Snowflake groups, Perron–Frobenius eigenvalues, and isoperimetric spectra

NOEL BRADY
MARTIN R BRIDSON
MAX FORESTER
KRISHNAN SHANKAR

The k–dimensional Dehn (or isoperimetric) function of a group bounds the volume of efficient ball-fillings of k–spheres mapped into k–connected spaces on which the group acts properly and cocompactly; the bound is given as a function of the volume of the sphere. We advance significantly the observed range of behavior for such functions. First, to each non-negative integer matrix P and positive rational number r, we associate a finite, aspherical 2–complex $X_{r,P}$ and determine the Dehn function of its fundamental group $G_{r,P}$ in terms of r and the Perron–Frobenius eigenvalue of P. The range of functions obtained includes $\delta(x) = x^s$, where $s \in \mathbb{Q} \cap [2, \infty)$ is arbitrary. Next, special features of the groups $G_{r,P}$ allow us to construct iterated multiple HNN extensions which exhibit similar isoperimetric behavior in higher dimensions. In particular, for each positive integer k and rational $s \geqslant (k+1)/k$, there exists a group with k–dimensional Dehn function x^s. Similar isoperimetric inequalities are obtained for fillings modeled on arbitrary manifold pairs $(M, \partial M)$ in addition to (B^{k+1}, S^k).

20F65; 20F69, 20E06, 57M07, 57M20, 53C99

Introduction

Given a k–connected complex or manifold one wants to identify functions that bound the volume of efficient ball-fillings for spheres mapped into that space. The purpose of this article is to advance the understanding of which functions can arise when one seeks optimal bounds in the universal cover of a compact space. Despite the geometric nature of both the problem and its solutions, our initial impetus for studying isoperimetric problems comes from algebra, more specifically the *word problem* for groups.

The quest to understand the complexity of word problems has been at the heart of combinatorial group theory since its inception. When one attacks the word problem for a finitely presented group G directly, the most natural measure of complexity is the Dehn function $\delta(x)$ which bounds the number of defining relations that one must apply to a word $w =_G 1$ to reduce it to the empty word; the bound is a function of word-length $|w|$. The function $\delta(x)$ recursive if and only if G has a solvable word problem.

Progress in the last ten years has led to a fairly complete understanding of which functions arise as Dehn functions of finitely presented groups. The most comprehensive information comes from [18] where, modulo issues associated to the $P = NP$ question, Birget, Rips and Sapir essentially provide a characterisation of the Dehn functions greater than x^4. In particular they show that the following *isoperimetric spectrum* is dense in the range $[4, \infty)$.

$$\mathrm{IP} = \{\alpha \in [1, \infty) \mid f(x) = x^\alpha \text{ is equivalent to a Dehn function}\}.$$

Gromov proved that $\mathrm{IP} \cap (1, 2)$ is empty and that word hyperbolic groups can be characterised as those which have linear Dehn functions. In [3] Brady and Bridson completed the understanding of the *coarse* structure of IP by providing a dense set of exponents in $\mathrm{IP} \cap [2, \infty)$. What remains unknown is the fine structure of $\mathrm{IP} \cap (2, 4)$. In particular, it has remained unknown whether $\mathbb{Q} \cap (2, 4) \subset \mathrm{IP}$. There has, however, been recent progress on understanding Dehn functions below x^4 that are not of the form x^α. For instance, Ol′shanskii and Sapir [16] have constructed groups with Dehn function $x^2 \log(x)$, and Ol′shanskii [15] has constructed examples with more exotic, almost-quadratic behavior.

What Brady and Bridson actually do in [3] is associate to each pair of positive integers $p \geqslant q$ a finite aspherical 2–complex whose fundamental group $G_{p,q}$ has Dehn function $x^{2\log_2 2p/q}$. These complexes are obtained by attaching a pair of annuli to a torus, the attaching maps being chosen so as to ensure the existence of a family of discs in the universal cover that display a certain *snowflake geometry* (cf Figure 4 below). In the present article we present a more sophisticated version of the snowflake construction that yields a much larger class of isoperimetric exponents.

Theorem A *Let P be an irreducible non-negative integer matrix with Perron–Frobenius eigenvalue $\lambda > 1$, and let r be a rational number greater than every row sum of P. Then there is a finitely presented group $G_{r,P}$ with Dehn function $\delta(x) \simeq x^{2\log_\lambda(r)}$.*

Here, $\simeq$ denotes coarse Lipschitz equivalence of functions. By taking P to be the 1×1 matrix (2^{2q}) and $r = 2^p$ (for integers $p > 2q$) we obtain the Dehn function $\delta(x) \simeq x^{p/q}$ and deduce the following corollary.

Corollary B $\mathbb{Q} \cap (2, \infty) \subset \mathrm{IP}$.

The influential work of M. Gromov [11], [12] embedded the word problem in the broader context of filling problems for Riemannian manifolds and combinatorial complexes. For example, Gromov's Filling Theorem [5] states that given a compact Riemannian manifold M, the smallest function bounding the area of least-area discs in M as a function of their boundary length is coarsely Lipschitz equivalent to the Dehn function of $\pi_1 M$. In the geometric context, it is natural to extend questions about the size of optimal fillings to higher-dimensional spheres, exploring higher-dimensional isoperimetric functions that bound the volume of optimal ball-fillings of spheres mapped into the manifold (or complex). Correspondingly, one defines higher-dimensional Dehn functions $\delta^{(k)}(x)$ for finitely presented groups G that have a classifying space with a compact $(k+1)$–skeleton (see Section 2). The equivalence class of $\delta^{(k)}$ is a quasi-isometry invariant of G, by Alonso–Wang–Pride [2].

In contrast to the situation of ordinary Dehn functions, Papasoglu [17] has shown that $\delta^{(2)}(x)$ is always bounded by a recursive function. This is not the case in higher dimensions, however. For each $k > 2$, Young [23] constructs a group for which $\delta^{(k)}(x)$ is not subrecursive.

For each positive integer k one has the *k–dimensional isoperimetric spectrum*

$$\mathrm{IP}^{(k)} = \{\alpha \in [1, \infty) \mid f(x) = x^{\alpha} \text{ is equivalent to a } k\text{–dimensional Dehn function}\}.$$

We do not yet have as detailed a knowledge of the structure of these sets as we do of $\mathrm{IP} = \mathrm{IP}^{(1)}$. Indeed knowledge until now has been remarkably sparse even for $\mathrm{IP}^{(2)}$: the results of Alonso *et al* [1], Wang and Pride [22], and Wang [21] provide infinite sets of exponents in the range $[3/2, 2)$ and provide evidence for the existence of exponents in the range $[2, \infty)$; the snowflake construction of Brady and Bridson [3] provides a dense set of exponents in the interval $[3/2, 2)$; and in Bridson [6] it is was proved that $2, 3 \in \mathrm{IP}^{(2)}$ (see also Burillo [8]). Gromov and others have investigated the isoperimetric behavior of lattices [12].

Our second theorem relieves the dearth of knowledge about the coarse structure of $\mathrm{IP}^{(k)}$, $k \geqslant 2$.

Theorem C *Let P be an irreducible non-negative integer matrix with Perron–Frobenius eigenvalue $\lambda > 1$, and let r be an integer greater than every row sum of P. Then for every $k \geqslant 2$ there is a group $\Sigma^{k-1} G_{r,P}$ of type $\mathcal{F}_{k+1}$ with k–dimensional Dehn function $\delta^{(k)}(x) \simeq x^{2 \log_{\lambda}(r)}$. There are also groups $\Sigma^{k-1}\mathbb{Z}^2$ of type $\mathcal{F}_{k+1}$ with k–dimensional Dehn function $\delta^{(k)}(x) \simeq x^2$.*

By taking P to be the 1×1 matrix (2^{2q}) and $r = 2^p$ we see that $\mathbb{Q} \cap [2, \infty) \subset \mathrm{IP}^{(k)}$; in particular $\mathrm{IP}^{(k)}$ is dense in the range $[2, \infty)$. But that falls short of one's expectations: as in the case $k = 1$, one anticipates that $\mathrm{IP}^{(k)}$ should be dense in the range that begins with the exponent $(k+1)/k$ corresponding to the isoperimetric inequality for spheres in Euclidean space. In order to fulfil this expectation, we investigate the higher Dehn functions of products $G \times \mathbb{Z}$ and prove the following theorem.

Theorem D *If P, λ and r are as in Theorem C, then for all $q, \ell \in \mathbb{N}$, the $(q+\ell)$–dimensional Dehn function of $\Sigma^{q-1} G_{r,P} \times \mathbb{Z}^\ell$ is equivalent to x^s, where $s = \frac{(\ell+1)\alpha - \ell}{\ell\alpha - (\ell-1)}$ and $\alpha = 2\log_\lambda(r)$. The $(q+\ell)$–dimensional Dehn function of $\Sigma^{q-1}\mathbb{Z}^2 \times \mathbb{Z}^\ell$ is equivalent to x^s, where $s = \frac{\ell+2}{\ell+1}$.*

By holding q and ℓ fixed and varying r and P, one obtains a dense set of exponents s in the interval $[\frac{\ell+2}{\ell+1}, \frac{\ell+1}{\ell}]$ including all rationals in this range. By varying q and ℓ with $k = q + \ell$ and taking account of Theorem C we deduce the following result, shown pictorially in Figure 1.

Corollary E $\mathbb{Q} \cap [(k+1)/k, \infty) \subset \mathrm{IP}^{(k)}$.

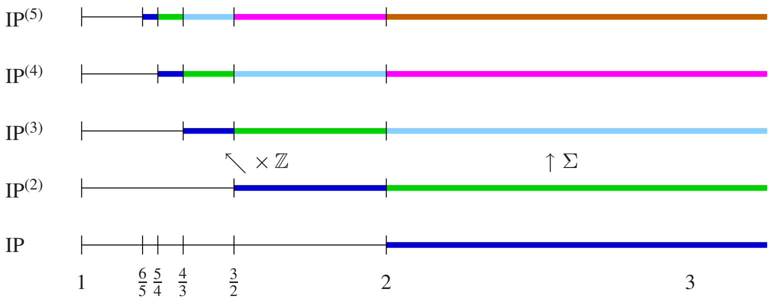

Figure 1: Isoperimetric exponents of $\Sigma^{q-1} G_{r,P} \times \mathbb{Z}^\ell$. Colors correspond to fixed values of q.

The main aim of Brady and Bridson's initial construction of snowflake groups [3] was to prove that the closure of $\mathrm{IP}^{(1)}$ is $\{1\} \cup [2, \infty)$. Corollary E implies that the closure of $\mathrm{IP}^{(k)}$ contains $\{1\} \cup [(k+1)/k, \infty)$. Building on this result, Brady and Forester [4] have recently shown that the closure of $\mathrm{IP}^{(k)}$ is in fact equal to $[1, \infty)$ for $k \geqslant 2$. Other examples, known earlier, include solvable and nilpotent groups whose two-dimensional Dehn functions appear to be $x \log x$ and $x^{4/3}$ respectively, by Wang [21] and Coulhon–Saloff-Coste [9] (the latter was pointed out to us by Robert Young). It should be noted, however, that in both cases the upper bound is derived using the Sobolev inequality from Varopoulos–Saloff-Coste–Coulhon [20]. The resulting isoperimetric inequality concerns embedded fillings only, which do not *a priori* suffice for our definitions (which allow singular maps and fillings).

This article is organised as follows. In Section 1 we outline the construction of the snowflake groups $G_{r,P}$ and their HNN extensions $\Sigma G_{r,P}$, deferring a detailed account to Sections 4 and 6. In Section 2 we define the class of maps with which we shall be working and record some pertinent properties; we also recall those elements of Perron–Frobenius theory that we require. The groups $G_{r,P}$ are fundamental groups of graphs of groups; in Section 3 we analyze the geometry of the vertex groups in these decompositions. The snowflake geometry of $G_{r,P}$ is described in Section 4 and this is analyzed in further detail in Section 5 to prove Theorem A. In Section 6 we turn our attention to higher Dehn functions and establish the lower bounds required for Theorem C by analyzing the geometry of an explicit sequence of embedded $(k+1)$–balls in the universal cover of a $(k+1)$–dimensional classifying space for $\Sigma^{k-1}G_{r,P}$. In Section 7 we establish the complementary upper bounds. The proof proceeds by induction, slicing balls into slabs based of lower-dimensional fillings. A lack of control on the topology of these slabs obliges one to prove a stronger result: instead of establishing bounds only on the behavior of ball-fillings for spheres, one must establish isoperimetric inequalities for all pairs of compact manifolds $(M^{(k+1)}, \partial M)$ mapping to the space in question. In Section 8 we analyze the isoperimetric behaviour of products $G \times \mathbb{Z}$ and complete the proof of Theorem D.

Acknowledgments We would like to thank Josh Barnard for helpful discusssions on admissible maps, and Robert Young for the observation mentioned above. We also thank the referee for many detailed and helpful comments.

Brady's work was funded in part by NSF grant no. DMS-0505707, Bridson's by a Royal Society Wolfson Research Merit Award and an EPSRC Senior Fellowship, Forester's by NSF grant no. DMS-0605137, and Shankar's by NSF grant no. DMS-0336681. Brady and Forester thank the Centre de Recerca Matemàtica for their support and hospitality.

1 An outline of the basic construction

The groups $G_{r,P}$ we consider are fundamental groups of graphs of groups whose underlying graphs are determined by a non-negative integer matrix P. The edge groups are infinite cyclic, with attaching maps determined by a rational number r. The vertex groups V_m have many properties in common with free abelian groups of rank m. Indeed, for the purposes of this summary, the reader may take $V_m = \mathbb{Z}^m$ (cf Remark 1.1). There is a distinguished element $c \in V_m$, corresponding to the diagonal element (the product of the standard generators) in $\mathbb{Z}^m$. The precise definition of V_m is given in Section 3.

The key geometric idea behind Theorem A is that efficient van Kampen diagrams for the groups $G_{r,P}$ exhibit the *snowflake geometry* illustrated in Figure 4. The essential features of such diagrams are these: the diagram is composed of polygonal subdiagrams joined across strips so that the dual to the decomposition is a tree T; and each of the polygonal subdiagrams is a van Kampen diagram in one of the vertex groups V_m (typically it is an $(m+1)$–gon with a base labeled by a power of the distinguished $c \in V_m$ and m other sides labeled by powers of the m standard generators of V_m).

The most important class of diagrams are those that are as symmetric as possible, having the property that as one moves from the circumcenter of the dual tree to the boundary of the diagram, the joining strips are all oriented in such a way that the length of the side strip *decreases* by a factor of r as one journeys towards the boundary. The labels on the outer sides of the strips are powers of the diagonal elements in various vertex groups V_m, and a crucial feature of our construction is that the cyclic subgroups $\langle c \rangle \subset G_{r,P}$ are distorted in a precisely understood manner, with distortion funtion $\simeq x^\alpha$ where $\alpha = \log_\lambda(r)$ and λ is the Perron–Frobenius eigenvalue of P. This distortion is determined through the analysis of certain paths, called *snowflake paths*, which play the role of quasi-geodesics in $G_{r,P}$. These snowflake paths are the result of a curve shortening process; the dynamics of this process are at the heart of our calculations and this is where the Perron–Frobenius theory enters – see Section 4.

If the tree T has radius d, then arguing by induction on d in a suitable class of diagrams, one calculates the length of the boundary to be $\sim d^{k/\alpha}$ if the central polygon has base $\sim d^k$. One has a precise understanding of the quadratic Dehn functions of the vertex groups V_m, and this leads to an area estimate of $\sim d^{2k}$ on these diagrams of diameter $\sim d^k$. Thus we obtain a family of diagrams with area $\sim d^{2k}$ and perimeter $\sim d^{k/\alpha}$, and an elementary manipulation of logs provides the required lower bound of $x^{2\log_\lambda(r)}$ for the Dehn function of $G_{r,P}$. The complementary upper bound is established in Section 5.

A key feature in our construction of $G_{r,P}$ is that when r is an integer, the snowflake diagrams admit a precise scaling by a factor of r, induced by a monomorphism. The ascending HNN extension $G^*_{r,P}$ of $G_{r,P}$ corresponding to this monomorphism is a group in which one can *stack* scaled snowflake diagrams (see Figure 8). By putting together two such stacks (using two stable letters) one obtains a *snowflake ball*, having the same proportions as its equatorial snowflake disk. That is, the interior volume and surface area of the ball are comparable to the area and boundary length, respectively, of the equatorial disk. In this way, one discovers the higher dimensional isoperimetric behavior of the multiple HNN extension $\Sigma G_{r,P}$ which is obtained by amalgamating two copies of $G^*_{r,P}$ along $G_{r,P}$.

The snowflake balls just described determine a lower bound for the 2–dimensional Dehn function mentioned in Theorem C; upper bounds in this case can be deduced by using Wang–Pride [22]. To proceed in higher dimensions we iterate the suspension procedure described above. Lower bounds can be determined as before. However, in dimensions greater than 2, upper bounds require new techniques. In particular, we need to consider isoperimetric inequalities for compact manifolds $(M, \partial M)$ other than (B^k, S^{k-1}). Using this perspective, we establish general upper bounds for ascending HNN extensions. This is achieved in Theorem 7.2 and is further refined in Theorem 8.1.

Remark 1.1 The actual vertex groups V_m of $G_{r,P}$ are themselves fundamental groups of graphs of groups with vertex groups $\mathbb{Z}^2$ and edge groups $\mathbb{Z}$. It turns out that this structure is compatible with the larger $G_{r,P}$ graph of groups structure. That is, $G_{r,P}$ itself may be viewed as the fundamental group of an aspherical 2–complex assembled from a finite collection of tori and annuli. With respect to a fixed framing on the tori, the attaching maps of the annuli are all powers of the slopes $\{1/0,\ 0/1,\ 1/1\}$. From this perspective, it is perhaps surprising that one can encode the range of isoperimetric exponents stated in Theorem A.

An explicit example

We conclude this outine with an explicit example illustrating Theorem A. The example that we present here has Dehn function $x^{p/q}$, where $p > 2q$ are positive integers (common factors are allowed).

Let P be the 1×1 matrix with entry $2^{2q} = 4^q$ and let $r = 2^p$. Then $G_{r,P}$ is the fundamental group of a graph of groups $\mathcal{G}$ with one vertex group and 4^q infinite cyclic edge groups. The single vertex group V_{4^q} is the fundamental group of a tree of groups that we shall describe in a moment. V_{4^q} has generators $a_1, \ldots, a_{4^q}$; the product of these generators $c = a_1 \cdots a_{4^q}$ plays a special role.

The ith edge group of $\mathcal{G}$ has two monomorphisms to the vertex group V_{4^q}. One maps the generator to c and the other maps the generator to $a_i^{2^p}$. Thus we have a relative presentation

$$G_{p/q} \;=\; G_{r,P} \;=\; \langle\, V_{4^q}, s_1, \ldots, s_{4^q} \mid s_i^{-1} a_i^{2^p} s_i = c \ \ (i = 1, \ldots, 4^q) \,\rangle.$$

It remains to elucidate the structure of the group V_{4^q}. This is the fundamental group of a tree of groups in which each of the vertex groups is isomorphic to $\mathbb{Z}^2$ and each of the edge groups is infinite cyclic. The underlying tree is a segment with 4^q-2 edges and 4^q-1 vertices. A basis $\{a_i, b_i\}$ is fixed for each vertex group, and the generator of each edge group maps to the generator a_i of the left-hand vertex group, and to the *diagonal element* $a_{i+1}b_{i+1}$ of the right-hand vertex group.

The generators $a_1, \ldots, a_{4^q}$ mentioned above are the generators a_i of these vertex groups together with $a_{4^q} = b_{4^q-1}$. The distinguished element c is the diagonal a_1b_1 of the leftmost vertex group $\mathbb{Z}^2$ (see Figure 2(a)).

Theorem A tells us that the Dehn function of $G_{p/q}$ is x^α where $\alpha = 2\log_{4^q} 2^p = p/q$. Consider, for example, the group $G_{5/2}$ with Dehn function $x^{5/2}$. In this case, the tree described above is a segment of length 14 and the above description of V_{4^q} yields the presentation

$$\langle a_1, b_1, a_2, b_2, \ldots, a_{15}, b_{15} \mid [a_i, b_i]\ (i = 1, \ldots, 15),\ b_i = a_{i+1}b_{i+1}\ (i = 1, \ldots, 14)\,\rangle.$$

Eliminating the superfluous generators $b_1, \ldots, b_{14}$ and relabelling b_{15} as a_{16}, as in the description of V_{4^q} above, we get

$$V_{16} \;=\; \langle\, a_1, \ldots, a_{16} \mid \theta \in \mathcal{C}_{16}\,\rangle$$

where $\mathcal{C}_{16}$ is the following set of commutators:

$$[a_1, a_2 \cdots a_{16}],\ [a_2, a_3 \cdots a_{16}], \cdots, [a_{14}, a_{15}a_{16}],\ [a_{15}, a_{16}].$$

Thus we obtain the explicit presentation

$$G_{5/2} = \langle\, a_1, \ldots, a_{16}, s_1, \ldots, s_{16} \mid \mathcal{C}_{16};\ s_i^{-1}a_i^{32}s_i = a_1 \cdots a_{16}\ (i = 1, \ldots, 16)\,\rangle.$$

We have just described a 32–generator, 31–relator presentation of $G_{5/2}$. The corresponding presentation for $G_{p/q}$ has 2^{2q+1} generators and $2^{2q+1}-1$ relations.

2 Preliminaries

In the first part of this section we recall the basic definitions associated to Dehn functions. We then gather those elements of Perron–Frobenius theory that will be needed in the sequel.

Dehn functions

Given a finitely presented group $G = \langle \mathcal{A} \mid \mathcal{R} \rangle$ and a word w in the generators $\mathcal{A}^{\pm 1}$ that represents $1 \in G$, one defines

$$\mathrm{Area}(w) \;=\; \min\{N \in \mathbb{N} \mid \exists \text{ equality } w = \prod_{j=1}^{N} u_j r_j u_j^{-1} \text{ freely, where } r_j \in \mathcal{R}^{\pm 1}\}\,.$$

The *Dehn function* $\delta(x)$ of the finite presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ is defined by

$$\delta(x) \;=\; \max\{\mathrm{Area}(w) \mid w \in \ker(F(\mathcal{A}) \to G),\ |w| \leqslant x\}$$

where $|w|$ denotes the length of the word w. It is straightforward to show that the Dehn functions of any two finite presentations of the same group are equivalent in the following sense (and modulo this equivalence relation it therefore makes sense to talk of “the” Dehn function of a finitely presented group).

Given two functions $f, g\colon\ [0,\infty) \to [0,\infty)$ we define $f \preccurlyeq g$ if there exists a positive constant C such that

$$f(x) \leqslant C\, g(Cx) + Cx$$

for all $x \geqslant 0$. If $f \preccurlyeq g$ and $g \preccurlyeq f$ then f and g are said to be *equivalent*, denoted $f \simeq g$.

Remark 2.1 In order to establish the relation $f \preccurlyeq g$ between non-decreasing functions, it suffices to consider relatively sparse sequences of integers. For if (n_i) is an unbounded sequence of integers for which there is a constant $C > 0$ such that $n_0 = 0$ and $n_{i+1} \leqslant Cn_i$ for all i, and if $f(n_i) \leqslant g(n_i)$ for all i, then $f \preccurlyeq g$. Indeed, given $x \in [0,\infty)$ there is an index i such that $n_i \leqslant x \leqslant n_{i+1}$, whence $f(x) \leqslant f(n_{i+1}) \leqslant g(n_{i+1}) \leqslant g(Cn_i) \leqslant g(Cx)$.

We refer to Bridson [5] for general facts about Dehn functions, in particular the interpretation of $\mathrm{Area}(w)$ in terms of *van Kampen diagrams* over $\langle \mathcal{A} \mid \mathcal{R} \rangle$. Recall that a van Kampen diagram for w is a labeled, contractible, planar 2–complex with a basepoint and boundary label w. Associated to such a diagram D one has a cellular map $\tilde{D}$ from D to the universal cover $\tilde{K}$ of the standard 2–complex of $\langle \mathcal{A} \mid \mathcal{R} \rangle$, respecting labels and basepoint. The diagram is said to be *embedded* if this map in injective.

Remark 2.2 If the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ is aspherical and the diagram D is embedded, then D has the smallest area among all diagrams with the same boundary label. To see this, note that if $\tilde{\Delta}$ is a diagram with the same boundary circuit as $\tilde{D}$, then $\tilde{D} - \tilde{\Delta}$ defines a 2–cycle in $\tilde{K}$, which must be zero since $H_2(\tilde{K};\mathbb{Z}) = 0$ and there are no 3–cells. Thus each 2–cell in the image of $\tilde{D}$ must also occur in the image of $\tilde{\Delta}$. And since $\tilde{D}$ is an embedding, the number of 2–cells in the image (hence domain) of $\tilde{\Delta}$ is at least $\mathrm{Area}(D)$.

Higher-dimensional Dehn functions

Our treatment of higher-dimensional Dehn (isoperimetric) functions is similar to that of Bridson [6], which is an interpretation of the more algebraic treatment of Alonso *et al* [2]. See Section 5 of [6] for an explanation of the differences with the approaches of other authors, in particular Gromov [12], Epstein *et al* [10], and Hatcher–Vogtmann [13].

The k–dimensional Dehn function is a function $\delta^{(k)}\colon\ \mathbb{N} \to \mathbb{N}$ defined for any group G that is of type $\mathcal{F}_{k+1}$ (that is, has a $K(G,1)$ with finite $(k+1)$–skeleton). Up to equivalence, $\delta^{(k)}(x)$ is a quasi-isometry invariant. Roughly speaking, $\delta^{(k)}(x)$ measures the number of $(k+1)$–cells that one needs in order to fill any singular k–sphere in $K(G,1)$ comprised of at most x k–cells. The reader who is happy with this description can skip the technicalities in the remainder of this subsection. However, to be precise one has to be careful about the classes of maps that one considers and the way in which one counts cells. To this end, we make the following definitions.

If W is a compact k–dimensional manifold and X a CW complex, an *admissible map* is a continuous map $f\colon\ W \to X^{(k)} \subset X$ such that $f^{-1}(X^{(k)} - X^{(k-1)})$ is a disjoint union of open k–dimensional balls, each mapped by f homeomorphically onto a k–cell of X.

If $f\colon\ W \to X$ is admissible we define the *volume* of f, denoted $\mathrm{Vol}^k(f)$, to be the number of open k–balls in W mapping to k–cells of X. This notion is useful because of the abundance of admissible maps:

Lemma 2.3 *Let W be a compact manifold (smooth or piecewise-linear) of dimension k and let X be a CW complex. Then every continuous map $f\colon\ W \to X$ is homotopic to an admissible map. If $f(\partial W) \subset X^{(k-1)}$ then the homotopy may be taken rel ∂W.*

Proof We prove the lemma in the smooth case; analogous methods apply in the piecewise-linear category (cf the transversality theorem of Buoncristiano–Rourke–Sanderson [7]).

First arrange that $f(W) \subset X^{(k)}$ using cellular approximation. Next consider $X^{(k)} - X^{(k-1)}$ as a smooth manifold and perturb f to be smooth on the preimage of this open set. Let $C \subset X^{(k)}$ be a set consisting of one point in the interior of each k–cell of X. By Sard's theorem we can choose each point of C to be a regular value of f. The preimage $f^{-1}(C)$ is now a codimension k submanifold of W (ie a finite set of points) and f is a local diffeomorphism at each of these points, by the inverse function theorem. Thus there is a neighborhood V of C consisting of a small open ball around each point, whose preimage in W is a disjoint union of open balls, each mapping diffeomorphically to a component of V. Now modify f by composing it with a map of X (homotopic to the identity) that stretches each component of V across the k–cell containing it, and pushes its complement into $X^{(k-1)}$. The resulting map is admissible. □

Given a group G of type $\mathcal{F}_{k+1}$, fix an aspherical CW complex X with fundamental group G and finite $(k+1)$–skeleton. Let $\widetilde{X}$ be the universal cover of X. If $f\colon\ S^k \to \widetilde{X}$ is an admissible map, define the *filling volume* of f to be the minimal volume of an extension of f to B^{k+1}:

$$\mathrm{FVol}(f) \;=\; \min\{\, \mathrm{Vol}^{k+1}(g) \mid g\colon\ B^{k+1} \to \widetilde{X},\ g|_{\partial B^{k+1}} = f\,\}.$$

Note that the maps g must be admissible for volume to be defined. Such extensions exist by Lemma 2.3, since $\pi_k(\widetilde{X})$ is trivial. Next we define the *k–dimensional Dehn function* of X to be

$$\delta^{(k)}(x) \;=\; \sup\{\, \mathrm{FVol}(f) \mid f\colon\ S^k \to \widetilde{X},\ \mathrm{Vol}^k(f) \leqslant x\,\}.$$

Again, the maps f are assumed to be admissible. We will also write $\delta^{(k)}(x)$ as $\delta^{(k)}_G(x)$ (recall that G is the fundamental group of X).

Remarks 2.4 (1) In these definitions one could equally well use X in place of $\widetilde{X}$, since maps $S^k \to X$ (or $B^{k+1} \to X$) and their lifts to $\widetilde{X}$ have the same volume. There are reasons to prefer $\widetilde{X}$, however, as we shall see in the next definition below.

(2) It is not difficult to show that the Dehn function $\delta^{(k)}_G(x)$ agrees with the notion defined by Alonso *et al* in [2]. A discussion along these lines is given in Bridson [6, Section 5]. Moreover it is proved in [2] that, up to equivalence, $\delta^{(k)}_G(x)$ depends only on G (and in fact is a quasi-isometry invariant); hence we refer to it as "the" k–dimensional Dehn function of G. It is also proved in [2] that the supremum in the definition of $\delta^{(k)}_G(x)$ is attained.

More general Dehn functions

The definition of $\delta^{(k)}(x)$ generalizes in a natural way to give Dehn functions modeled on manifolds other than B^{k+1}. For example, Gromov has defined *genus g filling invariants* based on surfaces other than the disk [12]. Here we need to consider arbitrary compact manifolds.

Let $(M, \partial M)$ be a compact manifold pair (smooth or piecewise-linear) with $\dim M = k+1$. If $f\colon\ \partial M \to \widetilde{X}$ is an admissible map define

$$\textbf{(2.5)} \qquad \mathrm{FVol}^M(f) \ = \ \min\{\, \mathrm{Vol}^{k+1}(g) \mid g\colon\ M \to \widetilde{X},\ g|_{\partial M} = f\,\}$$

and

$$\delta^M(x) \ = \ \sup\{\, \mathrm{FVol}^M(f) \mid f\colon\ \partial M \to \widetilde{X},\ \mathrm{Vol}^k(f) \leqslant x\,\}.$$

The *dimension* of $\delta^M(x)$ is k, the dimension of ∂M (when $\partial M \neq \emptyset$). In general we do not assume that M is connected or that $\partial M \neq \emptyset$. Note that if M is closed then $\delta^M(x)$ is identically zero, since M may be mapped to a point, of zero volume. We will also use the notation $\delta^M_G(x)$ for $\delta^M(x)$.

Remarks 2.6 (1) In the definition of $\delta^M(x)$ it is important that we use maps into $\widetilde{X}$, which is contractible, since maps $f\colon\ \partial M \to X$ need not have extensions to M. Note that if $(M, \partial M) = (B^{k+1}, S^k)$ then the definitions of $\delta^M(x)$ and $\delta^{(k)}(x)$ agree.

(2) The omission of X from the notation and the adoption of the alternative notation $\delta^M_G(x)$ suggest an implicit claim that, as in the case $M = B^{k+1}$, the equivalence class of $\delta^M(x)$ depends only on G. We shall address this issue elsewhere, as it would take us too far afield in the context of the current paper. The structure of the arguments in Sections 7 and 8 requires us to work with specific choices of X anyway.

(3) Also to be addressed elsewhere is whether the supremum in the definition of $\delta^M(x)$ is attained. The main difficulty arises when M is 3–dimensional, as we shall explain in a moment. In the current paper this issue plays no role because none of the bounds that we establish require *a priori* finiteness.

(4) If $\dim M = k+1 \geqslant 4$ then $\delta^M(x) \leqslant \delta^{(k)}(x)$ provided ∂M is connected or $\delta^{(k)}(x)$ is superadditive. In particular, $\delta^M(x)$ is finite. The key point to observe here is that if $N = \partial M$ is connected and $f : N \to \widetilde{X}$ has volume V, then there is an admissible homotopy with $(k+1)$–dimensional volume at most $\delta^{(k)}(V)$ from f to an admissible map $f' : N \to \widetilde{X}$ whose image lies $\widetilde{X}^{(k-1)}$; one can then fill f' by a map $M \to X$ with zero $(k+1)$–dimensional volume.

To see that this homotopy exists, one considers a $(k-1)$–sphere S in N that encloses a ball D containing all of the open discs that contribute to the volume of f. The restriction of f to S is trivial in $H_{k-1}(\widetilde{X}^{(k-1)})$ and hence in $\pi_{k-1}(\widetilde{X}^{(k-1)})$ (recall that $\widetilde{X}^{(k-1)}$ is $(k-2)$–connected, and $k > 2$). The null-homotopy $H : B^k \to \widetilde{X}^{(k-1)}$ of $f|_S$ furnished by this observation can be adjoined to $f|_D$ to give an admissible map $S^k \to \widetilde{X}$ of volume V. This can then be filled by an admissible map $B^{k+1} \to \widetilde{X}$ of volume at most $\delta^{(k)}(V)$. The desired map f' is defined to be the adjunction of $f|_{N-D}$ and H.

If $\dim M = 2$ then the same statement holds; this is proved below in Lemma 7.4. The case $\dim M = 3$ is different: Young [23] has constructed a group G such that if M is a 3–manifold with boundary $S^1 \times S^1$, then $\delta^M(x)$ is strictly larger than $\delta^{(2)}(x)$.

Remark 2.7 An obvious adaptation of the argument in Remark 2.2 shows that if X is an aspherical $(k+1)$–dimensional CW complex, $g\colon M^{k+1} \to X$ is an embedding, and $f = g|_{\partial M}$ (with f and g admissible) then $\mathrm{FVol}^M(f) = \mathrm{Vol}^{k+1}(g)$. That is, the embedding g has minimal volume among all extensions of f to the manifold M. We shall use this fact in particular in the case of high-dimensional balls to estimate $\delta^{(k)}(x)$ from below.

Perron–Frobenius Theory

A square non-negative matrix P is said to be *irreducible* if for every i and j there exists $k \geqslant 1$ such that the ij–entry of P^k is positive. The basic properties of irreducible matrices are summarized in the Perron–Frobenius theorem below. See Seneta [19] and Katok–Hasselblatt [14] for a more thorough treatment of this theory and its applications.

Proposition 2.8 (Perron–Frobenius theorem) *Let P be an irreducible non-negative $R \times R$ matrix. Then P has one (up to a scalar) eigenvector with positive coordinates and no other eigenvectors with non-negative coordinates. Moreover, the corresponding eigenvalue λ is simple, positive, and is greater than or equal to the absolute value of all other eigenvalues. If m and M are the smallest and largest row sums of P, then $m \leqslant \lambda \leqslant M$, with equality on either side implying equality throughout.*

Lemma 2.9 *Let P be an irreducible non-negative $R \times R$ matrix with Perron–Frobenius eigenvalue λ. Let $\{v_1, \ldots, v_R\}$ be a generalized eigenbasis for P, with v_1 a positive eigenvector for λ, and with corresponding inner product $\langle \cdot\,, \cdot \rangle$. Then $\langle u, v_1 \rangle > 0$ for every non-negative vector $u \in \mathbb{R}^R - \{0\}$.*

Proof Decompose $\mathbb{R}^R$ as $W_1 \oplus \cdots \oplus W_k$ where each W_i is a generalized eigenspace for P, with $W_1 = \langle v_1 \rangle$. Each W_i is P–invariant, as is the non-negative orthant $\mathcal{N}$, since P is non-negative. The intersection $(W_2 \oplus \cdots \oplus W_k) \cap \mathcal{N}$ must then be trivial, for otherwise it contains an eigenvector for P other than v_1 (or a scalar multiple), by the Brouwer fixed point theorem. Hence $\langle u, v_1 \rangle \neq 0$ for every $u \in \mathcal{N} - \{0\}$. Since $\mathcal{N} - \{0\}$ is connected and contains v_1, $\langle u, v_1 \rangle$ is positive. □

Proposition 2.10 (Growth rate) *Let P be an irreducible non-negative $R \times R$ matrix with Perron–Frobenius eigenvalue λ. Let $\|\cdot\|$ be a norm on $\mathbb{R}^R$. Then there are positive constants A_0, A_1 such that for every non-negative vector u in $\mathbb{R}^R$ and every integer $k > 0$, $A_0 \lambda^k \|u\| \leqslant \left\|P^k u\right\| \leqslant A_1 \lambda^k \|u\|$.*

Proof First, it is clear that by varying the constants, it suffices to consider any single norm $\|\cdot\|$. Consider a generalized eigenbasis $\{v_1, \ldots, v_R\}$ as in Lemma 2.9 (with v_1 a positive eigenvector for λ). Let $\langle \cdot\,, \cdot \rangle$ and $\|\cdot\|$ be the corresponding inner product and norm on $\mathbb{R}^R$. Let $\pi\colon \mathbb{R}^R \to \langle v_1 \rangle$ be orthogonal projection ($\pi(u) = \langle u, v_1 \rangle v_1$).

Define $A_0 = \inf\{\|\pi(u)\| / \|u\| \mid u \in \mathcal{N} - \{0\}\}$. Note that $A_0 > 0$ by Lemma 2.9 and compactness of $\mathcal{N} - \{0\}$ modulo homothety. For every $u \in \mathcal{N} - \{0\}$ we now have $\lambda^k A_0 \|u\| \leqslant \lambda^k \|\pi(u)\| = \left\|P^k \pi(u)\right\| \leqslant \left\|P^k u\right\|$. We also have $\left\|P^k u\right\| \leqslant \lambda^k \|u\|$ since λ is the spectral radius of P; hence $A_1 = 1$ will work. □

3 The vertex groups V_m

In this section we define groups V_m for each integer $m \geqslant 2$. We begin with a very brief overview of the construction of the groups $G_{r,P}$ so that the reader knows where the groups V_m fit into the overall picture.

An irreducible matrix P determines a directed graph (whose transition matrix is P). This graph is the underlying graph in a graph of groups description of the $G_{r,P}$ in Theorem A. The vertex groups in this graph of groups are precisely the groups V_m which we define and study in this section.

The groups V_m satisfy a number of the properties that the free abelian groups $\mathbb{Z}^m$ do, but they have geometric dimension 2. In particular, V_m has generators $a_1, \ldots, a_m$ and has the following scaling property (cf equation (**3.2**)): for any integer $N > 0$, the equality $a_1^N \cdots a_m^N = (a_1 \cdots a_m)^N$ holds. Moreover, this equality requires on the order of N^2 relations of V_m. This follows as a special case of Lemma 3.5, which gives careful estimates on the areas of certain words in V_m.

The groups V_m

Begin with $m-1$ copies of $\mathbb{Z}\times\mathbb{Z}$, the ith copy having generators $\{a_i, b_i\}$. The group V_m is formed by successively amalgamating these groups along infinite cyclic subgroups by adding the relations

$$b_1 = a_2 b_2, \quad b_2 = a_3 b_3, \quad \ldots, \quad b_{m-2} = a_{m-1} b_{m-1}.$$

Thus V_m is the fundamental group of a graph of groups whose underlying graph is a segment having $m-2$ edges and $m-1$ vertices. We define two new elements: $c = a_1 b_1$ and $a_m = b_{m-1}$. Then $a_1, \ldots, a_m$ generate V_m and the relation $a_1 \cdots a_m = c$ holds; see Figure 2(a). The element c is called the *diagonal element* of V_m. The additional relations $b_{m-2} = a_{m-1}a_m, \ \ldots, \ b_{m-k} = a_{m-k+1}\cdots a_m$ are also evident from Figure 2(a).

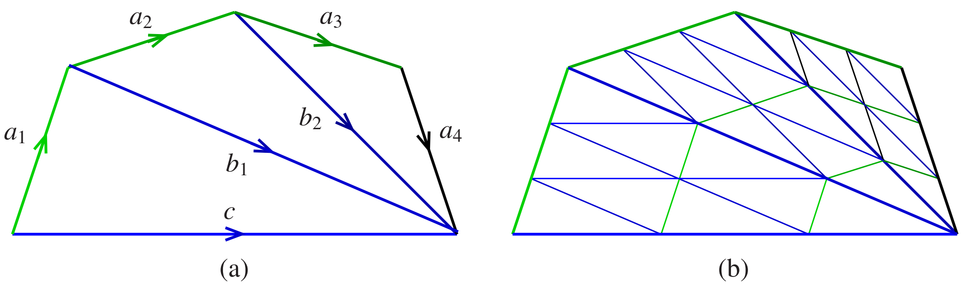

Figure 2: Some relations in V_4: $c = a_1a_2a_3a_4$ and $c^3 = (a_1)^3(a_2)^3(a_3)^3(a_4)^3$

If $m=1$ then we define V_m to be the infinite cyclic group $\langle a_1\rangle$ and we set $c = a_1$. Lemmas 3.1 and 3.5 below clearly hold in this case.

Lemma 3.1 (Shuffling Lemma) *Let $w = w(a_1, \ldots, a_m, c)$ be a word representing c^N in V_m for some integer N. Let n_i be the exponent sum of a_i in w, and n_c the exponent sum of c in w. Then the words $a_1^{n_1}\cdots a_m^{n_m}c^{n_c}$ and $c^{n_c}a_m^{n_m}\cdots a_1^{n_1}$ also represent c^N in V_m and $n_i = N - n_c$ for all i.*

Proof First we prove the second statement. The abelianization $V_m/[V_m, V_m] \cong \mathbb{Z}^m$ has $\{a_1, \ldots, a_m\}$ as a basis and the image of w is $a_1^{n_1+n_c}\cdots a_m^{n_m+n_c}$. Since c^N abelianizes to $a_1^N\cdots a_m^N$, we must have $n_i = N - n_c$ for all i.

To prove the first statement it now suffices to establish the following set of equalities for any integer N:

(3.2) $$(a_1\cdots a_m)^N \;=\; a_1^N\cdots a_m^N \;=\; a_m^N\cdots a_1^N \;=\; (a_m\cdots a_1)^N.$$

In fact we shall prove the following equalities, by induction on k:

$$(a_{m-k+1}\cdots a_m)^N \;=\; {a_{m-k+1}}^N\cdots a_m^N \;=\; a_m^N\cdots {a_{m-k+1}}^N \;=\; (a_m\cdots a_{m-k+1})^N.$$

The case $k = 1$ is evidently true. Suppose the equations hold for a given $k \geqslant 1$. By the induction hypothesis $a_{m-k}^N a_{m-k+1}^N\cdots a_m^N = a_{m-k}^N(a_{m-k+1}\cdots a_m)^N$. Then since $b_{m-k} = a_{m-k+1}\cdots a_m$ and this element commutes with a_{m-k}, we conclude that $a_{m-k}^N(a_{m-k+1}\cdots a_m)^N = (a_{m-k}\cdots a_m)^N$. The same commutation relation also yields

$$\begin{aligned} a_{m-k}^N(a_{m-k+1}\cdots a_m)^N \;&=\; (a_{m-k+1}\cdots a_m)^N a_{m-k}^N \\ &=\; (a_m\cdots a_{m-k+1})^N a_{m-k}^N \\ &=\; a_m^N\cdots a_{m-k+1}^N a_m^N. \end{aligned}$$

Finally we have $(a_m\cdots a_{m-k+1})^N a_{m-k}^N = (a_m\cdots a_{m-k+1}a_{m-k})^N$, again because a_{m-k} and b_{m-k} ($= a_m\cdots a_{m-k+1}$) commute. $\square$

Remark 3.3 (Scaling in V_m) Equation (**3.2**) plays a key role in this article. It shows that the basic relation shown in Figure 2(a) holds at larger scales as well. Figure 2(b) illustrates how these larger relations follow from the triangular relations $b_{i-1} = a_ib_i$ and $b_{i-1} = b_ia_i$.

The spaces X_m

To compute area in V_m we shall use a specific aspherical 2–complex X_m with fundamental group V_m. This complex is a union of $m-1$ tori, each triangulated with two 2–cells realizing the relations $a_ib_i = b_{i-1}$ and $b_ia_i = b_{i-1}$ (where $b_0 = c$ in the case $i = 1$). Thus the ith torus has standard generators given by the 1–cells a_i and b_i, and its diagonal is joined to the 1–cell b_{i-1} of the previous torus. In all there is one vertex, 1–cells $a_1, \ldots, a_{m-1}, b_0, \ldots, b_{m-1}$, and $2(m-1)$ triangular 2–cells.

The universal cover $\widetilde{X}_m$ is a union of planes, each covering one of the tori below. Each plane contains three families of parallel lines covering the 1–cells a_i, b_i, and b_{i-1}. The plane intersects neighboring planes along the b_j–lines for $j \neq 0, m-1$. These planes are the *vertex spaces* of $\widetilde{X}_m$ corresponding to the graph of groups decomposition of V_m described earlier. The incidence graph of the vertex spaces is the Bass–Serre tree for this decomposition, with edges corresponding to b_j–lines ($j \neq 0, m-1$).

Remark 3.4 Figure 2(b) shows an embedded disk in $\widetilde{X}_m$ with boundary word of the form $c^N = a_1^N\cdots a_m^N$ ($N = 3$). The triangles shown are 2–cells of $\widetilde{X}_m$. Each large triangular region lies in a vertex space of $\widetilde{X}_m$. There are similar embedded disks with boundary word $c^N = a_m^N\cdots a_1^N$ as well. All of these disks have area $(m-1)N^2$.

Throughout this article we usually work with the *standard generators* $\{a_1, \ldots, a_m\}$ for V_m. However in the area computation below we allow words involving the elements b_i as well.

Lemma 3.5 (Area in V_m) *Let $w(a_1, \ldots, a_{m-1}, b_0, \ldots, b_{m-1})$ be a word representing the element x^N for some N, where x is a generator a_i or b_i. Let w be expressed as $w_1 \cdots w_k$ where each w_i is a power of a generator. Then $N \leqslant |w|$ and $\mathrm{Area}(wx^{-N}) \leqslant 3\sum_{i<j} |w_i|\,|w_j|$.*

Note that if the sum included diagonal terms of the form $(3/2)\,|w_i|^2$ then the area bound would simply be $(3/2)\,|w|^2$. The leeway afforded by the absence of these terms will be exploited in the proof of Theorem A. (In particular, it would not suffice to know only that V_m has quadratic Dehn function.) Also the statement $N \leqslant |w|$ implies that every vertex space is a totally geodesic subspace of $\widetilde{X}_m$.

Proof First we prove that $N \leqslant |w|$ and then we establish the area bound. Both proofs are by induction on the complexity of the word w, defined as follows. Let p be a path in the 1–skeleton of $\widetilde{X}_m$ whose edge labels read w. Since w represents x^N, the endpoints of p lie in a single vertex space. Hence the induced path $\widehat{p}$ in the Bass–Serre tree is a closed path. The *complexity* of w is the length of $\widehat{p}$. Note that vertices of $\widehat{p}$ correspond to edges of p (or letters of w) and edges correspond to transitions between certain pairs of generators. Thus the complexity is also the number of such transitions occurring in w.

If w has complexity zero then p lies in a plane. The statement $N \leqslant |w|$ amounts to saying that x^N is a geodesic, which is clear. If $\widehat{p}$ has positive length then there is a non-trivial proper subpath $p' \subset p$ with endpoints on a single b_j–line. (These endpoints correspond to edges in $\widehat{p}$ that map to the same edge of the Bass–Serre tree, crossing and returning.) The subword $w' \subset w$ corresponding to p' represents an element of the form ${b_j}^M$. Let u be the word obtained from w by substituting ${b_j}^M$ for w'. Then u and w' both have complexity strictly smaller than that of w. By the induction hypothesis, $M \leqslant |w'|$ and $N \leqslant |u| = (|w| - |w'|) + M$. Therefore $N \leqslant |w|$.

Next we establish the area bound when w has complexity zero. Since p then lies entirely within a vertex space of $\widetilde{X}_m$, we may assume without loss of generality that $V_m = V_1$ and $x = b_0$, so that $w(a_1, b_0, b_1) = {b_0}^N$ in $V_1 = \langle a_1, b_1, b_0 \mid a_1 b_1 = b_0 = b_1 a_1 \rangle$. Since this group is abelian we can successively transpose adjacent subwords w_i and cancel pairs of the form xx^{-1}, to obtain $v = a_1^n b_1^n b_0^{N-n}$ for some n. Each transposition of letters contributes 2 to $\mathrm{Area}(wv^{-1})$, so we have $\mathrm{Area}(wv^{-1}) \leqslant$

$2\sum_{i<j}|w_i|\,|w_j|$. Next let I_a and I_b be the sets of indices for which w_i is a power of a_1 and b_1 respectively. Then $\sum_{i\in I_a}|w_i| \geqslant |n|$ and $\sum_{i\in I_b}|w_i| \geqslant |n|$, and therefore $\sum_{i<j}|w_i|\,|w_j| \geqslant n^2 = \mathrm{Area}(vb_0^{-N})$. Then we have $\mathrm{Area}(wb_0^{-N}) \leqslant \mathrm{Area}(wv^{-1}) + \mathrm{Area}(vb_0^{-N}) \leqslant 3\sum_{i<j}|w_i|\,|w_j|$ as desired.

Now suppose w has positive complexity. Define $w' \subset w$ and u as before, so that w' represents ${b_j}^M$, u is obtained from w by substituting ${b_j}^M$ for w', and both u and w' have smaller complexity than w. Note that $w' = w_{i_0}\cdots w_{i_1} \subset w_1\cdots w_k$ for some i_0 and i_1, and so $u = w_1\cdots w_{i_0-1}{b_j}^M w_{i_1+1}\cdots w_k$. Let $I = \{i_0,\ldots,i_1\}$. Applying the induction hypothesis to u and w' we obtain

$$\mathrm{Area}(ux^{-N}) \;\leqslant\; 3\sum_{\substack{i<j\\ i,j\notin I}}|w_i|\,|w_j| \;+\; 3\sum_{i\notin I}|w_i|\,M \tag{3.6}$$

and

$$\mathrm{Area}(w'{b_j}^{-M}) \;\leqslant\; 3\sum_{\substack{i<j\\ i,j\in I}}|w_i|\,|w_j|\,. \tag{3.7}$$

Since $M \leqslant |w'| = \sum_{j\in I}|w_j|$, inequality (3.6) becomes

$$\mathrm{Area}(ux^{-N}) \;\leqslant\; 3\sum_{\substack{i<j\\ i,j\notin I}}|w_i|\,|w_j| \;+\; 3\Big(\sum_{i\notin I}|w_i|\Big)\Big(\sum_{j\in I}|w_j|\Big). \tag{3.8}$$

Adding together (3.7) and (3.8) yields

$$\mathrm{Area}(w'{b_j}^{-M}) + \mathrm{Area}(ux^{-N}) \;\leqslant\; 3\sum_{i<j}|w_i|\,|w_j|$$

which proves the lemma because $\mathrm{Area}(wx^{-N}) \leqslant \mathrm{Area}(wu^{-1}) + \mathrm{Area}(ux^{-N})$ and $\mathrm{Area}(wu^{-1}) = \mathrm{Area}(w'{b_j}^{-M})$. □

4 The groups $G_{r,P}$ and snowflake words

The groups $G_{r,P}$

Start with a non-negative square integer matrix $P = (p_{ij})$ with R rows. Let m_i be the sum of the entries in the ith row and let $n = \sum_i m_i$, the sum of all entries. Form a directed graph Γ with vertices $\{v_1,\ldots,v_R\}$ and having p_{ij} directed edges from v_i to v_j. Label the edges as $\{e_1,\ldots,e_n\}$ and define two functions $\rho,\sigma\colon \{1,\ldots,n\} \to \{1,\ldots,R\}$ indicating the initial and terminal vertices of the edges, so that e_i is a directed edge from $v_{\rho(i)}$ to $v_{\sigma(i)}$ for each i. These functions also indicate the row and column of the matrix entry accounting for e_i. Partition the set $\{1, \dots, n\}$ as $\bigcup_i I_i$ by setting $I_i = \rho^{-1}(i)$. Note that $|I_i| = m_i$.

Let $M = \max\{m_i\}$ and choose a rational number $r = p/q$ with $p > Mq > 0$. We define a graph of groups $\mathcal{G}_{r,P}$ with underlying graph Γ as follows. The vertex group G_{v_i} at v_i will be V_{m_i} and all edge groups will be infinite cyclic. Relabel the standard generators of these vertex groups as $\{a_1, \dots, a_n\}$ in such a way that the standard generating set for G_{v_i} is $\{a_j \mid j \in I_i\}$. Let c_i be the diagonal element of the vertex group G_{v_i}. Then the inclusion maps are defined by mapping the generator of the infinite cyclic group G_{e_i} to the elements ${a_i}^p \in G_{v_{\rho(i)}}$ and ${c_{\sigma(i)}}^q \in G_{v_{\sigma(i)}}$.

Let s_i be the stable letter associated to the edge e_i. The fundamental group $G_{r,P}$ of $\mathcal{G}_{r,P}$ is obtained from the presentation

$$\langle\, G_{v_1}, \dots, G_{v_R}, s_1, \dots, s_n \mid s_i^{-1}{a_i}^p s_i = {c_{\sigma(i)}}^q \text{ for all } i \,\rangle$$

by adding relations $s_i = 1$ for each edge e_i in a maximal tree in Γ. However, we shall continue to use the generating set $\{a_1, \dots, a_n, s_1, \dots, s_n\}$ for $G_{r,P}$ even though some of these generators are trivial.

The spaces $X_{r,P}$

We define aspherical 2–complexes $X_{r,P}$ by forming graphs of spaces modeling $\mathcal{G}_{r,P}$. Namely, take the disjoint union of the spaces $X_{v_i} \approx X_{m_i}$ (one for each vertex v_i) and attach annuli A_i, one for each edge e_i of the graph. The two boundary curves of A_i are attached to the paths labeled ${a_i}^p$ in $X_{v_{\rho(i)}}$ and ${c_{\sigma(i)}}^q$ in $X_{v_{\sigma(i)}}$. The resulting 2–complex $X_{r,P}$ has fundamental group $G_{r,P}$ and it is aspherical because it is the total space of a graph of aspherical spaces.

The universal cover $\widetilde{X}_{r,P}$ is a union of copies of the universal covers $\widetilde{X}_{v_i}$ and infinite strips $\mathbb{R} \times [-1, 1]$ covering the annuli A_i. Each strip is tiled by 2–cells whose boundary labels read $s_i^{-1}{a_i}^p s_i {c_{\sigma(i)}}^{-q}$; the two sides $\mathbb{R} \times \{\pm 1\}$ consist of edges labeled a_i and $c_{\sigma(i)}$ respectively. Note that if a path crosses a strip along an edge labeled s_i and returns over s_i^{-1} then the power of a_i represented by the path is divisible by p.

Snowflake words

For each group element of the form c_i^N we will define two types of words in the generators $\{a_1, \dots, a_n, s_1, \dots, s_n\}$ representing that element, called positive and negative snowflake words. The structure of these words is governed by the dynamics of the matrix P. Some snowflake words are close to geodesics, and these are useful in determining the large scale geometry of $G_{r,P}$.

We define snowflake words recursively on $|N| \in \mathbb{N}$ as follows. Let

$$N_0 \;=\; \frac{p(M(q+2+p))}{p-Mq} + p.$$

Note for future reference that $N_0 > p$. Let c be the diagonal element of a vertex group with standard ordered generating set $\{a_{i_1}, \ldots, a_{i_m}\}$. A word w representing c^N is a *positive snowflake word* if either

(i) $|N| \leqslant N_0$ and $w = a_{i_1}^N \cdots a_{i_m}^N$, or

(ii) $|N| > N_0$ and $w = (s_{i_1} u_1 s_{i_1}^{-1})(a_{i_1}^{N_1}) \cdots (s_{i_m} u_m s_{i_m}^{-1})(a_{i_m}^{N_m})$ where each u_j is a positive snowflake word representing a power of $c_{\sigma(i_j)}$ and $|N_j| < p$ for all j.

In the second case note that each subword $(s_{i_j} u_j s_{i_j}^{-1})(a_{i_j}^{N_j})$ represents a power of a_{i_j}, and by Lemma 3.1 this power is N. Then since $|N_j| < p$, the word $(s_{i_j} u_j s_{i_j}^{-1})$ represents either ${a_{i_j}}^{\lfloor N/p \rfloor p}$ or ${a_{i_j}}^{\lceil N/p \rceil p}$. Consequently, the word u_j represents either ${c_{\sigma(i_j)}}^{\lfloor N/p \rfloor q}$ or ${c_{\sigma(i_j)}}^{\lceil N/p \rceil q}$. Recall that $\lfloor x \rfloor$ and $\lceil x \rceil$ denote the integers closest to x such that $\lfloor x \rfloor \leqslant x \leqslant \lceil x \rceil$, and so $\lfloor N/p \rfloor p$ and $\lceil N/p \rceil p$ are the multiples of p nearest to N.

A *negative snowflake word* is defined similarly, with the ordering of the terms representing powers of a_{i_j} reversed. More specifically, w satisfies either

(i′) $|N| \leqslant N_0$ and $w = a_{i_m}^N \cdots a_{i_1}^N$, or

(ii′) $|N| > N_0$ and $w = (a_{i_m}^{N_m})(s_{i_m} u_m s_{i_m}^{-1}) \cdots (a_{i_1}^{N_1})(s_{i_1} u_1 s_{i_1}^{-1})$ where u_j is a negative snowflake word representing a power of $c_{\sigma(i_j)}$ and $|N_j| < p$ for all j.

As with positive snowflake words, each word u_j will represent either ${c_{\sigma(i_j)}}^{\lfloor N/p \rfloor q}$ or ${c_{\sigma(i_j)}}^{\lceil N/p \rceil q}$.

To see that the recursion is well-founded note that the definition describes an iterated curve shortening process in which subwords of the form c^N are replaced by the words described in case (ii) or (ii′), with appropriate powers of $c_{\sigma(i_j)}$ in place of u_j; see Figure 3. Writing $|N| = Ap + B$ with $0 \leqslant B < p$, the new word representing c^N has length at most

$$M \cdot \max\{Aq + 2 + B, (A+1)q + 2 + (p - B)\} \;\leqslant\; M((A+1)q + 2 + p).$$

The latter quantity is stictly less than $|N| = Ap + B$ provided $A(p - Mq) > M(q+2+p)$. Since $B < p$, this occurs whenever $|N| \geqslant N_0$. Thus, the new curve is strictly shorter than c^N if $|N| \geqslant N_0$. Eventually the subwords c^N all have length at most N_0 and the

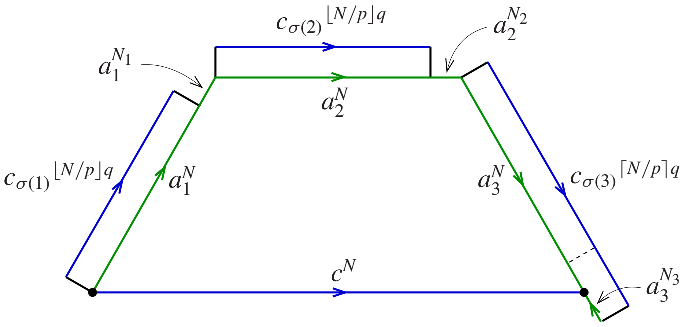

Figure 3: One way of shortening c^N. Here $\{a_1, a_2, a_3\}$ is the generating set for a vertex group V_3 with diagonal element c. The exponents N_1 and N_2 are both $N - \lfloor N/p \rfloor p$ and N_3 is $N - \lceil N/p \rceil p$. The short black edges are labeled s_1, s_2, s_3.

shortening procedure terminates. See also Figure 4 for the end result of this process. In this figure the top and bottom halves of the boundary are positive and negative snowflake words representing c^N.

Note that every snowflake word has a nested structure in which various subwords are themselves snowflake words. These are the subwords u_j arising at each stage. The minimal such subwords are those given by (i) and (i$'$) and these will be called *terminal subwords*. The *depth* of a snowflake subword is the number of snowflake subwords of type (ii) or (ii$'$) properly containing it, including the original snowflake word itself. Equivalently, it is the number of matching s_j, s_j^{-1} pairs enclosing it. Note that a snowflake word w contains a depth zero terminal subword if and only if w has the form (i) or (i$'$).

It is worth emphasizing that the curve shortening process is not canonically determined, but allows many choices. In each "remainder" term $a_i^{N_i}$ the exponent N_i may be positive or negative; the two possible values for N_i are $N - \lfloor N/p \rfloor p$ and $N - \lceil N/p \rceil p$. Figure 3 shows both possibilities occurring in a single step, for example. For this reason, a single snowflake word may have terminal subwords of different depths. However, Lemma 4.2 below shows that these depths will not differ substantially.

Remark 4.1 A special type of snowflake word plays a key role in the proof of Theorem C. If r is an integer (that is, $r = p/1$) and $N = r^k$ for some k, then the positive (resp. negative) snowflake word representing c_i^N is unique. What happens is that the exponents N_j in the expressions (ii) or (ii$'$) at each stage are always zero; there are no "remainder" terms $a_{i_j}^{N_j}$. Each subword u_j represents $c_{\sigma(i_j)}{}^{N/r}$, and N/r is again a power of r. Furthermore, all terminal subwords will have the form $a_{i_1} \cdots a_{i_m}$ or $a_{i_m} \cdots a_{i_1}$.

Lemma 4.2 (Terminal subword depth) *Given r and P there are positive constants B_0, B_1 with the following property. If a non-trivial snowflake word w representing c^N contains a terminal subword of depth d then $B_0 r^d \leqslant |N| \leqslant B_1 r^d$.*

Proof If $d = 0$ then w has the form (i) or (i$'$) and $1 \leqslant |N| \leqslant N_0$. Thus we need to arrange that $B_0 \leqslant 1$ and $B_1 \geqslant N_0$ for the lemma to hold in this case.

If $d > 0$ then we will show by induction on d that

(4.3) $$N_0 r^{d-1} - p(r^{d-2} + \cdots + r + 1) \;\leqslant\; |N| \;\leqslant\; N_0 r^d + p(r^{d-1} + \cdots + r + 1).$$

The lower bound then gives

$$|N| \;\geqslant\; N_0 r^{d-1} - p\left(\frac{r^{d-1} - 1}{r - 1}\right) \;\geqslant\; \frac{1}{r}\left(N_0 - \frac{p}{r-1}\right) r^d.$$

Recall that $N_0 > p$ and $r \geqslant 2$, which imply $N_0 > p/(r-1)$. Now we may find $B_0 > 0$ so that $B_0 \leqslant r^{-1}(N_0 - p/(r-1))$ and $B_0 \leqslant 1$, giving the desired bound.

The upper bound in (4.3) gives

$$|N| \;\leqslant\; N_0 r^d + p\left(\frac{r^d - 1}{r - 1}\right) \;\leqslant\; (N_0 + p) r^d$$

where the last inequality uses the fact that $r - 1 \geqslant 1$. Now choose $B_1 \geqslant N_0 + p$ to obtain the desired bound.

Next we prove (4.3) by induction on d. If $d = 1$ then $|N| > N_0$ and w is of the form (ii) or (ii$'$) where some u_j has the form (i) or (i$'$). Then u_j represents ${c_{\sigma(i_j)}}^{N'}$ with $N' \leqslant N_0$, and so $(s_{i_j} u_j s_{i_j}^{-1})$ represents ${a_{i_j}}^{rN'}$. This implies $|N| = |rN' + N_j| \leqslant rN_0 + p$.

For $d > 1$ write w in the form (ii) or (ii$'$). Then the terminal subword has depth $d - 1$ in u_j for some j. By the induction hypothesis u_j represents ${c_{\sigma(i_j)}}^{N'}$ where

(4.4) $$N_0 r^{d-2} - p(r^{d-3} + \cdots + 1) \;\leqslant\; \left|N'\right| \;\leqslant\; N_0 r^{d-1} + p(r^{d-2} + \cdots + 1).$$

Then $(s_{i_j} u_j s_{i_j}^{-1})$ represents ${a_{i_j}}^{rN'}$ and $rN' - p \leqslant |N| \leqslant rN' + p$. These bounds and (4.4) together imply (4.3). □

Proposition 4.5 (Snowflake word length) *Given r and P there are positive constants C_0, C_1 with the following property. If c is the diagonal element of one of the vertex groups and w is a snowflake word representing c^N then $C_0 |w|^\alpha \leqslant |N| \leqslant C_1 |w|^\alpha$, where $\alpha = \log_\lambda(r)$ and λ is the Perron–Frobenius eigenvalue of P.*

Proof If w is non-trivial and has the form (i) or (i$'$) then $1 \leqslant |N| \leqslant N_0$ and $|N| \leqslant |w| \leqslant r\,|N|$. Then $|w|^\alpha \leqslant (rN_0)^\alpha$, which implies

$$(rN_0)^{-\alpha}\,|w|^\alpha \;\leqslant\; |N| \;\leqslant\; |w|^\alpha\,.$$

Thus we need to arrange that $C_0 \leqslant (rN_0)^{-\alpha}$ and $C_1 \geqslant 1$ to cover this case.

Next assume that w is of type (ii) or (ii$'$), which implies that the depth of every terminal subword is at least one. Equivalently, w contains the letters s_j, s_j^{-1} for some j. Let $s(w)$ be the number of letters s_j or s_j^{-1} in w (for all indices j). Note that a subword of w containing no such letters has length at most rN_0. Since $s(w) \neq 0$, this implies

(4.6) $$s(w) \;\leqslant\; |w| \;\leqslant\; 2(rN_0+1)s(w).$$

Hence $s(w)$ gives an approximate measure of the length of w. It can be computed explicitly, by following the evolution of the curve shortening process, which in turn is governed by the matrix P. Note that matched s, s^{-1} pairs enclose snowflake subwords representing powers of c_j for various j. These subwords will be called *c_j–subwords*.

We claim that if w represents a power of c_k, and every terminal subword has depth i or greater, then the number of c_j–subwords of depth i is given by the kj–entry of P^i, denoted $p_{kj}^{(i)}$.

If $i = 1$ then the claim is evident from expressions (ii) and (ii$'$), since the entry p_{kj} of P gives the number of directed edges from vertex v_k to vertex v_j (and hence the number of occurrences of j among the indices $\sigma(i_1), \ldots, \sigma(i_m)$). Similarly, for $i > 1$, each c_ℓ–subword of depth $i-1$ contains $p_{\ell j}$ c_j–subwords of depth i, by (ii) and (ii$'$). The claim now follows by induction on i: summing over all snowflake subwords of depth $i-1$ and applying the induction hypothesis, the total number of c_j–subwords of depth i is $\sum_\ell p_{k\ell}^{(i-1)} p_{\ell j} = p_{kj}^{(i)}$.

Let $x_1, \ldots, x_R$ be the standard basis vectors of $\mathbb{R}^R$. Also let $\|\cdot\|_1$ denote the ℓ_1 norm on $\mathbb{R}^R$: $\|v\|_1$ is the sum of the absolute values of the entries of the vector v. Let P^T be the transpose of P.

The kj–entry of P^i is equal to the j–entry of the column vector $(P^T)^i(x_k)$. Suppose for the moment that every terminal subword of w has depth d. Then for $i \leqslant d$, the total number of s, s^{-1} pairs enclosing snowflake subwords of depth i is given by $\left\|(P^T)^i(x_k)\right\|_1$. Hence we have

$$s(w) \;=\; 2\left(\left\|P^T(x_k)\right\|_1 + \left\|(P^T)^2(x_k)\right\|_1 + \cdots + \left\|(P^T)^d(x_k)\right\|_1\right).$$

If we let d_0 and d_1 denote the smallest and largest depths of terminal subwords of w then we obtain

$$2\sum_{i=1}^{d_0} \left\|(P^T)^i(x_k)\right\|_1 \;\leqslant\; s(w) \;\leqslant\; 2\sum_{i=1}^{d_1} \left\|(P^T)^i(x_k)\right\|_1 .$$

Applying Proposition 2.10 with the norm $\|\cdot\|_1$ we have

$$2A_0\sum_{i=1}^{d_0}\lambda^i \;\leqslant\; s(w) \;\leqslant\; 2A_1\sum_{i=1}^{d_1}\lambda^i \;=\; \frac{2A_1\lambda}{\lambda-1}(\lambda^{d_1}-1)$$

which implies

$$2A_0\lambda^{d_0} \;\leqslant\; s(w) \;\leqslant\; \frac{2A_1\lambda}{\lambda-1}\lambda^{d_1}.$$

Hence by (**4.6**) we have

$$(2A_0)\lambda^{d_0} \;\leqslant\; |w| \;\leqslant\; \left(\frac{4(rN_0+1)A_1\lambda}{\lambda-1}\right)\lambda^{d_1}. \tag{4.7}$$

We complete the proof by applying Lemma 4.2 separately for the upper and lower bounds. Using $d=d_1$ we obtain

$$|N| \;\geqslant\; B_0r^{d_1} \;=\; B_0(\lambda^{d_1})^{\log_\lambda(r)} \;\geqslant\; B_0\left(\frac{4(rN_0+1)A_1\lambda}{\lambda-1}\right)^{-\log_\lambda(r)}|w|^{\log_\lambda(r)}.$$

Now choose $C_0>0$ satisfying $C_0\leqslant B_0\left(\frac{4(rN_0+1)A_1\lambda}{\lambda-1}\right)^{-\alpha}$ and $C_0\leqslant (rN_0)^{-\alpha}$ to obtain the desired lower bound.

Applying Lemma 4.2 with $d=d_0$ gives

$$|N| \;\leqslant\; B_1r^{d_0} \;=\; B_1(\lambda^{d_0})^{\log_\lambda(r)} \;\leqslant\; B_1(2A_0)^{-\log_\lambda(r)}|w|^{\log_\lambda(r)}$$

so choose C_1 with $C_1\geqslant B_1(2A_0)^{-\alpha}$ and $C_1\geqslant 1$. □

5 Proof of Theorem A

Throughout this section $G_{r,P}$ is fixed, with $r=p/q$ greater than all the row sums of P, and $\alpha=\log_\lambda(r)$, where λ is the Perron–Frobenius eigenvalue of P. Unless otherwise stated, all words use the generating set $\{a_1,\ldots,a_n,s_1,\ldots,s_n\}$ for $G_{r.P}$.

The lower bound

To establish the lower bound $\delta(x) \succcurlyeq x^{2\alpha}$ we will show that $\delta(n_i) \geqslant (C_0{}^2 4^{-\alpha})\, n_i{}^{2\alpha}$ for certain integers n_i tending to infinity. This is sufficient by Remark 2.1, provided the sequence (n_i) grows at most exponentially.

Note also that to establish a single inequality $\delta(n) \geqslant A$, it is enough to exhibit an embedded disk in $\widetilde{X}_{r,P}$ with boundary length n and area A or greater, by Remark 2.2. Here we are using the facts that $X_{r,P}$ is aspherical and 2–dimensional.

Choose a vertex group V_m in $G_{r,P}$ with $m \geqslant 2$ and let c be its diagonal element. There must be at least one vertex group of this type, for otherwise P would be a permutation matrix with Perron–Frobenius eigenvalue 1. For each i choose positive and negative snowflake words w_i^+ and w_i^- representing c^i. Then define $w_i = w_i^+(w_i^-)^{-1}$ and $n_i = |w_i|$. Note that $C_0\, 2^{-\alpha}\, |w_i|^\alpha \leqslant i \leqslant C_1\, 2^{-\alpha}\, |w_i|^\alpha$ by Proposition 4.5. It follows that the sequence (n_i) tends to infinity, and that it is exponentially bounded:

$$\frac{n_{i+1}}{n_i} \leqslant \left(\frac{(i+1)C_1}{iC_0}\right)^{1/\alpha} \leqslant \left(\frac{2C_1}{C_0}\right)^{1/\alpha}$$

for $i \geqslant 1$.

Next we find embedded disks Δ_i in $\widetilde{X}_{r,P}$ with boundary words w_i and estimate their areas. Each Δ_i is made of two disks Δ_i^+ and Δ_i^- with boundary words $w_i^+ c^{-i}$ and $c^i(w_i^-)^{-1}$ respectively, joined along the boundary arcs labeled c^{-i}, c^i. After joining, the arc labeled c^i will be called the *diameter* of Δ_i.

The disk $\Delta_i^\pm$ is a union of embedded disks in vertex spaces $\widetilde{X}_{m_i}$ and pieces of strips joining them. Consider the curve shortening process that transforms c^i into $w_i^\pm$. To build $\Delta_i^\pm$ simply fill the central region shown in Figure 3 with the embedded disk from Figure 2(b). Then fill each strip with either $\lfloor i/p \rfloor$ or $\lceil i/p \rceil$ copies of the 2–cell with the appropriate boundary word $s_j c_{\sigma(j)}{}^q s_j^{-1} a_j^{-p}$, and repeat the procedure. The resulting disk is a union of embedded disks in $\widetilde{X}_{r,P}$ joined along boundary arcs, with no folding along these arcs. Since each strip separates $\widetilde{X}_{r,P}$, one can see inductively (on the number of strips crossed by $\Delta_i^\pm$) that $\Delta_i^\pm$ is embedded. For the same reason, it suffices to note that no folding occurs when Δ_i^+ and Δ_i^- are joined together to conclude that Δ_i is embedded. Figure 4 shows an example of a disk Δ_i with boundary word w_i.

To estimate the area of Δ_i consider the central region in Δ_i^+ adjacent to Δ_i^-. By Remark 3.4 this subdisk of Δ_i has area $(m-1)i^2 \geqslant i^2$. Then since $i \geqslant C_0\, 2^{-\alpha} n_i{}^\alpha$ (as observed above) we conclude that

(5.1) $$\mathrm{Area}(\Delta_i) \geqslant (C_0{}^2 4^{-\alpha}) n_i{}^{2\alpha}$$

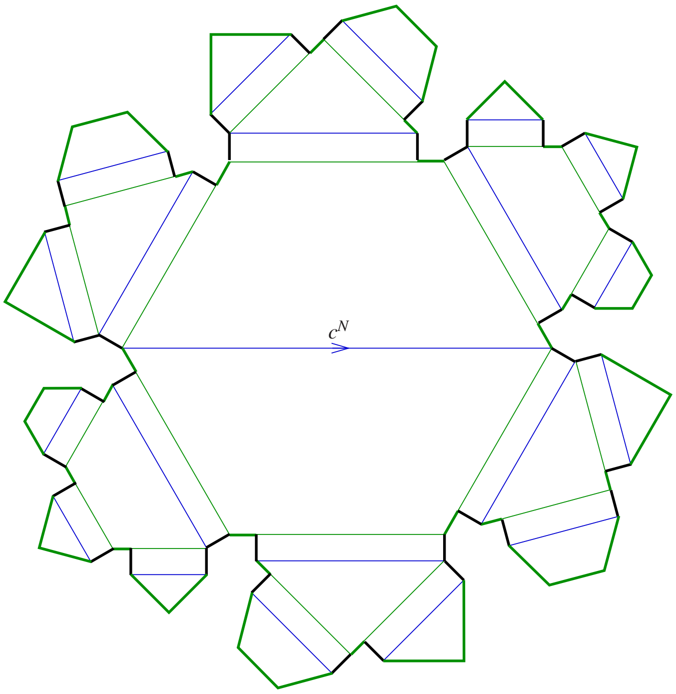

Figure 4: A snowflake disk based on the matrix $P = \left(\begin{smallmatrix} 1 & 1 \\ 2 & 1 \end{smallmatrix}\right)$. The upper and lower halves of the boundary curve are positive and negative snowflake paths representing c^N.

and therefore $\delta(n_i) \geqslant ({C_0}^2 4^{-\alpha}){n_i}^{2\alpha}$.

The upper bound

Suppose a word w represents an element of a vertex group V_m. The graph of groups structure of $G_{r,P}$ yields a decomposition of w as $w_1 \cdots w_k$ where each w_i is either an element of V_m, or begins with $s_j^{\pm}$ and ends with $s_j^{\mp}$ for some j. These latter cases occur when the path described by w leaves the vertex space $\widetilde{X}_m$ and then returns again over a strip in $\widetilde{X}_{r,P}$.

Recall that a strip in $\widetilde{X}_{r,P}$ has sides labeled a_i and $c_{\sigma(i)}$. The next lemma shows that a geodesic (in the generators $\{a_1, \ldots, a_n, s_1, \ldots, s_n\}$) can only enter a strip from (and return to) the a_i–side.

Lemma 5.2 *Let w be a geodesic in $G_{r,P}$ representing an element of a vertex group V_m. Then w is a product of subwords $w_1 \cdots w_k$ where each w_i is a power of a generator*

a_j, or begins with s_j and ends with s_j^{-1} (for some j) and represents a power of a_j.

Proof Let $w' \subset w$ be an innermost word that begins with s_ℓ^{-1} and ends with s_ℓ (for some ℓ) and whose corresponding path in $\widetilde{X}_{r,P}$ has endpoints in the same vertex space $\widetilde{X}_{v_{\sigma(\ell)}}$. Thus $w' = s_\ell^{-1} u s_\ell$ crosses a strip from the $c_{\sigma(\ell)}$–side, and the subword u only crosses strips from (and returns to) a_i–sides. That is, u can be written as $u_1 \cdots u_k$ where each u_i is a power of a generator a_j, or begins with s_j and ends with s_j^{-1} and represents a power of a_j.

Note that u has both endpoints on an a_ℓ–line in the vertex space $\widetilde{X}_{v_{\rho(\ell)}}$ across a strip from $\widetilde{X}_{v_{\sigma(\ell)}}$. Hence u represents a_ℓ^N for some N. Let u' be the word in the standard generators of $G_{v_{\rho(\ell)}} \cong V_m$ obtained by replacing each u_i by the appropriate power of a_j that it represents. Consider the word $u' a_\ell^{-N}$ which represents the trivial element c^0 in V_m. Since u' does not involve c, Lemma 3.1 implies that every a_j–exponent of $u' a_\ell^{-N}$ is zero. Hence u' has a_ℓ–exponent N and a_j–exponent zero for every $j \neq \ell$.

If any of the subwords u_i of u represent a power of a_j with $j \neq \ell$, then by Lemma 3.1 one could rearrange the subwords (preserving the property that u represents a_ℓ^N) so that those representing powers of a_j are adjacent. Then these adjacent subwords cancel in V_m and can be deleted, shortening w. Therefore every u_i represents a power of a_ℓ.

If none of the subwords u_i begins with s_ℓ and ends with s_ℓ^{-1} then $u = a_\ell^N$, but then w' could be replaced by a word $a_{i_1}^{N/r} \cdots a_{i_m}^{N/r}$ representing $c_{\sigma(\ell)}{}^{N/r}$. The new word is shorter than w because of the hypothesis that $m < r$, and therefore some u_i must have the form $s_\ell v s_\ell^{-1}$ after all. Now rearrange the subwords so that $s_\ell v s_\ell^{-1}$ occurs last. Again w can be shortened by replacing u with this rearranged word and then cancelling $s_\ell^{-1} s_\ell$ at the end. □

Proposition 5.3 *Let c be the diagonal element of one of the vertex groups in $G_{r,P}$. Then for every N there is a snowflake word w_{sf} and a geodesic w_{geo}, both representing c^N, with $|w_{sf}| \leqslant rN_0 \left|w_{geo}\right|$.*

Proof The proof is by induction on $|N|$. Let w be a geodesic representing c^N. We shall apply Lemma 3.1 inductively to rearrange and modify w into two words, a geodesic w_{geo} and a positive snowflake word w_{sf}. The two constructions are identical except at the base of the induction, which involves only certain segments of length at most rN_0.

Let $a_{i_1}, \ldots, a_{i_m}$ be the standard generators (in order) of the vertex group V_m containing c. If $|N| \leqslant N_0$ then define $w_{\mathrm{geo}} = w$ and $w_{\mathrm{sf}} = a_{i_1}^N \cdots a_{i_m}^N$. The desired conclusion holds in this case since $r > m$.

Suppose next that $|N| > N_0$. By Lemma 5.2 we can write w as $w_1 \cdots w_k$ where each subword has the form $a_j^{N_j}$ or $s_j u_j s_j^{-1}$. In the latter case $s_j u_j s_j^{-1}$ represents a power of a_j.

By Lemma 3.1 we can permute the subwords w_ℓ of w to arrange that those representing powers of a_{i_1} come first, those representing powers of a_{i_2} occur next, and so on. The resulting word is still a geodesic representing c^N. Note that two subwords cannot both be of the form $s_{i_j} u_j s_{i_j}^{-1}$ since they could be made adjacent, and then a cancellation of $s_{i_j}^{-1} s_{i_j}$ would be possible. Hence we can arrange for w to have the form

$$w = (s_{i_1} u_1 s_{i_1}^{-1})(a_{i_1}^{N_1})(s_{i_2} u_2 s_{i_2}^{-1})(a_{i_2}^{N_2}) \cdots (s_{i_m} u_m s_{i_m}^{-1})(a_{i_m}^{N_m}) \tag{5.4}$$

where each $s_{i_j} u_j s_{i_j}^{-1}$ represents a power of a_{i_j}. Next observe that $|N_j| < p$ for all j, since otherwise a subword of the form $s_{i_j}^{-1} a_{i_j}^{\pm p}$ could be replaced by a word of the form $a_{\ell_1}^{\pm q} \cdots a_{\ell_{m'}}^{\pm q} s_{i_j}^{-1}$ (that is, $c_{\sigma(i_j)}{}^{\pm q} s_{i_j}^{-1}$ expressed in the standard generators). Here m' is a row sum of P and so $r > m'$, making the new word shorter than w.

Recall that u_j represents a power of $c_{\sigma(j)}$. By Lemma 3.1 the power of a_{i_j} represented by $s_{i_j} u_j s_{i_j}^{-1}$ is $N - N_j$, and so u_j represents $c_{\sigma(j)}{}^{(N-N_j)/r}$. Recall that $N_0 > p$, hence $|N| > p > |N_j|$. Then since $r > 2$ it follows that $|(N - N_j)/r| < |N|$.

By induction $c_{\sigma(j)}{}^{(N-N_j)/r}$ is represented by a geodesic $(u_j)_{\mathrm{geo}}$ and a positive snowflake word $(u_j)_{\mathrm{sf}}$ satisfying the conclusion of the lemma. Define w_{geo} and w_{sf} by replacing each subword u_j in (**5.4**) by $(u_j)_{\mathrm{geo}}$ or $(u_j)_{\mathrm{sf}}$ accordingly. Then the desired conclusion also holds for w_{geo} and w_{sf}, since they agree except in the subwords $(u_j)_{\mathrm{geo}}$ and $(u_j)_{\mathrm{sf}}$. □

Corollary 5.5 (Edge group distortion) *Given r and P there is a positive constant D with the following property. If c is a diagonal element and w is a word representing c^N then $|N| \leqslant D\,|w|^\alpha$.*

Proof It suffices to consider the case when w is a geodesic. Apply Proposition 5.3 to obtain the geodesic w_{geo} and snowflake word w_{sf} representing c^N with $|w_{\mathrm{sf}}| \leqslant rN_0\,|w_{\mathrm{geo}}|$. Then Proposition 4.5 implies $|N| \leqslant C_1\,|w_{\mathrm{sf}}|^\alpha \leqslant C_1\,(rN_0)^\alpha\,|w_{\mathrm{geo}}|^\alpha$. □

The statement and proof of the next proposition are similar to those of Brady–Bridson [3, Proposition 3.2]. The case $N = 0$ establishes the upper bound of Theorem A.

Proposition 5.6 (Area bound) *Given r and P there is a positive constant E with the following property. If w is a word in $G_{r,P}$ representing x^N for some N, where x is either a generator a_i or the diagonal element of one of the vertex groups, then* $\mathrm{Area}(wx^{-N}) \leqslant E\,|w|^{2\alpha}$.

Proof We argue by induction on $|w|$. We shall prove the statement with $E = (3/2)r^2D^2$ (D given by Corollary 5.5). Let c denote the diagonal element of the vertex group V_m containing x.

Write w as $w_1 \cdots w_k$ where each w_i has the form $a_{j_i}^{N_i}$ or is a word beginning in $s_{j_i}^{\pm 1}$ and ending in $s_{j_i}^{\mp 1}$. In the latter cases w_i represents an element of the form c^{N_i} or $a_{j_i}^{N_i}$. Let I_c and I_a be the sets of indices for which these two cases occur, and let w' be the word obtained from w by replacing each subword w_i of this type with the appropriate word c^{N_i} or $a_{j_i}^{N_i}$. Then w' is a word in the standard generators of V_m (and the diagonal element) representing x^N, of length $\sum_i N_i$.

By Lemma 3.5 we have $\text{Area}(w'x^{-N}) \leqslant 3\sum_{i<j} N_i N_j$. To estimate each N_i we use Corollary 5.5 as follows. If $i \in I_c$ then w_i represents c^{N_i} and Corollary 5.5 gives $N_i \leqslant D\,|w_i|^{\alpha}$. If $i \in I_a$ then $w_i = s_{j_i} u_i s_{j_i}^{-1}$ for some u_i representing ${c_{\sigma(j_i)}}^{N_i/r}$ (because w_i represents $a_{j_i}^{N_i}$). Then by Corollary 5.5 we have $N_i/r \leqslant D(|w_i|-2)^{\alpha} \leqslant D\,|w_i|^{\alpha}$, so $N_i \leqslant rD\,|w_i|^{\alpha}$. Finally if $i \notin (I_c \cup I_a)$ then $N_i = |w_i| \leqslant |w_i|^{\alpha}$. Putting these observations together we have

$$\text{Area}(w'x^{-N}) \;\leqslant\; 3r^2D^2 \sum_{i<j} |w_i|^{\alpha}\,|w_j|^{\alpha}\,. \tag{5.7}$$

Next we use the induction hypothesis and Corollary 5.5 to bound $\text{Area}(ww'^{-1})$. First note that $\text{Area}(ww'^{-1}) \leqslant \sum_{i\in I_c} \text{Area}(w_i c^{-N_i}) + \sum_{i \in I_a} \text{Area}(w_i a_{j_i}^{-N_i})$.

If $i \in I_c$ then $w_i = s_{j_i}^{-1} u_i s_{j_i}$ where u_i represents ${a_{j_i}}^{rN_i}$. Applying the induction hypothesis to u_i we have $\text{Area}(u_i {a_{j_i}}^{-rN_i}) \leqslant (3/2)r^2D^2(|w_i|-2)^{2\alpha}$. The strip $s_{j_i}^{-1} {a_{j_i}}^{rN_i} s_{j_i} c^{-N_i}$ has area $N_i/q \leqslant (D/q)\,|w_i|^{\alpha} \leqslant D\,|w_i|^{\alpha}$, by Corollary 5.5. Thus

$$\begin{aligned} \text{Area}(w_i c^{-N_i}) &\leqslant (3/2)r^2D^2(|w_i|-2)^{2\alpha} + D\,|w_i|^{\alpha} \\ &\leqslant (3/2)r^2D^2((|w_i|-2)^{2\alpha} + |w_i|^{\alpha}) \\ &\leqslant (3/2)r^2D^2\,|w_i|^{2\alpha}\,. \end{aligned} \tag{5.8}$$

The last inequality above uses the fact that for numbers $x \geqslant 0$ one has $(x+2)^{2\alpha} \geqslant x^{\alpha}(x+2)^{\alpha} + 2^{\alpha}(x+2)^{\alpha} \geqslant x^{2\alpha} + (x+2)^{\alpha}$.

If $i \in I_a$ then $w_i = s_{j_i} u_i s_{j_i}^{-1}$ where u_i represents ${c_{\sigma(j_i)}}^{N_i/r}$. Applying the induction hypothesis to u_i we have $\text{Area}(u_i {c_{j_i}}^{-N_i/r}) \leqslant (3/2)r^2D^2(|w_i|-2)^{2\alpha}$. The strip $s_{j_i} {c_{j_i}}^{N_i/r} s_{j_i}^{-1} a_{j_i}^{-N_i}$ has area $(N_i/r)/q \leqslant (D/q)(|w_i|-2)^{\alpha} \leqslant D(|w_i|-2)^{\alpha}$, by Corollary 5.5. Therefore

$$\begin{aligned} \text{Area}(w_i a_{j_i}^{-N_i}) &\leqslant (3/2)r^2D^2(|w_i|-2)^{2\alpha} + D(|w_i|-2)^{\alpha} \\ &\leqslant (3/2)r^2D^2((|w_i|-2)^{2\alpha} + (|w_i|-2)^{\alpha}) \\ &\leqslant (3/2)r^2D^2\,|w_i|^{2\alpha}\,. \end{aligned} \tag{5.9}$$

Combining (**5.8**) and (**5.9**) we then have

$$\textbf{(5.10)}\qquad \mathrm{Area}(ww'^{-1}) \leqslant \sum_{i\in I_c\cup I_a} (3/2)r^2D^2\,|w_i|^{2\alpha} \leqslant \sum_i (3/2)r^2D^2\,|w_i|^{2\alpha}\,.$$

Finally, adding (**5.7**) and (**5.10**) together gives the desired result:

$$\begin{aligned}\mathrm{Area}(wx^{-N}) &\leqslant (3/2)r^2D^2\Big(\sum_i |w_i|^{\alpha}\Big)^2\\ &\leqslant (3/2)r^2D^2\Big(\sum_i |w_i|\Big)^{2\alpha} = (3/2)r^2D^2\,|w|^{2\alpha}\,.\end{aligned}$$

□

6 Suspension and snowflake balls

Throughout this section P denotes a non-negative $R\times R$ integer matrix with Perron–Frobenius eigenvalue λ, and r is an integer which is strictly greater than the largest row sum of P. In this section, we give an explicit description of the suspended snowflake groups $\Sigma G_{r,P}$ and the 3–dimensional $K(\Sigma G_{r,P},1)$ spaces $X^3_{r,P}$. Then we describe *snowflake balls* B^3_i which embed in the universal cover of $X^3_{r,P}$ and estimate their boundary areas. We show how to iterate this suspension procedure to obtain groups $\Sigma^k G_{r,P}$ and $(k+2)$–dimensional spaces $X^{k+2}_{r,P}$. Lastly we define higher-dimensional snowflake balls and estimate their boundary volumes.

Remark 6.1 In order to realize the exponents $(k+1)/k$ (the endpoints of the intervals in Figure 1, which are omitted otherwise) we add the free abelian group $\mathbb{Z}^2$ to the class of snowflake groups $G_{r,P}$. We endow $\mathbb{Z}^2$ with snowflake structure as follows

$$\mathbb{Z}^2 = \langle\, a_1, a_2, c \mid a_1a_2 = c,\; c = a_2a_1\,\rangle$$

and use the corresponding presentation 2–complex X in place of $X_{r,P}$. There is no matrix P associated to the group $\mathbb{Z}^2$, and so the only condition that we impose on the integer r is that $r\geqslant 2$. Since there are no stable letters s_i, we define the snowflake words to be the commutators $w_i = [a_1^{r^i}, a_2^{r^i}]$ and define the snowflake disks $B^2_i = \Delta_{r^i}$ to be the unique embedded disks in X with boundary w_i.

In the discussions that follow, whenever we talk about snowflake groups $G_{r,P}$, we shall always include $\mathbb{Z}^2$, and whenever we use the complexes $X_{r,P}$ we shall always include the presentation 2–complex X for $\mathbb{Z}^2$ described above.

The groups $\Sigma G_{r,P}$

Let $\phi\colon\ G_{r,P} \to G_{r,P}$ be the monomorphism which takes each a_i to a_i^r and each s_i to itself. The group $\Sigma G_{r,P}$ is defined to be the associated multiple HNN extension with stable letters u_1 and v_1:

$$\Sigma G_{r,P} \;=\; \langle\, G_{r,P}, u_1, v_1 \mid u_1 g u_1^{-1} = \phi(g),\ v_1 g v_1^{-1} = \phi(g)\ (g \in G_{r,P})\,\rangle.$$

The spaces $X^3_{r,P}$

These spaces will have fundamental group $\Sigma G_{r,P}$. Recall that $X_{r,P}$ is a 2–dimensional $K(G_{r,P}, 1)$ space. There is a cellular map $\Phi\colon\ X_{r,P} \to X_{r,P}$ which induces the map ϕ on the fundamental group. It maps the 1–cells labeled s_i homeomorphically to themselves, maps the 1–cells labeled a_i to themselves by degree r maps, and maps each 2–cell in the obvious manner; the image of each triangular 2–cell has combinatorial area r^2, and the image of the remaining 2–cells (which have an s_i edge in their boundaries) have combinatorial area r. The 3–complex $X^3_{r,P}$ with fundamental group $\Sigma G_{r,P}$ is obtained by taking two copies of the mapping torus of the map Φ and identifying them along a copy of $X_{r,P}$. From this perspective it is easy to see that $X^3_{r,P}$ is aspherical; each mapping torus is aspherical since $X_{r,P}$ is an aspherical 2–complex, and since Φ induces the monomorphism ϕ in π_1. We give more details of the cell structure of $X^3_{r,P}$ below.

Start with the 2–complex $X_{r,P}$ and form two copies of $X_{r,P} \times [0,1]$. Each copy is given the product cell structure, in which each k–cell of $X_{r,P}$ gives rise to a $(k+1)$–cell in $X_{r,P} \times (0,1)$. The “bottom” side $X_{r,P} \times \{0\}$ keeps its original cell structure and the “top” $X_{r,P} \times \{1\}$ is subdivided by pulling back under Φ the cell structure of $\Phi(X_{r,P})$. That is, each triangular 2–cell in a vertex space of $X_{r,P}$ is subdivided into r^2 triangles, and each edge space 2–cell (bearing the boundary label $s_j c_{\sigma(j)} s_j^{-1} a_j^r$) is subdivided into r copies of the same cell.

The vertical 1–cells of the two copies of $X_{r,P} \times [0,1]$ are labeled u_1 and v_1 respectively, oriented from $X_{r,P} \times \{1\}$ to $X_{r,P} \times \{0\}$. Finally to form $X^3_{r,P}$ one attaches the bottom of each piece to $X_{r,P}$ by the identity, and the top by the map Φ. Figures 5 and 6 illustrate the two types of 3–cell occurring in $X^3_{r,P}$.

Snowflake balls

We define embedded 3–dimensional balls B^3_j in $\widetilde{X}^3_{r,P}$ in a similar fashion to the snowflake disks constructed in Section 5. An essential difference, however, is that

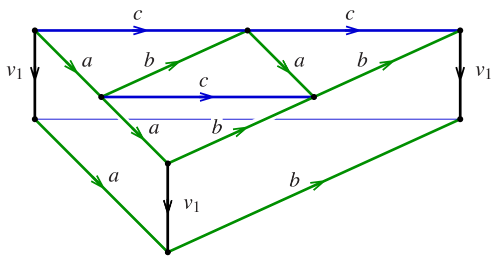

Figure 5: A triangular 3–cell (with $r = 2$)

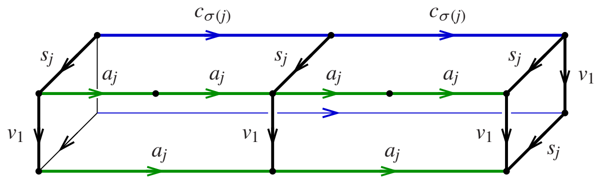

Figure 6: A rectangular 3–cell

now r is an integer, and the observations of Remark 4.1 apply. That is, snowflake disks with diameter labeled c^{r^i} are unique, and the corresponding snowflake words have no "remainder" terms.

As in the proof of Theorem A we let c be the diagonal element of a vertex group V_m in $G_{r,P} \subset \Sigma G_{r,P}$ where $m \geqslant 2$. We let w_i^+ and w_i^- denote respectively the (unique) positive and negative snowflake words representing c^{r^i}. (Note that the indexing here differs from that in Section 5, where these words would be called $w_{r^i}^{\pm}$.) Let B_i^2 be the snowflake disk bounded by $w_i = w_i^+(w_i^-)^{-1}$, with diameter labeled c^{r^i}. Note that B_i^2 is the same as the snowflake disk Δ_{r^i} of Section 5.

For each positive integer j, we shall use a *stack of thickened van Kampen disks* to define an embedded 3–ball B_j^3 in the universal cover of $X_{r,P}^3$. Note that the universal cover of $X_{r,P}^3$ contains infinitely many embedded copies of the universal cover of $X_{r,P}$; one for each coset of $G_{r,P}$ in $\Sigma G_{r,P}$. We call two such copies *adjacent* if the cosets have representatives which differ by right multiplication by $u_1^{\pm 1}$ or $v_1^{\pm 1}$.

The map $\Phi\colon X_{r,P} \longrightarrow X_{r,P}$ lifts to a map of universal covers which we also denote by Φ. Consider the image $\Phi(B_i^2)$ of the embedded snowflake disk B_i^2. This image is again embedded, but its boundary word is $\phi(w_i)$. If we apply the curve shortening procedure once to the subword $\phi(w_i^+)$ we obtain w_{i+1}^+, which is the positive snowflake word for $c^{r^{i+1}}$. Similarly, if we apply curve shortening once to the subword $\phi(w_i^-)$ we obtain the negative snowflake word for $c^{r^{i+1}}$. Thus $\Phi(B_i^2)$ is a sub-diagram of B_{i+1}^2. The *top half* of the ball B_j^3 is defined to be the union of the mapping cylinders of Φ with domain B_i^2 and codomain B_{i+1}^2 where i ranges from 1 to j; the copies of B_i^2 are identified. This embeds in the universal cover of $X_{r,P}^3$ as follows. The disk B_1^2 embeds in some copy of the universal cover of $X_{r,P}$, B_2^2 embeds in the adjacent copy obtained by right multiplying by u_1^{-1}, and the mapping cylinder of $\Phi\colon B_1^2 \to B_2^2$ embeds in the universal cover of $X_{r,P}^3$ to interpolate between the images of B_1^2 and B_2^2. Note that this embedding is possible since the universal covering of $X_{r,P}^3$ can be described as an infinite union of mapping cylinders of $\Phi\colon \widetilde{X}_{r,P} \to \widetilde{X}_{r,P}$ which is encoded by the Bass–Serre tree T corresponding to the multiple HNN description of $\Sigma G_{r,P}$.

We continue to add mapping cylinders of $\Phi\colon B_i^2 \to B_{i+1}^2$ for $i = 2, \ldots, j$, as indicated in the top half of the schematic diagram in Figure 7. The image of the union of the

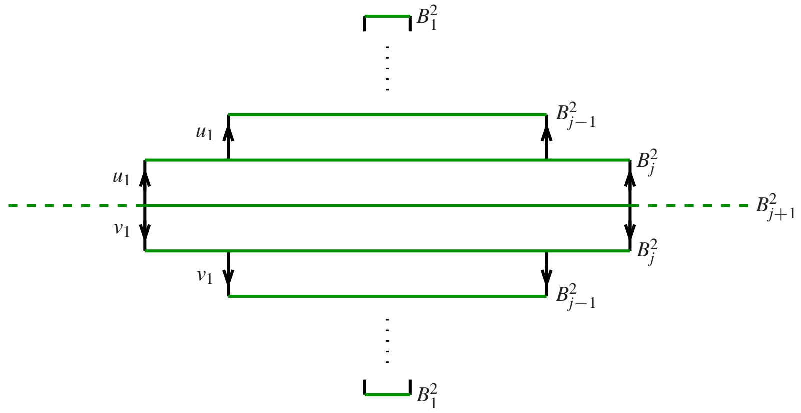

Figure 7: A schematic diagram of the embedded ball B_j^3

first few embedded layers is shown in Figure 8. In a similar fashion, we can embed a second copy of the union of mapping cylinders of $\Phi\colon B_i^2 \to B_{i+1}^2$. However, this time we start from the copy of B_j^2 in the image of the previous union, and add the mapping cylinders in descending order (so $i = j, \ldots, 1$) and require that new copies of the universal cover of $X_{r,P}$ differ by right multiplication by v_1^{+1}. The image of this family is indicated in the lower half of the schematic diagram of Figure 7, and the total union is the embedded ball B_j^3. It is easy to see that the union embeds, since each mapping cylinder embeds, and distinct mapping cylinders correspond to distinct layers in the 3–complex $\widetilde{X}_{r,P}^3$. These layers are distinct, since they map to distinct edges of the Bass–Serre tree T. Finally, there is a 2–dimensional “fringe” at the equator B_{j+1}^2 level. We remove this fringe by simply replacing the two embeddings of $\Phi\colon B_j^2 \to B_{j+1}^2$ by

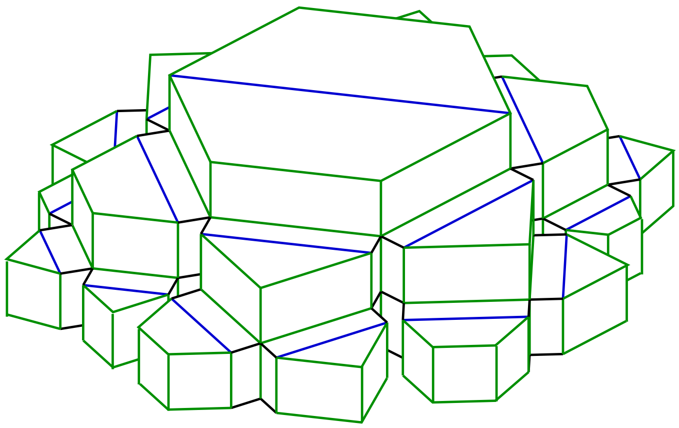

Figure 8: A few layers of B_j^3

embeddings of $\Phi\colon\ B_j^2 \to \Phi(B_j^2)$.

Lemma 6.2 *Given r and P there is a positive constant F_0 such that $\left|\partial B_j^2\right| \leqslant \mathrm{Area}(\partial B_j^3) \leqslant F_0 \left|\partial B_j^2\right|$ for every j.*

Proof The ball B_j^3 is a union of $2j$ mapping cylinders. See Figure 7 for a schematic representation. Its boundary area is twice the area of the *upper hemisphere*. This latter area is estimated as follows.

For each $1 \leqslant i \leqslant j$, there are $\left|\partial B_i^2\right|$ *vertical* (conjugation by u_1) 2–cells, which interpolate between ∂B_i^2 and $\Phi(\partial B_i^2)$. This proves the first inequality, $\left|\partial B_j^2\right| \leqslant \mathrm{Area}(\partial B_j^3)$.

For each $1 \leqslant i \leqslant j$ there are *horizontal* 2–cells which interpolate between $\Phi(\partial B_{i-1}^2)$ and ∂B_i^2. In the case $i = 1$ there is no loop $\Phi(\partial B_0^2)$, and the horizontal 2–cells just fill the van Kampen diagram B_1^2. For any i, the horizontal 2–cell contribution to the area is bounded above by $\left|\partial B_i^2\right|$. To see this, note that the horizontal interpolation is a union of pieces of the form $s_j a_{i_1} \cdots a_{i_m} s_j^{-1} a_j^{-r}$ where $\{a_1, \ldots, a_m\}$ generates a vertex group V_m, and the stable letter s_j conjugates the diagonal element of this vertex group to some generator a_j of $G_{r,P}$. The area of this piece is m, and its contribution to $\left|\partial B_i^2\right|$ is $m + 2$.

Counting vertical and horizontal 2–cells for both hemispheres we obtain

$$\mathrm{Area}(\partial B_j^3) \leqslant 4 \sum_{i=1}^{j} \left|\partial B_i^2\right| .$$

Proposition 4.5 implies that $|w_i^+| \leqslant C_0^{-1/\alpha} r^{i/\alpha}$ and so

$$4\sum_{i=1}^{j} \left|\partial B_i^2\right| \;=\; 8\sum_{i=1}^{j} \left|w_i^+\right| \;\leqslant\; 8C_0^{-1/\alpha}\sum_{i=1}^{j}(r^{1/\alpha})^i\,.$$

The last term is a geometric series, and so is bounded above by $F_0'(r^{1/\alpha})^j$ for a positive constant F_0' (independent of j). Proposition 4.5 also gives $C_1^{-1/\alpha} r^{j/\alpha} \leqslant \left|w_j^+\right|$ and so

$$\mathrm{Area}(\partial B_j^3) \;\leqslant\; F_0' r^{j/\alpha} \;\leqslant\; \frac{F_0'}{2} C_1^{1/\alpha} \left|\partial B_j^2\right|\,.$$

Now the desired (second) inequality holds by taking $F_0 = (F_0'/2)C_1^{1/\alpha}$. □

The inductive supension procedure

Having discussed $\Sigma G_{r,P}$ we define further suspensions $\Sigma^k G_{r,P}$ having $(k+2)$–dimensional Eilenberg–MacLane spaces $X_{r,P}^{k+2}$, and $(k+2)$–dimensional snowflake balls $B_j^{k+2} \subset \widetilde{X}_{r,P}^{k+2}$. We assume that the group $\Sigma^{k-1} G_{r,P}$, the space $X_{r,P}^{k+1}$, and snowflake balls $B_j^{k+1} \subset \widetilde{X}_{r,P}^{k+1}$ have already been constructed.

First we define the groups $\Sigma^k G_{r,P}$. Let $\phi_k \colon \Sigma^{k-1} G_{r,P} \to \Sigma^{k-1} G_{r,P}$ be the monomorphism which sends a_i to a_i^r and which leaves fixed the stable letters s_i, u_i, and v_i. We define $\Sigma^k G_{r,P}$ to be the multiple ascending HNN extension with two stable letters u_k and v_k, each acting by ϕ_k:

$$\Sigma^k G_{r,P} \;=\; \langle\, \Sigma^{k-1} G_{r,P}, u_k, v_k \mid u_k g u_k^{-1} = \phi_k(g),\ v_k g v_k^{-1} = \phi_k(g)\ (g \in \Sigma^{k-1} G_{r,P})\,\rangle.$$

Next we define the spaces $X_{r,P}^{k+2}$. The homomorphism ϕ_k is induced by a cellular map $\Phi_{k+1} \colon X_{r,P}^{k+1} \to X_{r,P}^{k+1}$. We define $X_{r,P}^{k+2}$ to be the double mapping torus with monodromy Φ_{k+1}. That is, take two copies of $X_{r,P}^{k+1} \times [0,1]$, identify the "bottom" sides $X_{r,P}^{k+1} \times \{0\}$ to $X_{r,P}^{k+1}$ by the identity, and attach the "top" sides $X_{r,P}^{k+1} \times \{1\}$ to $X_{r,P}^{k+1}$ by the map Φ_{k+1}. The vertical 1–cells of the copies of $X_{r,P}^{k+1} \times [0,1]$ are labeled u_k and v_k respectively, oriented from $X_{r,P}^{k+1} \times \{1\}$ to $X_{r,P}^{k+1} \times \{0\}$. The resulting space $X_{r,P}^{k+2}$ is given a cell structure analogous to that of $X_{r,P}^3$. As before, $X_{r,P}^{k+2}$ is aspherical, has dimension $k+2$, and has fundamental group $\Sigma^k G_{r,P}$.

Now we define the higher-dimensional snowflake balls. The map Φ_{k+1} lifts to a map $\widetilde{X}_{r,P}^{k+1} \to \widetilde{X}_{r,P}^{k+1}$ which we will also call Φ_{k+1}. We define $(k+2)$–dimensional balls B_j^{k+2} of diameter r^j for each j as unions of mapping cylinders (called *layers*) of the map Φ_{k+1} restricted to the $(k+1)$–dimensional balls B_i^{k+1}. These mapping cylinders are assembled as shown in Figure 7, with B_i^{k+1} in place of B_i^2. More specifically, we assume inductively that Φ_{k+1} maps B_i^{k+1} into a subcomplex of B_{i+1}^{k+1} for each i. Then the upper hemisphere of B_j^{k+2} is the union of the mapping cylinders of $\Phi_{k+1}\colon B_i^{k+1} \to B_{i+1}^{k+1}$ where i ranges from 1 to $j-1$, and the mapping cylinder of $\Phi_{k+1}\colon B_j^{k+1} \to \Phi_{k+1}(B_j^{k+1})$. The lower hemisphere is defined similarly, and the two are identifed along $\Phi_{k+1}(B_j^{k+1})$. Note that the subspaces $B_i^{k+1} - \Phi_{k+1}(B_{i-1}^{k+1})$ of the domains of these mapping cylinders lie in the boundary of B_j^{k+2}.

Recall that Φ_{k+1} maps B_i^{k+1} to a subcomplex of B_{i+1}^{k+1}. There is an induced map Φ_{k+2} from the mapping cylinder of $\Phi_{k+1}\colon B_i^{k+1} \to B_{i+1}^{k+1}$ to the mapping cylinder of $\Phi_{k+1}\colon B_{i+1}^{k+1} \to B_{i+2}^{k+1}$; use $\Phi_{k+1} \times \mathrm{id}$ on $B_i^{k+1} \times I$ and Φ_{k+1} on B_{i+1}^{k+1}. Then Φ_{k+2} maps layer i of B_j^{k+2} to layer $i+1$ of B_{j+1}^{k+2} for any $i \leqslant j$ (in either hemisphere). These maps defined on the layers of B_j^{k+2} join together to define the map $\Phi_{k+2}\colon B_j^{k+2} \to B_{j+1}^{k+2}$.

The balls B_j^{k+2} embed into $\widetilde{X}_{r,P}^{k+2}$ exactly as the balls B_j^3 embed into $\widetilde{X}_{r,P}^3$. That is, we consider $\widetilde{X}_{r,P}^{k+2}$ as a union of copies of the mapping cylinder of $\Phi_{k+1}\colon \widetilde{X}_{r,P}^{k+1} \to \widetilde{X}_{r,P}^{k+1}$ with the mapping parameter corresponding to right multiplication by u_k^{-1} or v_k^{-1}. Then the embedding $B_j^{k+2} \to \widetilde{X}_{r,P}^{k+2}$ is assembled from the embeddings $B_i^{k+1} \to \widetilde{X}_{r,P}^{k+1}$ (for $i \leqslant j$) as shown in Figure 7, with the upper hemisphere extending in the u_k direction and the lower hemisphere in the v_k direction. Under this embedding, the map $\Phi_{k+2}\colon B_j^{k+2} \to B_{j+1}^{k+2}$ described above is simply the restriction of $\Phi_{k+2}\colon \widetilde{X}_{r,P}^{k+2} \to \widetilde{X}_{r,P}^{k+2}$ to B_j^{k+2}.

For any k, we define the *shell* of a snowflake ball B_j^k to be the subspace $B_j^k - \Phi_k(B_{j-1}^k)$, or simply B_j^k in the case $j = 1$.

Lemma 6.3 $\mathrm{Vol}^k(\mathrm{shell}(B_j^k)) \leqslant \mathrm{Vol}^{k-1}(\partial B_j^k)$.

Proof It suffices to show that every k–cell of the shell has a $(k-1)$–dimensional face contained in ∂B_j^k. Recall that B_j^k is a union of layers, so consider the intersection of the shell with layer i (in either hemisphere). This layer is a mapping cylinder $\mathcal{M}(\Phi_{k-1}\colon B_i^{k-1} \to B_{i+1}^{k-1})$ and its preimage in B_{j-1}^k under Φ_k is layer $i-1$ of this smaller ball (or is empty in the case $i = 1$). Hence the intersection of the shell with layer i is

$$\begin{aligned} &\mathcal{M}(\Phi_{k-1}\colon B_i^{k-1} \to B_{i+1}^{k-1}) \ - \ \Phi_k(\mathcal{M}(\Phi_{k-1}\colon B_{i-1}^{k-1} \to B_i^{k-1})) \\ &= \mathcal{M}(\Phi_{k-1}\colon B_i^{k-1} \to B_{i+1}^{k-1}) \ - \ \mathcal{M}(\Phi_k\colon \Phi_{k-1}(B_{i-1}^{k-1}) \to \Phi_{k-1}(B_i^{k-1}))) \\ &= \mathcal{M}(\Phi_{k-1}\colon (B_i^{k-1} - \Phi_{k-1}(B_{i-1}^{k-1})) \to (B_{i+1}^{k-1} - \Phi_{k-1}(B_i^{k-1}))) \end{aligned}$$

if $i > 1$, and is $\mathcal{M}(\Phi_{k-1}\colon B_i^{k-1} \to B_{i+1}^{k-1})$ in the case $i = 1$. Either way, this part of shell(B_j^k) is the mapping cylinder of the restriction of Φ_{k-1} to shell(B_i^{k-1}). Hence each k–cell has a $(k-1)$–dimensional face in shell(B_i^{k-1}), which is contained in ∂B_j^k. □

The next result is a higher-dimensional analogue of Lemma 6.2.

Lemma 6.4 *Given r, P, and $k \geqslant 3$ there is a positive constant F_k such that $\mathrm{Vol}^{k-2}(\partial B_j^{k-1}) \leqslant \mathrm{Vol}^{k-1}(\partial B_j^k) \leqslant F_k \mathrm{Vol}^{k-2}(\partial B_j^{k-1})$ for every j.*

Proof We prove, for $k \geqslant 3$, the following two statements: there exist positive constants E_k, F_k such that

(1) $(2C_1^{-1/\alpha})(r^{1/\alpha})^j \leqslant \mathrm{Vol}^{k-2}(\partial B_j^{k-1}) \leqslant E_k(r^{1/\alpha})^j$, and

(2) $\mathrm{Vol}^{k-2}(\partial B_j^{k-1}) \leqslant \mathrm{Vol}^{k-1}(\partial B_j^k) \leqslant F_k \mathrm{Vol}^{k-2}(\partial B_j^{k-1})$

for all j (with C_1 given by Proposition 4.5). Statement (1) is a higher-dimensional analogue of Proposition 4.5 and (2) is the main statement of the lemma. The two statements are proved together by induction on k.

If $k = 3$ then (1) follows from Proposition 4.5, with $E_3 = 2C_0^{-1/\alpha}$. Statement (2) is given by Lemma 6.2 (with $F_3 = F_0$).

For $k > 3$ we prove (1) as follows. The induction hypothesis implies that

$$\mathrm{Vol}^{k-2}(\partial B_j^{k-1}) \leqslant F_{k-1} \mathrm{Vol}^{k-3}(\partial B_j^{k-2})$$

by (2) and $\mathrm{Vol}^{k-3}(\partial B_j^{k-2}) \leqslant E_{k-1}(r^{1/\alpha})^j$ by (1). Hence $\mathrm{Vol}^{k-2}(\partial B_j^{k-1}) \leqslant E_k(r^{1/\alpha})^j$ with $E_k = F_{k-1}E_{k-1}$. We also have (by induction) $\mathrm{Vol}^{k-2}(\partial B_j^{k-1}) \geqslant \mathrm{Vol}^{k-3}(\partial B_j^{k-2}) \geqslant (2C_1^{-1/\alpha})(r^{1/\alpha})^j$ by (2) and (1). This establishes (1).

To prove (2) we count vertical and horizontal $(k-1)$–cells of ∂B_j^k as in the proof of Lemma 6.2. In each hemisphere of B_j^k, layer i is a copy of the mapping cylinder of $\Phi_{k-1}\colon B_i^{k-1} \to B_{i+1}^{k-1}$. This layer meets ∂B_j^k in *horizontal* cells which are the $(k-1)$–cells of shell(B_i^{k-1}), and *vertical* cells, each of which is the product of a $(k-2)$–cell in ∂B_i^{k-1} with I. This latter observation implies the first inequality of (2) (taking $i = j$) and also that the number of vertical cells in layer i is at most $\mathrm{Vol}^{k-2}(\partial B_i^{k-1})$. The number of horizontal cells is at most $\mathrm{Vol}^{k-2}(\partial B_i^{k-1})$ by Lemma 6.3. Adding the contributions from all layers in both hemispheres, we obtain

$$\mathrm{Vol}^{k-1}(\partial B_j^k) \leqslant 4 \sum_{i=1}^{j} \mathrm{Vol}^{k-2}(\partial B_i^{k-1}).$$

Statement (1) implies $4\sum_{i=1}^{j} \mathrm{Vol}^{k-2}(\partial B_i^{k-1}) \leqslant 4E_k \sum_{i=1}^{j}(r^{1/\alpha})^i$ and the latter sum is a geometric series. Hence $\mathrm{Vol}^{k-1}(\partial B_j^k) \leqslant F_k'(r^{1/\alpha})^j$ for some constant F_k'. Now (1) implies that $\mathrm{Vol}^{k-1}(\partial B_j^k) \leqslant (F_k'/2)(C_1^{1/\alpha}) \mathrm{Vol}^{k-2}(\partial B_j^{k-1})$, establishing (2) with $F_k = (F_k'/2)C_1^{1/\alpha}$. □

7 Proof of Theorem C

We will establish upper and lower bounds for the k–dimensional Dehn functions $\delta^{(k)}(x)$ of the groups $\Sigma^{k-1}G_{r,P}$ and these will be equivalent. As usual λ denotes the Perron–Frobenius eigenvalue of P and $\alpha = \log_\lambda(r)$. In the case of $\Sigma^{k-1}\mathbb{Z}^2$ we define $\alpha = 1$.

The lower bound

As in the proof of Theorem A, we show that the embedded snowflake balls $B_i^{k+1} \subset \widetilde{X}_{r,P}^{k+1}$ have the correct proportions and are numerous enough to determine $\delta^{(k)}(x)$ from below.

First we show that for every $k \geqslant 1$ there is a constant G_k such that

(7.1) $$\mathrm{Vol}^{k+1}(B_i^{k+1}) \geqslant G_k \mathrm{Vol}^k(\partial B_i^{k+1})^{2\alpha}$$

for all i. The case $k=1$ was proved in (**5.1**) with $G_1 = (C_0)^2 4^{-\alpha}$. For $k > 1$ we proceed by induction. Note that $\mathrm{Vol}^{k+1}(B_i^{k+1}) \geqslant \mathrm{Vol}^k(B_i^k)$ since the latter is the volume of the mapping cylinder of $\Phi_k\colon B_i^k \to \Phi_k(B_i^k)$ inside B_i^{k+1}. We also have $\mathrm{Vol}^k(B_i^k) \geqslant G_{k-1}\mathrm{Vol}^{k-1}(\partial B_i^k)^{2\alpha}$ by the induction hypothesis. Lemma 6.4 implies that $G_{k-1}\mathrm{Vol}^{k-1}(\partial B_i^k)^{2\alpha} \geqslant G_{k-1}F_{k+1}^{-2\alpha}\mathrm{Vol}^k(\partial B_i^{k+1})^{2\alpha}$. Equation (**7.1**) now follows by taking $G_k = G_{k-1}F_{k+1}^{-2\alpha}$.

Next we show that for each $k \geqslant 2$ the sequence $(\mathrm{Vol}^k(\partial B_i^{k+1}))_i$ is exponentially bounded and tends to infinity. Consider first the case $k=2$. Then we have

$$\frac{\mathrm{Vol}^2(\partial B_{i+1}^3)}{\mathrm{Vol}^2(\partial B_i^3)} \leqslant \frac{F_0\,|\partial\Delta_{r^{i+1}}|}{|\partial\Delta_{r^i}|} \leqslant F_0\left(\frac{r^{i+1}C_1}{r^iC_0}\right)^{1/\alpha} = F_0\left(\frac{rC_1}{C_0}\right)^{1/\alpha}$$

where the first inequality holds by Lemma 6.2, and the second by Proposition 4.5. Thus, the sequence is exponentially bounded. For $k > 2$ we have

$$\frac{\mathrm{Vol}^k(\partial B_{i+1}^{k+1})}{\mathrm{Vol}^k(\partial B_i^{k+1})} \leqslant \frac{F_{k+1}\mathrm{Vol}^{k-1}(\partial B_{i+1}^k)}{\mathrm{Vol}^{k-1}(\partial B_i^k)}$$

by Lemma 6.4 and so $(\mathrm{Vol}^k(\partial B_i^{k+1}))_i$ is exponentially bounded, by induction on k. It tends to infinity because

$$\mathrm{Vol}^k(\partial B_i^{k+1}) \geqslant \mathrm{Vol}^2(\partial B_i^3) \geqslant |\partial \Delta_{r^i}| \geqslant 2C_1^{-1/\alpha}(r^{1/\alpha})^i$$

by Lemma 6.4, Lemma 6.2, and Proposition 4.5. Now, using Remarks 2.1 and 2.7, we conclude from (**7.1**) that $\delta^{(k)}(x) \succcurlyeq x^{2\alpha}$.

The upper bound

To establish the upper bound we must work with Dehn functions $\delta_G^M(x)$ modeled on arbitrary manifolds M with boundary, as defined in Section 2. Recall that the *dimension* of $\delta_G^M(x)$ is the dimension of ∂M, and $\delta_G^M(x)$ agrees with the usual k–dimensional Dehn function when M is the $(k+1)$–dimensional ball.

A function $F\colon \mathbb{N} \to \mathbb{N}$ is *superadditive* if $F(a+b) \geqslant F(a) + F(b)$ for all a, b. Recall that the *geometric dimension* of a group G is the smallest dimension of a $K(G,1)$ complex.

Theorem 7.2 *Let G be a group of type $\mathcal{F}_n$ and geometric dimension at most n, and fix a finite aspherical n–complex X with fundamental group G. Suppose that the Dehn function $\delta_G^M(x)$ (defined with respect to X) satisfies*

$$\delta_G^M(x) \leqslant F(x)$$

for every n–manifold M, where $F\colon \mathbb{N} \to \mathbb{N}$ is non-decreasing. Let H be a multiple ascending HNN extension of G. Then H is of type $\mathcal{F}_{n+1}$, has geometric dimension at most $n+1$, and

$$\delta_H^M(x) \leqslant F(x)$$

for every $(n+1)$–manifold M.

In the hypotheses we are including Dehn functions $\delta_G^M(x)$ where M has more than one connected component (otherwise we should add that F is superadditive). **Stipulation:** the n–dimensional Dehn functions in the conclusion are defined with respect to a fixed complex Y constructed in the proof of the theorem.

Proof First we define the finite $(n+1)$–dimensional complex Y with fundamental group H in the usual way. Suppose the multiple ascending extension has k stable letters. Form k copies of $X \times [-1,1]$, give each the product cell structure, and attach each copy of $X \times \{-1\}$ to X by the identity map. Then attach each copy of $X \times \{1\}$ to

X by the appropriate monodromy map, and call the resulting space Y. Let $Z \subset Y$ be the union of the spaces $X \times \{0\}$. There are natural projections along the fibers $p_0\colon\ Z \to X$ and $p_1\colon\ Z \to X$ which factor through $Z \times \{-1\}$ and $Z \times \{1\}$ respectively. Let $\widetilde{Y}$ be the universal cover of Y and let $\widetilde{X}$ and $\widetilde{Z}$ be the preimages of X and Z in $\widetilde{Y}$. The projections p_i lift to projections $p_i\colon\ \widetilde{Z} \to \widetilde{X}$ along fibers. Note that each component of $\widetilde{X}$ and $\widetilde{Z}$ is a copy of the universal cover of X, and in fact $p_0\colon\ \widetilde{Z} \to \widetilde{X}$ is a homeomorphism.

Each open k–cell σ^k in $\widetilde{Z} \times (-1,1) \subset \widetilde{Y}$ has the form $\sigma^{k-1} \times (-1,1)$ where σ^{k-1} is a $(k-1)$–cell in $\widetilde{X}$, and the restriction of p_0 to $\sigma^k \cap \widetilde{Z}$ is simply projection onto the first factor. Since $\widetilde{Z}$ is not a subcomplex of $\widetilde{Y}$, we measure volume in $\widetilde{Z}$ by passing to $\widetilde{X}$ via p_0. The description of p_0 just given leads to the following observation: if $f\colon\ M^k \to \widetilde{Y}$ is an admissible map transverse to $\widetilde{Z}$ and $\widetilde{X}$, and $N = f^{-1}(\widetilde{Z})$ and $M_0 = f^{-1}(\widetilde{X})$, then $p_0 \circ f|_N$ and $f|_{M_0}$ are admissible and

(7.3) $$\mathrm{Vol}^k(f) = \mathrm{Vol}^{k-1}(p_0 \circ f|_N) + \mathrm{Vol}^k(f|_{M_0})$$

where the left hand side is volume in $\widetilde{Y}$ and the right hand side is volume in $\widetilde{X}$.

Now suppose that M is a compact $(n+1)$–manifold with boundary and let $g\colon\ M \to \widetilde{Y}$ be a least-volume map with boundary $f = g|_{\partial M}$. We can arrange by a homotopy that $N = g^{-1}(\widetilde{Z})$ is a properly embedded codimension one submanifold with a product neighborhood $N \times [-1,1] \subset M$ such that $g^{-1}(\widetilde{Z} \times (-1,1)) = N \times (-1,1)$. (To do this, consider the composition $\pi\colon\ \widetilde{Y} \to Y \to S^1 \vee S^1 \to S^1$, where $S^1 \vee S^1$ is the underlying graph for the multiple HNN-description of $\pi_1(Y)$ and $S^1 \vee S^1 \to S^1$ is a fold. By a homotopy of g, $\pi \circ g$ can be made smooth in a neighborhood of $g^{-1}(\widetilde{Z})$. By a further homotopy, we can arrange that $\pi(\widetilde{Z})$ is a regular value of $\pi \circ g$; now use transversality.) The product structure on $N \times [-1,1]$ may be chosen so that $g|_{N \times (-1,1)}$ is the map $g|_N \times \mathrm{id}$. Note that N may have several connected components.

We claim that $\mathrm{Vol}^n(p_0 \circ g|_N)$ is smallest among all N–fillings of $p_0 \circ f|_{\partial N}\colon\ \partial N \to \widetilde{X}$. Assuming this for the moment, the theorem is proved as follows. We have $\mathrm{Vol}^{n+1}(g) = \mathrm{Vol}^n(p_0 \circ g|_N)$ by (7.3) because $\widetilde{X}$ has dimension n. Then $\mathrm{Vol}^n(p_0 \circ g|_N) = \mathrm{FVol}^N(p_0 \circ f|_{\partial N})$ by (2.5) and the claim, and the latter is at most $\delta^N_G(\mathrm{Vol}^{n-1}(p_0 \circ f|_{\partial N}))$ by the definition of δ^N_G. Equation (7.3) implies that $\delta^N_G(\mathrm{Vol}^{n-1}(p_0 \circ f|_{\partial N})) \leqslant \delta^N_G(\mathrm{Vol}^n(f))$. Then we have the desired bound

$$\mathrm{FVol}^M(f) \ = \ \mathrm{Vol}^{n+1}(g) \ \leqslant \ \delta^N_G(\mathrm{Vol}^n(f)) \ \leqslant \ F(\mathrm{Vol}^n(f))$$

by the main hypothesis and we conclude that $\delta^M_H(\mathrm{Vol}^n(f)) \leqslant F(\mathrm{Vol}^n(f))$. Since $\mathrm{Vol}^n(f)$ was arbitrary and F is non-decreasing, we have $\delta^M_H(x) \leqslant F(x)$ for all x.

Now we return to the claim that $\mathrm{Vol}^n(p_0 \circ g|_N) = \mathrm{FVol}^N(p_0 \circ f|_{\partial N})$. We show that if $p_0 \circ g|_N$ is not a least-volume filling of $p_0 \circ f|_{\partial N}$ then g can be modified rel ∂M to a map of smaller volume, contradicting the choice of g.

Let $M_0 = g^{-1}(\widetilde{X})$, and note that the frontier of M_0 in M is $N \times \{-1\} \cup N \times \{1\}$. These two subsets of ∂M_0 will be denoted M_0^- and M_0^+ respectively.

Suppose $\mathrm{Vol}^n(h) < \mathrm{Vol}^n(p_0 \circ g|_N)$ for some map $h\colon N \to \widetilde{X}$ with $h|_{\partial N} = p_0 \circ f|_{\partial N}$. Form a new copy of M in which $N \times (-1,1)$ is replaced by $N \times (-2,2)$. Define a new map $g'\colon M \to \widetilde{Y}$ by letting g' be g on M_0, $(p_0^{-1} \circ h) \times \mathrm{id}$ on $N \times (-1,1)$, and by extending to the remaining regions as follows. Note that $(p_0^{-1} \circ h) \times \mathrm{id}$ extends continuously to $N \times [-1,1]$ as h on $N \times \{-1\}$ and as $p_1 \circ p_0^{-1} \circ h$ on $N \times \{1\}$. Since each component of $\widetilde{X}$ is contractible the maps $p_1 \circ p_0^{-1} \circ h$ and $g|_{M_0^+}$ are homotopic rel ∂N. We let $g'|_{N\times[1,2]}\colon N \times [1,2] \to \widetilde{X}$ be such a homotopy. Similarly $g'|_{N\times[-2,-1]}$ is defined to be a homotopy in $\widetilde{X}$ from $g|_{M_0^-}$ to h, fixing ∂N pointwise. This defines the map $g'\colon M \to \widetilde{Y}$.

Now collapse each fiber of $\partial N \times [1,2]$ and $\partial N \times [-2,-1]$ to a point, to obtain a new copy of M with a map $g''\colon M \to \widetilde{Y}$ which agrees with g on ∂M. Note that all of $M - (N \times (-1,1))$ maps by g'' into $\widetilde{X}$ and $g''|_{N\times(-1,1)} = (p_0^{-1} \circ h) \times \mathrm{id}$. Hence by (7.3) we have $\mathrm{Vol}^{n+1}(g'') = \mathrm{Vol}^n(h) < \mathrm{Vol}^n(p_0 \circ g|_N) = \mathrm{Vol}^{n+1}(g)$, a contradiction. □

Lemma 7.4 *If G is finitely presented, $\delta_G(x) \leqslant F(x)$ with $F(x)$ superadditive, and M is a compact 2–manifold with boundary, then $\delta_G^M(x) \leqslant F(x)$.*

In particular if $\delta_G(x)$ is superadditive then $\delta_G^M(x) \leqslant \delta_G(x)$ for every compact 2–manifold M.

Proof If M is connected with one boundary component then let $q\colon M \to D^2$ be a quotient map which collapses the complement of a collar neighborhood of ∂M to a point. Then $\mathrm{Area}(g \circ q) = \mathrm{Area}(g)$ for any map $g\colon D^2 \to \widetilde{X}$, and we have $\delta_G^M(x) \leqslant \delta_G(x) \leqslant F(x)$.

If N is closed then $\delta_G^{M\sqcup N}(x) = \delta_G^M(x)$ since N may be assigned zero area by mapping it to a point. So without loss of generality assume that M has no closed components. For each component M' of M there is a quotient map to a connected, simply connected space Z' which is a union of disks (one for each boundary component of M') and arcs joining them. Taking a union of such spaces and maps, we have a quotient map $M \to Z$. Every map $D^2 \sqcup \cdots \sqcup D^2 \to \widetilde{X}$ extends to a map $Z \to \widetilde{X}$ which yields (by composition) a map $M \to \widetilde{X}$ with the same area. Hence $\delta_G^M(x) \leqslant \delta_G^{D^2\sqcup\cdots\sqcup D^2}(x)$. Now superadditivity of F implies $\delta_G^{D^2\sqcup\cdots\sqcup D^2}(x) \leqslant F(x)$. □

Theorem 7.5 *Let G be a finitely presented group of geometric dimension 2 with $\delta_G(x)$ equivalent to a superadditive function. Let H be obtained from G by performing n iterated multiple ascending HNN extensions. Then $\delta_H^{(n+1)}(x) \preccurlyeq \delta_G(x)$.*

The upper bound of Theorem C follows immediately, by Theorem A.

Proof Let $F_0(x)$ be superadditive where $F_0(x) \simeq \delta_G(x)$. Then $\delta_G(x) \leqslant F(x) = CF_0(Cx) + Cx$ for some C and $F(x)$ is superadditive. The result now follows directly from Lemma 7.4 and Theorem 7.2. □

The case $n = 1$ of Theorem 7.5 was proved by Wang and Pride [22], using a more direct method.

8 Products with $\mathbb{Z}$

In this section we determine higher Dehn functions of $G \times \mathbb{Z}$ for certain groups G. In these cases the geometry of $G \times \mathbb{Z}$ is accurately represented by embedded balls which are products of optimal balls in G with intervals, with suitably chosen lengths. We conclude the section by proving Theorem D.

To establish an upper bound for Dehn functions of $G \times \mathbb{Z}$ we need the following refinement of Theorem 7.2. The proof is based on Theorem 6.1 of Alonso *et al* [1].

Theorem 8.1 *Let G be a group of type $\mathcal{F}_n$ and geometric dimension at most n, and fix a finite aspherical n–complex X with fundamental group G. Suppose that the Dehn function $\delta_G^M(x)$ satisfies*

$$\delta_G^M(x) \leqslant Cx^s$$

for every n–manifold M, and fixed $C > 0$ and $s > 1$. Then

$$\delta_{G\times\mathbb{Z}}^M(x) \leqslant C^{1/s}x^{2-1/s}$$

for every $(n+1)$–manifold M.

Proof First note that we are in the situation of Theorem 7.2, which is valid, but no longer provides the best possible upper bound. Define Y, Z, p_0, and p_1 as in the proof of Theorem 7.2. Note that now the projections along fibers $p_0, p_1 \colon \widetilde{Z} \to \widetilde{X}$ are both homeomorphisms, and $\mathrm{Vol}^k(p_0 \circ f) = \mathrm{Vol}^k(p_1 \circ f)$ for any $f \colon N^k \to \widetilde{Z}$.

Given a compact $(n+1)$–manifold M with boundary, consider a map $f \colon \partial M \to \widetilde{Y}$. Arrange that $L = f^{-1}(\widetilde{Z})$ is a codimension one submanifold with a product neighborhood $L \times [-1,1] \subset \partial M$ such that $f^{-1}(\widetilde{Z} \times (-1,1)) = L \times (-1,1)$. As before, the product structure on $L \times [-1,1]$ can be chosen so that $f|_{L\times(-1,1)}$ is the map $f|_L \times \mathrm{id}$.

We will prove that $\delta^M_{G\times\mathbb{Z}}(x) \leqslant C^{1/s}x^{2-1/s}$ by induction on the number of connected components of L. If $L = \emptyset$ then $f(\partial M) \subset \widetilde{X}$. The components of ∂M may map into different components of $\widetilde{X}$. However, by joining these components with a minimal collection of embedded arcs in the 1–skeleton of $\widetilde{Y}$, one obtains a contractible subcomplex $T \subset \widetilde{Y}$ of dimension n containing $f(\partial M)$. Then f extends to a map $g\colon M \to T \subset \widetilde{Y}$ with $\mathrm{Vol}^{n+1}(g) = 0$.

Now assume that $L \neq \emptyset$. Let $\widetilde{Z}_0$ be a connected component of $\widetilde{Z}$ such that $L_0 = f^{-1}(\widetilde{Z}_0)$ is a non-empty union of components of L, and $f(L)$ lies entirely in one component of $\widetilde{Y} - p_1(\widetilde{Z}_0)$. (Think of L_0 as an *innermost* union of components of L.) Let $N_1 \subset \partial M - (L_0 \times (-1,1))$ be the union of components having boundary $L_0 \times \{1\}$. That is, N_1 and its complement N_{-1} in $\partial M - (L_0 \times (-1,1))$ map to opposite sides of $\widetilde{Z}_0 \times (-1,1)$ in $\widetilde{Y}$, and in fact $f(N_1) \subset p_1(\widetilde{Z}_0) \subset \widetilde{X}$, by the choice of $\widetilde{Z}_0$.

Our method now is to fill L_0 with a least-volume copy of N_1 and then fill the two sides of ∂M efficiently by M (using the induction hypothesis) and $N_1 \times I$. These fillings fit together to yield a filling of f by M having the required volume.

Let $v = \mathrm{Vol}^n(f)$ and $u = \mathrm{Vol}^{n-1}(p_0 \circ f|_{L_0})$ (which is equal to $\mathrm{Vol}^n(f|_{L_0\times(-1,1)})$ by (7.3)). Let $h\colon N_1 \to \widetilde{X}$ be a least-volume N_1–filling of $p_0 \circ f|_{L_0}$. Thus, $h|_{\partial N_1} = p_0 \circ f|_{L_0}$ and $\mathrm{Vol}^n(h) \leqslant Cu^s$. Define a new map $f'\colon \partial M \to \widetilde{Y}$ by first collapsing the fibers of $L_0 \times [-1,1]$ to points, and then sending N_{-1} by f and N_1 by h. Since h is least-volume and $L_0 \times [-1,1]$ was collapsed we have $\mathrm{Vol}^n(f') \leqslant v - u$. Also $(f')^{-1}(\widetilde{Z}) = L - L_0$, so by the induction hypothesis there is a map $g_{-1}\colon M \to \widetilde{Y}$ with $g_{-1}|_{\partial M} = f'$ such that

$$\mathrm{Vol}^{n+1}(g_{-1}) \leqslant C^{1/s}(v-u)^{2-1/s}.$$

Next let $g_1\colon N_1 \times [-1,1] \to \widetilde{Y}$ be a homotopy which begins with h on $N_1 \times \{-1\}$ and pushes across $\widetilde{Z}_0 \times (-1,1)$ and then deforms within $p_1(\widetilde{Z}_0)$ to $f|_{N_1}$, with the boundary fixed pointwise. This latter homotopy exists since $p_1(\widetilde{Z}_0)$ is contractible. Note that $\mathrm{Vol}^{n+1}(g_1) = \mathrm{Vol}^n(h)$ by (7.3) since $p_1(\widetilde{Z}_0)$ has dimension n.

Now join $N_1 \subset \partial M$ to $(N_1 \times \{-1\}) \subset N_1 \times [-1,1]$ to get a new copy of M and a map $g\colon M \to \widetilde{Y}$ extending g_{-1} and g_1. Then $g|_{\partial M} = f$ and

$$\mathrm{Vol}^{n+1}(g) \;\leqslant\; C^{1/s}(v-u)^{2-1/s} + v_h$$

where $v_h = \mathrm{Vol}^n(h)$. Now $s > 1$ and $v \geqslant u$ imply

$$\begin{aligned}\mathrm{Vol}^{n+1}(g) \;&\leqslant\; C^{1/s}(v-u)v^{1-1/s} + v_h\\ &=\; C^{1/s}v^{2-1/s}\left(1 - \frac{u}{v} + \frac{v^{(1/s)-1}v_h}{C^{1/s}v}\right).\end{aligned} \tag{8.2}$$

Recall that $v_h = \mathrm{Vol}^n(h) \leqslant \mathrm{Vol}^n(f|_{N_1}) \leqslant v$ because h is least-volume. Hence

$$\begin{aligned} 1 - \frac{u}{v} + \frac{v^{(1/s)-1} v_h}{C^{1/s} v} &\leqslant 1 - \frac{u}{v} + \frac{v_h^{(1/s)-1} v_h}{C^{1/s} v} \\ &= 1 - \frac{u}{v} + \frac{v_h^{1/s}}{C^{1/s} v}. \end{aligned} \tag{8.3}$$

The main hypothesis implies that $v_h \leqslant Cu^s$, or $v_h^{1/s} \leqslant C^{1/s} u$, again because h is least-volume. Thus

$$1 - \frac{u}{v} + \frac{v_h^{1/s}}{C^{1/s} v} \leqslant 1 - \frac{u}{v} + \frac{u}{v} = 1. \tag{8.4}$$

By equations (**8.2**), (**8.3**), and (**8.4**) we have $\mathrm{Vol}^{n+1}(g) \leqslant C^{1/s} v^{2-1/s}$ where $v = \mathrm{Vol}^n(g|_{\partial M})$, which completes the proof. □

Definition 8.5 Let G be a group of type $\mathcal{F}_{k+1}$ and geometric dimension at most $k+1$. The k–dimensional Dehn function $\delta_G^{(k)}(x)$ *has embedded representatives* if there is a finite aspherical $(k+1)$–complex X, a sequence of embedded $(k+1)$–dimensional balls $B_i \subset \widetilde{X}$, and a function $F(x) \simeq \delta_G^{(k)}(x)$, such that the sequence given by $(n_i) = (\mathrm{Vol}^k(\partial B_i))$ tends to infinity and is exponentially bounded, and $\mathrm{Vol}^{k+1}(B_i) \geqslant F(n_i)$ for each i.

The lower bounds established in this article for various Dehn functions are all obtained by constructing embedded representatives and applying Remarks 2.1 and 2.7. In particular the k–dimensional Dehn functions of $\Sigma^{k-1} G_{r,P}$ and $\Sigma^{k-1}\mathbb{Z}^2$ have embedded representatives.

The next result generalizes [1, Theorem 6.3] to higher dimensions.

Proposition 8.6 *Let G be a group of type $\mathcal{F}_{k+1}$ and geometric dimension at most $k+1$. Suppose the k–dimensional Dehn function $\delta^{(k)}(x)$ of G is equivalent to x^s and has embedded representatives. Then $G \times \mathbb{Z}$ has $(k+1)$–dimensional Dehn function $\delta^{(k+1)}(x) \succcurlyeq x^{2-1/s}$, with embedded representatives.*

Proof We establish the lower bound $\delta^{(k+1)}(x) \succcurlyeq x^{2-1/s}$ for $G \times \mathbb{Z}$ as follows. Since $\delta_G^{(k)}(x)$ has embedded representatives, let X, $F(x)$, B_i, and (n_i) be as in Definition 8.5; without loss of generality suppose that $F(x) = Cx^s$ for some $C > 0$. Define $m_i = 3\,\mathrm{Vol}^{k+1}(B_i)$. The space $Y = X \times S^1$ has fundamental group $G \times \mathbb{Z}$ and universal cover $\widetilde{Y} = \widetilde{X} \times \mathbb{R}$. Consider the $(k+2)$–dimensional balls

$$C_i = B_i \times [0, m_i/3n_i] \subset \widetilde{Y}.$$

The boundary of C_i is $\partial B_i \times [0, m_i/3n_i] \cup B_i \times \partial[0, m_i/3n_i]$ which implies that

$$\mathrm{Vol}^{k+1}(\partial C_i) = m_i.$$

We also have $\mathrm{Vol}^{k+2}(C_i) = \mathrm{Vol}^{k+1}(B_i)m_i/3n_i = (m_i)^2/9n_i$ for each i. Since $m_i = 3\,\mathrm{Vol}^{k+1}(B_i) \geqslant 3C(n_i)^s$ we have $(3C)^{-1/s}(m_i)^{1/s} \geqslant n_i$. Then

$$\mathrm{Vol}^{k+2}(C_i) = \frac{(m_i)^2}{9n_i} \geqslant \left(\frac{C^{1/s}}{3^{2-1/s}}\right)(m_i)^{2-1/s}.$$

Note that $\widetilde{Y}$ is aspherical and has dimension $k+2$, and so C_i is a least-volume ball (cf Remark 2.7). Therefore $\delta^{(k+1)}(m_i) \geqslant (C^{1/s}/3^{2-1/s})(m_i)^{2-1/s}$ for each i. Now it remains to check that the sequence (m_i) has the required properties. It tends to infinity since $m_i \geqslant 3C(n_i)^s$. Also each ball $B_i \subset \widetilde{X}$ is least-volume, so there is a constant D such that $m_i \leqslant D(n_i)^s$ for all i.[1] Then $m_{i+1}/m_i \leqslant (D/C)(n_{i+1}/n_i)^s$, which is bounded. Now Remark 2.1 implies that $\delta^{(k+1)}(x) \succcurlyeq x^{2-1/s}$. □

We are now in a position to prove Theorem D.

Proof of Theorem D Fix r, P, and q, let $s(\ell) = \frac{2(\ell+1)\alpha-\ell}{2\ell\alpha-(\ell-1)}$, and let G_ℓ be the group $\Sigma^{q-1}G_{r,P} \times \mathbb{Z}^\ell$. (Or let $s(\ell) = \frac{\ell+2}{\ell+1}$ and $G_\ell = \Sigma^{q-1}\mathbb{Z}^2 \times \mathbb{Z}^\ell$.) We verify by induction on ℓ the following statements for G_ℓ:

(1) $\delta^M(x) \leqslant Cx^{s(\ell)}$ for all $(q+\ell+1)$–manifolds M and some constant $C > 0$,

(2) $\delta^{(q+\ell)}(x) \succcurlyeq x^{s(\ell)}$, and

(3) $\delta^{(q+\ell)}(x)$ has embedded representatives.

The first two statements together imply $\delta^{(q+\ell)}(x) \simeq x^{s(\ell)}$.

If $\ell = 0$ then (1) follows from Theorem 7.2 and Lemma 7.4. Statement (2) holds by Theorem C, and we have already observed that (3) holds for these groups.

For $\ell > 0$ note first that $s(\ell) = 2 - 1/s(\ell-1)$. Then statement (1) holds by Theorem 8.1 and property (1) of $G_{\ell-1}$. Proposition 8.6 implies (2) and (3) by properties (1)–(3) of $G_{\ell-1}$. □

[1] Here we are using the upper bound for $\delta^{(k)}_G(x)$.

References

[1] J. M. Alonso, W. A. Bogley, R. M. Burton, S. J. Pride, and X. Wang, *Second order Dehn functions of groups*, Quart. J. Math. Oxford Ser. (2), 49 (1998), pp. 1–30.

[2] J. M. Alonso, X. Wang, and S. J. Pride, *Higher-dimensional isoperimetric (or Dehn) functions of groups*, J. Group Theory, 2 (1999), pp. 81–112.

[3] N. Brady and M. R. Bridson, *There is only one gap in the isoperimetric spectrum*, Geom. Funct. Anal., 10 (2000), pp. 1053–1070.

[4] N. Brady and M. Forester, *Density of isoperimetric spectra*. Preprint, 2008.

[5] M. R. Bridson, *The geometry of the word problem*, in Invitations to geometry and topology, vol. 7 of Oxf. Grad. Texts Math., Oxford Univ. Press, Oxford, 2002, pp. 29–91.

[6] ——, *Polynomial Dehn functions and the length of asynchronously automatic structures*, Proc. London Math. Soc. (3), 85 (2002), pp. 441–466.

[7] S. Buoncristiano, C. P. Rourke, and B. J. Sanderson, *A geometric approach to homology theory*, Cambridge University Press, Cambridge, 1976. London Mathematical Society Lecture Note Series, No. 18.

[8] J. Burillo, *Lower bounds of isoperimetric functions for nilpotent groups*, in Geometric and computational perspectives on infinite groups (Minneapolis, MN and New Brunswick, NJ, 1994), vol. 25 of DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Amer. Math. Soc., Providence, RI, 1996, pp. 1–8.

[9] T. Coulhon and L. Saloff-Coste, *Isopérimétrie pour les groupes et les variétés*, Rev. Mat. Iberoamericana, 9 (1993), pp. 293–314.

[10] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston, *Word processing in groups*, Jones and Bartlett Publishers, Boston, MA, 1992.

[11] M. Gromov, *Hyperbolic groups*, in Essays in group theory, vol. 8 of Math. Sci. Res. Inst. Publ., Springer, New York, 1987, pp. 75–263.

[12] ——, *Asymptotic invariants of infinite groups*, in Geometric group theory, Vol. 2 (Sussex, 1991), vol. 182 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1993, pp. 1–295.

[13] A. Hatcher and K. Vogtmann, *Isoperimetric inequalities for automorphism groups of free groups*, Pacific J. Math., 173 (1996), pp. 425–441.

[14] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, vol. 54 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.

[15] A. Y. Ol′shanskii, *Groups with quadratic-non-quadratic Dehn functions*, Internat. J. Algebra Comput., 17 (2007), pp. 401–419.

[16] A. Y. Ol′shanskii and M. V. Sapir, *Groups with small Dehn functions and bipartite chord diagrams*, Geom. Funct. Anal., 16 (2006), pp. 1324–1376.

[17] P. Papasoglu, *Isodiametric and isoperimetric inequalities for complexes and groups*, J. London Math. Soc. (2), 62 (2000), pp. 97–106.

[18] M. V. Sapir, J.-C. Birget, and E. Rips, *Isoperimetric and isodiametric functions of groups*, Ann. of Math. (2), 156 (2002), pp. 345–466.

[19] E. Seneta, *Nonnegative matrices and Markov chains*, Springer Series in Statistics, Springer-Verlag, New York, second ed., 1981.

[20] N. T. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups*, vol. 100 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 1992.

[21] X. Wang, *Second order Dehn functions of finitely presented groups and monoids*, 1996. PhD Thesis, University of Glasgow.

[22] X. Wang and S. J. Pride, *Second order Dehn functions and HNN-extensions*, J. Austral. Math. Soc. Ser. A, 67 (1999), pp. 272–288. Group theory.

[23] R. Young, *A note on higher-order filling functions*. Preprint, arxiv:0805.0584.

Mathematics Department, University of Oklahoma
Norman, OK 73019, USA

Mathematical Institute, 24-29 St Giles'
Oxford, OX1 3LB, UK

Mathematics Department, University of Oklahoma
Norman, OK 73019, USA

Mathematics Department, University of Oklahoma
Norman, OK 73019, USA

nbrady@math.ou.edu, bridson@maths.ox.ac.uk, forester@math.ou.edu, shankar@math.ou.edu