1}(a_{i-1})^{-1} \theta^{-2}(a_{i-1})^{-1} \dots \theta^{-r}(a_{i-1})^{-1}$$

in  $F(a_1, \dots, a_i)$ .

**Proposition 4.1.** *The group  $G_k$  can be presented as*

$$\langle \alpha_1, \dots, \alpha_k, s \mid \alpha_1^{-1} s \alpha_1 = s, \alpha_i^{-1} s \alpha_i = \alpha_{i-1} \ (i > 1) \rangle.$$

*Proof.* For each  $r = 1, \dots, k$ , let  $\mathcal{P}_k(r)$  be the presentation with generators

$$a_1, \dots, a_r, \alpha_{r+1}, \dots, \alpha_k, s$$

and relators

$$\begin{aligned} sa_i s^{-1} &= \theta(a_i) & (i = 1, \dots, r), \\ \alpha_{r+1}^{-1} s \alpha_{r+1} &= \theta^{r-k}(a_r) s, \\ \alpha_i^{-1} s \alpha_i &= \alpha_{i-1} & (i = r+2, \dots, k). \end{aligned}$$

We will show that the  $\mathcal{P}_k(r)$  all present  $G_k$ . Indeed,

$$\mathcal{P}_k(k) = \langle a_1, \dots, a_k, s \mid sa_i s^{-1} = \theta(a_i) \ (i \geq 1) \rangle$$

is equivalent to our original presentation (1) under the substitution  $s = t^{-1}$ . We will show that, for each  $r = 2, \dots, k$ , the presentation  $\mathcal{P}_k(r)$  can be converted to  $\mathcal{P}_k(r-1)$  by the substitution  $\alpha_r = \theta^{r-k}(a_r) s$ .

It follows from (25) that  $\theta^{r-k}(a_r) = a_r w_r$  where  $w_r = \theta^{-1}(a_{r-1})^{-1} \dots \theta^{-k}(a_{r-1})^{-1}$ . Since  $w_r$  is a word on  $a_1, \dots, a_{r-1}$ , we may eliminate  $a_r$  from  $\mathcal{P}_k(r)$  using the identity  $a_r = \alpha_r s^{-1} w_r^{-1}$ . Observe that  $a_r$  occurs in precisely two of the defining relators of  $\mathcal{P}_k(r)$ , namely  $\alpha_{r+1}^{-1} s \alpha_{r+1} = \theta^{r-k}(a_r) s$  and  $sa_r s^{-1} = a_r a_{r-1}$ . Substituting  $\alpha_r = \theta^{r-k}(a_r) s$  converts the first to  $\alpha_{r+1}^{-1} s \alpha_{r+1} = \alpha_r$ . Substituting  $a_r = \alpha_r s^{-1} w_r^{-1}$  converts the second to  $s \alpha_r s^{-1} w_r^{-1} s^{-1} = \alpha_r s^{-1} w_r^{-1} a_{r-1}$ , which can be rearranged to give  $\alpha_r^{-1} s \alpha_r = s^{-1} w_r^{-1} a_{r-1} s w_r s$ .

To complete the proof that  $\mathcal{P}_k(r)$  and  $\mathcal{P}_k(r-1)$  are equivalent, it is therefore enough to show that the relators

$$sa_i s^{-1} = \theta(a_i) \quad (i = 1, \dots, r-1)$$

suffice to convert  $s^{-1} w_r^{-1} a_{r-1} s w_r$  to  $\theta^{r-k-1}(a_{r-1})$ . Well, these relators define the group  $G_{r-1}$ , which can be presented as

$$\langle a_1, \dots, a_{r-1} \mid s^{-1} a_i s = \theta^{-1}(a_i) \ (\forall i \geq 1) \rangle.$$

And in  $G_{r-1}$  we calculate that

$$\begin{aligned} s^{-1} w_r^{-1} a_{r-1} s w_r &= \theta^{-1}(w_r^{-1} a_{r-1}) w_r \\ &= \theta^{-1}(\theta^{r-k}(a_{r-1}) \theta^{r-k+1}(a_{r-1}) \dots \theta^{-1}(a_{r-1}) a_{r-1}) w_r \\ &= \theta^{r-k-1}(a_{r-1}) \theta^{r-k}(a_{r-1}) \dots \theta^{-1}(a_{r-1}) w_r \\ &= \theta^{r-k-1}(a_{r-1}). \end{aligned}$$

We have thus shown that  $G_k$  can be presented by

$$\mathcal{P}_k(1) = \langle a_1, \alpha_2, \dots, \alpha_k, s \mid sa_1 s^{-1} = a_1, \alpha_2^{-1} s \alpha_2 = a_1 s, \alpha_i^{-1} s \alpha_i = \alpha_{i-1} \ (i > 2) \rangle.$$

The substitution  $\alpha_1 = a_1 s$  converts this into the required form.  $\square$

## 5. FREENESS OF THE SUBGROUPS $H$ AND $H_k$

In this section we prove:

**Proposition 5.1.** *The subgroup  $H_k$  of  $G_k$  is free with free basis  $a_1 t, \dots, a_k t$ , and the subgroup  $H$  of  $G$  is free with free basis  $a_1 t, a_2 t, \dots$*

To facilitate an induction argument, we will prove the following more elaborate proposition. Proposition 5.1 will follow because if  $w = w(a_1 t, \dots, a_k t)$  is freely reduced and represents 1 in  $G_k$  (or, equivalently, in  $G$ ), then  $w = \varepsilon$  by conclusion (i), and so  $a_1 t, \dots, a_k t$  are each not the identity and satisfy no non-trivial relations.

**Proposition 5.2.** *Let  $u = u(a_1 t, \dots, a_k t)$  be a freely reduced word with free-by-cyclic normal form  $v t^r$  — that is,  $u = v t^r$  in  $G_k$ ,  $v = v(a_1, \dots, a_k)$  is reduced, and  $r \in \mathbb{Z}$ .*

(i) *If  $v = \varepsilon$ , then  $u = \varepsilon$ .*

- (ii) If  $v = \theta(a_{k+1}^{-1})\theta^{1-r}(a_{k+1})$  in  $F(a_1, a_2, \dots)$ , then  $u = \varepsilon$ .
- (iii) If  $v$  is positive, then  $u$  is positive.

We emphasise that we are considering  $u$  as a word on the  $a_i t$  — it is freely reduced if and only if it contains no subword  $(a_i t)^{\pm 1}(a_i t)^{\mp 1}$ .

*Proof of Proposition 5.2.* We first show that for all fixed  $k \geq 1$ , if (iii) holds, then so do (i) and (ii).

For (i), note that if  $u = t^r$  in  $G$ , then  $u^{-1} = t^{-r}$ . Thus (iii) implies that both of the freely reduced words  $u$  and  $u^{-1}$  are positive. Hence  $u = \varepsilon$ .

For (ii), we will separately consider the cases  $r = 0$ ,  $r < 0$ , and  $r > 0$ . If  $r = 0$ , then  $u = 1$  in  $G$  and hence  $u = \varepsilon$  by (i). If  $r < 0$ , then  $1 - r \geq 1$  and so  $\theta^{1-r}(a_{k+1}) = a_{k+1}a_k w$  in  $F(a_1, a_2, \dots)$  for some positive word  $w = w(a_1, \dots, a_k)$ . It follows that  $v$  is positive and therefore (iii) implies that  $u$  is positive. Thus  $r \geq 0$ , giving a contradiction. If  $r > 0$ , one calculates that  $u^{-1} = t^{-r}\theta^{1-r}(a_{k+1}^{-1})\theta(a_{k+1}) = \theta(a_{k+1}^{-1})\theta^{1+r}(a_{k+1})t^{-r}$  in  $F(a_1, a_2, \dots)$ . Since  $1 + r \geq 1$ , the reduced form of  $\theta(a_{k+1}^{-1})\theta^{1+r}(a_{k+1})$  is positive, and so (iii) implies that  $u^{-1}$  is positive. Thus  $-r \geq 0$ , giving a contradiction.

We now prove (iii) by induction on  $k$ . Since  $G_1$  is free abelian with basis  $a_1, t$ , it is easy to check that (iii) holds in the case  $k = 1$ . As an inductive hypothesis, assume that assertions (i), (ii) and (iii) all hold for smaller values of  $k$ . If  $u$  contains no occurrence of an  $(a_k t)^{\pm 1}$ , then we are done. Otherwise, write  $u = \sigma_0(a_k t)^{\epsilon_1}\sigma_1(a_k t)^{\epsilon_2}\dots(a_k t)^{\epsilon_m}\sigma_m$ , where each  $\sigma_i = \sigma_i(a_1 t, \dots, a_{k-1} t)$  and each  $\epsilon_i \in \{\pm 1\}$ .

Each  $\sigma_i$  has free-by-cyclic normal form  $\tau_i t^{s_i}$  for some  $\tau_i = \tau_i(a_1, \dots, a_{k-1})$  and some  $s_i \in \mathbb{Z}$ . Direct calculation of the normal form of  $u$  — moving all the  $t^{\pm 1}$  to the right-hand end and applying the automorphism  $\theta^{\mp 1}$  whenever a  $t^{\pm 1}$  is moved past a letter  $a_i$  — gives that  $v$  freely equals

$$v' := \tau_0 \theta^{\lambda_1}(a_k^{\epsilon_1}) \theta^{\mu_1}(\tau_1) \theta^{\lambda_2}(a_k^{\epsilon_2}) \dots \theta^{\lambda_m}(a_k^{\epsilon_m}) \theta^{\mu_m}(\tau_m),$$

where

$$\lambda_i = \begin{cases} -(s_0 + \dots + s_{i-1} + \epsilon_1 + \dots + \epsilon_{i-1}) & \text{if } \epsilon_i = 1, \\ -(s_0 + \dots + s_{i-1} + \epsilon_1 + \dots + \epsilon_i) & \text{if } \epsilon_i = -1, \end{cases}$$

$$\mu_i = -(s_0 + \dots + s_{i-1} + \epsilon_1 + \dots + \epsilon_i).$$

We claim that  $\epsilon_i = 1$  for all  $i$ . For a contradiction, suppose otherwise. Observe that, for each  $s \in \mathbb{Z}$ , there are words  $w_s = w_s(a_1, \dots, a_{k-1})$  and  $w'_s = w'_s(a_1, \dots, a_{k-1})$  such that  $\theta^s(a_k) = a_k w_s$  and  $\theta^s(a_k^{-1}) = w'_s a_k^{-1}$ . Since  $v$  is positive, there must be a subword  $a_k^{\pm 1} \chi a_k^{\mp 1}$  in  $v'$  which freely equals the empty word and in which  $\chi = \chi(a_1, \dots, a_{k-1})$ . The way this subword must arise is that for some  $i$ , either

- (a)  $\epsilon_i = -1$ ,  $\epsilon_{i+1} = 1$  and  $\theta^{\mu_i}(\tau_i) = 1$ , or
- (b)  $\epsilon_i = 1$ ,  $\epsilon_{i+1} = -1$  and  $\theta^{\lambda_i}(a_k) \theta^{\mu_i}(\tau_i) \theta^{\lambda_{i+1}}(a_k^{-1}) = 1$ .

In the first case  $\tau_i = 1$  and hence the induction hypothesis (assertion (i)) gives that  $\sigma_i = \varepsilon$ . But this contradicts the supposition that  $u$  is freely reduced. In the second case, one calculates that  $\lambda_i - \mu_i = 1$  and  $\lambda_{i+1} - \mu_i = 1 - s_i$ , and so  $\tau_i = \theta(a_k^{-1})\theta^{1-s_i}(a_k)$ . The induction hypothesis (assertion (ii)) implies that  $\sigma_i = \varepsilon$ , but again this contradicts the supposition that  $u$  is freely reduced.

To complete our proof of (iii), we will show that all the  $\sigma_i$  are positive. Since  $v$  is positive and each  $\epsilon_i = 1$ , we have that  $\tau_0$  is positive and each  $\theta^{\lambda_i}(a_k) \theta^{\mu_i}(\tau_i)$  is positive. The inductive

hypothesis (assertion (iii)) immediately gives that  $\sigma_0$  is positive. Suppose we have shown that  $\sigma_0, \dots, \sigma_{j-1}$  are positive, for some  $j$ . It follows that  $s_0, \dots, s_{j-1} \geq 0$ , whence  $\lambda_j \leq 0$ . Note that if  $w = w(a_1, \dots, a_k)$  is positive and  $s \geq 0$ , then  $\theta^s(w)$  is positive. Hence  $a_k \theta^{\mu_j - \lambda_j}(\tau_j) = a_k \theta^{-1}(\tau_j)$  is positive. Since  $\theta^{-1}(\tau_j)$  is a word on  $a_1^{\pm 1}, \dots, a_{k-1}^{\pm 1}$ , it follows that  $\theta^{-1}(\tau_j)$  is positive, and hence that  $\tau_j$  is positive. Applying the induction hypothesis (assertion (iii)) gives that  $\sigma_j$  is positive.  $\square$

## 6. A LOWER BOUND ON THE DISTORTION OF $H_k$ IN $G_k$

In the following lemma we see the battle between Hercules and the hydra manifest in  $G_k$ .

**Lemma 6.1.** *For all  $k, n \geq 1$ , there is a positive word  $u_{k,n} = u_{k,n}(a_1 t, \dots, a_k t)$  of length  $\mathcal{H}_k(n)$  that equals  $a_k^n t^{\mathcal{H}_k(n)}$  in  $G_k$ .*

*Proof.* Consider the following calculation in which successive  $t$  are moved to the front and paired off with the  $a_i$ . [We illustrate the calculation in the case  $k \geq 3$  and  $n \geq 2$  — for  $k = 2$ , the letters  $a_{k-2}$  would not appear and for  $k = 1$ , neither would the  $a_{k-1}$ .]

$$\begin{aligned} a_k^n t^{\mathcal{H}_k(n)} &= (a_k t) t^{-1} a_k^{n-1} t t^{\mathcal{H}_k(n)-1} \\ &= (a_k t) (a_k a_{k-1})^{n-1} t^{\mathcal{H}_k(n)-1} \\ &= (a_k t) (a_k t) t^{-1} a_{k-1} (a_k a_{k-1})^{n-2} t t^{\mathcal{H}_k(n)-2} \\ &= (a_k t) (a_k t) a_{k-1} a_{k-2} (a_k a_{k-1} a_{k-1} a_{k-2})^{n-2} t^{\mathcal{H}_k(n)-2} \\ &\vdots \end{aligned}$$

A van Kampen diagram displaying this calculation in the case  $k = 2$  and  $n = 4$  is shown in Figure 1.

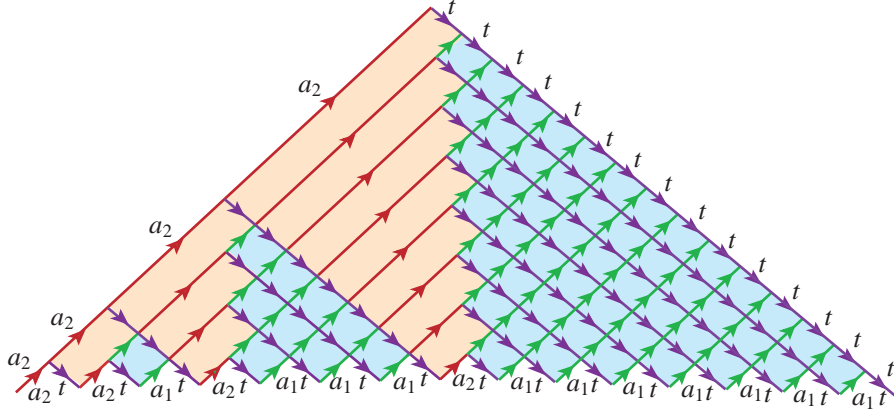


FIGURE 1. A van Kampen diagram showing that  $a_2^4 t^{15} = u_{2,4}$  in  $G_2$  where  $u_{2,4} = a_2 t a_2 t a_1 t a_2 t (a_1 t)^3 a_2 t (a_1 t)^7$ .

One sees the Hercules-versus-the-hydra battle

$$a_k^n \rightarrow (a_k a_{k-1})^{n-1} \rightarrow a_{k-1} a_{k-2} (a_k a_{k-1} a_{k-1} a_{k-2})^{n-2} \rightarrow \dots$$

being played out in this calculation. The pairing off of a  $t$  with an  $a_i$  corresponds to a decapitation, and the conjugation by  $t$  that moves that  $t$  into place from the right-hand end causes a hydra-regeneration for the intervening subword. So by Proposition 1.1, after  $\mathcal{H}_k(n)$  steps we have a positive word on  $u_{k,n} = u_{k,n}(a_1 t, \dots, a_k t)$ , and its length is  $\mathcal{H}_k(n)$ .  $\square$

Our next proposition establishes that  $\text{Dist}_{H_k}^{G_k} \geq \mathcal{H}_k$  for all  $k \geq 2$ . The case  $k = 1$  is straightforward:  $H_1 \cong \mathbb{Z}$  is undistorted in  $G_1 \cong \mathbb{Z}^2$  and  $\mathcal{H}_1(n) = n$ . The calculation in the proof of the proposition is illustrated by a van Kampen diagram in Figure 2 in the case  $k = 2$  and  $n = 4$  — the idea is that a copy of the diagram from Figure 1 fits together with its mirror image along intervening  $a_1$ - and  $a_2$ -corridors to make a diagram demonstrating the equality of a freely reduced word of extreme length on  $a_1 t, \dots, a_k t$  with a short word on  $a_1, \dots, a_k, t$ .

**Proposition 6.2.** *For all  $k \geq 2$  and  $n \geq 1$ , there is a reduced word of length  $2\mathcal{H}_k(n) + 3$  on the free basis  $a_1 t, \dots, a_k t$  for  $H_k$  which, in  $G_k$ , equals a word of length  $2n + 4$  on  $a_1, \dots, a_k, t$ .*

*Proof.* The relation  $t^{-1}a_2t = a_2a_1$  can be expressed as  $a_2^{-1}ta_2 = ta_1^{-1}$ , from which one obtains  $a_2^{-1}t^{\mathcal{H}_k(n)}a_2 = (ta_1^{-1})^{\mathcal{H}_k(n)}$ . Combined with Lemma 6.1 this gives

$$(ta_1^{-1})^{\mathcal{H}_k(n)} = a_2^{-1}a_k^{-n}u_{k,n}a_2.$$

So, as  $a_1$  commutes with  $t$ , we deduce that  $ta_1$  commutes with  $a_2^{-1}a_k^{-n}u_{k,n}a_2$ , and therefore

$$a_k^n a_2 ta_1 a_2^{-1} a_k^{-n} = u_{k,n} a_2 ta_1 a_2^{-1} u_{k,n}^{-1}$$

in  $G_k$ . The word on the left-hand side of this equation has length  $2n + 4$ . The word on the right-hand side freely equals  $u_{k,n}(a_2t)(a_1t)(a_2t)^{-1}u_{k,n}^{-1}$ , which, viewed as a word on  $a_1t, \dots, a_k t$ , is freely reduced and has length  $2\mathcal{H}_k(n) + 3$ , since  $u_{k,n}$  is a positive word.  $\square$

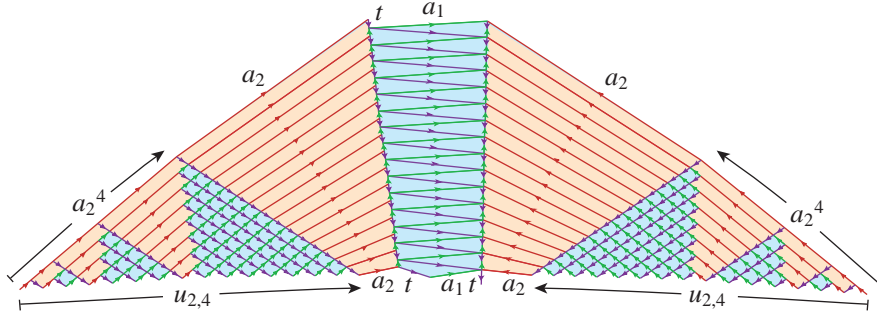


FIGURE 2. A van Kampen diagram demonstrating the equality  $a_2^4 a_2 t a_1 a_2^{-1} a_2^{-4} = u_{2,4} (a_2 t) (a_1 t) (a_2 t)^{-1} u_{2,4}^{-1}$  in  $G_2$ .

## 7. RECURSIVE STRUCTURE OF WORDS

This section contains preliminaries that will feed into the proof, presented in Section 9, that  $\text{Dist}_{H_k}^{G_k} \leq A_k$ . In outline, we will bound the distortion of  $H_k$  in  $G_k$  by first supposing  $u = u(t, a_1, \dots, u_k)$  represents an element of  $H_k$ . We will shuffle all the  $t^{\pm 1}$  in  $u$  to the start, with the effect of applying  $\theta^{\pm 1}$  to each  $a_i^{\pm 1}$  they pass. After freely reducing, we will have a word  $t^r w$  where  $w = w(a_1, a_2, \dots)$ . We will then look to carry the  $t^r$  back through  $w$  from left to right, converting all it passes to a word on  $a_1 t, a_2 t, \dots$  which will represent an element of  $G_k$  since the indices of the  $a_i$  will be no higher than those occurring in  $w$ . Estimating the length of this word will give an upper bound on  $\text{Dist}_{H_k}^{G_k}$ .

For convenience, we work with the group  $G$  and its subgroup  $H$  defined in Section 1.2.

When carrying the power of  $t$  through  $w$  we will face the problem of whether a word  $t^r w$ , where  $w = w(a_1, a_2, \dots)$ , represents an element of a coset  $Ht^s$  in  $G$  for some  $s \in \mathbb{Z}$ . We will see that the answer is not always affirmative — these cosets do not cover  $G$ . However,

if  $t^r w = \sigma t^s$  for some  $\sigma = \sigma(a_1 t, a_2 t, \dots)$  and some  $s \in \mathbb{Z}$ , then  $\sigma$  is unique up to free-equivalence since  $H$  is free (Proposition 5.1) and  $s$  is unique by our next lemma. Indeed, this implies that  $Ht^{s_1}$  and  $Ht^{s_2}$  are equal precisely when  $s_1 = s_2$ .

**Lemma 7.1.** *If  $\ell \in \mathbb{Z}$  and  $t^\ell \in H$ , then  $\ell = 0$ .*

*Proof.* Were  $t^\ell \in H$  for some integer  $\ell \neq 0$ , then  $\mathbb{Z}^2 \cong \langle a_1 t, t^\ell \rangle$  would be a subgroup of  $H$  contra to the freeness of  $H$  established in Proposition 5.1.  $\square$

Our next lemma will be the crux of our method for establishing an upper bound on distortion. It identifies recursive structure that will allow us to analyse the process of passing a power of  $t$  through a word  $w = w(a_1, a_2, \dots)$ , so as to leave behind a word on  $a_1 t, a_2 t, \dots$ .

For a non-empty freely-reduced word  $w = w(a_1, a_2, \dots)$ , define the *rank of  $w$*  to be the highest  $k$  such that  $a_k^{\pm 1}$  occurs in  $w$ . We define the empty word to have rank 0. For an integer  $k \geq 1$ , define a *piece of rank  $k$*  to be a freely-reduced word  $a_k^{\epsilon_1} \pi a_k^{-\epsilon_2}$  where  $\pi = \pi(a_1, \dots, a_{k-1})$  and  $\epsilon_1, \epsilon_2 \in \{0, 1\}$ . Notice that a piece of rank  $k$  will always also be a piece of rank  $k + 1$  and that the empty word is a piece of rank  $k$  for every  $k$ .

For a non-empty freely-reduced word  $w$  of rank  $k$ , define the *number of pieces in  $w$*  to be the least integer  $m$  such that  $w$  can be expressed as a concatenation  $w_1 \dots w_m$  of subwords  $w_i$  each of which is a piece of rank  $k$ . (We say the empty word is composed of 0 pieces.) Observe that

- (i) each  $a_k$  and  $a_k^{-1}$  in  $w$  is the first or last letter of some  $w_i$ , respectively;
- (ii) for  $i = 1, \dots, m - 1$ , either the final letter of  $w_i$  is  $a_k^{-1}$  or the first of  $w_{i+1}$  is  $a_k$ , but never both; and
- (iii) if  $a_k^{-1} \chi a_k$  is a subword of  $w$  and  $\chi = \chi(a_1, \dots, a_{k-1})$ , then  $\chi = w_i$  for some  $i$ .

In particular,  $w_1, \dots, w_m$  are uniquely determined by the locations of the  $a_k^{\pm 1}$  in  $w$ , and so we call the list of subwords  $w_1, \dots, w_m$  the *partition of  $w$  into pieces*.

For example,  $w := a_3^{-1} a_1 a_2 a_3 a_2^{-1} a_3 a_1^{-1} a_3^{-1}$  has rank 3 and its partition into pieces is  $w = w_1 w_2 w_3 w_4$  where  $w_1 = a_3^{-1}$ ,  $w_2 = a_1 a_2$ ,  $w_3 = a_3 a_2^{-1}$ , and  $w_4 = a_3 a_1^{-1} a_3^{-1}$ .

**Lemma 7.2.** *Suppose  $w = w(a_1, \dots, a_k)$  is a non-empty freely-reduced word of rank  $k$  and  $r$  and  $s$  are integers such that  $t^r w \in Ht^s$ . Let  $w = w_1 \dots w_n$  be the partition of  $w$  into pieces. Then there exist integers  $r = r_0, r_1, \dots, r_n = s$  such that  $t^{r_i} w_{i+1} \in Ht^{r_{i+1}}$  for each  $i$ .*

*Proof.* As  $t^r w \in Ht^s$ , there is some reduced word  $v = v(a_1 t, \dots, a_k t)$  such that  $t^r w = vt^s$ . Form the analogue of a *partition into pieces* for  $v$  — that is, express  $v$  as a concatenation  $v_1 \dots v_m$  of subwords  $v_i$  each of the form  $(a_k t)^{\epsilon_1} \tau (a_k t)^{-\epsilon_2}$  where  $\tau = \tau(a_1 t, \dots, a_{k-1} t)$  and  $\epsilon_1, \epsilon_2 \in \{0, 1\}$  and  $m$  is minimal.

Note that  $v$  is non-empty as otherwise  $w$  would equal  $t^{s-r}$  in  $G$  and so be the empty word by the free-by-cyclic structure of  $G$ . Note also that no  $v_i$  is the empty word since  $m$  is minimal.

One can obtain  $t^r w$  from  $vt^s$  by carrying all the  $t^{\pm 1}$  to the left and freely reducing. More particularly, the  $t^s$  at the end of  $vt^s$  and all the  $t^{\pm 1}$  in  $v_m$  can be collected immediately to the left of  $v_m$ , and then those  $t^{\pm 1}$  and the  $t^{\pm 1}$  in  $v_{m-1}$  can be carried to the left of  $v_{m-1}$ , and so on. Accordingly, inductively define  $w'_m, \dots, w'_1$  and  $r_m, \dots, r_0$  by setting  $r_m := s$  and then for  $i = m, \dots, 1$  taking  $r_{i-1}$  and  $w'_i = w'_i(a_1, \dots, a_k)$  to be the unique integer and reduced word such that  $v_i t^{r_i} = t^{r_{i-1}} w'_i$ . Then  $r_0 = r$  and  $w$  is (a priori) the freely reduced form of  $w'_1 \dots w'_m$ . We claim that, in fact,  $w'_1 \dots w'_m$  is the partition of  $w$  into pieces of rank  $k$  — that is,  $m = n$  and  $w'_i = w_i$  for all  $i$ . This will suffice to establish the lemma.

To prove this claim, we will show that for all  $i$ , if  $v_i = (a_k t)^{\epsilon_1} \tau (a_k t)^{-\epsilon_2}$  where  $\tau = \tau(a_1 t, \dots, a_{k-1} t)$  and  $\epsilon_1, \epsilon_2 \in \{0, 1\}$ , then  $w'_i$  is a reduced word  $a_k^{\epsilon_1} \pi a_k^{-\epsilon_2}$  for some  $\pi = \pi(a_1, \dots, a_{k-1})$ . Moreover, if  $\epsilon_1 = \epsilon_2 = 0$ , then  $\pi$  is not the empty word. In particular, no  $w'_i$  is the empty word.

Well,  $v_i t^{r_i} = t^{r_i-1} w'_i$ . Consider the process of carrying each  $t^{\pm 1}$  in  $v_i t^{r_i}$  to the front of the word, applying  $\theta^{\pm 1}$  to each  $a_j$  they pass and then freely reducing, to give  $t^{r_i-1} w'_i$ . Throughout this process, no new  $a_k^{\pm 1}$  are produced and, such is  $\theta$ , no  $a_l$  appear to the left of the  $a_k$  in  $v_i$  (if present) or the the right of the  $a_k^{-1}$  (if present) — see (2) and (24). This means that the only way  $w'_i$  could fail to be a reduced word of the form  $a_k^{\epsilon_1} \pi a_k^{-\epsilon_2}$  where  $\pi = \pi(a_1, \dots, a_{k-1})$ , would be for  $\epsilon_1$  and  $\epsilon_2$  to both be 1 and  $\pi$  be the empty word. But in that case,  $w'_i$  would be the empty word and so  $v_i$  would equal  $t^{r_i-1-r_i}$  in  $G_k$  and  $r_{i-1} - r_i$  would be 0 by Lemma 7.1. But then  $v_i$  would be the empty word by Proposition 5.1 which, as we observed, is not the case. Likewise, when  $\epsilon_1 = \epsilon_2 = 0$ , it cannot be the case that  $\pi = w'_i$  is the empty word, as otherwise  $v_i$  would again be the empty word.

So properties (i), (ii) and (iii) all apply to  $w'_1, \dots, w'_m$  as they are inherited the corresponding properties for  $v_1, \dots, v_m$ . It follows from these properties together with the fact that each  $w'_i$  is reduced, that  $w'_1 \dots w'_m$  is reduced and is the partition of  $w$  into pieces of rank  $k$ .  $\square$

## 8. PASSING POWERS OF $t$ THROUGH $\theta^n(a_k^{\pm 1})$ .

The words  $\theta^n(a_k^{\pm 1})$  will play a crucial role in our proof that  $\text{Dist}_{H_k}^{G_k} \leq A_k$ . The next lemma reveals their recursive structure. The first part is proved by an induction on  $n$ . The second part is then an immediate consequence.

**Lemma 8.1.**

$$\theta^n(a_k) = \begin{cases} a_k \theta^0(a_{k-1}) \theta^1(a_{k-1}) \dots \theta^{n-1}(a_{k-1}) & n > 0 \\ a_k & n = 0 \\ a_k \theta^{-1}(a_{k-1}^{-1}) \theta^{-2}(a_{k-1}^{-1}) \dots \theta^n(a_{k-1}^{-1}) & n < 0, \end{cases}$$

$$\theta^n(a_k^{-1}) = \begin{cases} \theta^{n-1}(a_{k-1}^{-1}) \theta^{n-2}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1}) a_k^{-1} & n > 0 \\ a_k^{-1} & n = 0 \\ \theta^n(a_{k-1}) \theta^{n+1}(a_{k-1}) \dots \theta^{-1}(a_{k-1}) a_k^{-1} & n < 0. \end{cases}$$

When attempting to carry a power of  $t$  through a word  $w = w(a_1, a_2, \dots)$ , we will frequently be faced with the special case where  $w$  is of the form  $\theta^n(a_k^{\pm 1})$ . We now focus on this situation.

**Definition 8.2.** Define

$$\Lambda = \bigcup_{i \in \mathbb{Z}} H t^i.$$

For each integer  $k \geq 1$ , define

$$S_k = \{n \in \mathbb{Z} : \theta^n(a_k) \in \Lambda\}$$

and define the function  $\phi_k : S_k \rightarrow \mathbb{Z}$  by setting  $\phi_k(n)$  to be the unique integer satisfying

$$\theta^n(a_k) t^{\phi_k(n)} \in H.$$

Note that this extends the previous definition of the functions  $\phi_k$  given in Section 3 since  $\phi_k(n) = \mathcal{H}(\theta^n(a_k))$  for  $n \geq 0$ .

**Lemma 8.3.**

- (i)  $S_1 = \mathbb{Z}$  and  $\phi_1(n) = 1$  for all  $n \in S_1$ .
- (ii)  $S_2 = \mathbb{Z}$  and  $\phi_2(n) = n + 1$  for all  $n \in S_2$ .
- (iii) If  $k \geq 3$ , then  $S_k = \mathbb{N}$ .

*Proof.* It is easy to check that  $S_1 = S_2 = \mathbb{Z}$ ,  $\phi_1(n) = 1$ ,  $\phi_2(n) = n + 1$  and that  $\mathbb{N} \subseteq S_k$  for all  $k$ .

Let  $k \geq 3$  and suppose that  $n < 0$  lies in  $S_k$ . Since  $\theta^n(a_k)t^{\phi_k(n)}$  lies in  $H$ , so does  $(a_k t)^{-1}\theta^n(a_k)t^{\phi_k(n)} = a_{k-1}^{-1}\theta^{-1}(a_{k-1}^{-1}) \dots \theta^{n+1}(a_{k-1}^{-1})t^{\phi_k(n)-1}$ , and hence, by Lemma 7.2,  $a_{k-1}^{-1}$  lies in  $Ht^r$  for some  $r$ . It follows that  $\theta^{-r}(a_{k-1})t^r \in H$  and so  $r = \phi_{k-1}(-r)$ . If  $k = 3$ , this is contradiction, since it implies  $r = -r + 1$ . If  $k > 3$ , then  $-r \in S_{k-1}$ , and so, by the induction hypothesis,  $r \leq 0$ . But then  $\phi_{k-1}(-r) \geq 1$ , by (18), and hence  $r \geq 1$ , a contradiction.  $\square$

Let  $d_H$  denote the word metric on  $H$  with respect to the generating set  $a_1 t, a_2 t, \dots$

**Lemma 8.4.** *If  $n \in S_k$  and  $h = \theta^n(a_k)t^{\phi_k(n)}$ , then  $d_H(1, h) = \phi_k(|n|)$ .*

*Proof.* If  $k = 1$ , then the result is obvious. If  $k = 2$ , then  $h = a_2 a_1^n t^{n+1} = (a_2 t)(a_1 t)^n$  so  $d_H(1, h) = 1 + |n| = \phi_k(|n|)$ . If  $k \geq 3$ , then  $n \geq 0$ . Thus the word  $\theta^n(a_k)$  is positive and hence  $d_H(1, h) = \phi_k(n) = \phi_k(|n|)$ .  $\square$

**Lemma 8.5.**

- (i) *Let  $h = t^r \theta^i(a_k)t^{-s}$ . Then  $h \in H$  if and only if  $i - r \in S_k$  and  $s = r - \phi_k(i - r)$ .*
- (ii) *Let  $h = t^r \theta^i(a_k^{-1})t^{-s}$ . Then  $h \in H$  if and only if  $i - s \in S_k$  and  $r = s - \phi_k(i - s)$ .*

*Proof.* For (i), note that  $h = \theta^{i-r}(a_k)t^{-s}$  and apply Definition 8.2. For (ii), note that  $h^{-1} = t^s \theta^i(a_k)t^{-r}$  and apply (i).  $\square$

**Lemma 8.6.** *If  $k \geq 3$  and  $t^r \theta^i(a_k^{-1}) \in \Lambda$ , then  $r < i$ .*

*Proof.* If  $t^r \theta^i(a_k^{-1}) \in Ht^s$ , then Lemmas 8.3 and 8.5 give that  $i - s \geq 0$  and  $s - r = \phi_k(i - s) \geq 1$ . Thus  $i - r \geq 1$ .  $\square$

The exceptional nature of  $S_1$  and  $S_2$  highlighted by Lemma 8.3 means that small values of  $k$  will have to be treated separately in our proof. This motivates the inclusion of the following result, a special case of Lemma 8.5. Note in particular that (ii) implies that  $t^r \theta^i(a_2^{-1}) \in \Lambda$  if and only if  $r + i$  is odd.

**Lemma 8.7.**

- (i) *Let  $h = t^r \theta^i(a_2)t^{-s}$ . Then  $h \in H$  if and only if  $s = 2r - i - 1$ .*
- (ii) *Let  $h = t^r \theta^i(a_2^{-1})t^{-s}$ . Then  $h \in H$  if and only if  $s = \frac{1}{2}(r + i + 1)$ .*

*Proof.* This follows immediately from Lemma 8.5 and the fact, given in Lemma 8.3, that  $\phi_2(n) = n + 1$ .  $\square$

The following result concerns passing a power of  $t$  through a sequence of terms of the form  $\theta^i(a_2^{\pm 1})$ . The statement is made neater by the use of the following formula, which is a consequence of Lemma 8.1 :

$$\theta^a(a_3^{-1})\theta^b(a_3) = \begin{cases} \theta^a(a_2) \dots \theta^{b-1}(a_2) & a < b, \\ 1 & a = b, \\ \theta^{a-1}(a_2^{-1}) \dots \theta^b(a_2^{-1}) & a > b. \end{cases}$$

**Lemma 8.8.** *Let  $\sigma = t^r \theta^a(a_3^{-1}) \theta^b(a_3)$  and  $s = 2^{b-a}(r - a - 2) + b + 2$  for some integers  $r, a, b$ . Then  $\sigma \in \Lambda$  if and only if  $s$  is an integer. Furthermore, in this case,  $\sigma \in Ht^s$ .*

*Proof.* We split the proof into two claims. The first claim is that if  $\sigma \in Ht^{s'}$  for some integer  $s'$ , then  $s = s'$ . In particular, this implies that if  $\sigma \in \Lambda$ , then  $s$  is an integer. If  $a = b$ , then clearly  $s' = r = s$ . If  $a < b$ , then  $\theta^a(a_3^{-1}) \theta^b(a_3) = \theta^a(a_2) \dots \theta^{b-1}(a_2)$ . By the Lemma 7.2, there exist integers  $r = r_0, r_1, \dots, r_{b-a} = s'$  such that  $t^{r_i} \theta^{a+i}(a_2) \in Ht^{r_{i+1}}$ . By Lemma 8.7,  $r_{i+1} = 2r_i - a - i - 1$ , which solves to give  $r_i = 2^i(r - a - 2) + i + a + 2$ . Substituting  $i = b - a$  gives  $s' = s$ . On the other hand, suppose that  $a > b$ . Note that  $t^r \theta^a(a_3^{-1}) \theta^b(a_3) \in Ht^{s'}$  implies that  $t^{s'} \theta^b(a_3^{-1}) \theta^a(a_3) \in Ht^r$ . Since  $b < a$ , we can substitute into the above solution to obtain  $r = 2^{a-b}(s' - b - 2) + a + 2$ , which rearranges to give  $s' = s$ . This completes the proof of our first claim.

The second claim is that if  $s$  is an integer, then  $\sigma \in Ht^s$ . If  $a = b$ , then this clearly holds. Suppose that  $a < b$ . Then  $\sigma = t^r \theta^a(a_2) \dots \theta^{b-1}(a_2)$ , so certainly  $\sigma \in \Lambda$  since all the letters  $a_2^{\pm 1}$  that appear are positive. Therefore  $\sigma \in Ht^s$  by the first claim. Now suppose that  $a > b$ . Since  $s$  is an integer, we can define  $\tau = t^s \theta^b(a_3^{-1}) \theta^a(a_3) = t^s \theta^b(a_2) \dots \theta^{a-1}(a_2)$ . Then certainly  $\tau \in \Lambda$  — say  $\tau \in Ht^{r'}$ . By the first claim,  $r' = 2^{a-b}(s - b - 2) + a + 2 = r$ . Therefore  $t^s \theta^b(a_3^{-1}) \theta^a(a_3) \in Ht^r$ , whence  $t^r \theta^a(a_3^{-1}) \theta^b(a_3) \in Ht^s$ , and the second claim is proved.  $\square$

## 9. AN UPPER BOUND ON THE DISTORTION OF $H_k$ IN $G_k$

Next we turn to estimates associated with pushing a power of  $t$  from left to right through a word  $w = w(a_1, \dots, a_k)$  or through a piece of  $w$ , so as to leave a word on  $a_1 t, \dots, a_k t$  times a power of  $t$ . We will need to keep track of both the length of that word on the  $a_1 t, \dots, a_k t$  and the power of  $t$  that emerges to its right. Accordingly, let us define four families of functions,  $\psi_{k,l}(n), \Psi_{k,l,p}(n), \kappa_{k,l}(n), K_{k,l,p}(n)$  for integers  $k \geq 1$  and  $l, p, n \geq 0$ .

- $\psi_{k,l}(n)$  is the least integer  $N$  such that if  $h \in H$  is represented by a word  $t^r \pi t^{-s}$  with  $\pi$  a piece of rank  $k$ , with  $\ell(\pi) \leq l$ , and with  $|r| \leq n$ , then  $d_H(1, h) \leq N$ .
- $\Psi_{k,l,p}(n)$  is the least integer  $N$  such that if  $h \in H$  is represented by a word  $t^r w t^{-s}$  with  $w = w(a_1, \dots, a_k)$  a word of at most  $p$  pieces, with  $\ell(w) \leq l$ , and with  $|r| \leq n$ , then  $d_H(1, h) \leq N$ .
- $\kappa_{k,l}(n)$  is the least integer  $N$  such that if  $\pi$  is a piece of rank  $k$  with  $\ell(\pi) \leq l$  and  $r$  is an integer with  $|r| \leq n$  and  $t^r \pi \in \Lambda$ , then  $t^r \pi \in Ht^s$  for some  $s$  with  $|s| \leq N$ .
- $K_{k,l,p}(n)$  is the least integer  $N$  such that if  $w$  is a word of rank at most  $k$  with at most  $p$  pieces and with  $\ell(w) \leq l$  and  $r$  is an integer with  $|r| \leq n$  and  $t^r w \in \Lambda$ , then  $t^r w \in Ht^s$  for some  $s$  with  $|s| \leq N$ .

We will frequently make use, without further comment, of the fact that each of these functions is increasing in  $k, l, p$  and  $n$ .

The main technical result of this section is the following proposition. In the corollary that follows it we explain how the upper bound it gives on  $\Psi_{k,l,p}(n)$  leads to our desired bound  $\text{Dist}_{H_k}^{G_k} \leq Ak$ .

**Proposition 9.1.** *For all  $k \geq 1$ , there exist integers  $C_k \geq 1$  such that for all  $l, p, n \geq 0$ ,*

$$\begin{aligned}\kappa_{k,l}(n) &\leq A_{k-1}(C_k n + C_k l), \\ K_{k,l,p}(n) &\leq A_{k-1}^{(p)}(C_k n + C_k l), \\ \psi_{k,l}(n) &\leq A_{k-1}(C_k n + C_k l), \\ \Psi_{k,l,p}(n) &\leq A_{k-1}^{(3p)}(C_k n + C_k l).\end{aligned}$$

**Corollary 9.2.** *For all  $k \geq 1$ , the distortion function of  $H_k$  in  $G_k$  satisfies*

$$\text{Dist}_{H_k}^{G_k} \leq A_k.$$

*Proof of Corollary 9.2.* Since  $G_1 \cong \mathbb{Z}^2$  and  $H_1 \cong \mathbb{Z}$ ,  $H_1$  is undistorted in  $G_1$  and  $\text{Dist}_{H_1}^{G_1} \leq A_1$ . Now suppose that  $k \geq 2$  and that  $u = u(a_1, \dots, a_k, t)$  is a word of length at most  $n$  representing an element of  $H$ . By carrying each  $t^{\pm 1}$  to the front, we see that  $u$  is equal in  $G_k$  to  $t^r w$  for some integer  $r$  and some freely reduced word  $w = w(a_1, \dots, a_k)$ . These satisfy  $|r| \leq n$  and  $\ell(w) \leq Cn^k$  for some integer  $C > 0$  depending only on  $k$  — see, for example, Section 3.3 of [8].

We first show that the number of pieces of  $w$  is at most  $n + 1$ . Indeed, the process of carrying each  $t^{\pm 1}$  to the front of  $u$  has the effect of applying  $\theta^{\pm 1}$  to each  $a_i$  it passes. The form of the automorphism  $\theta$  ensures that no new  $a_k^{\pm 1}$  are created by this process. The number of occurrences of  $a_k^{\pm 1}$  in  $w$ , which we denote by  $\ell_k(w)$ , is therefore at most  $n$ . Let  $w = w_1 \dots w_p$  be the partition of  $w$  into pieces. Say  $w_i = a_k^{\epsilon_i} \pi_i a_k^{-\epsilon_i}$  where  $\epsilon_i^-, \epsilon_i^+ \in \{0, 1\}$  and  $\pi_i = \pi_i(a_1, \dots, a_{k-1})$ . Observe that, for each  $i$ , precisely one of  $\epsilon_i^+$  and  $\epsilon_{i+1}^-$  is equal to 1. Indeed, if  $\epsilon_i^+ = \epsilon_{i+1}^- = 0$ , then the pieces  $w_i$  and  $w_{i+1}$  could be concatenated to form a single piece, contradicting the minimality of  $p$ , and if  $\epsilon_i^+ = \epsilon_{i+1}^- = 1$ , then  $w$  would not be freely reduced. So

$$\ell_k(w) = \sum_{i=1}^p (\epsilon_i^- + \epsilon_i^+) = \epsilon_1^- + \sum_{i=1}^{p-1} (\epsilon_i^+ + \epsilon_{i+1}^-) + \epsilon_n^+ = \epsilon_1^- + p - 1 + \epsilon_n^+,$$

whence  $p \leq \ell_k(w) + 1 \leq n + 1$ .

Now,

$$d_H(1, u) = d_H(1, t^r w) \leq \Psi_{k, \ell(w), p}(|r|) \leq \Psi_{k, Cn^k, n+1}(n),$$

which is at most

$$A_{k-1}^{(3n+3)}(C_k C n^k + C_k n)$$

by Proposition 9.1. Choose an integer  $N$  large enough that  $n^k \leq 2^n$  for  $n \geq N$ . Then, for  $n \geq \max\{N, 1\}$ ,

$$\begin{aligned}d_H(1, u) &\leq A_{k-1}^{(3n+3)}(C_k C A_2(n) + C_k n) && \text{by (8)} \\ &\leq A_{k-1}^{(3n+3)}(C_k C A_k(n) + C_k n) && \text{by (7), (8)} \\ &\leq A_{k-1}^{(3n+3)}(A_k(C_k C n) + C_k n) && \text{by (8), (10)} \\ &\leq A_{k-1}^{(3n+3)}(A_k((C_k C + C_k)n)) && \text{by (8), (13)} \\ &= A_k((C_k C + C_k + 3)n + 3) && \text{by (4)}.\end{aligned}$$

□

Proposition 9.1 will follow from the relationships between  $\psi_{k,l}(n)$ ,  $\Psi_{k,l,p}(n)$ ,  $\kappa_{k,l}(n)$  and  $K_{k,l,p}(n)$  set out in the next proposition. Of its claims, (28) and (31) are the most challenging to establish; we postpone their proof to Proposition 9.4, which itself will draw on Lemmas 9.5, 9.6 and 9.7.

**Proposition 9.3.** For integers  $k \geq 1$  and  $l, p, n \geq 0$ ,

$$(26) \quad \kappa_{1,l}(n) \leq n + 1,$$

$$(27) \quad K_{k,l,p}(n) \leq \max_{\substack{q \leq p \\ l_1 + \dots + l_q \leq l}} \left\{ \kappa_{k,l_1}(\dots \kappa_{k,l_{q-1}}(\kappa_{k,l_q}(n)) \dots) \right\},$$

$$(28) \quad \kappa_{k+1,l}(n) \leq 2K_{k,l,l}(2\phi_{k+1}(n)),$$

$$(29) \quad \psi_{1,l}(n) \leq 1,$$

$$(30) \quad \Psi_{k,l,p}(n) \leq p\psi_{k,l}(K_{k,l,p}(n)),$$

$$(31) \quad \psi_{k+1,l}(n) \leq 3K_{k,l,l}(2\phi_{k+1}(n)) + \Psi_{k,l,l}(2\phi_{k+1}(n)).$$

*Proof.* We first establish (26) and (29). Consideration of the empty word gives that  $\kappa_{k,0}(n) = n$  and  $\psi_{k,0} = 0$ . Now suppose that  $l \geq 1$  and note that the only pieces of rank 1 are  $a_1^{\pm 1}$ . If  $h = t^r a_1^{\pm 1} t^{-s}$  lies in  $H$ , then  $d_H(1, h) = 1$  and  $r - s = \pm 1$ , whence  $|s| \leq |r| + 1$ . Thus  $\kappa_{k,l}(n) \leq n + 1$  and  $\psi_{k,l}(n) = 1$ .

For (27) and (30), let  $h = t^r w t^{-s}$  where  $w = w(a_1, \dots, a_k)$  is a word of length at most  $l$  with at most  $p$  pieces and  $|r| \leq n$ . Let  $w = w_1 \dots w_q$  be the partition of  $w$  into pieces, where  $q \leq p$ . If  $h \in H$ , then Lemma 7.2 implies that there exist integers  $r = r_0, r_1, \dots, r_q = s$  and elements  $h_1, \dots, h_q$  in  $H$  such that  $t^{r_{i-1}} w_i = h_i t^{r_i}$ . Thus  $|r_i| \leq \kappa_{k,\ell(w_i)}(|r_{i-1}|)$ , whence

$$|s| \leq \kappa_{k,\ell(w_q)}(\dots (\kappa_{k,\ell(w_1)}(|r|)) \dots) \leq \kappa_{k,\ell(w_q)}(\dots (\kappa_{k,\ell(w_1)}(n)) \dots)$$

and we obtain inequality (27). For inequality (30), note that  $|r_i| \leq K_{k,\ell(w_1 \dots w_i),i}(|r|) \leq K_{k,l,p}(n)$ , whence

$$d_H(1, h) \leq \sum_{i=1}^q d_H(1, h_i) \leq \sum_{i=1}^q \psi_{k,\ell(w_i)}(|r_{i-1}|) \leq p\psi_{k,l}(K_{k,l,p}(n)).$$

Finally, (28) and (31) will follow from Proposition 9.4.  $\square$

We now derive Proposition 9.1 from Proposition 9.3. We first use (26), (27) and (28) to obtain bounds on  $\kappa_{k,l}(n)$  and  $K_{k,l,p}(n)$  in terms of Ackermann's functions. We then derive bounds on  $\psi_{k,l}(n)$  and  $\Psi_{k,l,p}(n)$  from (29), (30) and (31), having fed in our bounds on  $\kappa_{k,l}(n)$  and  $K_{k,l,p}(n)$ .

*Proof of Proposition 9.1.* We will need the inequality, established in Lemma 3.2, that for  $n \geq 0$  and  $k \geq 2$ ,

$$(32) \quad \phi_k(n) \leq A_{k-1}(n + k).$$

We first prove that there exist integers  $D_k \geq 1$  such that

$$(33) \quad \kappa_{k,l}(n) \leq A_{k-1}(D_k n + D_k l),$$

$$(34) \quad K_{k,l,p}(n) \leq A_{k-1}^{(p)}(D_k n + D_k l).$$

Inequalities (26) and (27) together imply that  $K_{1,l,p}(n) \leq n + p$ . Thus (33) and (34) hold in the case  $k = 1$  with  $D_1 = 1$ . Now suppose that  $k \geq 2$  and that (33) and (34) hold for smaller

values of  $k$ . If  $l = 0$ , then, using (9), we calculate that  $\kappa_{k,l}(n) = n \leq A_{k-1}(n)$ . If  $l \geq 1$ , then

$$\begin{aligned}
\kappa_{k,l}(n) &\leq 2K_{k-1,l,l}(2\phi_k(n)) && \text{by (28)} \\
&\leq 2K_{k-1,l,l}(2A_{k-1}(n+k)) && \text{by (32)} \\
&\leq 2A_{k-2}^{(l)}(2D_{k-1}A_{k-1}(n+k) + D_{k-1}l) \\
&\leq 2A_{k-2}^{(l)}(A_{k-1}(2D_{k-1}n + D_{k-1}l + 2D_{k-1}k)) && \text{by (8), (10), (13)} \\
&= 2A_{k-1}(2D_{k-1}n + (D_{k-1} + 1)l + 2D_{k-1}k) && \text{by (4)} \\
&\leq A_{k-1}(4D_{k-1}n + 2(D_{k-1} + 1)l + 4D_{k-1}k) && \text{by (10)} \\
&\leq A_{k-1}(4D_{k-1}n + [2(D_{k-1} + 1) + 4D_{k-1}k]l) && \text{by (8)}.
\end{aligned}$$

Taking  $D_k = \max\{2(D_{k-1} + 1) + 4D_{k-1}k, 1\}$ , we obtain (33).

For (34) we calculate that

$$\begin{aligned}
K_{k,l,p}(n) &\leq \max_{\substack{q \leq p \\ l_1 + \dots + l_q \leq l}} \left\{ \kappa_{k,l_1}(\dots \kappa_{k,l_{q-1}}(\kappa_{k,l_q}(n)) \dots) \right\} && \text{by (27)} \\
&\leq \max_{\substack{q \leq p \\ l_1 + \dots + l_q \leq l}} \left\{ A_{k-1}(\dots A_{k-1}(A_{k-1}(D_k n + D_k l_q) + D_k l_{q-1}) \dots) \right\} && \text{by (8)} \\
&\leq \max_{\substack{q \leq p \\ l_1 + \dots + l_q \leq l}} \left\{ A_{k-1}^{(q)} \left( D_k n + D_k \sum_{i=1}^q l_i \right) \right\} && \text{by (8), (13)} \\
&\leq \max_{q \leq p} \left\{ A_{k-1}^{(q)}(D_k n + D_k l) \right\} && \text{by (8)} \\
&\leq A_{k-1}^{(p)}(D_k n + D_k l) && \text{by (9)}.
\end{aligned}$$

Next, we combine (29), (30) and (31) with (33) and (34) to deduce that there exist integers  $E_k, F_k \geq 1$  such that

$$(35) \quad \psi_{k,l}(n) \leq A_{k-1}(E_k n + E_k l),$$

$$(36) \quad \Psi_{k,l,p}(n) \leq A_{k-1}^{(3p)}(F_k n + F_k l).$$

It follows from (29) and (30) that  $\Psi_{1,l,p}(n) \leq p$ . Thus (35) and (36) hold in the case  $k = 1$  with  $E_k = F_k = 1$ . Now suppose that  $k \geq 2$  and that (35) and (36) hold for smaller values of  $k$ . If  $l = 0$ , then  $\psi_{k,l}(n) = 0 \leq A_{k-1}(0)$ . If  $l \geq 1$ , then

$$\begin{aligned}
\psi_{k,l}(n) &\leq 3K_{k-1,l,l}(2\phi_k(n)) + \Psi_{k-1,l,l}(2\phi_k(n)) && \text{by (31)} \\
&\leq 3K_{k-1,l,l}(2A_{k-1}(n+k)) + \Psi_{k-1,l,l}(2A_{k-1}(n+k)) && \text{by (32)} \\
&\leq 3A_{k-2}^{(l)}(2D_{k-1}A_{k-1}(n+k) + D_{k-1}l) \\
&\quad + A_{k-2}^{(3l)}(2F_{k-1}A_{k-1}(n+k) + F_{k-1}l) && \text{by (34)} \\
&\leq 3A_{k-2}^{(l)}(A_{k-1}(2D_{k-1}(n+k) + D_{k-1}l)) \\
&\quad + A_{k-2}^{(3l)}(A_{k-1}(2F_{k-1}(n+k) + F_{k-1}l)) && \text{by (8), (10), (13)} \\
&= 3A_{k-1}(2D_{k-1}(n+k) + (D_{k-1} + 1)l) \\
&\quad + A_{k-1}(2F_{k-1}(n+k) + (F_{k-1} + 3)l) && \text{by (4)} \\
&\leq A_{k-1}(6D_{k-1}(n+k) + 3(D_{k-1} + 1)l) \\
&\quad + A_{k-1}(2F_{k-1}(n+k) + (F_{k-1} + 3)l) && \text{by (10)} \\
&\leq A_{k-1}(2(3D_{k-1} + F_{k-1})(n+k) + (3D_{k-1} + F_{k-1} + 4)l) && \text{by (12)} \\
&\leq A_{k-1}(2(3D_{k-1} + F_{k-1})n + (3(2k+1)D_{k-1} + (2k+1)F_{k-1} + 4)l).
\end{aligned}$$

Taking  $E_k = 3(2k+1)D_{k-1} + (2k+1)F_{k-1} + 4$ , we obtain (35).

If  $p = 0$  or  $l = 0$ , then, using (9), we calculate that  $\Psi_{k,l,p}(n) = 0 \leq A_{k-1}^{(3p)}(0)$ . If  $l, p \geq 1$ , then

$$\begin{aligned}
\Psi_{k,l,p}(n) &\leq p\psi_{k,l}(K_{k,l,p}(n)) && \text{by (30)} \\
&\leq p\psi_{k,l}(A_{k-1}^{(p)}(D_k n + D_k l)) && \text{by (34)} \\
&\leq pA_{k-1}(E_k A_{k-1}^{(p)}(D_k n + D_k l) + E_k l) \\
&\leq pA_{k-1}^{(p+1)}(D_k E_k n + (D_k + 1)E_k l) && \text{by (8), (9), (10), (13)} \\
&\leq A_{k-1}^{(2p+1)}(D_k E_k n + (D_k + 1)E_k l) && \text{by (11),} \\
&\leq A_{k-1}^{(3p)}(D_k E_k n + (D_k + 1)E_k l) && \text{by (9).}
\end{aligned}$$

Taking  $F_k = (D_k + 1)E_k$ , we obtain (36).

Finally, the proof is completed by taking  $C_k = \max\{D_k, E_k, F_k\}$  and applying (8).  $\square$

The remainder of this section is devoted to establishing (28) and (31). This is done in Proposition 9.4, which draws on Lemmas 9.5, 9.6 and 9.7 that follow. We now outline our strategy.

Suppose that  $t^r a_k^{\epsilon_1} w a_k^{-\epsilon_2} t^{-s}$ , where  $r, s \in \mathbb{Z}$ ,  $\epsilon_1, \epsilon_2 \in \{0, 1\}$  and  $w = w(a_1, \dots, a_{k-1})$ , represents an element  $h \in H$ . Our approach will be to find elements  $h_1, h_2 \in H$ , integers  $r', s'$  and a word  $w' = w'(a_1, \dots, a_{k-1})$  such that  $h$  is represented by  $h_1 t^{r'} w' t^{-s'} h_2$ . The functions  $K_{k-1,*,*}$  and  $\Psi_{k-1,*,*}$  will then control the behaviour of the subword  $t^{r'} w' t^{-s'}$ . Together with estimates for  $d_H(1, h_i)$ ,  $|r'|$ ,  $|s'|$  and  $\ell(w')$ , this will allow us to derive bounds on  $|s|$  and  $d_H(1, h)$ .

As indicated by Lemma 8.3, the case  $k = 2$  is exceptional and so will be treated separately. For  $k \geq 3$ , the  $h_1, h_2, r', s'$  and  $w$  will be produced by Lemma 9.5. This lemma takes integers  $k, n$  and  $\epsilon$ , with  $k \geq 3$  and  $\epsilon \in \{0, 1\}$ , and gives an integer  $n'$ , an element  $h \in H$  and a word  $u = u(a_1, \dots, a_{k-1})$  such that  $t^n a_k = h t^{n'} u$  in  $G$ . Applying Lemma 9.5 to  $k, r$  and  $\epsilon_1$  will produce  $r', h_1$  and a word  $u_1$ . Applying Lemma 9.5 to  $k, s$  and  $\epsilon_2$  will produce  $s', h_2^{-1}$  and a word  $u_2$ . The word  $w'$  will then be defined to be the free reduction of  $\tilde{w} := u_1 w u_2^{-1}$ .

The relationship between the input and output of Lemma 9.5 is determined by which of the following holds:

- (i)  $\epsilon = 0$ ,
- (ii)  $\epsilon = 1$  and  $n \leq 0$ , or
- (iii)  $\epsilon = 1$  and  $n > 0$ .

*A priori*, this would lead to us having to consider nine distinct cases, depending on the values of  $\epsilon_1$  and  $\epsilon_2$  and the signs of  $r$  and  $s$ . To streamline the process, Lemma 9.5 packages (i) and (ii) together: it considers the cases that either  $n\epsilon \leq 0$  or  $n\epsilon > 0$ . As such, we need now only consider four cases, depending on the signs of  $r\epsilon_1$  and  $s\epsilon_2$ .

The form of  $\tilde{w}$  will depend on which of (i), (ii) or (iii) applies to  $r$  and  $\epsilon_1$  and to  $s$  and  $\epsilon_2$ . Lemmas 9.6 and 9.7 will be brought to bear to ensure that enough cancellation occurs to obtain a sufficiently strong bound on  $\ell(w')$ .

**Proposition 9.4.** *Let  $h = t^r a_k^{\epsilon_1} w a_k^{-\epsilon_2} t^{-s}$  where  $\epsilon_1, \epsilon_2 \in \{0, 1\}$ ,  $w = w(a_1, \dots, a_{k-1})$  and  $k \geq 2$ . Let  $n$  and  $l$  be integers with  $|r| \leq n$  and  $\ell(w) \leq l$ . If  $h \in H$ , then*

$$\begin{aligned}
|s| &\leq 2K_{k-1,l,l}(2\phi_k(n)), \\
d_H(1, h) &\leq 3K_{k-1,l,l}(2\phi_k(n)) + \Psi_{k-1,l,l}(2\phi_k(n)).
\end{aligned}$$

*Proof.* We claim that there exist  $h_1, h_2 \in H$ ,  $r', s' \in \mathbb{Z}$  and  $w' = w'(a_1, \dots, a_{k-1})$  such that  $h = h_1 t^{r'} w' t^{-s'} h_2$  in  $G$  and

$$(37) \quad |r'| \leq 2\phi_k(n),$$

$$(38) \quad |s| \leq |s'| + 1,$$

$$(39) \quad d_H(1, h_1) \leq |r'| + 1,$$

$$(40) \quad d_H(1, h_2) \leq |s'| + 1,$$

$$(41) \quad \ell(w') \leq l.$$

The result follows from the claim by direct calculation. Indeed, since the number of pieces of a word is bounded by its length,

$$(42) \quad |s'| \leq K_{k-1, \ell(w'), \ell(w')}(|r'|),$$

$$(43) \quad d_H(1, t^{r'} w' t^{-s'}) \leq \Psi_{k-1, \ell(w'), \ell(w')}(|r'|).$$

We will also need the inequality

$$(44) \quad K_{k, l, p}(n) \geq n,$$

which follows immediately from consideration of the empty word. We can now calculate that

$$\begin{aligned} |s| &\leq |s'| + 1 && \text{by (38)} \\ &\leq K_{k-1, \ell(w'), \ell(w')}(|r'|) + 1 && \text{by (42)} \\ &\leq K_{k-1, l, l}(2\phi_k(n)) + 1 && \text{by (37), (41)} \\ &\leq 2K_{k-1, l, l}(2\phi_k(n)) && \text{by (18), (44)} \end{aligned}$$

$$\begin{aligned} d_H(1, h) &\leq d_H(1, h_1) + d_H(1, t^{r'} w' t^{-s'}) + d_H(1, h_2) \\ &\leq |r'| + 1 + \Psi_{k-1, \ell(w'), \ell(w')}(|r'|) + |s'| + 1 && \text{by (39), (40), (43)} \\ &\leq 2\phi_k(n) + 1 + \Psi_{k-1, l, l}(2\phi_k(n)) + K_{k-1, \ell(w'), \ell(w')}(|r'|) + 1 && \text{by (37), (41), (42)} \\ &\leq 4\phi_k(n) + \Psi_{k-1, l, l}(2\phi_k(n)) + K_{k-1, l, l}(2\phi_k(n)) && \text{by (18), (37), (41)} \\ &\leq 3K_{k-1, l, l}(2\phi_k(n)) + \Psi_{k-1, l, l}(2\phi_k(n)) && \text{by (44)}. \end{aligned}$$

We first prove the claim for  $k = 2$ . Since  $t^p a_2 = (a_2 t)(a_1 t)^{-p} t^{2p-1}$ , we can take  $w'$  to be  $w$  and define  $h_1, h_2, r'$  and  $s'$  by

$$\begin{aligned} h_1 &= \begin{cases} 1 & \epsilon_1 = 0, \\ (a_2 t)(a_1 t)^{-r} & \epsilon_1 = 1, \end{cases} & r' &= \begin{cases} r & \epsilon_1 = 0, \\ 2r - 1 & \epsilon_1 = 1, \end{cases} \\ h_2 &= \begin{cases} 1 & \epsilon_2 = 0, \\ (a_1 t)^{-s}(a_2 t)^{-1} & \epsilon_2 = 1, \end{cases} & s' &= \begin{cases} s & \epsilon_2 = 0, \\ 2s - 1 & \epsilon_2 = 1. \end{cases} \end{aligned}$$

Inequalities (38) and (41) are immediate. For (37), use the fact, from Lemma 8.3, that  $\phi_2(n) = n + 1$ . Inequality (39) is immediate if  $\epsilon_1 = 0$ . If  $\epsilon_1 = 1$ , then  $r = \frac{1}{2}(r' + 1)$ , whence  $|r| \leq \frac{1}{2}(|r'| + 1)$ . But  $r' \neq 0$ , so  $|r| \leq |r'|$  and  $d_H(1, h_1) = |r| + 1 \leq |r'| + 1$ . Inequality (40) is derived similarly.

We now prove the claim for  $k \geq 3$ . First apply Lemma 9.5 to  $k, r, \epsilon_1$  to produce  $r', h_1$  and a word  $u_1$ . Then apply it to  $k, s, \epsilon_2$  to produce  $s', h_2^{-1}$  and a word  $u_2$ . Defining  $\tilde{w} := u_1 w u_2^{-1}$ , we have that  $h$  is represented by  $h_1 t^{r'} \tilde{w} t^{-s'} h_2$  and hence that  $t^{r'} \tilde{w} t^{-s'} \in H$ . It is immediate from the bounds given in Lemma 9.5 that (37)–(40) hold. Finally, we define  $w'$  to be the free reduction of  $\tilde{w}$ . To establish (41), we consider four cases.

*Case:*  $r\epsilon_1 \leq 0, s\epsilon_2 \leq 0$ . We have that  $\tilde{w} = w$  and so it is immediate that  $\ell(w') \leq \ell(w)$ .

*Case:*  $r\epsilon_1 > 0$ ,  $s\epsilon_2 \leq 0$ . We have that  $\tilde{w} = \theta^{r-1}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1})w$ . Since  $t^{r'}\theta^{r-1}(a_{k-1}^{-1})$  does not lie in  $\Lambda$ , applying Lemma 7.2 to  $t^{r'}\tilde{w}t^{-s'}$  shows that, when  $\tilde{w}$  is freely reduced, each  $a_{k-1}^{-1}$  in  $\theta^{r-1}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1})$  cancels into  $w$ . It follows from Lemma 9.7 that  $\ell(w') \leq \ell(w)$ .

*Case:*  $r\epsilon_1 \leq 0$ ,  $s\epsilon_2 > 0$ . We have that  $\tilde{w} = w\theta^0(a_{k-1}) \dots \theta^{s-1}(a_{k-1})$ . Since  $t^{s'}\theta^{s-1}(a_{k-1}^{-1})$  does not lie in  $\Lambda$ , applying Lemma 7.2 to  $t^{s'}\tilde{w}^{-1}t^{-r'}$  shows that, when  $\tilde{w}$  is freely reduced, each  $a_{k-1}$  in  $\theta^0(a_{k-1}) \dots \theta^{s-1}(a_{k-1})$  cancels into  $w$ . It follows from Lemma 9.7 that  $\ell(w') \leq \ell(w)$ .

*Case:*  $r\epsilon_1 > 0$ ,  $s\epsilon_2 > 0$ . We have that  $\tilde{w} = \theta^{r-1}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1})w\theta^0(a_{k-1}) \dots \theta^{s-1}(a_{k-1})$ . Since neither  $t^{r'}\theta^{r-1}(a_{k-1}^{-1})$  nor  $t^{s'}\theta^{s-1}(a_{k-1}^{-1})$  lies in  $\Lambda$ , we are in a position to apply Lemma 9.6. If case (i) of Lemma 9.6 occurs, then, when  $\tilde{w}$  is freely reduced, each  $a_{k-1}^{-1}$  in  $\theta^{r-1}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1})$  and each  $a_{k-1}$  in  $\theta^0(a_{k-1}) \dots \theta^{s-1}(a_{k-1})$  cancels into  $w$ . Applying Lemma 9.7 gives that  $\ell(w') \leq \ell(w)$ . On the other hand, suppose that case (ii) of Lemma 9.6 occurs, so  $w'$  is the free reduction of  $\theta^{r-1}(a_{k-1}^{-1})\theta^{s-1}(a_{k-1})$ . We will show that  $r = s$ , whence  $w'$  is the empty word and trivially  $\ell(w') \leq l$ . If  $k = 3$ , then  $t^{r'}w't^{-s'} = t^{r-1}\theta^{r-1}(a_2^{-1})\theta^{s-1}(a_2)t^{1-s} = t^{r-s}a_1^{s-r}$  in  $G$ . Since this element lies in  $H$ ,  $r - s = s - r$ , whence  $r = s$ . If  $k = 4$ , then  $t^{r'}w't^{-s'}$  is freely equal to  $t^{r-1}\theta^{r-1}(a_3^{-1})\theta^{s-1}(a_3)t^{1-s}$ . Since this lies in  $H$ , applying Lemma 8.8 and solving the resulting equation gives  $r = s$ . Finally, suppose that  $k > 4$ . Lemma 8.1 gives that

$$t^{r'}w't^{-s'} \stackrel{\text{fr}}{=} \begin{cases} t^{r-1}\theta^{r-1}(a_{k-2}) \dots \theta^{s-2}(a_{k-2})t^{1-s} & r < s, \\ t^{r-s} & r = s, \\ t^{r-1}\theta^{r-2}(a_{k-2}^{-1}) \dots \theta^{s-1}(a_{k-2}^{-1})t^{1-s} & r > s. \end{cases}$$

By Lemma 8.6, neither  $t^{r-1}\theta^{r-2}(a_{k-2}^{-1})$  nor  $t^{s-1}\theta^{s-2}(a_{k-2}^{-1})$  lies in  $\Lambda$ , since  $k-2 \geq 3$ . Thus, by Lemma 7.2, both  $r < s$  and  $s > r$  lead to a contradiction. Hence  $r = s$  as required.  $\square$

**Lemma 9.5.** *Given integers  $k, n, \epsilon$ , with  $k \geq 3$  and  $\epsilon \in \{0, 1\}$ , there exists an integer  $n'$ , an element  $h \in H$  and a word  $u = u(a_1, \dots, a_{k-1})$  such that  $t^n a_k^\epsilon = ht^{n'}u$  in  $G$ ,*

$$|n| - 1 \leq |n'| \leq 2\phi_k(|n|), \text{ and} \\ d_H(1, h) \leq \max\{|n'|, 1\}.$$

*Furthermore,*

- (i) *if  $n\epsilon \leq 0$ , then  $u$  is the empty word;*
- (ii) *if  $n\epsilon > 0$ , then  $n' = n - 1$ ,  $u = \theta^{n-1}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1})$  and  $t^{n'}\theta^{n-1}(a_{k-1}^{-1}) \notin \Lambda$ .*

*Proof.* We consider three cases.

*Case:*  $\epsilon = 0$ . We trivially obtain an instance of conclusion (i) by taking  $n' = n$ ,  $h = 1$  and  $u$  to be the empty word. The upper bound on  $|n'|$  follows from (18) and (21).

*Case:*  $\epsilon = 1$  and  $n \leq 0$ . Following the calculation

$$t^n a_k = \theta^{-n}(a_k)t^n = \theta^{-n}(a_k)t^{\phi_k(|n|)}t^{n-\phi_k(|n|)},$$

we obtain an instance of conclusion (i) by taking  $n' = n - \phi_k(|n|)$ ,  $h = \theta^{-n}(a_k)t^{\phi_k(|n|)}$  and  $u$  to be the empty word. It follows immediately from the definition of the function  $\phi_k$  that  $h \in H$  and from Lemma 8.4 that  $d_H(1, h) = \phi_k(|n|)$ . By (18),  $\phi_k(|n|)$  is positive whence  $|n'| = |n| + \phi_k(|n|)$  and  $d_h(1, h) \leq |n'|$ . Applying (18) and (21) gives  $|n| + 1 \leq |n'| \leq 2\phi_k(|n|)$ .

*Case:*  $\epsilon = 1$  and  $n > 0$ . Following the calculation

$$t^n a_k = a_k a_k^{-1} t^n a_k = a_k t^n \theta^n(a_k^{-1}) a_k = (a_k t) t^{n-1} \theta^{n-1}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1})$$

we obtain an instance of conclusion (ii) by taking  $n' = n - 1$ ,  $h = (a_k t)$  and

$$u = \theta^{n-1}(a_{k-1}^{-1}) \dots \theta^0(a_{k-1}^{-1}).$$

The upper bound on  $|n'|$  follows from (18) and (21). The fact that  $t^{n'} \theta^{n-1}(a_{k-1}^{-1})$  does not lie in  $\Lambda$  follows from Lemmas 8.6 and 8.7.  $\square$

**Lemma 9.6.** *Let  $\sigma = t^r \theta^a(a_k^{-1}) \dots \theta^0(a_k^{-1}) w \theta^0(a_k) \dots \theta^b(a_k) t^{-s}$  where  $w = w(a_1, \dots, a_k)$  is freely reduced and  $a, b \geq 0$ . Suppose  $\sigma$  represents an element of  $H$  but  $t^r \theta^a(a_k^{-1}) \notin \Lambda$  and  $t^s \theta^b(a_k^{-1}) \notin \Lambda$ . Then either*

- (i)  $w$  has a prefix  $\theta^0(a_k) \dots \theta^{a-1}(a_k) a_k$  and suffix  $a_k^{-1} \theta^{b-1}(a_k^{-1}) \dots \theta^0(a_k^{-1})$ , or
- (ii)  $w = \theta^0(a_k) \dots \theta^{a-1}(a_k) \theta^{b-1}(a_k^{-1}) \dots \theta^0(a_k^{-1})$ .

*Proof.* Write  $l_1$  for the letter  $a_k^{-1}$  of the term  $\theta^a(a_k^{-1})$  of  $\sigma$  and write  $l_2$  for the letter  $a_k$  of the term  $\theta^b(a_k)$  of  $\sigma$ . Lemma 7.2 implies that, when  $\sigma$  is freely reduced, both  $l_1$  and  $l_2$  cancel. Let  $l'$  be the letter  $a_k$  that cancels with  $l_1$

If  $l'$  lies in  $w$ , then  $l_2$  must cancel with a letter to the right of  $l'$  in  $w$ , and we have case (i).

On the other hand, suppose that  $l'$  lies in the subword  $\theta^0(a_k) \dots \theta^b(a_k)$ . If  $l'$  is distinct from  $l_2$ , then  $l_2$  must cancel with an  $a_k^{-1}$  lying to the right of  $l'$ . But this is a contradiction, since all the occurrences of  $a_k^{\pm 1}$  in  $\theta^0(a_k) \dots \theta^b(a_k)$  are positive. Thus  $l' = l_2$ . Now  $\theta^{a-1}(a_k^{-1}) \dots \theta^0(a_k^{-1}) w \theta^0(a_k) \dots \theta^{b-1}(a_k)$  must be freely trivial and we have case (ii).  $\square$

**Lemma 9.7.** *Let  $w = \theta^0(a_k) \dots \theta^r(a_k)$  where  $r \geq 0$ . Let  $l$  be the last  $a_k$  appearing in  $w$  and partition  $w$  as  $w = uv$  where  $u$  is the prefix of  $w$  ending with  $l$  and  $v$  is the suffix of  $w$  coming after  $l$ . Then  $\ell(u) \geq \ell(v)$ .*

*Proof.* Note that  $u = \theta^0(a_k) \dots \theta^{r-1}(a_k) a_k$  and, by Lemma 8.1,  $v = \theta^0(a_{k-1}) \dots \theta^{r-1}(a_{k-1})$ . It thus suffices to prove that  $\ell(\theta^i(a_k)) \geq \ell(\theta^i(a_{k-1}))$  for  $i \geq 0$ . But this follows by an easy induction on  $k$  from the structures of  $\theta^i(a_k)$  and  $\theta^i(a_{k-1})$  respectively given by Lemma 8.1.  $\square$

## 10. GROUPS WITH ACKERMANNIAN DEHN FUNCTIONS

Recall that  $\Gamma_k$  is the HNN extension of  $G_k$  over  $H_k$  in which the stable letter commutes with all elements of  $H_k$ :

$$\Gamma_k := \langle a_1, \dots, a_k, t, p \mid t^{-1} a_1 t = a_1, t^{-1} a_i t = a_i a_{i-1} \ (i > 1), [p, a_i t] = 1 \ (i > 0) \rangle.$$

**Proposition 10.1.** *The group  $\Gamma_1$  has Dehn function  $\simeq$ -equivalent to  $n \mapsto n^2$ .*

*Proof.* Making the substitution  $\alpha = a_1 t$  shows that  $\Gamma_1$  is a right-angled Artin group with presentation  $\langle \alpha, t, p \mid [t, \alpha], [p, \alpha] \rangle$ . It follows that  $\Gamma_1$  is CAT(0) [11] whence it has Dehn function  $\simeq$ -equivalent to  $n^2$  by [9, Proposition 1.6.III.Γ].  $\square$

**Proposition 10.2.** *For all  $k \geq 2$ , the group  $\Gamma_k$  has Dehn function  $\simeq$ -equivalent to  $A_k$ .*

*Proof.* Let  $k \geq 2$ . The Dehn function of a CAT(0) group is either linear or quadratic [7, Theorem 6.2.1], with the linear case occurring precisely when the group is hyperbolic [7, Theorem 6.1.5]. By Theorem 1.3, the group  $G_k$  is CAT(0). However, since it contains an embedded copy of  $\mathbb{Z}^2$  it is not hyperbolic [7, Theorem 6.1.10]. The Dehn function of  $G_k$  is therefore quadratic. By Theorem 1.3, the distortion function of  $H_k$  in  $G_k$  is  $\simeq$ -equivalent to  $A_k$ . Plugging these two functions into [9, Theorem 6.20.III.Γ] gives lower and upper bounds for the Dehn function of  $\Gamma_k$  of  $\max\{n^2, nA_k(n)\}$  and  $nA_k(n)^2$  respectively, up to  $\simeq$ -equivalence. By (9), this lower bound is equal to  $A_k(n)^2$  and the upper bound is at most

$A_k(n)^3$ . But (14) implies that, for any  $C \geq 1$ , the function  $n \mapsto A_k(n)^C$  is  $\approx$ -equivalent to  $A_k$ .  $\square$

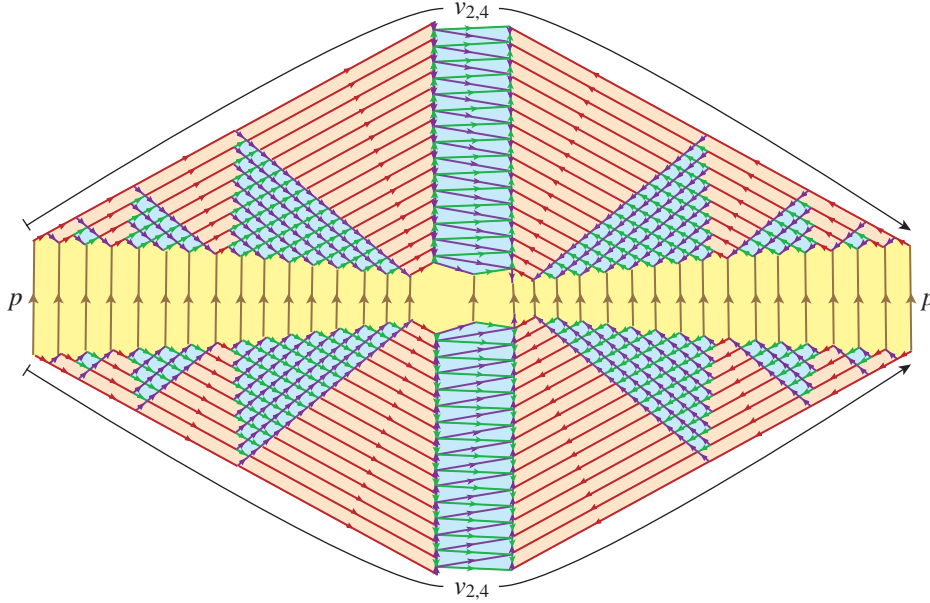


FIGURE 3. A van Kampen diagram for  $[v_{2,4}, p]$  — an example of a word which represents the identity in  $\Gamma_k$  but can only be filled by a large area diagram.

The ideas behind [9, Theorem 6.20.III.Γ] used here are most transparent via the tools of van Kampen diagrams and corridors. For example, towards the lower bound, consider the words

$$v_{k,n} := a_k^n a_2 t a_1 a_2^{-1} a_k^{-n}$$

of Section 6, which equal

$$w_{k,n} := u_{k,n} (a_2 t) (a_1 t) (a_2 t)^{-1} u_{k,n}^{-1}$$

in  $G_k$ . Observe that  $[v_{k,n}, p] = 1$  in  $\Gamma_k$  and that in any van Kampen diagram for  $[v_{k,n}, p]$ , there must be a  $p$ -corridor connecting the two boundary edges labelled by  $p$ . (Figure 3 is an example of such a diagram when  $k = 2$  and  $n = 4$ .) The word on  $a_1 t, \dots, a_k t$  written along each side of this corridor must equal  $v_{k,n}$  in  $G_k$  and so freely equals  $w_{k,n}$ . It follows that any van Kampen diagram for  $[v_{k,n}, p]$  has area at least the length of  $w_{k,n}$ , which is  $2\mathcal{H}_k(n) + 3$ . So, as the length of  $[v_{k,n}, p]$  is  $4n + 10$ , this leads to a lower bound of  $A_k(n) \approx \mathcal{H}_k$  on the Dehn function of  $G_k$ .

#### REFERENCES

- [1] G. Arzhantseva and D. Osin. Solvable groups with polynomial Dehn functions. *Trans. Amer. Math. Soc.*, 354(8):3329–3348, 2002.
- [2] J. Barnard, N. Brady, and P. Dani. Super-exponential distortion of subgroups of  $\text{CAT}(-1)$  groups. *Algebr. Geom. Topol.*, 7:301–308, 2007.
- [3] G. Baumslag. A non-cyclic one-relator group all of whose finite quotients are cyclic. *J. Austral. Math. Soc.*, 10:497–498, 1969.
- [4] G. Baumslag, M. R. Bridson, C. F. Miller, and H. Short. Finitely presented subgroups of automatic groups and their isoperimetric functions. *J. London Math. Soc. (2)*, 56(2):292–304, 1997.
- [5] L. D. Beklemishev. The worm principle. In *Logic Colloquium '02*, volume 27 of *Lect. Notes Log.*, pages 75–95. Assoc. Symbol. Logic, La Jolla, CA, 2006.

- [6] A. A. Bernasconi. *On HNN-extensions and the complexity of the word problem for one-relator groups*. PhD thesis, University of Utah, 1994.  
<http://www.math.utah.edu/~sg/Papers/bernasconi-thesis.pdf>.
- [7] M. R. Bridson. The geometry of the word problem. In M. R. Bridson and S. M. Salamon, editors, *Invitations to Geometry and Topology*, pages 33–94. O.U.P., 2002.
- [8] M. R. Bridson. Polynomial Dehn functions and the length of asynchronously automatic structures. *Proc. London Math. Soc.*, 85(2):441–465, 2002.
- [9] M. R. Bridson and A. Haefliger. *Metric Spaces of Non-positive Curvature*. Number 319 in Grundlehren der mathematischen Wissenschaften. Springer Verlag, 1999.
- [10] W. Buchholz. An independence result for  $(\Pi_1^1\text{-CA})+\text{Bl}$ . *Ann. Pure Appl. Logic*, 33(2):131–155, 1987.
- [11] R. Charney and M. W. Davis. Finite  $K(\pi, 1)$ s for Artin groups. In *Prospects in topology (Princeton, NJ, 1994)*, volume 138 of *Ann. of Math. Stud.*, pages 110–124. Princeton Univ. Press, Princeton, NJ, 1995.
- [12] D. E. Cohen. The mathematician who had little wisdom: a story and some mathematics. In *Combinatorial and geometric group theory (Edinburgh, 1993)*, volume 204 of *London Math. Soc. Lecture Note Ser.*, pages 56–62. Cambridge Univ. Press, Cambridge, 1995.
- [13] D. E. Cohen, K. Madlener, and F. Otto. Separating the intrinsic complexity and the derivational complexity of the word problem for finitely presented groups. *Math. Logic Quart.*, 39(2):143–157, 1993.
- [14] B. Farb. The extrinsic geometry of subgroups and the generalised word problem. *Proc. London Math. Soc.* (3), 68(3):577–593, 1994.
- [15] S. M. Gersten. Isodiametric and isoperimetric inequalities in group extensions. Preprint, University of Utah, 1991.
- [16] S. M. Gersten. Isoperimetric and isodiametric functions. In G. Niblo and M. Roller, editors, *Geometric group theory I*, number 181 in LMS lecture notes. Camb. Univ. Press, 1993.
- [17] S. M. Gersten. Quadratic divergence of geodesics in CAT(0) spaces. *Geom. Funct. Anal.*, 4(1):37–51, 1994.
- [18] S. M. Gersten and H. Short. Small cancellation theory and automatic groups. *Invent. Math.*, 102(2):305–334, 1990.
- [19] 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.
- [20] M. Hamano and M. Okada. A relationship among Gentzen’s proof-reduction, Kirby-Paris’ hydra game and Buchholz’s hydra game. *Math. Logic Quart.*, 43(1):103–120, 1997.
- [21] P. Jolissaint. Rapidly decreasing functions in reduced  $C^*$ -algebras of groups. *Trans. Amer. Math. Soc.*, 317(1):167–196, 1990.
- [22] L. Kirby and J. Paris. Accessible independence results for Peano arithmetic. *Bull. London Math. Soc.*, 14(4):285–293, 1982.
- [23] N. Macura. Quadratic isoperimetric inequality for mapping tori of polynomially growing automorphisms of free groups. *Geom. Funct. Anal.*, 10(4):874–901, 2000.
- [24] N. Macura. Detour functions and quasi-isometries. *Q. J. Math.*, 53(2):207–239, 2002.
- [25] K. Madlener and F. Otto. Pseudonatural algorithms for the word problem for finitely presented monoids and groups. *J. Symbolic Comput.*, 1(4):383–418, 1985.
- [26] M. Mitra. Cannon–Thurston maps for trees of hyperbolic metric spaces. *J. Diff. Geom.*, 48(1):135–164, 1998.
- [27] G. A. Niblo and L. D. Reeves. The geometry of cube complexes and the complexity of their fundamental groups. *Topology*, 37(3):621–633, 1998.
- [28] A. Yu. Ol’shanskii and M. V. Sapir. Length and area functions on groups and quasi-isometric Higman embeddings. *Internat. J. Algebra Comput.*, 11(2):137–170, 2001.
- [29] Ch. Pittet. *Géométrie des groupes, inégalités isopérimétriques de dimension 2 et distorsions*. PhD thesis, Université de Genève, 1992.
- [30] A. N. Platonov. An isoperimetric function of the Baumslag–Gersten group. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, 3:12–17, 70, 2004. Translation in *Moscow Univ. Math. Bull.* 59 (2004).
- [31] H. E. Rose. *Subrecursion: functions and hierarchies*, volume 9 of *Oxford Logic Guides*. The Clarendon Press Oxford University Press, New York, 1984.

WILL DISON

Department of Mathematics, University Walk, Bristol, BS8 1TW, UK  
w.dison@bristol.ac.uk, <http://www.maths.bris.ac.uk/~mawjd/>

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/>