Isoperimetric inequalities for nilpotent groups

S. M. Gersten, D. F. Holt and T. R. Riley

January 16, 2002, Revised July 7, 2002
Appeared in *Geom. Funct. Analysis*, 13, pages 795-814, 2003.

Abstract

We prove that every finitely generated nilpotent group of class c admits a polynomial isoperimetric function of degree $c+1$ and a linear upper bound on its filling length function.

1991 Mathematics Subject Classification: 20F05, 20F32, 57M07
Key words and phrases: nilpotent group, isoperimetric function, filling length

1 Introduction

The main result of this article is a proof of what has been known as the *"$c+1$-conjecture"* (see [3] or $5\mathrm{A}_5'$ of [13]).

Theorem A. *Every finitely generated nilpotent group G admits a polynomial isoperimetric inequality of degree $c+1$, where c is the nilpotency class of G.*

An isoperimetric inequality for a finite presentation $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ of a group G concerns *null-homotopic* words w – that is, words that evaluate to the identity in G. It gives an upper bound (an *"isoperimetric function"*), in terms of the length of w, on the number of times one has to *apply relators* from $\mathcal{R}$ to w in a process of reducing it to the empty word. (More details are given in §2.) The following couple of remarks are important in making sense of the statement of Theorem A. Nilpotent groups are *coherent*, by which we mean that their finitely generated subgroups are finitely presentable. In particular, all finitely generated nilpotent groups have a finite presentation. Further, an isoperimetric function $f_{\mathcal{P}}(n)$ concerns a fixed finite presentation $\mathcal{P}$ for G; however, as we will explain in §2, if $\mathcal{Q}$ is another finite presentation for G then there is an isoperimetric function $f_{\mathcal{Q}}(n)$ for $\mathcal{Q}$ that satisfies $f_{\mathcal{P}} \simeq f_{\mathcal{Q}}$, where $\simeq$ is a well-known equivalence relation. In particular, if $f_{\mathcal{P}}$ is polynomial of some given degree ≥ 1, then we can always take $f_{\mathcal{Q}}$ to be a polynomial of the same degree.

Our strategy will be to prove the $c+1$-conjecture by an induction on the class c. However we use an induction argument in which we keep track of more than an isoperimetric function. In fact, we prove the following stronger theorem.

Theorem B. *Finitely generated nilpotent groups G of class c admit (n^{c+1}, n) as an $(\mathrm{Area}, \mathrm{FL})$-pair.*

The terminology used above is defined carefully in §2, but what this theorem says is essentially the following. Suppose w is a null-homotopic word w of length n in some finite presentation for G. We can reduce w to the empty word by applying at most $O(n^{c+1})$ relators from the presentation, and in such a way that in the process the intermediate words have length at most $O(n)$.

Let

$$G \;=\; \Gamma_1 > \Gamma_2 > \cdots > \Gamma_{c+1} \;=\; \{1\}\,,$$

defined inductively by $\Gamma_1 := G$ and $\Gamma_{i+1} := [G, \Gamma_i]$, be the lower central series for a class c nilpotent group G.

We will use a generating set $\mathcal{A}$ for G which will be a disjoint union of sets $\mathcal{A}_i$, where, for each i, $\mathcal{A}_i$ is a generating set for Γ_i modulo Γ_{i+1}. For $\mathcal{A}_1$, we take inverse images in G of generators of cyclic invariant factors of the abelian group G/Γ_2. Then, inductively, for $i > 0$, we define $\mathcal{A}_{i+1} := \{[x, y] \mid x \in \mathcal{A}_1, y \in \mathcal{A}_i\}$.

The idea of the proof of Theorem B is to start by reducing the word w to the identity modulo Γ_c. The quotient G/Γ_c is nilpotent of class $c-1$ and so, by induction hypothesis, we can reduce w to the identity in a presentation for G/Γ_c, by applying $O(n^c)$ defining relators. When we carry out the corresponding reduction in a presentation for G we introduce $O(n^c)$ extra generators from $\mathcal{A}_c$. We use the definition of these generators as commutator words z to *compress* their powers z^m with $m = O(n^c)$ to words of length $O(n)$. This compression process is handled in Proposition 3.1 and Corollaries 3.2 and 3.3, and we show that it can be accomplished by applying $O(n^{c+1})$ relators.

In addition to proving the $c+1$-conjecture, Theorem B yields the following corollary.

Corollary B.1. *If $\mathcal{P}$ is a finite presentation for a nilpotent group then there exists $\lambda > 0$ such that the filling length function* FL *of $\mathcal{P}$ satisfies* $\mathrm{FL}(n) \leq \lambda n$ *for all $n \in \mathbb{N}$.*

This result was proved by the third author in [20] via the rather indirect technique of using asymptotic cones: finitely generated nilpotent groups have simply connected asymptotic cones (see Pansu [17]) and groups with simply connected asymptotic cones have linearly bounded filling length. This article provides a direct combinatorial proof. Note, also, that the filling length function for a finite presentation $\mathcal{P}$ is an isodiametric function for $\mathcal{P}$. So we also have a direct combinatorial proof that finitely generated nilpotent groups admit linear isodiametric functions.

There is a computer science reinterpretation of Theorem B because Area and FL can be recognised to be measures of computational complexity in the following context.

Suppose $\mathcal{P}$ is a finite presentation for a group G. One can attempt to solve the word problem using a non-deterministic Turing machine as follows. Initially the input word w of length n is displayed on the Turing tape. A *step* in the operation of the machine is an application of a relator, a free reduction, or a free expansion (see §2). The machine searches for a *proof* that $w = 1$ in G – that is, a sequence of steps that reduces w to the empty word. The running time of a *proof* is the number of steps, and its space is the length of the portion of the Turing tape used in the course of the *proof*.

The time $\mathrm{Time}(w)$ for a word w such that $w = 1$ in G is the minimum running time amongst *proofs* for w, and the time function $\mathrm{Time} : \mathbb{N} \to \mathbb{N}$ is defined by $\mathrm{Time}(n) := \max \{\mathrm{Time}(w) \mid \ell(w) \leq n \text{ and } w = 1 \text{ in } G\}$. Similarly we define the space function Space by setting $\mathrm{Space}(w)$ to be the minimal space amongst all proofs for w.

On comparing with the definitions of §2, we quickly recognise that the space function Space is precisely the same as the filling function FL. Also the time function Time is closely related to the function Area. (In fact Time is precisely the function Height of Remark 2.5, where it is explained how Height relates to Area.) Thus we have the following corollary to Theorem B.

Corollary B.2. *Given a finite presentation for a nilpotent group G of class c, the non-deterministic Turing machine described above solves the word problem with* Space(n) *bounded above by a linear function of n and* Time(n) *bounded above by a polynomial in n of degree $c+1$. Moreover, given an input word w such that $w = 1$ in G, there is a* proof *for w that runs within both these bounds simultaneously. [In fact, using the details of our proof of Theorem B, it is possible to make the Turing machine* deterministic.*]*

[Note that these complexity bounds are not the best possible. The torsion subgroup T of a finitely generated nilpotent group G is finite, and by a result of S.A. Jennings (see [14] or Theorem 2.5 of [21]), the torsion-free group G/T can be embedded in an upper unitriangular matrix group U over $\mathbb{Z}$. It is easy to see that the integer entries of a matrix in U representing a word of length n in G/T are at most $O(n^{d-1})$, where d is the degree of the matrix. Using this and the fact that two k-digit numbers can be multiplied in time $O(k\,(\log k)^2)$, we can obtain a deterministic solution to the word problem in G in time $O(n\,(\log n)^2)$ and space $O(\log n)$.

Even this is not the best possible result on the time complexity. In the unpublished manuscript [5], Cannon, Goodman and Shapiro show that finitely generated nilpotent groups admit a Dehn algorithm if we adjoin some extra symbols to the group generators, and this leads to a linear time solution of the word problem in G.]

The proof of the $c+1$-conjecture is the culmination of a number of results spanning the last decade or so. The first author proved in [8] that G admits a polynomial isoperimetric inequality of degree 2^h, where h is the Hirsch length of G. The degree was improved to $2 \cdot 3^c$ by G. Conner [6], and then to $2c$ by C. Hidber [15].

We also mention Ch. Pittet [18], who proved that a lattice in a simply connected homogeneous nilpotent Lie group of class c admits a polynomial isoperimetric function of degree $c+1$. (The nilpotent Lie group is called *homogeneous* if its Lie algebra is graded.) Gromov suggested a possible generalisation to the non-homogeneous case by perturbing the structure constants [13] 5.A_5, and it remains a challenge to carry out this plan.

The isoperimetric inequality proved in this article is the best possible bound in terms of the class in general. For example if G is a free nilpotent group of class c then its minimal isoperimetric function (*a.k.a.* its Dehn function) is polynomial of degree $c+1$ (see [2] or [10] for the lower bound and [18] for the upper bound). However it is not best possible for individual nilpotent groups: for example D. Allcock proved in [1] that the $2n+1$ dimensional integral Heisenberg groups admit quadratic isoperimetric functions for $n > 1$; these groups are all nilpotent of class 2 (see [1] for a history of the problem along with a discussion of other approaches to it). By way of contrast, the 3-dimensional integral Heisenberg group has a cubic minimal isoperimetric polynomial ([7] and [9]).

We give two corollaries of Theorem A concerning the cohomology of groups and differential geometry respectively.

The first corollary is about the growth of cohomology classes. Let G be a finitely generated group with word metric determined by a finite set of generators and denoted $d(1,g) = |g|$. Recall that a real valued 2-cocycle (for the trivial G-action on $\mathbb{R}$) on the bar construction is a function $f : G \times G \to \mathbb{R}$ satisfying $f(x,yz) + f(y,z) = f(x,y) + f(xy,z)$ for all $x, y, z \in G$.

Corollary A.1. *Let G be a finitely generated nilpotent group of class c and let $\zeta \in H^2(G,\mathbb{R})$. Then there is a 2-cocycle f in the class ζ which satisfies $|f(x,y)| \leq M(|x|+|y|)^{c+1}$ for all $x, y \in G$ and constant $M > 0$.*

In an appendix to this article we give a general result relating isoperimetric functions to growth of cocycles that allows us to deduce Corollary A.1 from Theorem A.

The terminology (due to Gromov) of an *isoperimetric function* for a group is motivated by the analogous notion with the same name from differential geometry. One can draw parallels between van Kampen diagrams (see Remark 2.3) filling edge-circuits in the Cayley 2-complex associated to a finite presentation of a group and homotopy discs for rectifiable loops in the universal cover of a Riemannian manifold.

Corollary A.2. *Let M be a closed Riemannian manifold whose fundamental group is nilpotent of class c. Then there is a polynomial f of degree $c+1$ such that for every rectifiable loop of length L in the universal cover $\widetilde{M}$, there is a singular disc filling it of area at most $f(L)$.*

For the proof one considers a piecewise C_1-map of the presentation complex K of a finite presentation of G into M and the induced map of universal covers $\widetilde{K} \to \widetilde{M}$. An edge-circuit p in $\widetilde{K}$ defines a null-homotopic word w in $\pi_1 M$, and a van Kampen diagram D for w gives a Lipschitz filling for p in $\widetilde{K}$. For a general Lipschitz loop p in $\widetilde{M}$, one uses Theorem 10.3.3 of [7] to homotop p by a Lipschitz homotopy to an edge-circuit in $\widetilde{K}$, which is then filled to realise the desired isoperimetric inequality for p.

2 Isoperimetric functions and filling length functions

Let $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ be a finite presentation for a group G. A *word* w is an element of the free monoid $(\mathcal{A} \cup \mathcal{A}^{-1})^\star$. Denote the length of w by $\ell(w)$. We say w is *null-homotopic* when $w = 1$ in G. For a word $w = a_1{}^{\varepsilon_1} a_2{}^{\varepsilon_2} \ldots a_s{}^{\varepsilon_s}$, where each $a_i \in \mathcal{A}$ and each $\varepsilon_i = \pm 1$, the *inverse word* w^{-1} is $a_s{}^{-\varepsilon_s} \ldots a_2{}^{-\varepsilon_2} a_1{}^{-\varepsilon_1}$.

Definition 2.1. A *$\mathcal{P}$-sequence* S is a finite sequence of words $w_0, w_1, \ldots, w_m$ such that each w_{i+1} is obtained from w_i by one of three moves.

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

2. *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.

We refer to m as the *height* $\mathrm{Height}(S)$ of S and we define the *filling length* $\mathrm{FL}(S)$ of S by

$$\mathrm{FL}(S) \ := \ \max\left\{\ell(w_i) \mid 0 \leq i \leq m\right\}.$$

The *area* $\mathrm{Area}(S)$ of S is defined to be the number of i such that w_{i+1} is obtained from w_i by an *application of a relator* move.

If $w = w_0$ and w_m is the empty word then we say that S is a *null-$\mathcal{P}$-sequence* for w (or when non-ambiguous just a "*null-sequence*"). We remark that Definition 2.1 gives the moves in what is called a *rewriting system* for $\mathcal{P}$ in some related literature – see [16] or [19] for example.

In this article we are concerned with two *filling functions*, that measure "*area*" and "*filling length*" – two different aspects of the *geometry of the word problem* for G.

Definition 2.2. Let w be a null-homotopic word in $\mathcal{P}$.

We define the *area* $\mathrm{Area}(w)$ of w by

$$\mathrm{Area}(w) \ := \ \min\left\{\mathrm{Area}(S) \mid \text{null-}\mathcal{P}\text{-sequences } S \text{ for } w\right\},$$

that is, the minimum number of relators that one has to apply to reduce w to the empty word. Similarly, the *filling length* of w is

$$\mathrm{FL}(w) \ := \ \min\left\{\mathrm{FL}(S) \mid \text{null-}\mathcal{P}\text{-sequences } S \text{ for } w\right\}.$$

We define the *Dehn function* $\mathrm{Area} : \mathbb{N} \to \mathbb{N}$ (also known as the *minimal isoperimetric function*) and the *filling length function* $\mathrm{FL} : \mathbb{N} \to \mathbb{N}$ for $\mathcal{P}$ by

$$\begin{aligned} \mathrm{Area}(n) &:= \max\left\{\mathrm{Area}(w) \mid \text{null-homotopic words } w \text{ with } \ell(w) \leq n\right\} \\ \mathrm{FL}(n) &:= \max\left\{\mathrm{FL}(w) \mid \text{null-homotopic words } w \text{ with } \ell(w) \leq n\right\}. \end{aligned}$$

Remark 2.3. The formulations of the definitions above are those we will use in this article, but we mention some equivalent alternatives that occur elsewhere in the literature.

We could equivalently define $\mathrm{Area}(w)$ to be the minimal number of 2-cells needed to construct a van Kampen diagram for w, or the minimal N such that 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 $r_i \in \mathcal{R}^{\pm 1}$ and words u_i.

Similarly, $\mathrm{FL}(w)$ can be defined in terms of *shellings* of van Kampen diagrams D_w for w. A *shelling* of D_w is a combinatorial null-homotopy down to the base vertex. The filling length of a shelling is the maximum length of the boundary loops of the diagrams one encounters in the course of the null-homotopy. The filling length of D_w is defined to be the minimal filling length amongst shellings of D_w and then $\mathrm{FL}(w)$ can be defined to be

$$\min\left\{\mathrm{FL}(D_w) \mid \text{van Kampen diagrams } D_w \text{ for } w\right\}.$$

For detailed definitions and proofs we refer the reader to [11].

The following is a technical lemma that we will use in §3.

Lemma 2.4. *Let S be a $\mathcal{P}$-sequence $w_0, w_1, \ldots, w_m$ as defined above. Let $C := \max\left\{\ell(r) \mid r \in \mathcal{R}\right\}$. There is another $\mathcal{P}$-sequence $\hat{w}_0, \hat{w}_1, \ldots, \hat{w}_{\hat{m}}$, which we will call $\hat{S}$, such that*

- $\hat{w}_0 = w_0$, $\hat{w}_{\hat{m}} = w_m$
- $\mathrm{Area}(\hat{S}) = \mathrm{Area}(S)$
- $\mathrm{FL}(\hat{S}) \leq \mathrm{FL}(S) + C$

and such that every time $\hat{w}_{i+1}$ is obtained from a word $\hat{w}_i$ in the sequence $\hat{S}$ by an application of a relator *the whole of a cyclic conjugate of an element of $\mathcal{R}$ or its inverse is inserted into $\hat{w}_i$.*

Proof. Suppose that w_{i+1} is obtained from a word w_i in the sequence S by an *application of a relator*: that is, $w_i = \alpha u \beta$ and $w_{i+1} = \alpha v \beta$ for some words α, β, u, v where uv^{-1} is a cyclic conjugate some element of $\mathcal{R}^{\pm 1}$. We can obtain w_{i+1} from w_i by inserting $u^{-1}v$ into $w_i = \alpha u \beta$ to get $\alpha u u^{-1} v \beta$ and then using at most C free reductions to retrieve $\alpha v \beta$. ■

Remark 2.5. We mention that the area of a null-homotopic word w is closely related to the *height* of its null-sequences. Define

$$\text{Height}(w) \;:=\; \min\left\{\text{Height}(S) \mid \text{null-sequences } S \text{ for } w\right\}.$$

Then for all null-homotopic words w,

$$\text{Area}(w) \;\leq\; \text{Height}(w) \;\leq\; (C+1)\text{Area}(w) + \ell(w),$$

where $C := \max\left\{\ell(r) \mid r \in \mathcal{R}\right\}$.

The inequality $\text{Area}(w) \leq \text{Height}(w)$ follows from the definitions since free reductions and expansions do not contribute to the area. To obtain the inequality $\text{Height}(w) \leq (C+1)\text{Area}(w) + \ell(w)$ first take a van Kampen diagram D_w for w with $\text{Area}(D_w) = \text{Area}(n)$. Take any shelling

$$D_w \;=\; D_0, D_1, \ldots, D_m \;=\; \star$$

of D_w down to its base vertex $\star$ in which each D_{i+1} is obtained from D_i by a *2-cell collapse* or a *1-cell collapse* (but never a *1-cell expansion*) and let w_j be the boundary word of D_j. Then $w_0, w_1, \ldots, w_m$ is a null-sequence for w, where w_{i+1} is obtained from w_i by applying a relator if D_{i+1} is obtained from D_i by a *2-cell collapse*, and by a free reduction if D_{i+1} is obtained from D_i by a *1-cell collapse*. We obtain the required inequality by observing that the number of *2-cell collapse* moves in the shelling is $\text{Area}(D_w) = \text{Area}(w)$ and the total number of *1-cell collapse* moves is at most the total number if 1-cells in the 1-skeleton of D_w, which is at most $C\text{Area}(w) + \ell(w)$.

An *isoperimetric inequality* for $\mathcal{P}$ is provided by any function $f : \mathbb{N} \to \mathbb{N}$ such that $\text{Area}(n) \leq f(n)$ for all n. We refer to f as an *isoperimetric function* for $\mathcal{P}$.

It is important to note that $\text{Area} : \mathbb{N} \to \mathbb{N}$ and $\text{FL} : \mathbb{N} \to \mathbb{N}$ are both defined for a fixed finite presentation. However each function is a group invariant in the sense that each is well behaved under change of finite presentation as we now explain.

Definition 2.6. For two functions $f, g : \mathbb{N} \to \mathbb{N}$ we say that $f \preceq g$ when there exists $C > 0$ such that $f(n) \leq Cg(Cn + C) + Cn$ for all n, and we say $f \simeq g$ if and only if $f \preceq g$ and $g \preceq f$.

Let $\mathcal{P}$ and $\mathcal{Q}$ be two presentations of the same group G, let $\text{Area}_{\mathcal{P}}$ and $\text{Area}_{\mathcal{Q}}$ be their Dehn functions, and let $\text{FL}_{\mathcal{P}}$ and $\text{FL}_{\mathcal{Q}}$ be their filling length functions. Then it is proved in [12] that $\text{Area}_{\mathcal{P}} \simeq \text{Area}_{\mathcal{Q}}$ and in [11] that $\text{FL}_{\mathcal{P}} \simeq \text{FL}_{\mathcal{Q}}$.

Definition 2.7. We say that a pair (f, g) of functions $\mathbb{N} \to \mathbb{N}$ is an (Area, FL)-pair for $\mathcal{P}$ when there exists a constant $\lambda > 0$ such that, for any null-homotopic word w, there exists a null-$\mathcal{P}$-sequence S with

$$\begin{aligned} \text{Area}(S) &\leq \lambda f(\ell(w)) \\ \text{FL}(S) &\leq \lambda g(\ell(w)). \end{aligned}$$

The first part of the following lemma justifies the deduction of Theorem A and Corollary B.1 from Theorem B, and second part justifies the terminology used in the statement of Theorem B.

Lemma 2.8. *If (f, g) is an (Area, FL)-pair for $\mathcal{P}$ then f is an isoperimetric function for $\mathcal{P}$ and g is an upper bound for its filling length function.*

Moreover, if (n^α, n^β) is an (Area, FL)-pair for $\mathcal{P}$ where $\alpha, \beta \geq 1$, then (n^α, n^β) is also an (Area, FL)-pair for any other finite presentation $\mathcal{Q}$ of the same group G.

Proof. The first part of the lemma is immediate from the definitions.

It follows from the proofs given in [11] and [12] that if (f, g) is an (Area, FL)-pair for $\mathcal{P}$, then (f', g') is an (Area, FL)-pair for $\mathcal{Q}$, where $f'(n) = Cf(Cn) + Cn$ and $g'(n) = Cg(Cn) + Cn$ for some $C > 0$.

But if $f(n) = n^\alpha$ with $\alpha \geq 1$, then $Cf(Cn) + Cn \leq C'n^\alpha$ for suitable constant $C' > 0$. It follows that if (n^α, n^β) is an (Area, FL)-pair for $\mathcal{P}$, then it is also an (Area, FL)-pair for $\mathcal{Q}$. ■

3 Proof of Theorem B

We say that an element of a nilpotent group G has *weight* k if it lies in $\Gamma_k \smallsetminus \Gamma_{k+1}$, where $\{\Gamma_i\}$ is the lower central series of G, as defined in §1. Before we come to the proof of Theorem B we give a proposition and corollary concerning *compressing* powers of elements of weight c in a finite presentation for a nilpotent group of class c.

We use the following conventions for commutator words:

$$\begin{aligned} [a] &:= a \\ [\,a\,,\,b\,] &:= a^{-1}\,b^{-1}\,a\,b \\ [\,a_1\,,\,a_2\,,\,\ldots,\,a_{c-1}\,,\,a_c\,] &:= [\,a_1\,,\,[\,a_2\,,\,\ldots,[\,a_{c-1}\,,\,a_c\,]\ldots]]. \end{aligned}$$

Let $\mathcal{X} = \{x_1, x_2, \ldots, x_c\}$ be an alphabet and for $k = 1, 2, \ldots, c$ define

$$\begin{aligned} \mathcal{X}_k &:= \{x_k, x_{k+1}, \ldots, x_c\} \\ \mathcal{R}_k &:= \left\{[y_k, y_{k+1}, \ldots, y_{c+1}] \mid y_j \in \mathcal{X}_k^{\pm 1},\ k \leq j \leq c+1\right\}. \end{aligned}$$

Let $\mathcal{P}_k := \langle \mathcal{X}_k \mid \mathcal{R}_k \rangle$, which is a finite presentation for a free nilpotent group G_k of class $c + 1 - k$. Define z_k to be the word $[\,x_k\,,\,x_{k+1}\,,\,\ldots,\,x_c\,]$. In particular, we have $z_c = x_c$.

One can regard the first part of the forthcoming proposition as comparing the use of commutators in a free nilpotent group with the representation of positive integers s by their n-ary expansion $s = s_0 + s_1 n + s_2 n^2 + \ldots$. The second part gives upper bounds on the area and filling length of a $\mathcal{P}$-sequence between the commutators corresponding to the n-ary representations for s and $s+1$.

The reason this proposition will be important to us is that is allows us to replace certain words of length $O(n^c)$ by words of length $O(n)$. This will be essential in our induction proof of Theorem B because it allows us to maintain control on the filling length.

Let us fix $n \in \mathbb{N}$, which will be the length of a null-homotopic word w in our induction proof of Theorem B. We make some further definitions. For $0 \leq s \leq n^c - 1$ express s as a sum:

$$s \;=\; s_0 \;+\; s_1\,n \;+\; \ldots \;+\; s_{c-1}\,n^{c-1},$$

where each $s_i \in \{0, 1, \ldots, n-1\}$. Then for $k = 1, 2, \ldots, c$ define the word

$$u(k,s) \;:=\; {z_k}^{s_{k-1}}\,[\,{x_k}^{n}\,,\,{z_{k+1}}^{s_k}\,[{x_{k+1}}^{n}\,,\,\ldots,\,{z_{c-1}}^{s_{c-2}}\,[\,{x_{c-1}}^{n}\,,\,{z_c}^{s_{c-1}}\,]\ldots]\,].$$

Also define

$$u(k, n^c) \ := \ [\, {x_k}^n\, ,\, {x_{k+1}}^n\, ,\, \ldots,\, {x_c}^n\,].$$

Proposition 3.1. *Fix $n \in \mathbb{N}$ and $0 \leq s \leq n^c - 1$. Then with the definitions above, $u(1, s) = {z_1}^s$ and $u(1, n^c) = {z_1}^{n^c}$ in G_1.*

Moreover, there exists $\kappa > 0$ such that for $0 \leq s \leq n^c - 1$ we can transform $z_1\, u(1, s)$ to $u(1, s+1)$ via a $\mathcal{P}_1$-sequence of filling length at most κn and of area at most κn^{k+1} where k is the positive integer such that $n^k \mid (s+1)$ but $n^{k+1} \nmid (s+1)$.

Proof. Our proof is by induction on c. The base case of $c = 1$ is straightforward because $u(1, s) = {z_1}^s = {x_1}^s$ and we can take $\kappa = 1$.

We now prove the induction step. Suppose $c \in \mathbb{N} \smallsetminus \{0, 1\}$. Since the free reduction of $u(1, 0)$ is the empty word, we have $u(1, 0) = 1$ in G_1. Therefore it suffices to show that for $s, n \in \mathbb{N}$ with $1 \leq s \leq n^c - 1$, there is a $\mathcal{P}_1$-sequence from $z_1\, u(1, s)$ to $u(1, s+1)$ within the required filling length and area bounds.

Express s as a sum:

$$s \ = \ s_0 \ + \ s_1\, n \ + \ \ldots \ + \ s_{c-1}\, n^{c-1},$$

where each $s_i \in \{0, 1, \ldots, n-1\}$. Define

$$t \ := \ s_1 \ + \ s_2\, n \ + \ \ldots \ + \ s_{c-2}\, n^{c-2}.$$

Then $u(1, s) \ = \ {z_1}^{s_0}\, [\, {x_1}^n\, ,\, u(2, t)\,]$ as words.

If $s_0 + 1 < n$ then $z_1\, u(1, s) = u(1, s+1)$ as words and there is a trivial $\mathcal{P}_1$-sequence from $z_1\, u(1, s)$ to $u(1, s+1)$. If $s_0 + 1 = n$ then we calculate that $z_1\, u(1, s) = u(1, s+1)$ in G_1 as follows:

$$\begin{aligned}
z_1\, u(1, s) \ &= \ {z_1}^n\, [\, {x_1}^n\, ,\, u(2, t)\,] && (1)\\
&= \ {z_1}^n\, {x_1}^{-n}\, u(2, t)^{-1}\, {x_1}^n\, u(2, t) && (2)\\
&= \ {z_1}^n\, {x_1}^{-n}\, u(2, t)^{-1}\, {z_2}^{-1}\, z_2\, {x_1}^n\, u(2, t) && (3)\\
&= \ {x_1}^{-n}\, u(2, t)^{-1}\, {z_2}^{-1}\, {x_1}^n\, z_2\, u(2, t) && (4)\\
&= \ [\, {x_1}^n\, ,\, z_2\, u(2, t)\,] && (5)\\
&= \ [\, {x_1}^n\, ,\, u(2, t+1)\,] && (6)\\
&= \ u(1, s+1). && (7)
\end{aligned}$$

In the step from (3) to (4) we use the fact that $z_1 \;=\; [\,x_1\,,\,z_2\,]$ and is central. For the step from (5) to (6) we invoke the induction hypothesis to tell us that $z_2\,u(2,t) \;=\; u(2,t{+}1)$ in G_2, from which it follows that $z_2\,u(2,t)\,u(2,t+1)^{-1}$ is central in G_1 and $[\,{x_1}^n\,,\,z_2\,u(2,t)\,] \;=\; [\,{x_1}^n\,,\,u(2,t+1)\,]$.

The course of the calculation above dictates how to construct a $\mathcal{P}_1$-sequence from $z_1\,u(1,s)$ to $u(1,s+1)$. The steps that require some further explanation are those at which *application of relator* moves are used: that is, from (3) to (4) and from (5) to (6).

The step from (3) to (4) is performed by a $\mathcal{P}_1$-sequence in which we introduce n commutator words ${z_1}^{-1} = [x_1, z_2]^{-1}$ to move z_2 past ${x_1}^n$. More precisely, we repeat the following n times. Introduce the relator ${z_1}^{-1}x_1z_2{x_1}^{-1}{z_2}^{-1}$ immediately before a subword z_2x_1, and cancel the resulting subword ${x_1}^{-1}{z_2}^{-1}z_2x_1$. So the z_2x_1 has now been replaced by ${z_1}^{-1}x_1z_2$. Use commutator relators to move the ${z_1}^{-1}$ to the left of the word, and cancel it with one of the z_1 letters there. The number of letters each z_1 has to be moved past is bounded by n up to a multiplicative constant that depends only on $\mathcal{P}_1$ (since $\ell(u(2,t)) \leq \overline{\kappa}n$, where $\overline{\kappa} > 0$ is a constant depending only on $\mathcal{P}_2$). So, up to a multiplicative constant, the area of this $\mathcal{P}_1$-sequence is at most n^2 and its filling length is at most n.

Now let us explain how to construct a $\mathcal{P}_1$-sequence for the step from (5) to (6). Suppose that $n^k \mid (s+1)$ but $n^{k+1} \nmid (s+1)$. Then $n^{k-1} \mid (t+1)$ but $n^k \nmid (t+1)$. By induction hypothesis there is a $\mathcal{P}_2$-sequence $\overline{S}$ from $z_2\,u(2,t)$ to $u(2,t+1)$ with area at most $\overline{\kappa}n^k$ and filling length at most $\overline{\kappa}n$.

Lemma 2.4 allows us to assume (by suitably adjusting the constant $\overline{\kappa}$) that in every instance of an *application of a relator* move in the sequence $\overline{S}$ a whole cyclic conjugate of a relator is inserted.

We now explain how to use $\overline{S}$ to induce a $\mathcal{P}_1$-sequence that transforms $[\,{x_1}^n\,,\,z_2\,u(2,t)\,]$ into $[\,{x_1}^n\,,\,u(2,t+1)\,]$. First define $\overline{S}'$ to be the $\mathcal{P}_2$-sequence that transforms $u(2,t)^{-1}\,{z_2}^{-1}$ to $u(2,t+1)^{-1}$ and is obtained by inverting every word in $\overline{S}$. Now

$$[\,{x_1}^n\,,\,z_2\,u(2,t)\,] = {x_1}^{-n}\,u(2,t)^{-1}\,{z_2}^{-1}\,{x_1}^n\,z_2\,u(2,t).$$

Consider running the $\mathcal{P}_2$-sequences $\overline{S}$ and $\overline{S}'$ concurrently on the subwords $z_2\,u(2,t)$ and $u(2,t)^{-1}\,{z_2}^{-1}$ respectively in $[\,{x_1}^n\,,\,z_2\,u(2,t)\,]$: that is, we do the first move in $\overline{S}$ and then the first move in $\overline{S}'$, then the second move in $\overline{S}$ and then the second move in $\overline{S}'$, and so on. However we want to construct a $\mathcal{P}_1$-sequence, not a $\mathcal{P}_2$-sequence.

Suppose a move in $\overline{S}$ is the insertion of a word $r \in \mathcal{R}_2$. Then the corresponding move in $\overline{S}'$ inserts the word r^{-1}. Use free expansion moves to insert the word $r^{-1}r$ in the place where $\overline{S}$ dictated that r was to be inserted; then use relators in $\mathcal{R}_1$ to move r^{-1} through the word to the place where $\overline{S}'$ dictated r^{-1} was to be inserted. (Recall that r represents a central element of G_1 and the appropriate commutator to move it past letters from $\mathcal{X}_1$ are in $\mathcal{R}_1$.)

The number of relators from $\mathcal{R}_1$ that have to be applied to move each r^{-1} to its appropriate place is at most n up to a multiplicative constant. It follows that we can find a constant $\kappa > 0$ such that there is a $\mathcal{P}_1$-sequence that transforms the word $z_1\, u(1,s)$ into $u(1,s+1)$, with area at most κn^{k+1} and filling length at most κn. This completes the proof. ∎

Recall from §1 that we define a set of generators $\mathcal{A}$ for a finitely generated nilpotent group G of class c, as follows. The set $\mathcal{A}$ is a disjoint union of sets $\mathcal{A}_i$, where, for each i, $\mathcal{A}_i$ is a generating set for Γ_i modulo Γ_{i+1}. For $\mathcal{A}_1$, we take inverse images in G of generators of cyclic invariant factors of the abelian group G/Γ_2. Then, inductively, for $i > 0$, we define $\mathcal{A}_{i+1} := \{[x,y] \mid x \in \mathcal{A}_1, y \in \mathcal{A}_i\}$.

Corollary 3.2. *Let G be a finitely generated nilpotent group of class c, and let $\mathcal{A}$ be the generating set defined above. Then, for any finite presentation $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ for G, there is a constant $\xi > 0$ depending only on $\mathcal{P}$ with the following properties.*

*Fix $n \in \mathbb{N}$. For each $z \in \mathcal{A}_c$ there are "*compression *words" $u_z(s)$ for $0 \leq s \leq n^c$ with the property that $u_z(s) = z^s$ in G.*

Moreover, we can transform each $z\, u_z(s)$ to $u_z(s+1)$ via a $\mathcal{P}$-sequence, which when all are concatenated:

$$z^{n^c} \to z^{n^c}\, u_z(0) \to z^{n^c-1}\, u_z(1) \to z^{n^c-2}\, u_z(2) \to \cdots \to z\, u_z(n^{c-1}) \to u_z(n^c),$$

gives a $\mathcal{P}$-sequence that converts z^{n^c} to $u_z(n^c)$, and has area at most ξn^{c+1} and filling length at most ξn.

Proof. First observe that if we can prove the corollary for any one finite set of relations $\mathcal{R}$ then it holds true for any finite set of relations.

By definition, each $z \in \mathcal{A}_c$ is a commutator of length c. Take any $\mathcal{R}$ such that $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ is a finite presentation for G and $\mathcal{R}$ includes all commutators involving $\mathcal{A}^{\pm 1}$ of length $c+1$. So the presentation for a free nilpotent group used in Proposition 3.1 is a subpresentation of $\mathcal{P}$. Take $z := z_1$ in the statement of Proposition 3.1 and define $u_z(s) := u(1, s)$. Then $u_z(s) = z^s$ in G as required.

Furthermore for $0 \leq s \leq n^c - 1$ Proposition 3.1 allows us to find a $\mathcal{P}$-sequence from $z\, u_z(s)$ to $u_z(s+1)$. Let k be the positive integer such that $n^k \mid (s+1)$ but $n^{k+1} \nmid (s+1)$. Then there is some $\kappa > 0$, depending only on $\mathcal{P}$ such that the area of this $\mathcal{P}$-sequence is at most κn^{k+1}, and its filling length is at most κn.

Since the free reduction of $u_z(0)$ is the empty word, we can transform z^{n^c} to $z^{n^c}\, u_z(0)$ by free expansion moves.

Now, the number of $s \in \{0, 1, \ldots, n^c - 1\}$ such that $n^k \mid (s+1)$ is at most n^{c-k}. For each such s the area of the sequence from $z\, u_z(s)$ to $u_z(s+1)$ is at most κn^{k+1}. So, taking $\xi := (c+1)\kappa$, the total area of the $\mathcal{P}$-sequence converting z^{n^c} to $u_z(n^c)$ is at most ξn^{c+1}. And its filling length is at most $\kappa n \leq \xi n$. ∎

Corollary 3.3. *Fix $n \in \mathbb{N}$. Let $\xi > 0$ and the words $u_z(s)$ for $0 \leq s \leq n^c$ be as in Corollary 3.2. Extend the sequence $u_z(s)$ of words to all $s \geq 0$ by defining:*

$$u_z(A + Bn^c) \ := \ u_z(A) u_z(n^c)^B$$

for $0 \leq A \leq n^c - 1$ and $B \geq 0$. Then $u_z(s) = z^s$ for all s. Also the $\mathcal{P}$-sequence that transforms z^{A+Bn^c} to $u_z(A+Bn^c)$, and extends the $\mathcal{P}$-sequence of Proposition 3.1 in the obvious way, has area at most $(B+1)\,\xi\, n^{c+1}$ and filling length at most $(B+1)\,\xi$.

Before we come to the proof of Theorem B we give two technical lemmas.

Lemma 3.4. *Suppose that the abelian group H is generated by the finite set $\mathcal{Y}$. Then there is a subset of $\mathcal{Y}$ that freely generates a free abelian group having finite index in H.*

Proof. This is by induction on $|\mathcal{Y}|$. There is nothing to prove if $|\mathcal{Y}| = 0$. If $|\mathcal{Y}| > 0$, let $\mathcal{Y} = \mathcal{Y}_1 \cup \{y\}$ with $|\mathcal{Y}_1| = |\mathcal{Y}| - 1$, and let $\mathcal{Y}_1$ generate K. By inductive hypothesis, $\mathcal{Y}_1$ has a subset $\mathcal{Y}_2$ that freely generates a free abelian subgroup L with $|K : L|$ finite. If $|H : L|$ is finite then we are done. Otherwise H/K is infinite cyclic and generated by yK, and then $\mathcal{Y}_2 \cup \{y\}$ has the required property. ∎

Lemma 3.5. *Suppose the finite group H of order t is generated by $\mathcal{Y}$. Then any word $w \in (\mathcal{Y} \cup \mathcal{Y}^{-1})^\star$ with $\ell(w) \geq t$ contains a subword u with $u = 1$ in H.*

Proof. Let w_i be the prefix of w of length i, for $0 \leq i \leq \ell(w)$. Then, if $\ell(w) \geq t$, there exist i, j with $0 \leq i < j \leq \ell(w)$ and $w_i = w_j$ in H, and then $w_j = w_i v$, where v is the required subword of w. ∎

We now give the proof of Theorem B. The following argument is expressed in terms of null-$\mathcal{P}$-sequences, however the reader familiar with van Kampen diagrams will be able to reinterpret the ideas in geometric terms.

Proof of Theorem B. It suffices to find any one finite presentation $\mathcal{P}$ for G which admits the required $(\mathrm{Area}, \mathrm{FL})$-pair.

Let $\mathcal{A}$ be the generating set for G defined both in §1 and before Corollary 3.2. Define $\overline{G} := G/\Gamma_c$, which is a finitely generated nilpotent group of class $c-1$.

Suppose $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ is a finite presentation for G. Then define $\overline{\mathcal{P}} = \langle \overline{\mathcal{A}} \mid \overline{\mathcal{R}} \rangle$ to be the finite presentation for $\overline{G}$ in which $\overline{\mathcal{A}}$ is obtained from $\mathcal{A}$ by removing all generators of weight c, and $\overline{\mathcal{R}}$ is obtained from $\mathcal{R}$ by deleting all occurrences of these generators from the words in $\mathcal{R}$.

We prove that (n^{c+1}, n) is an $(\mathrm{Area}, \mathrm{FL})$-pair for G by induction on c. The base case of $c = 1$ concerns abelian groups. The process of moving letters through words, collecting together and cancelling every instance of each particular generator, gives the required $(\mathrm{Area}, \mathrm{FL})$-pair of (n^2, n).

We now prove the induction step. Suppose w is a null-homotopic word in a finite presentation $\mathcal{P}$ for G. Define $n := \ell(w)$. Let $\overline{w}$ be the word obtained from w by removing all occurrences of generators in $\mathcal{A}_c$. By induction hypothesis, $\overline{\mathcal{P}}$ admits (n^c, n) as an $(\mathrm{Area}, \mathrm{FL})$-pair. Thus there is a constant $\overline{\lambda}$ depending only on $\overline{\mathcal{P}}$ such that there is a null-$\overline{\mathcal{P}}$-sequence $\overline{S}$:

$$\overline{w} \ = \ \overline{w}_0, \overline{w}_1, \ldots, \overline{w}_{\overline{m}} \ = \ 1$$

for $\overline{w}$ with $\mathrm{Area}(\overline{S}) \leq \overline{\lambda} n^c$ and $\mathrm{FL}(\overline{S}) \leq \overline{\lambda} n$. Our intention is to produce a null-$\mathcal{P}$-sequence S for w from $\overline{S}$ that demonstrates that G admits (n^{c+1}, n) as an $(\mathrm{Area}, \mathrm{FL})$-pair.

Now $\mathcal{A}_c$ generates the abelian group Γ_c, and it follows from Lemma 3.4, that we can write $\mathcal{A}_c = \mathcal{A}_{c1} \cup \mathcal{A}_{c2}$ where $\mathcal{A}_{c2}$ freely generates a free abelian group K having finite index in Γ_c. Define $t := |\Gamma_c/K|$. Let $\mathcal{A}_{c2} = \{z_1, z_2, \ldots, z_k\}$. For $s \geq 0$ let $u_{z_j}(s)$ denote the *compression words* of Corollary 3.3.

The elements of $\mathcal{A}_c$ are all central in G, so we may as well assume that $\mathcal{R}$ includes the commutators $[x, z]$ for all $x \in \mathcal{A}^{\pm 1}$ and $z \in {\mathcal{A}_c}^{\pm 1}$.

Lemma 2.4 tells us that, by suitably adjusting the constant $\overline{\lambda}$, we can assume that each time a word $\overline{w}_{i+1}$ is obtained from a word $\overline{w}_i$ in the sequence $\overline{S}$ by *applying a relator* $\overline{r} \in \overline{\mathcal{R}}$, the whole of some cyclic conjugate of $\overline{r}^{\pm 1}$ is inserted into $\overline{w}_i$.

Suppose a word u of length at most t in the letters ${\mathcal{A}_{c1}}^{\pm 1}$ represents an element of K. Then we can choose a word v_u in the letters ${\mathcal{A}_{c2}}^{\pm 1}$ such that $u = v_u$ in G. There are only finitely many words of length at most t so we may as well assume that each $u{v_u}^{-1}$ is in $\mathcal{R}$.

We now explain how to find a null-$\mathcal{P}$-sequence for w, by *expanding* $\overline{S}$. Suppose that the move from $\overline{w}_i$ to $\overline{w}_{i+1}$ is an instance of an *application of a relator* move in the null-$\overline{\mathcal{P}}$-sequence $\overline{S}$. So $\overline{w}_i = \alpha\beta$ for some words α and β. Also $\overline{w}_{i+1} = \alpha\overline{r}\beta$, where some cyclic conjugate of $\overline{r}$ is in $\overline{\mathcal{R}}^{\pm 1}$. There is some r which has a cyclic conjugate in $\mathcal{R}^{\pm 1}$ and from which one can obtain $\overline{r}$ by removing all occurrences of generators in ${\mathcal{A}_c}^{\pm 1}$. The idea is to apply an *expanded version* of the move: instead of inserting $\overline{r}$, we insert r. The newly introduced generators from ${\mathcal{A}_c}^{\pm 1}$ are then collected at the ends of the word and *compressed* before the $\mathcal{P}$-sequence is continued as dictated by $\overline{\mathcal{P}}$.

More precisely, we construct a null-$\mathcal{P}$-sequence S for w as follows. Some of the words in S will be of the form

$$w_i \;=\; u_{z_k}(q_{ki}^-)^{-1} \,\ldots\, u_{z_2}(q_{2i}^-)^{-1}\, u_{z_1}(q_{1i}^-)^{-1}\, \overline{w}_i\, v_i\, u_{z_1}(q_{1i}^+)\, u_{z_2}(q_{2i}^+) \,\ldots\, u_{z_k}(q_{ki}^+),$$

for $0 \leq i \leq \overline{m}$. In w_i each q_{ji}^+ and q_{ji}^- is a non-negative integer, and v_i is a word of length less than t in the generators ${\mathcal{A}_{c1}}^{\pm 1}$.

The $\mathcal{P}$-sequence S starts by transforming w to w_0 by moving all generators from ${\mathcal{A}_c}^{\pm 1}$, using commutator relations (recall that we assumed these to be in $\mathcal{R}$) as follows: move all occurrences of the letters $z_1, z_2, \ldots, z_k$ of $\mathcal{A}_{c2}$ towards the (right-hand) end of the word, all occurrences of ${z_1}^{-1}, {z_2}^{-1}, \ldots, {z_k}^{-1}$ towards the start of the word. Move all generators from ${\mathcal{A}_{c1}}^{\pm 1}$ towards the right to make the subword v_0.

We will soon explain more precisely how these generators from $\mathcal{A}_{c1}$ and $\mathcal{A}_{c2}$ are *collected* and *compressed*. But first we begin to explain how to transform w_i to w_{i+1}. First start by transforming the subword $\overline{w}_i$ as dictated by $\overline{S}$: if the move is a free expansion or a free reduction then this move is performed and we immediately have w_{i+1}. However if the move is an *application of a relator* then we apply the *expanded version*. This introduces letters from ${\mathcal{A}_c}^{\pm 1}$. Move the $z_1, z_2, \ldots, z_k$ towards the end of the word, move the ${z_1}^{-1}, {z_2}^{-1}, \ldots, {z_k}^{-1}$ towards the start of the word, and move ${\mathcal{A}_{c1}}^{\pm 1}$ to the start of the subword v_i.

Now when a letter $z_j \in \mathcal{A}_c$ is moved towards the (right-hand) end of a word (in the circumstance of either of the two preceding paragraphs) it will arrive at some subword word $u_{z_j}(q)$, where $q \geq 0$. We then transform

$$z_j \, u_{z_j}(q) \ \rightarrow \ u_{z_j}(q+1)$$

via some $\mathcal{P}$-sequence S_{jq}. Similarly when a ${z_j}^{-1}$ is moved towards the start of the word it meets some $u_{z_j}(q)^{-1}$, where $q \geq 0$, and we then make the transformation

$$u_{z_j}(q)^{-1} \, {z_j}^{-1} \ \rightarrow \ u_{z_j}(q+1)^{-1}$$

using the $\mathcal{P}$-sequence obtained by inverting every word in S_{jq}.

Consider moving a letter z from ${\mathcal{A}_{c1}}^{\pm 1}$ to the location of v_i in w_i to produce a subword zv_i. If a subword of zv_i represents an element of K we immediately replace it by a word v_{i+1} in the letters ${\mathcal{A}_{c2}}^{\pm 1}$ and, by assumption, we can do this using a relator in $\mathcal{R}$. The resulting letters from ${\mathcal{A}_{c2}}^{\pm 1}$ are then moved and collected in the manner explained in the previous paragraph. It follows from Lemma 3.5 that the words v_i will always have length less than t. The same strategy is used to produce the subword v_0 of w_0 when transforming w to w_0.

Now $w_{\overline{m}}$ is a null-homotopic word in Γ_c, and so $v_{\overline{m}}$ represents an element of K. But in that case $v_{\overline{m}}$ must be the empty word. So $q^+_{j\overline{m}} = q^-_{j\overline{m}}$ for $j = 1, 2, \ldots, k$ because $\mathcal{A}_{c2}$ freely generates K. Thus we can complete our $\mathcal{P}$-sequence for w by freely reducing $w_{\overline{m}}$ to the empty word.

All that remains is to explain how the process above gives a $\mathcal{P}$-sequence such that

$$\begin{aligned} \mathrm{Area}(S) \ &\leq \ \lambda \, n^{c+1}, \\ \mathrm{FL}(S) \ &\leq \ \lambda \, n, \end{aligned}$$

where λ is a constant depending only on $\mathcal{P}$.

Let M be the maximum number of letters from ${\mathcal{A}_c}^{\pm 1}$ occurring in any one of the relators in $\mathcal{R}$. The powers q^+_{ji} and q^-_{ji} are all bounded by $(M + M^2)\mathrm{Area}(\overline{S}) + n + Mn$, as we see because letters from ${\mathcal{A}_{c2}}^{\pm 1}$ only arise in the following four ways:

(i). At most M letters from ${\mathcal{A}_{c2}}^{\pm 1}$ arise on account of each of the *application of a relator* moves in $\overline{S}$, of which there are at most $\text{Area}(\overline{S})$.

(ii). At most $M\text{Area}(\overline{S})$ letters from ${\mathcal{A}_{c1}}^{\pm 1}$ appear as a result of *expanding* the *application of a relator* moves in $\overline{\mathcal{P}}$. For each such letter $z \in {\mathcal{A}_{c1}}^{\pm 1}$, at most one relator is used to transform some zv_i into v_{i+1}, and at most M further letters from ${\mathcal{A}_{c2}}^{\pm 1}$ are then released: that is, a total of at most $M^2\text{Area}(\overline{S})$ letters.

(iii). There are at most n letters from ${\mathcal{A}_{c2}}^{\pm 1}$ in w.

(iv). Also there are at most n letters from ${\mathcal{A}_{c1}}^{\pm 1}$ in w and recall that in forming w_0 we collected these together to make v_0 – after each letter is moved to the location of v_0, a relation may be applied, releasing at most M letters from ${\mathcal{A}_{c2}}^{\pm 1}$.

Therefore Corollary 3.3 tells us that the length of each of the subwords $u_{z_j}(q_{ji}^+)$ and $u_{z_j}(q_{ji}^-)$ for powers of elements of $\mathcal{A}_{c2}$ is $O(n)$. Moreover, by that corollary we also know that the total contribution to the area of S arising in the compression process producing these subwords is $O(n^{c+1})$.

It follows that there is a constant $\lambda' > 0$, depending only on $\mathcal{P}$, such that $\text{FL}(S) \leq \lambda' n$. For the bound on $\text{Area}(S)$ notice that all that remains is to bound the contribution made by our use of commutators to move letters through words. Well, each of the $O(n^c)$ letters from ${\mathcal{A}_c}^{\pm 1}$ that arises in S is moved a distance at most $\lambda' n$. So indeed, we can find $\lambda \geq \lambda'$, depending only on $\mathcal{P}$, such that $\text{Area}(S) \leq \lambda\, n^{c+1}$ and $\text{FL}(S) \leq \lambda\, n$. ∎

A Isoperimetric functions and growth of cocycles

Corollary A.1 follows from a general result relating isoperimetric functions to growth of cocycles. In accordance with the referee's request we include a proof of this theorem.

Theorem C. *Let G be a finitely presented group and let $\zeta \in H^2(G, \mathbb{R})$. Let $h : \mathbb{N} \to \mathbb{N}$ be an isoperimetric function for G. Then there is a 2-cocycle $f : G \times G \to \mathbb{R}$ in the class ζ and a function $h' : \mathbb{N} \to \mathbb{N}$ equivalent to h so that $|f(x, y)| \leq h'(|x| + |y|)$ for all $x, y \in G$.*

Proof. Let $\mathcal{P}$ be a finite presentation for G and let $X_{\mathcal{P}}$ be the presentation 2-complex. Let X be a space of type $K(G,1)$ with $X^{(2)} = X_{\mathcal{P}}$; recall that X is obtained from $X_{\mathcal{P}}$ by attaching cells of dimension at least 3 to kill homotopy groups π_n, $n \geq 2$. Note that the 1-skeleton $\Gamma := \widetilde{X}^{(1)}$ of the universal cover $\widetilde{X}$ of X is the Cayley graph of $\mathcal{P}$, and $\widetilde{X}^{(2)}$ is the Cayley 2-complex.

It is known that the following 3 groups are canonically isomorphic [4]:

(i). $H^2(G, \mathbb{R})$,

(ii). $H^2(X, \mathbb{R})$, and

(iii). the equivariant cohomology ${H^2}_G(\widetilde{X}, \mathbb{R})$ for the left G-action on $\widetilde{X}$.

Thus $\zeta \in H^2(G, \mathbb{R})$ is represented by an *equivariant* 2-cocycle $F \in Z^2(\widetilde{X}, \mathbb{R})$. In particular since there are only finitely many orbits of 2-cells of $\widetilde{X}$ under the G-action, it follows that there exists $M > 0$ so that $|F(e^{(2)})| \leq M$ for all 2-cells $e^{(2)}$ of $\widetilde{X}$.

Choose a vertex v of Γ as the base point for $\widetilde{X}$, so we can identify $\widetilde{X}^{(0)}$ equivariantly with G; then v is identified with the neutral element 1 of G. Choose a geodesic combing $\{p(x) \mid x \in G\}$ for Γ from the base point; this means that $p(x)$ is an edge-path from 1 to x which is geodesic for the word metric. Using the left G-action we define $p(x,y) = x \cdot p(x^{-1}y)$, a geodesic edge-path from x to y in Γ. For each pair $(x,y) \in G^{(2)}$ let $\mathfrak{c}(x,y) = p(x) \cdot p(x,xy) \cdot p(xy)^{-1}$, an edge-circuit in Γ, and let $D(x,y)$ be a minimal area van Kampen diagram filling $\mathfrak{c}(x,y)$. Finally let $f(x,y) = F(D(x,y)) := \sum_{e^{(2)}} F(e^{(2)})$, where $e^{(2)}$ in the sum ranges over the oriented 2-cells of $D(x,y)$. (The cocycle condition for F implies that $f(D(x,y)) = F(D')$, where D' is any other van Kampen diagram filling $\mathfrak{c}(x,y)$. The minimality of $D(x,y)$ for area is used only for the purpose of estimating $|f(x,y)|$.)

The following properties hold for f and suffice to establish the theorem:

(1). f is a 2-cocycle on the bar construction,

(2). $f \in \zeta$, and

(3). $f(x,y) \leq M\,\mathrm{Area}(2(|x| + |y|))$ for all $x, y \in G$, where Area is the Dehn function for $\mathcal{P}$.

Claim (1) follows from the fact that F is a 2-cocycle on $\widetilde{X}$; this is illustrated in Figure 1. Claim (2) is a consequence of the construction of the

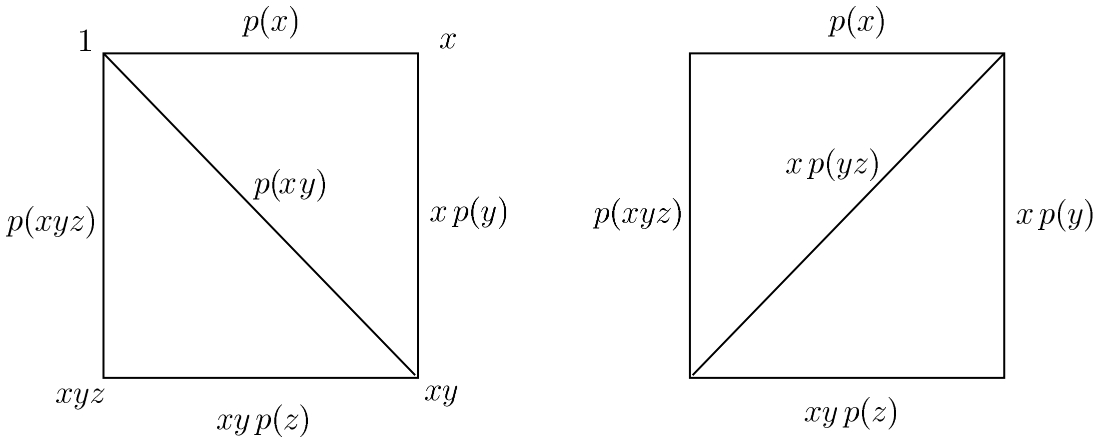

Figure 1: The cocycle condition

canonical isomorphism between $H^2(G,\mathbb{R})$ and $H^2{}_G(X,\mathbb{R})$. We now establish claim (3).

We have $|f(x,y)| = |F(D(x,y))| = |\sum_{e^{(2)} \text{ in } D(x,y)} F(e^{(2)})| \leq M\text{Area}(D(x,y))$. But $\ell(\mathfrak{c}(x,y)) = |x| + |y| + |xy| \leq 2(|x| + |y|)$, whence $\text{Area}(D(x,y)) \leq \text{Area}(2(|x| + |y|))$. It follows that $|f(x,y)| \leq M\,\text{Area}(2(|x| + |y|))$, and the proof is complete. ∎

Acknowledgement. We are grateful to an anonymous referee for some minor suggestions.

References

[1] D. Allcock. An isoperimetric inequality for the Heisenberg groups. *Geom. Funct. Anal.*, 8(2):219–233, 1998.

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

[3] M. R. Bridson. On the geometry of normal forms in discrete groups. *Proc. London Math. Soc.*, 67(3):596–616, 1993.

[4] K. S. Brown. *Cohomology of groups.* Number 87 in Graduate Texts in Mathematics. Springer Verlag, 1982.

[5] J. Cannon, O. Gooodman, and M. Shapiro. Dehn's algorithm for non-hyperbolic groups. Preprint.

[6] G. Conner. Isoperimetric functions for central extensions. In R. Charney, M. Davis, and M. Shapiro, editors, *Geometric Group Theory*, volume 3 of *Ohio State University, Math. Res. Inst. Publ.*, pages 73–77. de Gruyter, 1995.

[7] 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, 1992.

[8] S. M. Gersten. Isodiametric and isoperimetric inequalities in group extensions. Preprint, University of Utah, 1991.

[9] S. M. Gersten. Dehn functions and l_1-norms of finite presentations. In G. Baumslag and C. Miller, editors, *Algorithms and classification in combinatorial group theory (Berkeley, CA, 1989)*, volume 23 of *Math. Sci. Res. Inst. Publ.*, pages 195–224. Springer-Verlag, 1992.

[10] S. M. Gersten. Isoperimetric and isodiametric functions. In G. Niblo and M. Roller, editors, *Geometric group theory I*, number 182 in LMS lecture notes. Camb. Univ. Press, 1993.

[11] S. M. Gersten and T. R. Riley. Filling length in finitely presentable groups. *Geom. Dedicata*, 92:41–58, 2002.

[12] S. M. Gersten and H. Short. Some isoperimetric inequalities for free extensions. *Geom. Dedicata*, 92:63–72, 2002.

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

[14] P. Hall. Nilpotent groups. Notes of lectures given at the Canadian Mathematical Congress, University of Alberta, 1957.

[15] C. Hidber. Isoperimetric functions of finitely generated nilpotent groups. *J. Pure Appl. Algebra*, 144(3):229–242, 1999.

[16] D. F. Holt. Decision problems in finitely presented groups. In *Computational methods for representations of groups and algebras (Essen, 1997)*, pages 259–265. Birkhäuser, Basel, 1999.

[17] P. Pansu. Croissance des boules et des géodesiques fermées dans les nilvariétés. *Ergodic Theory Dynam. Systems*, 3:415–445, 1983.

[18] Ch. Pittet. Isoperimetric inequalities for homogeneous nilpotent groups. In R. Charney, M. Davis, and M. Shapiro, editors, *Geometric Group Theory*, volume 3 of *Ohio State University, Mathematical Research Institute Publications*, pages 159–164. de Gruyter, 1995.

[19] S. Pride. Geometric methods in combinatorial semigroup theory. In *Semigroups, formal languages and groups (York, 1993)*, pages 215–232. Kluwer Acad. Publ., Dordrecht, 1995.

[20] T. R. Riley. *Asymptotic invariants of infinite discrete groups.* DPhil thesis, Oxford University, 2002.

[21] B. A. F. Wehrfritz. *Infinite Linear Groups.* Springer-Verlag, 1973.

Steve M. Gersten
Mathematics Department, 155S. 1400E., Rm. 233, Univ. of Utah, Salt Lake City, UT 84112, USA
`gersten@math.utah.edu, http://www.math.utah.edu/~gersten`

Derek F. Holt
Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK
`dfh@maths.warwick.ac.uk, http://www.maths.warwick.ac.uk/~dfh`

Tim R. Riley
Mathematical Institute, 24-28 St. Giles', Oxford OX1 3LB, UK
`rileyt@maths.ox.ac.uk, http://www.maths.ox.ac.uk/~rileyt`