Filling Length in Finitely Presentable Groups

Steve M Gersten *

Mathematics Department, University of Utah, Salt Lake City, UT 84112, USA
`gersten@math.utah.edu`
`http://www.math.utah.edu/~gersten`

Tim R Riley †

Mathematical Institute, 24-29 St Giles, Oxford, OX1 3LB, UK
`rileyt@maths.ox.ac.uk`

Abstract. We study the filling length function for a finite presentation of a group Γ, and interpret this function as an optimal bound on the length of the boundary loop as a van Kampen diagram is collapsed to the basepoint using a combinatorial notion of a null-homotopy. We prove that filling length is well behaved under change of presentation of Γ.

We look at "AD-pairs" (f, g) for a finite presentation $\mathcal{P}$: that is, an isoperimetric function f and an isodiametric function g that can be realised simultaneously. We prove that the filling length admits a bound of the form $[g + 1][\log(f + 1) + 1]$ whenever (f, g) is an AD-pair for $\mathcal{P}$. Further we show that (up to multiplicative constants) if x^r is an isoperimetric function ($r \geq 2$) for a finite presentation then (x^r, x^{r-1}) is an AD-pair. Also we prove that for all finite presentations filling length is bounded by an exponential of an isodiametric function.

Keywords: finitely presented group, filling length, isoperimetric function

Mathematics Subject Classification (1991): 20F06 (Primary); 20F05, 20F32, 57M07

1. Introduction

Let $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ be a finite presentation for a group Γ. The filling length function for $\mathcal{P}$ is one of a number of functions that arise from considering *null-homotopic* words; that is, words that evaluate to 1 in Γ. (A *word* is an element of the free monoid $(\mathcal{A} \cup \mathcal{A}^{-1})^\star$ on the generators $\mathcal{A}$ and their formal inverses.) *Null-homotopic words* are characterised by the existence of a *null-sequence*, defined as follows.

DEFINITION. A *null-sequence* for a null-homotopic word w is a sequence of words $w = w_0, w_1, \ldots, w_m = 1$ such that each w_{i+1} is obtained from w_i by one of three moves.

1. *Elementary (or free) reduction.* Remove a subword aa^{-1} from w_i, where a is a generator or the inverse of a generator.

2. ***Elementary (or free) expansion.*** Insert a subword aa^{-1} into w_i, where a is a generator or the inverse of a generator. So $w_{i+1} = uaa^{-1}v$ for some words u, v such that $w_i = uv$ in $(\mathcal{A} \cup \mathcal{A}^{-1})^\star$.

3. ***Application of a relator.*** Replace $w_i = \alpha u\beta$ by $\alpha v\beta$, where uv^{-1} is a cyclic conjugate of one of the defining relators or its inverse and $v^{-1} = a_s^{-1}a_{s-1}^{-1} \ldots a_1^{-1}$ if $v = a_1 \ldots a_{s-1}a_s$, where a_i are generators or inverses of generators of $\mathcal{P}$.

All this takes place in $(\mathcal{A} \cup \mathcal{A}^{-1})^\star$. Here we have denoted the empty word by "1", the neutral element in this monoid. Let $\ell(w)$ denote length of a word w in $(\mathcal{A} \cup \mathcal{A}^{-1})^\star$; that is, $\ell(w)$ is the number of letters in w.

DEFINITION. Let w be a null-homotopic word in $\mathcal{P}$ and let S denote a null-sequence $w = w_0, w_1, \ldots, w_m = 1$ for w. We let $H(S)$ denote the maximum of $\ell(w_i)$, $0 \leq i \leq m$, and we define $h_0(w)$ to be the minimum of $H(S)$ as S ranges over all null-sequences for w. We define the ***filling length function*** $h_0 : \mathbb{N} \to \mathbb{N}$ by setting $h_0(n)$ to be the maximum of $h(w)$ as w ranges over all null-homotopic words of length at most n.

Two better known filling functions, whose definitions we will recall in the next section, are the Dehn function f_0 (also known as the optimal isoperimetric function) and the optimal isodiametric function g_0. We say that f and g are respectively isoperimetric and isodiametric functions for Γ if $f_0(n) \leq f(n)$ and $g_0(n) \leq g(n)$ for all n. We make the following definition.

DEFINITION. An ***AD-pair*** for the finite presentation $\mathcal{P}$ is a pair of functions (f, g) from $\mathbb{N}$ to $\mathbb{N}$ such that for every circuit w of length at most n in the Cayley graph there exists a van Kampen diagram $\mathcal{D}$ with area at most $f(n)$ and diameter at most $g(n)$. (So, in particular, f is an isoperimetric function and g is an isodiametric function.)

The similar notion of an AR-pair is discussed in [9], where radius is considered in place of diameter.

DEFINITION. Given two functions $f_1, f_2 : (0, \infty) \to (0, \infty)$ we say $f_1 \preceq f_2$ when there exists $C > 0$ such that for all $l \in (0, \infty)$, $f_1(l) \leq Cf_2(Cl + C) + Cl + C$. This yields the equivalence relation: $f_1 \simeq f_2$ if and only if $f_1 \preceq f_2$ and $f_2 \preceq f_1$.

It is natural to ask how f_0, g_0 and h_0 relate to each other. For a finitely presentation of a group the following hold. For all n

$$2g_0(n) \leq h_0(n) \leq 2Kf_0(n) + n,$$

where $K := \max\{\ell(r) : r \in \mathcal{R}\}$. Further

$$f_0(n) \preceq \exp h_0(n), \text{ and}$$
$$f_0(n) \preceq \exp\exp(g_0(n) + n).$$

D. E. Cohen and the first author proved a double exponential bound for f_0 in terms of g_0 (see [3] and [6]). In fact, moreover, for an isodiametric function g for a finite presentation $\mathcal{P}$, there exists $E > 1$ such that $(E^{E^{g(x)+x}}, g(x))$ is an AD-pair for $\mathcal{P}$. (See [6].) We will prove the remainder of the above results in the next section.

We remark that it follows from the inequalities above that the word problem in Γ is solvable if and only any one of f_0, g_0 and h_0 is a recursive function. See [7] for example.

It is an open question (to our knowledge first raised by John Stallings) whether the double exponential bound can be improved to single exponential. (In the most extreme known examples f_0 and g_0 differ exponentially - for example $\langle a, b \mid b^{-1}ab = a^2 \rangle$.) In the light of the single exponential bound on f_0 in terms of h_0, it would seem that studying h_0 is a natural approach to this question. Our first theorem offers a bound on h_0 in terms of an AD-pair for a finite presentation.

THEOREM 1. Let (f, g) be an AD-pair for the finite presentation $\mathcal{P}$. Then $h_0(n) \preceq [g(n) + 1][\log(f(n) + 1) + 1]$ for all n.

It is an open question whether this result can be improved to $h_0 \preceq g_0$. Gromov observed in [10], 5C that if such a bound could be found then using the bound $f_0 \preceq \exp h_0$ one could deduce a positive answer to Stallings' question.

Our second theorem gives a new example of an AD-pair.

THEOREM 2. Let Γ be a group admitting a polynomial isoperimetric function of degree $r \geq 2$. Then up to a common multiplicative constant (x^r, x^{r-1}) is an AD-pair for Γ.

Our third theorem is the result of applying Theorem 1 to the AD-pair $(E^{E^{g(x)+x}}, g(x))$ given above. (The factor involving $g(n) + 1$ that one obtains by applying Theorem 1 can always be absorbed into the exponential factor $E^{g(n)+n}$.)

THEOREM 3. For a general finite presentation the filling length function is bounded by an exponential in the optimal isodiametric function: $h_0(n) \preceq \exp(g_0(n) + n)$.

2. The geometric interpretation of filling length

As before, $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ is a finite presentation for a group Γ.

DEFINITION. A ***van Kampen diagram*** $\mathcal{D}$ for a null-homotopic word w in $\mathcal{P}$ is a finite, planar, contractible, combinatorial 2-complex; its 1-cells are directed and labelled by generators and the boundary labels of each of its 2-cells are cyclic conjugates of relators or inverse relators. Further the boundary label for $\mathcal{D}$ is w when read (by convention anticlockwise) from a base point $\star$ in $\partial \mathcal{D}$. (See page 155 of [2], or pages 233ff of [11].)

Equivalently one can define a van Kampen diagram $\mathcal{D}$ as a combinatorial cell structure on S^2 together with a distinguished 2-cell e_∞ and a combinatorial map f from $S^2 \setminus e_\infty$ to the presentation 2-complex of $\mathcal{P}$. The attaching map of e_∞ is then mapped by f to w.

We can consider $\mathcal{D}$ to be providing a *homotopy disc* for w. Then, in analogy with the study of the homotopy discs filling null-homotopic loops in a Riemannian manifold, we associate various filling invariants to the geometry of van Kampen diagrams for null-homotopic words w. (Many such filling invariants are discussed by Gromov in Chapter 5 of [10].)

The Dehn function arises from considering the (combinatorial) area Area$(\mathcal{D})$ of $\mathcal{D}$; that is, the number of 2-cells in $\mathcal{D}$. The optimal isodiametric function concerns the diameter Diam$(\mathcal{D})$, which is defined to be the maximum over all vertices v of $\mathcal{D}$ of the distance in the 1-skeleton of $\mathcal{D}$ from v to the base point of $\mathcal{D}$. (The 1-skeleton is endowed with the metric that uniformly gives each 1-cell length 1.) The Dehn function $f_0 : \mathbb{N} \to \mathbb{N}$ and the optimal isodiametric function $g_0 : \mathbb{N} \to \mathbb{N}$ for the finite presentation $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ of a group Γ are defined by

$$f_0(n) := \max \left\{ \text{Area}(w) : \text{words } w \text{ with } \ell(w) \le n \text{ and } w = 1 \text{ in } \Gamma \right\},$$

$$g_0(n) := \max \left\{ \text{Diam}(w) : \text{words } w \text{ with } \ell(w) \le n \text{ and } w = 1 \text{ in } \Gamma \right\}.$$

From an algebraic point of view the Dehn function $f_0(n)$ is the least N such that for every null-homotopic word w with $\ell(w) \le n$ there is an equality

$$w = \prod_{i=1}^{N} u_i^{-1} r_i u_i$$

in the free group $F(\mathcal{A})$ for some relators $r_i \in \mathcal{R}^{\pm 1}$, and words u_i. Similarly (up to an additive constant, $\frac{1}{2} \max \{ \ell(r) : r \in \mathcal{R} \}$), the optimal isodiametric function $g_0(n)$ is the minimal bound on the length of the conjugating words u_i, amongst such equalities in $F(\mathcal{A})$ for null-homotopic words w with $\ell(w) \le n$.

Let us pursue further the analogy of $\mathcal{D}$ with a homotopy disc. Filling length concerns the length of the boundary curve contracted across $\mathcal{D}$ in a combinatorial notion of a null-homotopy. The situation in a Riemannian manifold X is that one considers contracting a null-homotopic loop $\gamma : [0, 1] \to X$ based at $\star \in X$ to the constant loop at $\star$. By definition there is some continuous map $H : [0, 1] \times [0, 1] \to X$ denoted by $H_t(s) = H(t, s)$ with $H_0 = \gamma$, $H_1(s) = \star$ for all s, and $H_t(0) = H_t(1) = \star$ for all t. *Filling length* is the optimal bound on the length of the loops H_t. So the filling length of H is the supremum of the lengths of the loops H_t for $t \in [0, 1]$, and the filling length of γ is the infimum of the filling lengths of all possible null-homotopies H. Then (using Gromov's notation) define $\text{Fill}_0 \text{ Leng} \ell$ to be the supremum of the filling lengths of all null-homotopic loops $\gamma : [0, 1] \to X$ of length at most ℓ and based at $\star \in X$.

We give a combinatorial notion of a null-homotopy across a van Kampen diagram. A ***shelling*** (also known as an ***combinatorial null-homotopy***) of $\mathcal{D}$ is a sequence $\mathcal{D} = \mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_m = \star$ of van Kampen diagrams where $\mathcal{D}_{i+1}$ is obtained from $\mathcal{D}_i$ by one of the following three types of moves.

A. ***1-cell collapse.*** Remove a pair (e^1, e^0) such that $e^0 \in \partial e^1$ is a 0-cell in $\mathcal{D}_i$ which is not the base point $\star \in \mathcal{D}_i$, and e^1 is a 1-cell only attached to the rest of the diagram at one 0-cell which is not e^0.

B. ***1-cell expansion.*** Suppose (e^1, e^0) is a pair such that e^1 is a 1-cell in the interior of $\mathcal{D}_i$ and $e^0 \in \partial e^1 \cap \partial \mathcal{D}_i$. Make a cut along e^1 starting from e^0, so two copies of e^0 and e^1 are found in $\mathcal{D}_{i+1}$. This has the effect of introducing two new 1-cells into the boundary of the diagram.

C. ***2-cell collapse.*** Remove a pair (e^2, e^1) where e^2 is a 2-cell of $\mathcal{D}_i$ with e^1 a 1-cell of $\partial e^2 \cap \partial \mathcal{D}_i$ (note that the 0-skeleton of $\mathcal{D}_i$ is the same as that of $\mathcal{D}_{i+1}$).

Notice that the effect on the boundary word of performing A, B or C is an instance of move 1, 2, or 3 (respectively) from the definition of a null-sequence given in the introduction. In the case of move B, if $e^0 = \star$ then the free expansion of the boundary word occurs either at the start or at the end of the word, depending on which of the two copies of e^0 in $\mathcal{D}_{i+1}$ one takes to be the new base point.

For a shelling $S = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_m)$ of $\mathcal{D}$, let w_i be the boundary word of $\mathcal{D}_i$ and let

$$H(S) := \max \{\ell(w_i) : i = 0, 1, \ldots, m\}.$$

Define $FL(\mathcal{D})$, the filling length of $\mathcal{D}$, to be the minimum of $H(S)$ over all shellings S of $\mathcal{D}$. Define $FL(w)$ to be the minimum of $FL(\mathcal{D})$ over all van Kampen diagrams for w.

We claim that $FL(w)$ is the same as $h_0(w)$ where h_0 is the filling length function defined in the introduction using *null-sequences.*

PROPOSITION 1. *We have $h_0(w) = FL(w)$ for all null-homotopic words w.*

Proof. Let $\mathcal{D}$ be a van Kampen diagram for w that gives an optimal shelling $\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_m$ and let $w = w_0, w_1, \ldots, w_m = 1$ be the boundary labels of the diagrams in the shelling read as ordinary (not cyclic) words from the base point. Then $w_0, w_1, \ldots, w_m$ is a null-sequence and so $h(w) \leq \max\{\ell(w_i) : 0 \leq i \leq m\} = FL(w)$.

To obtain the opposite inequality, let S be a null-sequence $w = w_0, w_1, \ldots, w_m = 1$ for w with $H(S) = h(w)$. We need

LEMMA 1. *Under the assumptions above there is a van Kampen diagram $\mathcal{D}$ for w and a shelling of $\mathcal{D}$ whose sequence of intermediate boundary labels have length at most $h_0(w)$.*

Proof. We begin by drawing w_i, $0 \leq i \leq m$, as the horizontal arc in the Cartesian plane of length $\ell(w_i)$ beginning at $(0, i)$ and ending at $(\ell(w_i), i)$, subdivided at integral points, oriented from left to right, with the segment $[(i, j), (i, j + 1)]$ labelled by the j-th letter of w_i. We shall define maps f_i on subsets of the strip $Y_i = \{(x, y) \mid i \leq y \leq i + 1\}$ with contractible fibres, where the definition depends on the type of operation that is applied to go from w_i to w_{i+1}. There are three cases to consider.

1. The operation is an *elementary expansion,* so $w_i = uv \to uxx^{-1}v$, with x a generator or inverse generator, and with u of length j. In this case we shall construct an explicit retraction f_i of the strip Y_i onto the union of the line $y = i$ and the line segment $[(j, i), (j + 1, i + 1)]$. The retraction is vertical projection on (x, y) if $x \leq j$ and is a retraction downward along a line of slope $1/2$ if $x \geq j$. In the triangle $\Delta[(j, i), (j, i + 1), (j + 1, i + 1)]$ the retraction is vertical projection onto the side $[(j, i), (j + 1, i + 1)]$. And in the triangle $\Delta[(j, i), (j + 1, i + 1), (j + 2, i + 1)]$ the retraction is downward along lines of slope $1/2$ onto the side $[(j, i), (j + 1, i + 1)]$.

The crucial features of the map f_i that we need are

a) f_i folds the segment $[(j, i + 1), (j + 2, i + 1)]$ onto the segment $[(j, i), (j + 1, i + 1)]$, and

b) the fibres of the map f_i are finite trees.

The decomposition space of the map f_i on Y_i is in this case the result of an elementary expansion applied to the line $y = i$, where the edge adjoined, when oriented upward, has label x.

2. The operation is an *elementary reduction.* In this case we interchange the roles of i and $i+1$ in the definition of the map f_i. Again properties a) and b) hold for this map.

3. The operation is an *application of a relator.* Let the operation be $w_i = \alpha u\beta \to \alpha v\beta$, where uv^{-1} is a cyclic conjugate of a defining relator, which will be assumed to be a nonempty word. Let $\ell(\alpha) = j$, $\ell(u) = s$, and $\ell(v) = t$.

We define an equivalence relation on Y_i whose equivalence classes consist of segments $[(x, i), (x, i + 1)]$ if $x \le j$ and segments parallel to and to the right of the segment $[(i, j + s), (i + 1, j + t)]$. Thus each equivalence class consists of either a single segment, or, in case $s = 0$ or $t = 0$, the union of two segments with a point in common. The map f_i is the decomposition map for this equivalence relation, and it follows that all fibres of f_i are finite trees.

Observe that the effect of taking the decomposition space of the map f_i to Y_i is to attach to the line $y = i$ an arc of length t beginning at (j, i) and ending at $(j + s, i)$ with label v and to attach a 2-cell with boundary label uv^{-1}.

We now take the decomposition space Y of the plane determined by imposing the equivalence relations of the fibres of all the maps f_i, $0 \le i \le m - 1$, and let f be the corresponding map, $f : \mathbb{R}^2 \to Y$. Although the fibres of each f_i are contractible, it will not be the case in general that the fibres of f are contractible (for example, if we do an elementary expansion and then the same reduction, that will introduce a cycle in the fibre of f). There will be 1-dimensional cycles occurring in a fibre, and these can pinch off 2-sphere components of Y, among other things. After we throw away all the cycles, the result is a van Kampen diagram diagram $\mathcal{D}$ for w (precisely, if we take maximal cycles in the fibre of f, in the sense that these do not contain other cycles in their interior, and collapse their interiors to points, then the resulting decomposition space is $\mathbb{R}^2$, as an application of the Vietoris mapping theorem, since the fibres of the resulting map are contractible; the cell structure and labelling for the van Kampen diagram comes from the the cell structure and labelling exhibited for the decomposition spaces of the maps f_i above in cases 1 and 3).

Let $\bar{w}_i$ be the images of the w_i after the portions in the 2-sphere components are discarded. Then $w = \bar{w}_0, \bar{w}_1, \ldots, \bar{w}_m = 1$ is a null-sequence (except that it is possible that $\bar{w}_i = \bar{w}_{i+1}$ for some i). If $\bar{w}_{i+1}$ is

the result of an *elementary reduction* in $\bar{w}_i$ then $\mathcal{D}_{i+1}$ is the consequence of performing a *1-cell collapse* on $\mathcal{D}_i$. Similarly if $\bar{w}_{i+1}$ is obtained from $\bar{w}_i$ by a *free expansion* then $\mathcal{D}_{i+1}$ is the result of a *1-cell expansion* in $\mathcal{D}_i$. The relationship between $\mathcal{D}_i$ and $\mathcal{D}_{i+1}$ is more complicated when $\bar{w}_{i+1}$ is the result of an *applying a relator* to $\bar{w}_i$. Recall that $\bar{w}_i = \alpha u \beta$ and $\bar{w}_{i+1} = \alpha v \beta$, where uv^{-1} is a a cyclic conjugate of a relator or an inverse relator. If u is not the empty word then $\mathcal{D}_{i+1}$ is obtained from $\mathcal{D}_i$ by a *2-cell collapse* followed by a sequence of *1-cell collapses*. If u is the empty word let a be the first letter of v, and then $\mathcal{D}_{i+1}$ can be produced from $\mathcal{D}_i$ by a 1-cell expansion that inserts aa^{-1} into the boundary word of $\mathcal{D}_i$, followed by a 2-cell collapse.

The diagrams of the resulting shelling of $\mathcal{D}$ have boundary word of length at most $\max\{\ell(\bar{w}_i) : i = 0, 1, \ldots, m\} \leq h_0(w)$.

This completes the proof of Lemma 1.

We can now complete the proof of Proposition 1. It follows from Lemma 1 that

$$\mathrm{FL}(w) \leq \max\{\ell(\bar{w}_i) : 1 \leq i \leq m\} \leq \max\{\ell(w_i) : 1 \leq i \leq m\} = h(w).$$

Since we have already shown that $h(w) \leq \mathrm{FL}(w)$, it follows that $\mathrm{FL}(w) = h(w)$, and the proof of the Proposition 1 is complete.

We can now deduce two results given in the introduction:

COROLLARY 1. *For a finite presentation* $\langle \mathcal{A} \mid \mathcal{R} \rangle$, *for all* n,

$$2g_0(n) \leq h_0(n) \leq 2Kf_0(n) + n,$$

where $K := \max\{\ell(r) : r \in \mathcal{R}\}$.

Proof. Let $\mathcal{D}$ be a van Kampen diagram for a null-homotopic word w of length n. Let $S = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_m)$ be any shelling of $\mathcal{D}$. The boundary circuit of each $\mathcal{D}_i$ corresponds to a circuit in the 1-skeleton $\mathcal{D}$ of (which may involve back-tracking). Given any vertex v in $\mathcal{D}$, there must be some $\mathcal{D}_i$ whose circuit passes through v. But this provides two paths in the 1-skeleton of $\mathcal{D}$ from v to $\star$, from which we deduce the inequality that the diameter is bounded above by half the filling length. The second inequality holds because the boundary words of the diagrams $\mathcal{D}_i$ have length at most twice the total length of the 1-skeleton of $\mathcal{D}$, which is at most $2Kf_0(n) + n$.

COROLLARY 2. *For a finite presentation*

$$f_0 \preceq \exp h_0.$$

Proof. This bound is due to the first author and Gromov (see pages 100ff of [10]). Let w be a null-homotopic word. Take any van Kampen diagram $\mathcal{D}_0$ for w and any shelling $S = (\mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_m)$ for $\mathcal{D}_0$. For $i = 0, 1, \dots, m$ let w_i be the boundary word of $\mathcal{D}_i$. Each word w_i corresponds to a circuit p_i in $\mathcal{D}_0^{(1)}$ (possibly with backtracking). If there exists $k > j$ such that the words w_j and w_k are the same then perform the following surgery on $\mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_j$: cut along p_j and p_k, discard the portion between p_j and p_k and glue the remaining two portions along the equal words $w_j = w_k$. For $i = 0, 1, \dots, j$ define $\bar{\mathcal{D}}_i$ to be the diagram obtained in this way by surgery on $\mathcal{D}_i$. For $i = j+1, j+2, \dots, m-(k-j)$ define $\bar{\mathcal{D}}_i := \mathcal{D}_{i+(k-j)}$. The result is a new shelling $\bar{\mathcal{D}}_0, \bar{\mathcal{D}}_1, \dots, \bar{\mathcal{D}}_{m-(k-j)}$ of a van Kampen diagram $\bar{\mathcal{D}}_0$ for w, in which there is at least one repetition fewer in the sequence of boundary words.

Now repeat this procedure until we have a shelling for w in which the boundary labels of the diagrams are all distinct.

The area of a van Kampen diagram is the number of *2-cell collapse* moves in any of its shellings. There are at most $\exp(C\, h_0(n))$ words of length at most $h_0(n)$ for some constant $C > 0$. The bound on $f_0(n)$ follows and the proof of the corollary is complete.

The functions f_0, g_0 and h_0 are dependent on the choice of finite presentation $\mathcal{P}$ for Γ. However these functions do behave well on change of presentation in the following sense.

It is well known that if $\mathcal{P}$ and $\mathcal{Q}$ are finite presentations for the same group then the two Dehn functions defined with respect to the two presentations are $\simeq$-equivalent (as defined in the introduction). The same can be said of the optimal isodiametric function. (See, for example, [9].)

THEOREM 4. Suppose h_P and h_Q are the filling length functions defined with respect to two finite presentations $\mathcal{P}$ and $\mathcal{Q}$ for the same group. Then $h_P \simeq h_Q$.

Proof. Recall that Tietze's theorem [11] states that there is a finite sequence of Tietze operations and their inverses starting at $\mathcal{P}$ and terminating at $\mathcal{Q}$. There are two types of Tietze operations, types I and II, and their inverses. A type I operation $\mathcal{P} \to \mathcal{Q}$ adjoins a new free generator t and a new relator tu^{-1}, where u is a word in the generators of $\mathcal{P}$ (so u does not involve t), whereas a type II operation does not change the generators but adjoins a new relator R which is a consequence of the relations of $\mathcal{P}$.

Consider first the type I operation $\mathcal{P} \to \mathcal{Q}$ given by adjoining the new generator t and new relator tu^{-1}, where u does not involve t. In the argument that follows we shall assume that u is not the empty word.

A separate and easier argument must be given if u is empty, which we omit.

Suppose first that w is a null-homotopic word in the generators of $\mathcal{Q}$. The length of w increases by at most a constant factor $A = \ell(u)$ when one replaces all occurrences of t in S by u. Thus by at most $\ell(w)$ applications of a relator we obtain a new word w' in the generators of $\mathcal{P}$. Now take an optimal null-sequence for w' with respect to $\mathcal{P}$. It follows that $h_{\mathcal{Q}}(n) \leq h_{\mathcal{P}}(An)$ for all n.

Next suppose that w is a null-homotopic word in the generators of $\mathcal{P}$ (so w involves no t's) and let $w = w_0, w_1, \ldots, w_m$ be an extremal null-sequence for w in $\mathcal{Q}$. Thus $w_1, w_2, \ldots, w_{m-1}$ can involve t. Let $\bar{w}_i$ be the result of replacing each occurrence of $t^{\pm 1}$ by $u^{\pm 1}$, so $\bar{w}_i$ is a word in the generators of $\mathcal{P}$. The sequence $w_0, \bar{w}_1, \bar{w}_2, \ldots, \bar{w}_{m-1}, w_m$ can be refined to a null-sequence in $\mathcal{P}$ by inserting intermediate elementary expansions and reductions (for example, if w_i has a subword tt^{-1}, then $\bar{w}_i$ has the subword uu^{-1}; if $w_i \to w_{i+1}$ consists of removing tt^{-1}, then one can remove uu^{-1} from $\bar{w}_i$ by a sequence of $\ell(u)$ elementary reductions). Furthermore we have $\ell(\bar{w}_i) \leq \ell(u)\ell(w_i)$ (it is here that we use $\ell(u) \neq 0$). It follows from these observations that $h_{\mathcal{P}}(w) \leq \ell(u)h_{\mathcal{Q}}(w)$, and hence $h_{\mathcal{P}}(n) \leq \ell(u)h_{\mathcal{Q}}(n)$ for all n.

Combining the inequalities of the previous two paragraphs gives $h_{\mathcal{Q}}(n) \leq h_{\mathcal{P}}(An)$ and $h_{\mathcal{Q}}(n) \leq \ell(u)h_{\mathcal{Q}}(n)$, which proves the equivalence for type I Tietze operations.

Next assume that $\mathcal{Q}$ is obtained from $\mathcal{P}$ by a type II Tietze operation, where we adjoin as new relator the relation R.

It is clear that $h_{\mathcal{Q}} \leq h_{\mathcal{P}}$ because there are more null-sequences for null-homotopic words with respect to $\mathcal{Q}$.

Let us seek a bound in the other direction. The word R can be obtained from the empty word by a sequence $S_R = (u_0, u_1, \ldots, u_m)$ where $u_0 = 1$, $u_m = R$ and each u_{i+1} is obtained from u_i by either *free expansion*, *free reduction* or an *application of a relator* in the presentation $\mathcal{P}$. Similarly R^{-1} can be obtained via the sequence $S_R^{-1} := (u_0^{-1}, u_1^{-1}, \ldots, u_m^{-1})$. Let $M := \max\{\ell(u_i) : i = 0, 1, \ldots, m\}$.

Suppose S is a null-sequence for w with respect to the presentation $\mathcal{Q}$. We obtain a null-sequence for w with respect to the presentation $\mathcal{P}$ by using S_R in a process that replaces all applications of the relator R in the sequence S. Observe that if w_{i+1} is obtained from w_i by application of the relator R, then we can equally well produce w_{i+1} by inserting a copy of $R^{\pm 1}$ into w_i and then performing some free reductions. Further we can use the sequence $S_R^{\pm 1}$ to insert the copy of $R^{\pm 1}$. The result is a null-sequence for w with respect to the presentation $\mathcal{P}$, with an increase of at most $M + \ell(R)$ in the maximum length of the words in the sequence.

From the results above we deduce that $h_Q \leq h_P \leq h_Q + M + \ell(R)$. So $h_P \simeq h_Q$, and the proof of the theorem is complete.

3.   Examples and applications

1. ***Polynomial isoperimetric function.*** If Γ satisfies a polynomial isoperimetric function of degree $r \geq 2$ then (x^r, x^{r-1}) is an AD-pair for Γ up multiplicative constants. This is the conclusion of Theorem 2 and is proved in section 5.

It follows from Theorem 1 that $h_0(n) \preceq n^{r-1} \log(n+1)$. This contrasts with the inequality $h_0 \preceq f_0$ in the introduction.

2. ***Bridson's groups.*** The group Γ_m defined for $m \geq 2$ by

$$\langle a_1, \ldots, a_m, s, t, \tau \mid \text{for } i < m, \ s^{-1}a_i s = a_i a_{i+1},$$
$$[t, a_i] = [\tau, a_i] = [s, a_m] = [t, a_m] = [\tau, a_m t] = 1 \rangle$$

has AD-pair (x^{2m+1}, x^m) up to multiplicative constants. This family of examples is due to Bridson – see [1]. The group Γ_m is constructed as follows. Take the mapping torus of the free group automorphism whose action on homology is given by the matrix that has a single Jordan block of size m and eigenvalue 1. Then take a 2-step HNN extension: add a stable letter t which acts trivially on the free fibre, and add a further stable letter τ that commutes with the free subgroup generated by $\{a_1, a_2, \ldots, a_{m-1}, a_m t\}$. Bridson shows that in fact x^{2m+1} is the *optimal* isoperimetric function and x^m is the *optimal* isodiametric function.

Theorem 1 gives us bounds of $x^m \log x$ on their filling length functions, which is a significant improvement on the bounds x^{2m+1} obtained from the inequality $h_0 \preceq f_0$.

3. ***Arbitrary finite presentations.*** If $\mathcal{P}$ is an arbitrary finitely presentation with an isoperimetric function f, then $(f(x), f(x) + x)$ is AD-pair. (See Gersten [7], Lemma 2.2.)

Also, as discussed in the introduction, if g is an isodiametric function for $\mathcal{P}$ then there is some $E > 1$ such that $(E^{E^{g(x)+x}}, g(x))$ is an AD-pair. Applying Theorem 1 yields the conclusion of Theorem 3: that g_0 is bounded by an exponential in $g_0(n) + n$.

4. ***Asynchronously combable groups.*** Up to multiplicative constants, (E^x, x) for some $E > 1$ and $(xL(x), x)$ are both AD-pairs when Γ is an asynchronously combable group. Here the length function $L(n)$ is the maximum length of combing paths for group elements at distance at most n from the identity. That (E^x, x) is an AD-pair follows from

the linear bound on the filling length function and that $f_0 \preceq \exp h_0$; see Corollary 2. In particular (x^2, x) is an AD-pair when Γ is synchronously automatic since Γ then admits a combing in which the combing lines are quasi-geodesics (see [4], pages 84-86).

Asynchronously combable groups have linear bounds on their filling length functions. (Theorem 1 gives weaker bounds on the filling length.) This is a result of the first author - Theorem 3.1 on page 130 of [8], where the notation $LCNH_1$ is in this case what we call linearly bounded filling length. In essence the homotopy can be performed by contracting in the direction of the combing, so the contracting loop always remains normal to the combing lines. (See also [7] for definitions.)

4.   A bound on the filling length function

THEOREM 1.   *Let (f, g) be an AD-pair for the finite presentation $\mathcal{P}$. Then $h_0(n) \preceq [g(n) + 1][\log(f(n) + 1) + 1]$ for all n.*

First we give a lemma about rooted trees that we shall use in the proof of this theorem. Let $\mathcal{T}$ be a finite rooted tree in which each node has valence three except for the root (valence two) and the leaves (valence one).

Let $\mathcal{F}$ be a finite forest of such trees. The **visible nodes** of $\mathcal{F}$ are the roots. An **elementary shelling** is the removal of the root of one of its trees (together with the two edges that meet that root when the tree has more than one node). A **(complete) shelling** is a sequence of elementary shellings ending with the empty forest. The **visibility number** of a shelling of $\mathcal{F}$ is the maximum number of visible vertices occurring in the shelling. The visibility number $\mathrm{VN}(\mathcal{F})$ is the minimum visibility number of all shellings.

Let $\mathrm{N}(\mathcal{T})$ denote the number of nodes of $\mathcal{T}$.

LEMMA 2.   *Let the integer d be determined by $2^d - 1 < \mathrm{N}(\mathcal{T}) \leq 2^{d+1} - 1$. Then $\mathrm{VN}(\mathcal{T}) \leq d + 1$.*

Proof. To obtain this bound on $\mathrm{VN}(\mathcal{T})$ we shall perform each elementary shelling by always choosing a tree with the least number of nodes to shell first. This is what we call a **logarithmic shelling**.

We argue by induction on $\mathrm{N}(\mathcal{T})$, where the induction begins when $\mathrm{N}(\mathcal{T}) = 1$; in this case $d = 0$ and $\mathrm{VN}(\mathcal{T}) = 1$, as required.

For the induction step, assume that $\mathrm{N}(\mathcal{T}) > 1$ with $2^d - 1 < \mathrm{N}(\mathcal{T}) \leq 2^{d+1} - 1$. Removing the root of $\mathcal{T}$ produces two trees $\mathcal{T}_1, \mathcal{T}_2$. We let $\mathrm{N}(\mathcal{T}_1) \leq \mathrm{N}(\mathcal{T}_2)$. Let $2^{d_i} - 1 < \mathrm{N}(\mathcal{T}_i) \leq 2^{d_i + 1} - 1$ for $i = 1, 2$. By the induction hypothesis we have $\mathrm{VN}(\mathcal{T}_i) \leq d_i + 1$ for $i = 1, 2$. Since we shell

$\mathcal{T}_1$ first, we get $\mathrm{VN}(\mathcal{T}) \leq \max(\mathrm{VN}(\mathcal{T}_1)+1, \mathrm{VN}(\mathcal{T}_2)) \leq \max(d_1+2, d_2+1)$ by the induction hypothesis. There are now two cases, depending on whether $d_1 < d_2$ or $d_1 = d_2$.

Case 1. $d_1 < d_2$. In this case $\max(d_1 + 2, d_2 + 1) = d_2 + 1 \leq d + 1$, so we get $\mathrm{VN}(\mathcal{T}) \leq d + 1$ as required.

Case 2. $d_1 = d_2$. Here $\max(d_1 + 2, d_2 + 1) = d_1 + 2$. We have $2(2^{d_1} - 1) + 1 < \mathrm{N}(\mathcal{T}_1) + \mathrm{N}(\mathcal{T}_2) + 1 \leq 2(2^{d_1+1} - 1) + 1$, whence $2^{d_1+1} - 1 < \mathrm{N}(\mathcal{T}) \leq 2^{d_1+2} - 1$. It follows that $d = d_1 + 1$, and $\mathrm{VN}(\mathcal{T}) \leq d + 1$ as required.

This completes the induction, and the proof of Lemma 1 is complete.

COROLLARY 3. $\mathrm{VN}(\mathcal{T}) < \log_2(\mathrm{N}(\mathcal{T}) + 1) + 1$.

Proof. Write $2^d - 1 < \mathrm{N}(\mathcal{T}) \leq 2^{d+1} - 1$, so $\mathrm{VN}(\mathcal{T}) \leq d+1 < \log_2(\mathrm{N}(\mathcal{T}) + 1) + 1$, as required.

REMARK. Note that since $\mathrm{VN}(\mathcal{T})$ is an integer, the upper bound for $\mathrm{VN}(\mathcal{T})$ in the corollary can be replaced by the least integer bounding $\log_2(\mathrm{N}(\mathcal{T}) + 1)$ from above. Stated in this form, the result is sharp, as we see by taking T to be the complete rooted tree $\mathcal{T}(d)$ of depth d. In this case $\mathcal{T}(d)$ has $2^{d+1} - 1$ nodes, and the visibility number is $d + 1$.

We will deduce Theorem 1 from the following proposition.

PROPOSITION 2. *Suppose that Γ is the group of a finite triangular presentation $\mathcal{P}$ in which no generator represents identity element, and let w be a null-homotopic word with $n := \ell(w)$. Given a van Kampen diagram $\mathcal{D}$ for w with $\mathcal{D} := \mathrm{Diam}(\mathcal{D})$ and $A := \mathrm{Area}(\mathcal{D})$ we find*

$$\mathrm{FL}(\mathcal{D}) \leq (2D + 1)(\log_2(A + 1) + 1) + 4D + 1 + n.$$

REMARK. Any finite presentation $\langle \mathcal{A} \,|\, \mathcal{R} \rangle$ for a group Γ yields a finite *triangular* presentation for Γ. Such presentations are characterised by the length of relators being at most three. If $r \in \mathcal{R}$ is expressible in $F(\mathcal{A})$ as $w_1 w_2$ where $\ell(w_1), \ell(w_2) \geq 2$ then add a new generator a to $\mathcal{A}$, and in $\mathcal{R}$ replace r by $a^{-1} w_1$ and $a w_2$. A triangular presentation is achieved after a finite number of such transformations.

Furthermore the condition that no generator represents the identity element can be achieved by Tietze transformations [11]. For suppose that the generator a of $\mathcal{P}$ represents the identity element of the group Γ of the presentation. Then we may adjoin a as an additional relator and repeatedly use a combination of Tietze operations to remove all occurrences of a from the other relators. Finally, we can remove the

pair consisting of the generator a and relator a, thereby reducing both the number of generators and the number of relators by one.

Recall from section 2 that f_0, g_0 and h_0 are invariant up to $\simeq$-equivalence on change of finite presentation, and that we know from Proposition 1 that $\mathrm{FL}(w) = h_0(w)$ for all null-homotopic words w. So Proposition 2 is sufficient to prove Theorem 1.

Proof of Proposition 2. We start by taking a maximal geodesic tree $\mathcal{T}$ in the 1-skeleton of the van Kampen diagram $\mathcal{D}$, rooted at the base point $\star$ of $\mathcal{D}$. So from any vertex of $\mathcal{D}$ there is a path in $\mathcal{T}$ to $\star$ with length at most $\mathcal{D}$.

By cutting along paths in $\mathcal{T}$ we can decompose $\mathcal{D}$ into sub-diagrams $\mathcal{D}_i$ where only one edge from $\partial\mathcal{D} - \mathcal{T}$ occurs in each $\mathcal{D}_i$.

To achieve the bound claimed in the proposition we will shell $\mathcal{D}$ by shelling each $\mathcal{D}_i$ in turn. So let us first explain a means of shelling each $\mathcal{D}_i$. We will use six types of *2-cell collapse* moves. These are depicted in Figure 1, with the thicker lines representing edges in $\mathcal{T}$. The arrows indicate edges in the boundary of the van Kampen diagram. In each move a 1-cell in $\mathcal{D} - \mathcal{T}$ is removed. To see that these are all the 2-cell collapsing moves we require to shell $\mathcal{D}_i$, observe that $\mathcal{D}_i - \mathcal{T}$ is connected. Also recall that we assumed that in the presentation $\mathcal{P}$ no generator represents the identity - it follows that there are no degenerate triangles or bi-gons amongst the 2-cells of the diagram $\mathcal{D}$. That is, all triangles (resp. bi-gons) have three (resp. two) distinct 1-cell and 0-cells. Also note that there can be no 2-cell whose boundary consists of just one 1-cell and one 0-cell.
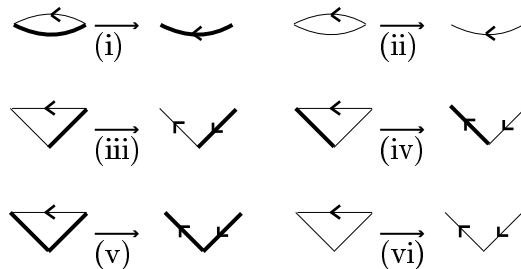

Figure 1. 2-cell collapse moves.

We now give the means of performing the shelling in a way that realises the bound on filling length. Repeatedly apply the following four steps:

1. 1-cell collapse (as defined in section 2),

2. moves (i) and (ii): bi-gon collapse,

3. moves (iii) and (iv),

4. moves (v) and (vi) in accordance with *logarithmic shelling*.

The first step in the list that is available is performed, and then we return to the start of the list. The effect is to shell the diagram $\mathcal{D}_i$ eventually leaving just the base point $\star$.

The means by which we use *logarithmic shelling* to choose which 2-cell to collapse when performing step 4 requires some explanation. Take the dual graph $\mathcal{G}$ of $\mathcal{D}_i - \mathcal{T}$, which is made up of vertices dual to faces in $\mathcal{D}_i$ and edges dual to edges in $\mathcal{D}_i^{(1)} - \mathcal{T}$. Now $\mathcal{G}$ is a tree for the following reason. If $\mathcal{G}$ contains a loop then it contains a simple loop. Consider $\mathcal{G}$ to be inscribed in $\mathcal{D}_i$ in the natural way. Then there would be a vertex of $\mathcal{D}_i$ in the interior of the simple loop and this could not be connected to $\partial \mathcal{D}_i$ by a path in $\mathcal{T}$, which would contradict maximality of $\mathcal{T}$.

The **shelling tree** is then obtained by reducing $\mathcal{G}$ in the following manner: for any vertex v which is dual to a 2-cell of the form (i), (ii), (iii) or (iv) of Figure 1, collapse one of the edges that meets v, at each stage reducing both the number of edges and the number of vertices by one. An example is illustrated in Figure 2 in which thicker lines represent edges in $\mathcal{T}$. The branching vertices of the shelling tree correspond to 2-cells in $\mathcal{D}$ of the form (vi) and the leaves correspond to 2-cells of the form (v). Observe that the shelling tree has a natural root: the vertex corresponding to the 2-cell in $\mathcal{D}_i$ of the form (vi) that is first reached on shelling $\mathcal{D}_i$ (and when there is no such 2-cell the shelling tree consists of just one vertex). Now the shelling tree is a finite rooted tree in which each node has valence three except for the root (valence two) and the leaves (valence one), *i.e.* a tree of form discussed in at the start of this section. When performing step 4 we chose the 2-cell to be pushed across in accordance with the process of logarithmic shelling of rooted trees used in Lemma 2. The number of nodes in the tree is at most A and so by Corollary 3 the visibility number is at most $\log_2(A + 1) + 1$.
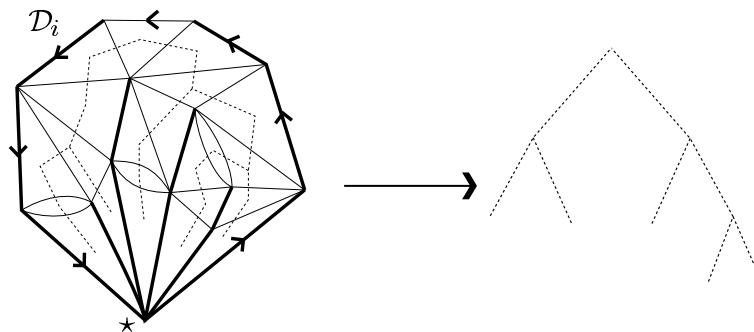

Figure 2. $\mathcal{D}_i$ and its shelling tree.

It remains to explain how to bound the filling length of the shelling of $\mathcal{D}_i$ given above, and then how to deduce a bound on the filling length of $\mathcal{D}$. Consider the situation when the next step to be applied is number 4. The visibility number associated to the shelling tree constructed above is at most $\log_2(A+1) + 1$. The boundary loop includes at most $\log_2(A+1) + 1$ edges of the type occurring in move (vi). These are separated by paths in $\mathcal{T}$ of length at most $2D$. The loop is closed by another path in $\mathcal{T}$ again of length at most $2D$. So this loop has length at most:

$$\log_2(A+1) + 1 + 2D\log_2(A+1) + 2D = (2D+1)(\log_2(A+1) + 1).$$

Now applying move (v) or (vi) increases the length of the loop by 1, creating two new *channels* where moves (i), (ii), (iii) and (iv) may be performed. (Channels correspond to paths in $\mathcal{G}$ between branching vertices.) Consider then applying steps 1, 2 and 3. Step 1 can only decrease the length, and step 2 leaves it unchanged. Step 3 can be applied at most $2D$ times in each of the two channels. Thus the increase in length before step 4 is next applied is at most $1 + 4D$. This gives the bound

$$\mathrm{FL}(\mathcal{D}_i) \ \leq \ (2D+1)(\log_2(A+1) + 1) + 4D + 1. \tag{1}$$

Now as one reads w around $\mathcal{D}$ from the basepoint one meets the subdiagrams $\mathcal{D}_1, \mathcal{D}_2, \ldots$ in turn. Shell the $\mathcal{D}_i$ in this order in the manner described above (except that when shelling $\mathcal{D}_i$, do not collapse the 1-cells of $\mathcal{D}_i \cap \mathcal{D}_{i+1}$). Consider the shelling of $\mathcal{D}$ at the stage where the subdiagram $\mathcal{D}_i$ is being shelled. The boundary loop consists of a portion of the boundary loop of the shelling of $\mathcal{D}_i$, a portion of the geodesic arc from $\star$ to $\partial\mathcal{D}$ that runs between $\mathcal{D}_i$ and $\mathcal{D}_{i+1}$, and a portion of $\partial\mathcal{D}$. These latter two portions have total length at most n, and so

$$\mathrm{FL}(\mathcal{D}) \leq \max_i \{\mathrm{FL}(\mathcal{D}_i)\} + n.$$

Combining this with (1) above we have our result.

REMARK. The shelling of $\mathcal{D}$ involves *1-cell collapse* and *2-cell collapse* moves, but no *1-cell expansion* moves.

5. An AD-pair for groups satisfying a polynomial isoperimetric inequality

In this section we prove:

THEOREM 2. *Let Γ be a group admitting a polynomial isoperimetric function of degree $r \geq 2$. Then up to a common multiplicative constant (x^r, x^{r-1}) is an AD-pair for Γ.*

Papasoglu gives this result for $r = 2$ in [12], page 799. It requires a small generalisation of his argument to obtain the result for all $r \geq 2$, as follows. (See also [10], page 100.)

DEFINITION. The *radius* of a van Kampen diagram $\mathcal{D}$ is

$$\operatorname{Rad}(\mathcal{D}) := \max \left\{ d(v, \partial \mathcal{D}) : v \text{ is a vertex of } \mathcal{D} \right\},$$

where $d(v, \partial \mathcal{D})$ is the combinatorial distance in the 1-skeleton.

DEFINITION. For a subcomplex K of $\mathcal{D}$ define $\operatorname{star}(K)$ to be the union of closed 2-cells meeting K. Define $\operatorname{star}_i(K)$ to be the i-th iterate of the star operation for $i \geq 1$; by convention $\operatorname{star}_0(K) = K$. So if Γ is triangularly presented then the 0-cells in $\operatorname{star}_i(\partial \mathcal{D})$ are precisely those a distance at most i from ∂D.

The substance of Theorem 2 is in the following lemma.

LEMMA 3. *Suppose $\langle \mathcal{A} \,|\, \mathcal{R} \rangle$ is a finite triangular presentation for a group Γ (see the paragraph following Proposition 2) and that $\mathcal{R}$ includes all null-homotopic words of length at most 3. Suppose further that there is $M > 0$ such that $f_0(n) \leq Mn^r$ for all n. Then for all null-homotopic w we have $\operatorname{Rad}(\mathcal{D}) \leq 12M\ell(w)^{r-1}$, where $\mathcal{D}$ is a minimal area van Kampen diagram for w.*

Recall that a change of finite presentation induces a $\simeq$-equivalence on isoperimetric and isodiametric functions. Note that there are only finitely many words of length at most 3 in a finitely presented group. Also observe that adding $n/2$ is sufficient to obtain a diameter bound from a radius bound. Thus this lemma is sufficient to prove Theorem 2.

Proof of Lemma 3. We proceed by induction on n. For $n \leq 3$ the result follows from our insistence that $\mathcal{R}$ includes all null-homotopic words of length at most 3.

For the induction step suppose w is null-homotopic and $\ell(w) = n$. Let $\mathcal{D}$ be a minimal area van Kampen diagram for w. Let $N_i := \operatorname{star}_i(\partial \mathcal{D})$. Now N_i, being a connected subcomplex of $\mathcal{D}$ with open discs removed from its interior. Define c_0 to be the boundary of the 2-dimensional portions of $\mathcal{D}$ and for $i \geq 1$ define $c_i := \partial N_i$. Each c_i is the union of simple closed curves any two of which meet at one point or not at all.

Now $\operatorname{Area}(N_{i+1}) - \operatorname{Area}(N_i) \geq l(c_i)/3$ because every 1-cell of c_i lies in the boundary of some 2-cell in $N_{i+1} - N_i$. For all i,

$$\operatorname{Area}(\mathcal{D}) \geq \operatorname{Area}(N_{i+1}) \geq l(c_i)/3 + l(c_{i-1})/3 + \cdots + l(c_0)/3.$$

Thus if $l(c_i) > n/2$ for all $i \leq 6Mn^{r-1}$ we get a contradiction of the area bound Mn^r for $\mathcal{D}$. So for some $i \leq 6Mn^{r-1}$ we find $l(c_i) \leq n/2$. We can appeal to the inductive hypothesis to learn that the diagrams enclosed by the simple closed curves constituting c_i have radius at most $12M(n/2)^{r-1} \leq 6Mn^{r-1}$. This is true since the minimality of the area of $\mathcal{D}$ implies that each c_i is filled by a diagram of minimal area.

So a vertex v of $\mathcal{D}$ either lies in N_i, in which case $d(v, \partial\mathcal{D}) \leq 6Mn^{r-1}$, or is in a diagram enclosed by one of the simple closed curves c of c_i. In the latter case $d(v, \partial\mathcal{D}) \leq d(v, c) + d(c, \partial\mathcal{D}) \leq 12Mn^{r-1}$ as required, thus completing the proof of the lemma.

6. Concluding remarks

OPEN QUESTION. In connection with the question of John Stallings, whether for a finite presentation there is always a simple exponential bound $f_0 \preceq \exp g_0$, it is natural to ask whether there is always an AD-pair of the form $(\exp g_0, g_0)$. This adds the requirement that the $\exp g_0$ bound on f_0 is always realisable on the same van Kampen diagram as g_0. Our main theorem gives a necessary condition that this be true, namely that $h_0 \preceq g_0{}^2$.

We shall now make some observations relevant to the single exponential question just stated.

PROPOSITION 3. *If $\mathcal{P}$ is a finite presentation, then for all integers $N \geq 3$ there exists $C(N) > 0$ such that for all van Kampen diagrams $\mathcal{D}$ in $\mathcal{P}$ all of whose vertices have valence at most N one has*

1. $\mathrm{Area}(\mathcal{D}) \leq N^{\mathrm{Diam}(\mathcal{D})+1} - 1$, *and*

2. $\mathrm{FL}(\mathcal{D}) \leq C(N) \cdot \mathrm{Diam}(\mathcal{D})^2 + n + 1$.

Proof. The number of vertices at a given distance i from the base point is at most $N(N-1)^{i-1}$, so it follows that the number of geometric edges $E(\mathcal{D})$ satisfies $E(\mathcal{D}) \leq N + N^2 + \cdots + N^D < \frac{N^{D+1}-1}{N-1}$, where $\mathcal{D} = \mathrm{Diam}(\mathcal{D})$. Since each edge is incident with at most 2 faces, we get $\mathrm{Area}(\mathcal{D}) \leq 2\frac{N^{D+1}-1}{N-1} \leq N^{D+1} - 1$, giving the first conclusion of the proposition.

From Proposition 2 it follows that $\mathrm{FL}(\mathcal{D}) \leq (2D+1)(D+1)\log_2(N) + 4D + 1 + n = C(N)D^2 + n + 1$, where $C(N)$ depends only on N, proving the second conclusion.

COROLLARY 4. *For every finite presentation $\mathcal{P}$ there is a constant $C > 0$ such that if $\mathcal{D}$ is an immersed topological disc diagram in $\mathcal{P}$, then $\mathrm{FL}(\mathcal{D}) \leq C \cdot \mathrm{Diam}(\mathcal{D})^2 + n + 1$.*

Proof. Since $\mathcal{D}$ is immersed, the valence of a vertex v is at most the number of edges incident at a vertex of the Cayley graph, namely, twice the number of generators of $\mathcal{P}$. The corollary follows from the second conclusion of the proposition.

REMARK. As we mentioned in the introduction, if $h_0 \preceq g_0$ then it follows that $f_0 \preceq \exp g_0$; one sees this as a consequence of the bound $f_0 \preceq \exp h_0$. We do not know an example from finitely presented groups where $h_0 \preceq g_0$ fails; however it is shown by Frankel and Katz in [5] that this can fail in a simply connected Riemannian context. (However their example does not amount to a properly discontinuous cocompact action by isometries on a simply connected Riemannian manifold, so it does not correspond to an example arising from finitely presented groups.)

References

1. Bridson, M. R.: 1999, 'Asymptotic cones and polynomial isoperimetric inequalities'. *Topology* **38**(3), 543–554.
2. Bridson, M. R. and A. Haefliger: 1999, *Metric Spaces of Non-positive Curvature*, Grundlehren der mathematischen Wissenschaften, Band 319, Springer Verlag, Berlin-New York.
3. Cohen, D. E.: 1991, 'Isoperimetric and isodiametric inequalities for group presentations'. *Int. J. of Alg. and Comp.* **1**(3), 315–320.
4. Epstein, D. B. A., J. W. Cannon, S. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston: 1992, *Word Processing in Groups*. Jones and Bartlett, Boston-London.
5. Frankel, S. and M. Katz: 1993, 'The Morse landscape of a Riemannian disc.'. *Ann. Inst. Fourier, Grenoble* **43**(2), 503–507.
6. Gersten, S. M.: 1991, 'The double exponential theorem for isoperimetric and isodiametric functions'. *Int. J. of Alg. and Comp.* **1**(3), 321–327.
7. Gersten, S. M.: 1993, 'Isoperimetric and isodiametric functions'. In: G. Niblo and M. Roller (eds.): *Geometric group theory I*, London Math. Soc. Lecture Note Series, vol. 181, Cambridge Univ. Press.
8. Gersten, S. M.: 1995, 'Asynchronously Automatic Groups'. In: Charney, Davis, and Shapiro (eds.): *Geometric group theory.* pp. 121–133, Ohio State Univ. Math. Research Institute Publ. vol 3, Walter de Gruyter, Berlin-New York.
9. Gersten, S. M. and H. Short: 2000, 'Some isoperimetric inequalities for free extensions'. Preprint, http://www.math.utah.edu/~gersten/grouptheory.htm/.
10. Gromov, M.: 1993, 'Asymptotic invariants of infinite groups'. In: G. Niblo and M. Roller (eds.): *Geometric group theory II*, London Math. Soc. Lecture Notes Series, vol. 182, Cambridge Univ. Press.
11. Lyndon, R. C. and P. E. Schupp: 1977, *Combinatorial Group Theory.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 89, Springer-Verlag, Berlin-New York.
12. Papasoglu, P.: 1996, 'On the asymptotic invariants of groups satisfying a quadratic isoperimetric inequality'. *J. Differential Geom.* **44**, 789–806.