DEHN FUNCTIONS AND L_1-NORMS OF FINITE PRESENTATIONS

S. M. Gersten

Abstract. An l_1-norm function is introduced as a computational tool to estimate the Dehn function of a finite presentation and settle certain imbedding questions for finitely presented groups.

§1. Introduction

This note represents an initial attempt to decide what are the finitely presented subgroups of an automatic group in the sense of Thurston [**CEHPT**]. We may start by comparing the situation with Gromov's hyperbolic groups [**Gr**]. A hyperbolic group is a finitely generated group whose Cayley graph in the word metric satisfies the thin triangle condition: there is a positive number $A > 0$ so that for every geodesic triangle XYZ and point $P \in [X, Y]$ the distance of P from the union of the sides $[X, Z] \cup [Z, Y]$ is bounded by A. An automatic group G is one possessing a finite set of semigroup generators S and a regular language $\mathcal{L}$ in the free monoid S^* generated by S so that $\mathcal{L}$ represents every element of G and so that the question whether any two elements of $\mathcal{L}$ represent elements of G at most a unit distance apart in the Cayley graph can be decided by a finite state automaton. For more information on hyperbolic groups and on automatic groups the reader should consult [**BGSS**] in this volume; other references are the original papers [**Gr**] and [**CEHPT**] as well as the two fine sets of notes, [**CDP**] and [**GH**].

A result of Gromov's is that a hyperbolic group contains no subgroup isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$ (see [**Gr**] Corollary 8.2.C, page 212); this turns out to be a powerful necessary criterion for a group to be hyperbolic. The same techniques involving translation lengths along geodesics can be used to show that no Baumslag-Solitar group $B_{k,l} = G\langle x, y \mid yx^k\bar{y} = x^l \rangle$, $k \cdot l \neq 0$, imbeds as a subgroup of a hyperbolic group. Since it follows from the results of [GS2] that the groups $B_{k,\pm k}$ are automatic, the natural question to ask is whether $B_{k,l}$ imbeds as a subgroup of an automatic group if $|k| \neq |l|$. This question is open in general, but we shall prove the following result.

Theorem A. *If G is the fundamental group of a finite aspherical 2-complex and G is automatic, then $B_{k,l}$ does not imbed as a subgroup of G if $|k| \neq |l|$.*

Our techniques involve calculating a lower bound for the Dehn function δ_X of a finite connected 2-complex X. Here $\delta_X(n)$ is the maximum over all tuples of null homotopic circuits $\mathbf{w}$ of total length at most n of the sum of the minimum number of faces in van Kampen diagrams for the component circuits of $\mathbf{w}$. The Dehn function of a finite presentation is defined to be the Dehn function of the 2-complex canonically associated to the presentation. It turns out that the growth of the function δ_X depends only on the fundamental group of X. Furthermore Gromov [**Gr**] has shown that $\pi_1(X)$ is hyperbolic iff there exists a constant $A > 0$ such that $\delta_X(n) \leq An$ for all $n > 0$, while Thurston has shown that if $\pi_1(X)$ is automatic, then there exists a constant $A > 0$ such that $\delta_X(n) \leq An^2$ [**BGSS**].

To estimate the Dehn function we calculate the l_1-norm of the 2-chain $c(\tilde{f})$ determined by a van Kampen diagram $f : D \to X$ for the null homotopic circuit w in $X^{(1)}$; here $\tilde{f} : D \to \tilde{X}$ is a lift of $f : D \to X$ to the universal cover $\tilde{X}$ of X. The l_1-norm of the 2-chain $c(\tilde{f})$ is an invariant $|w|_1$ of w if X is aspherical. The l_1-norm function $\lambda_X(n)$ can then be defined as the supremum of $|\mathbf{w}|_1$ over all tuples of such circuits $\mathbf{w}$ in $X^{(1)}$ of total length at most n, where $|\mathbf{w}|_1$ is the sum of the l_1-norms $|w|_1$ of the component circuits w of $\mathbf{w}$. An analogous definition of λ_X is made in general taking into account the fact that the chain $c(\tilde{f})$ is determined only modulo spherical cycles. The growth of the function λ_X depends only on $\pi_1(X)$ and in general we have $\lambda_X \leq \delta_X$.

We can now refine the statement of theorem A as the consequence of theorems B and C below, which constitute the main results of this note.

Theorem B. *If X is the 2-complex canonically associated to the Baumslag-Solitar presentation $\langle\, x, y \mid yx^k\bar{y} = x^l \,\rangle$ where $|k| \neq |l|$, then $\lambda_X(n)$ grows faster than any polynomial function of n; that is, there are no constants $A, d > 0$ with $\lambda_X(n) \leq An^d$ for all $n > 0$.*

Theorem C. *Let $f : X \to Y$ be a map of finite aspherical 2-complexes such that the induced homomorphism on fundamental groups is injective. Then there exist positive constants $A, B,$ and C such that $\lambda_X(n) \leq A\lambda_Y(Bn) + Cn$ for all $n > 0$.*

By the results of [**BGSS**] the Baumslag-Solitar groups $B_{k,l}$ are all asynchronously automatic and hence their Dehn functions' growth is bounded by a simple exponential, where a simple exponential function of n is of the form A^n for some constant $A > 0$. In §6 we give a number of examples of the use of our results and we produce some examples of even more rapidly growing Dehn functions. In particular we make use of Dehn functions to settle negatively a question raised by Baumslag, whether a 1-relator group of deficiency at least 2 is an amalgam of finitely generated free groups amalgamating finitely generated subgroups; our example $\langle\, x, y, z \mid x^{x^y} = x^2 \,\rangle$, where $x^y = yx\bar{y}$, has a Dehn function which grows faster than any simple exponential.

This article is based on an address given in the Workshop on Algorithmic Problems held at MSRI in January, 1989. We should like to thank the referee for his careful reading of this paper and for his suggestions, which we have adopted to aid the exposition.

§2. THE FUNCTION δ_X

We shall consider combinatorial 2-complexes and combinatorial maps [**G**]. Recall that a combinatorial map is one whose restriction to each open cell of each dimension is a homeomorphism onto its image. A combinatorial 2-complex is one whose attaching maps are combinatorial maps. We shall use freely the terminology of van Kampen diagrams [**LS**]. Recall that a van Kampen diagram is a combinatorial map whose domain is a contractible plane point set. In addition we shall need at one point to consider the quotient complex of a combinatorial 2-complex by a subcomplex. There is a canonical way of giving the quotient the structure of a combinatorial 2-complex that is discussed in [**G**].

A presentation $\mathcal{P}$ gives rise to a combinatorial 2-complex $K(\mathcal{P})$ whose 0-skeleton consists of a single point, whose edges are generators and whose 2-cells are attached by maps that spell out the relators. Presentations and such 2-complexes are essentially the same thing, where the relators are determined up to cyclic permutation of letters and order of relators.

Definition 2.1. Let X be a finite connected combinatorial 2-complex. Let $\mathbf{w} = (w_1, w_2, \dots, w_r)$ be an ordered set (for variable r) of edge circuits in the 1-skeleton each of which is null homotopic in X. We define the length $\ell(\mathbf{w}) = \sum_{i=1}^{r} \ell(w_i)$, where $\ell(w_i)$ is the length of w_i. We define $\mathrm{Area}_X(w_i)$ to be the minimal number of faces (=2-cells) in a van Kampen diagram for w_i and we set $\mathrm{Area}_X(\mathbf{w}) = \sum_i \mathrm{Area}_X(w_i)$. We define the Dehn function δ_X (or *isoperimetric function*) on the natural numbers by setting $\delta_X(n) = \max_{\ell(\mathbf{w}) \leq n} \mathrm{Area}_X(\mathbf{w})$.

The motivation for this definition comes from differential geometry, where we span a null homotopic loop w by a minimal surface and calculate the area of this surface. The fact that we have to consider families of null homotopic loops here is a technicality to make our definitions homotopy invariant.

If $\mathcal{P}$ is a finite presentation, we set $\delta_{\mathcal{P}} = \delta_{K(\mathcal{P})}$.

The interest in the Dehn function of a finite presentation $\mathcal{P}$ is explained by the next result. Here $G(\mathcal{P}) = \pi_1(K(\mathcal{P}))$ is the group of the presentation.

Proposition 2.2. *If $\mathcal{P}$ is a finite presentation, then the group $G(\mathcal{P})$ has a solvable word problem iff $\delta_{\mathcal{P}}$ is recursive.*

Proof. Assume first that $\delta_{\mathcal{P}}$ is recursive and a word w in the generators is given. One can then calculate $\delta_{\mathcal{P}}(\ell(w))$, where w is identified with the 1-tuple (w), to obtain an effective bound on the number of relators and their inverses in an expression for w as a product of conjugates of them. We can also obtain an effective bound on the lengths of the words performing the conjugations on the relators as follows. If K is a minimal van Kampen diagram for w, we let $\mathrm{Area}(K)$ be the number of faces and let $\mathrm{girth}(K)$ be the length of the longest edge path in $K^{(1)}$ which does not contain a circuit. The lengths of words performing the conjugations are then bounded by $\mathrm{girth}(K)$ and we have the estimate

$$\mathrm{girth}(K) \leq L\ \mathrm{Area}(K) + \ell(w),$$

where L is the length of the longest relator in $\mathcal{P}$. Since $\mathrm{Area}(K)$ is effectively estimated by $\delta_{\mathcal{P}}(\ell(w))$, this gives an effective bound on the lengths of the conjugating words. It follows that the words of length at most n which represent 1 in $G(\mathcal{P})$ can be effectively listed and one can hence check whether or not a given word w lies on this list. Thus $G(\mathcal{P})$ has a solvable word problem if $\delta_{\mathcal{P}}$ is recursive.

Conversely assume that $G(\mathcal{P})$ has a solvable word problem. Then it is decidable whether each ordered r-tuple $\mathbf{w} = (w_1, w_2, \ldots, w_r)$ of words is such that each word w_i represents 1. In particular, for such tuples $\mathbf{w}$, one can calculate effectively an expression for each component as a product of conjugates of relators and inverses of relators. Thus one can also calculate a minimal such expression, involving the fewest relators and inverse relators (making use of the girth estimate in the preceding paragraph). Thus $\mathrm{Area}_{\mathcal{P}}(\mathbf{w})$ is calculable for such tuples $\mathbf{w}$ and consequently the maximal area of such tuples $\mathbf{w}$ with $\ell(\mathbf{w}) \leq n$ is effectively calculable. That is, $\delta_{\mathcal{P}}$ is effectively calculable.

Definition 2.3. Suppose that f and g are two functions defined on $\mathbb{N}$ taking non negative real values. We write $f \sim g$ if there are positive constants A, A', B, B', C and C' such that $f(n) \leq Ag(Bn) + Cn$ and $g(n) \leq A'f(B'n) + C'n$ for all $n \in \mathbb{N}$. This relation is clearly an equivalence relation, and we say that two equivalent functions have the same 'growth'.

Proposition 2.4. *If $\mathcal{P}$ and $\mathcal{Q}$ are finite presentations for isomorphic groups, then $\delta_{\mathcal{P}} \sim \delta_{\mathcal{Q}}$.*

Proof. We shall show that the Dehn functions for Tietze equivalent presentations have the same growth. Suppose that $\mathcal{P} = \langle \mathcal{X} \mid \mathcal{R} \rangle$ and suppose that $\mathcal{Q}$ is obtained from $\mathcal{P}$ by adjoining one additional relator R, where R is a consequence of $\mathcal{R}$. If all components of the tuple $\mathbf{w}$ represent 1, then we clearly have $\mathrm{Area}_{\mathcal{P}}(\mathbf{w}) \geq \mathrm{Area}_{\mathcal{Q}}(\mathbf{w})$, where $\mathbf{w}$ is a tuple of words in the generators $\mathcal{X}^{\pm 1}$. On the other hand, the word R is the boundary label of a van Kampen diagram in $\mathcal{P}$. If A is the number of faces in one such $\mathcal{P}$ diagram for R, then we see by subdividing the faces labelled R in $\mathcal{Q}$ diagrams that $\mathrm{Area}_{\mathcal{P}}(\mathbf{w}) \leq A \cdot \mathrm{Area}_{\mathcal{Q}}(\mathbf{w})$. It follows from these observations that $\delta_{\mathcal{P}} \sim \delta_{\mathcal{Q}}$.

Now suppose that $\mathcal{Q}$ is obtained from $\mathcal{P}$ by adjoining one free generator $t \notin \mathcal{X}$ and adjoining one relator tr^{-1} with r a word of length L in the generators $\mathcal{X}^{\pm 1}$. If $\mathbf{w}$ is a tuple of words in $\mathcal{X}$ each representing 1 in $G(\mathcal{P})$, then $\mathrm{Area}_{\mathcal{P}}(\mathbf{w}) = \mathrm{Area}_{\mathcal{Q}}(\mathbf{w})$. This is true since a face with label tr^{-1} can never have an edge labelled t as an interior edge in a minimal diagram for $\mathbf{w}$, for otherwise that diagram would not be reduced.

Suppose next that $\mathbf{u}$ is a tuple of words in the generators of $\mathcal{Q}$ representing 1 and $n = \ell(\mathbf{w})$. We have already seen that an edge labelled t cannot occur an an interior edge of a minimal diagram for a component of $\mathbf{u}$. The only occurrences of t in such a diagram occur on the boundary, and there are at most n of them in all the components of $\mathbf{u}$. Thus there are at most a total of n faces with label tr^{-1} in all the minimal diagrams for the components. Next replace each occurrence of t in a component word of $\mathbf{u}$ by the word r (without doing any cancellations) to obtain a tuple $\mathbf{w}$ in $\mathcal{X}^{\pm 1}$. Observe that $\ell(\mathbf{w}) \leq nL$ and all components of $\mathbf{w}$ represent 1. Given minimal $\mathcal{P}$ diagrams for the components of $\mathbf{w}$ we obtain $\mathcal{Q}$ diagrams for the components of $\mathbf{u}$ by attaching at most n faces in total with label tr^{-1} along the boundaries. Hence $\mathrm{Area}_{\mathcal{Q}}(\mathbf{u}) \leq \mathrm{Area}_{\mathcal{P}}(\mathbf{w}) + n$ and consequently $\delta_{\mathcal{Q}}(n) \leq \delta_{\mathcal{P}}(Ln) + n$. Taken with the result of the preceding paragraph, this yields $\delta_{\mathcal{P}} \sim \delta_{\mathcal{Q}}$.

It follows from these calculations and from Tietze's theorem that if $\mathcal{P}$ and $\mathcal{Q}$ are two finite presentations for isomorphic groups, then $\delta_{\mathcal{P}} \sim \delta_{\mathcal{Q}}$. This completes the proof of Proposition 2.4.

We want to extend this result about presentations to 2-complexes. The crucial step in the argument turns out to be an invariance under free product with an infinite cycle. First we observe

Lemma 2.5. *For any finite connected 2-complex X we have $\delta_X(m+n) \geq \delta_X(m) + \delta_X(n)$.*

Proof. Let $\mathbf{u} = (u_1, u_1, \ldots, u_r)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_s)$ be tuples of null homotopic edge circuits in $X^{(1)}$ with $\ell(\mathbf{u}) \leq m$ and $\ell(\mathbf{v}) \leq n$. We may assume that $\mathrm{Area}_X(\mathbf{u})$ and $\mathrm{Area}_X(\mathbf{v})$ are maximal among all tuples with lengths in these ranges. Let $\mathbf{w} = (u_1, u_2, \ldots, u_r, v_1, \ldots, v_s)$ Then $\mathrm{Area}_X(\mathbf{w}) = \mathrm{Area}_X(\mathbf{u}) + \mathrm{Area}_X(\mathbf{v}) = \delta_X(m) + \delta_X(n)$ and $\ell(\mathbf{w}) \leq m + n$. From this follows the desired inequality.

Proposition 2.6. *Let $\mathcal{P}$ be a finite presentation and let $\mathcal{Q}$ be obtained from $\mathcal{P}$ by adjoining one new free generator t and leaving the relators unchanged. Then $\delta_{\mathcal{P}} = \delta_{\mathcal{Q}}$.*

Proof. Since an edge labelled t cannot appear in the interior of a reduced diagram, it is clear that $\delta_{\mathcal{P}} \leq \delta_{\mathcal{Q}}$. We shall establish the opposite inequality $\delta_{\mathcal{Q}}(n) \leq \delta_{\mathcal{P}}(n)$ by induction on n. The induction starts trivially with $n = 0$. In the inductive step, assume the inequality is established for numbers less than n, $n > 0$, and let $\mathbf{w}$ be a tuple of words representing 1 in $\mathcal{Q}$ with $\ell(\mathbf{w}) = n$. If none of the component words of $\mathbf{w}$ contains t, then $\mathrm{Area}_{\mathcal{Q}}(\mathbf{w}) = \mathrm{Area}_{\mathcal{P}}(\mathbf{w})$. If the component word w_i involves t, then a minimal van Kampen diagram K for w_i cannot be a disc (since an edge labelled t is not in the boundary of any face). We replace w_i with the a set of words obtained by reading the boundary labels of the disc components of K; the disc subdiagram of K for each of these latter words must be minimal, for otherwise K would not be a minimal diagram for w_i. Thus the area is unchanged by replacing w_i by this collection, but the length is reduced by at least one. We replace $\mathbf{w}$ by a new tuple whose components in some ordering are the old components w_j for $j \neq i$ and an ordering of the collection of words with which we replaced w_i. This does not change the area but decreases the length, so the induction hypothesis can apply. Observe that this last step uses the fact that $\delta_{\mathcal{P}}$ is monotone increasing; this fact follows from Lemma 2.5.

Theorem 2.7. *If X and X' are finite connected 2-complexes with isomorphic fundamental groups, then $\delta_X \sim \delta_{X'}$.*

Proof. We shall establish that $\delta_X \sim \delta_{\mathcal{P}}$, where $\mathcal{P}$ is a finite presentation for $\pi_1(X, x)$. This reduces the Theorem to the earlier result, Proposition 2.4, about presentations.

Let T be a maximal tree in the 1-skeleton of X and let $p : X \to Y := X/T$ be the quotient map. We recall here that Y is given the structure of a combinatorial 2-complex by the method of [**G**]. Let $q : X \to X/X^{(0)} := Z$ be the quotient map. Note that p is a homotopy equivalence whereas q imbeds the fundamental group of X as a free factor of $\pi_1(Z)$, where the complementary factor is freely generated by a set of oriented edges of T. Since Y and Z are 1-vertex 2-complexes, they can be considered as presentations, and we shall do so. Let D denote the diameter of the tree T.

Suppose now that $\mathbf{w}$ is a tuple of null homotopic edge circuits in Y with $\ell(\mathbf{w}) = n$. We can lift each component and hence lift $\mathbf{w}$ to a tuple $\mathbf{u}$ of null homotopic edge circuits in X, where $\ell(\mathbf{u}) \leq (D+1)\ell(\mathbf{w})$. By considering minimal diagrams for the components of $\mathbf{u}$ we see that $\mathrm{Area}_Y(\mathbf{w}) \leq \mathrm{Area}_X(\mathbf{u}) \leq \delta_X((D+1)n)$. Hence $\delta_Y(n) \leq \delta_X((D+1)n)$.

Suppose now that $\mathbf{v}$ is a tuple of null homotopic edge circuits in X. Then we have $\mathrm{Area}_X(\mathbf{v}) = \mathrm{Area}_Z(q(\mathbf{v}))$, since identifying vertices does not change the class of diagrams. Here $q(\mathbf{v})$ is the image tuple in Z. But this means that $\delta_X \leq \delta_Z$. Recalling that Z is essentially a presentation, we see that $\delta_Y \sim \delta_Z$ by applying Propositions 2.6 and 2.4. It follows that $\delta_X \sim \delta_Y$. Since Y can be considered a presentation, this completes the proof of the Theorem.

Remark. If we were considering only presentations, we could have defined a simpler notion of the Dehn function, by considering only words representing 1 instead of tuples of words. This more complicated definition was needed to assure that Proposition 2.6 holds.

The next two assertions, 2.8 and 2.9, represent substantial theorems.

Theorem 2.8 (Gromov). *If $\mathcal{P}$ is a finite presentation, then $G(\mathcal{P})$ is hyperbolic iff there is a constant $A > 0$ such that for all $n > 0$ one has $\delta_{\mathcal{P}}(n) \leq An$.*

An elementary proof of the difficult 'if' direction due to H. Short can be found in the appendix of [**GS1**].

Theorem 2.9 (Thurston). *If $\mathcal{P}$ is a finite presentation such that the group $G(\mathcal{P})$ is automatic, then there is a constant $A > 0$ such that for all $n > 0$ one has $\delta_{\mathcal{P}}(n) \leq An^2$,* (the so called 'quadratic isoperimetric inequality') .

It is proved in [**CEHPT**] that if $G(\mathcal{P})$ is automatic, then there is a constant $A > 0$ such that for all words w representing 1 one has $\mathrm{Area}_{\mathcal{P}}(w) \leq A\ell(w)^2$. Using the convexity of the squaring function $x \to x^2$, it follows that $\mathrm{Area}_{\mathcal{P}}(\mathbf{w}) \leq A\ell(\mathbf{w})^2$. Hence $\delta_{\mathcal{P}}(n) \leq An^2$.

The converse to Theorem 2.9 is false. Thurston [**Th**] asserts that the 5-dimensional integral Heisenberg group $G(\langle x_1, x_2, y_1, y_2, z \mid [x_1, x_2] = z =$

$[y_1, y_2]$; $[x_i, y_j] = 1$ for all i, j; z central $\rangle)$ satisfies the quadratic isoperimetric inequality but is not automatic.

Theorem 2.10. *All finite small cancellation presentations $\mathcal{P}$ and all finitely presented subgroups of $G(\mathcal{P})$ satisfy the quadratic isoperimetric inequality.*

A small cancellation presentation $\mathcal{P}$ is one having the property that for any reduced van Kampen diagram D in $\mathcal{P}$, the associated diagram $\bar{D}$ is of *non positive curvature*; here $\bar{D}$ is obtained by removing interior vertices of valence two from D and declaring each face F of $\bar{D}$ to be a regular Euclidean n-gon of unit side, if F has n sides. To say that $\bar{D}$ is of non positive curvature is to say that the angle sum at each interior vertex of $\bar{D}$ is at least 2π. The quadratic isoperimetric inequality here follows from a fundamental result of Reshetnyak's [**Re**]. The result for finitely presented subgroups follows from a covering space argument.

Remark. It is an open question whether a group possessing a finite small cancellation presentation is automatic. It is shown in [**GS1**] and [**GS2**] that all groups possessing finite $C(p) - T(q)$ presentations are automatic, if $(p, q) = (3, 6)$, $(4, 4)$, or $(6, 3)$. Here the $C(p)$ condition means that each face of $\bar{D}$ has at least p sides and the $T(q)$ condition means that each interior vertex of $\bar{D}$ has valence at least q. One may also ask whether the fundamental group of a compact space of non positive curvature, in the general sense Gromov has defined it [**Gr**] in terms of the CAT(0) inequality, is automatic.

We end this section with a useful result that demonstrates why it is necessary to formulate the notion of Dehn function for 2-complexes and not just for finite presentations.

Proposition 2.11. *Let G be a finitely presented group and let $H < G$ be a subgroup of finite index. If $\mathcal{P}$ and $\mathcal{Q}$ are finite presentations for G and H respectively, then $\delta_{\mathcal{P}} \sim \delta_{\mathcal{Q}}$.*

Proof. Let X be the 2-complex canonically associated to $\mathcal{P}$ and let $p : Y \to X$ be the covering space corresponding to $H < G = \pi_1(X)$. If w is an edge circuit in $X^{(1)}$ which is null homotopic in X, then w is covered by an edge circuit $\tilde{w}$ in $Y^{(1)}$ and $\text{Area}_X(w) = \text{Area}_Y(\tilde{w})$; both of these assertions follow from the covering homotopy theorem, since a van Kampen diagram is simply connected. From this observation one deduces easily that $\delta_X \sim \delta_Y$. The result then follows from Theorem 2.7.

§3. THE FUNCTION λ_X

If A is a free abelian group with given basis $\{e_i \mid i \in I\}$ and if $a = \sum_{i\in I} n_i e_i \in A$, where $n_i = 0$ for all but finitely many indices $i \in I$, we set $|a|_1 = \sum_i |n_i|$, the l_1-norm of a. If Y is a 2-complex, then the group of 2-chains with integral coefficients $C_2(Y)$ has a preferred basis of 2-cells, so we have the l_1-norm defined.

Suppose now that X is a finite connnected 2-complex and w is a circuit in $X^{(1)}$ with w null homotopic in X. If we choose a van Kampen diagram $f : D \to X$, we can lift f to the universal cover $\tilde{X}$ to get $\tilde{f} : D \to \tilde{X}$ with $p \cdot \tilde{f} = f$; here $p : \tilde{X} \to X$ is the universal cover of X.

If we choose base points we get a chain $c(\tilde{f}) \in C_2(\tilde{X})$. The l_1-norm $|c(\tilde{f})|_1$ is independent of the choice of base points and of the lift $\tilde{f}$. In fact, if $g : \tilde{X} \to \tilde{X}$ is a deck transformation, then $|c(g \cdot \tilde{f})|_1 = |c(\tilde{f})|_1$.

A more serious matter is the dependence on f. If $f' : D' \to X$ is another van Kampen diagram for w, we can glue f and f' together to get a map $h : S^2 \to X$ with lift $\tilde{h} : S^2 \to \tilde{X}$. Then $c(\tilde{h}) = c(\tilde{f}) - c(\tilde{f}') \in Z_2(\tilde{X})$, where $Z_2(\tilde{X})$ is the cycle group in dimension 2. Recall that by the Hurewicz theorem $Z_2(\tilde{X}) = \pi_2(X)$. We set

$$|w|_1 = \inf_{z \in Z_2(\tilde{X})} |c(\tilde{f}) + z|_1. \tag{3.1}$$

We note the following result which follows easily from the definitions.

Lemma 3.2. *For each null homotopic edge circuit w in X and lift $\tilde{f}$ of the van Kampen diagram f for w we have $|w|_1 \le |c(\tilde{f})|_1$. In addition, $|c(\tilde{f})|_1$ is at most as large as the number of faces in the domain of f.* □

We note the next result which shows how to calculate $|w|_1$ in one case.

Lemma 3.3. *If $\pi_2(X) = 0$, then $|w|_1 = |c(\tilde{f})|_1$ for any van Kampen diagram for w.*

Proof. Since $\pi_2(X) = Z_2(\tilde{X}) = 0$, the indeterminacy z in the definition 3.1 of $|w|_1$ is zero.

Suppose now that $\mathbf{w} = (w_1, w_2, \ldots, w_r)$ is an r-tuple of edge circuits in $X^{(1)}$, each of which is null homotopic in X. We define $|\mathbf{w}|_1 = \sum_i |w_i|_1$ and we define $\lambda_X(n)$ as the maximum value of $|\mathbf{w}|_1$ over all tuples $\mathbf{w}$ satisfying $\ell(\mathbf{w}) \leq n$.

The next result is immediate from Lemma 3.2.

Lemma 3.4. *For all finite connected 2-complexes X we have $\lambda_X \leq \delta_X$.* □

Theorem 3.5. *If X and Y are finite connected 2-complexes with isomorphic fundamental groups, then $\lambda_X \sim \lambda_Y$.*

This result is proved by following the same steps leading to the proof of Theorem 2.7. One first defines $\lambda_{\mathcal{P}} = \lambda_{K(\mathcal{P})}$, where $\mathcal{P}$ is a finite presentation. The 'growth' of $\lambda_{\mathcal{P}}$ is shown to be invariant under Tietze transformation. One proves the analog of Proposition 2.6 and deduces from it that $\lambda_X \sim \lambda_{\mathcal{P}}$, where $\mathcal{P}$ is a finite presentation for $\pi_1(X, x)$. Since the arguments are copies of those in §2, we omit them.

Remark 3.6. Suppose that the Hurewicz map $\pi_2(X) \to H_2(X)$ is zero (such a 2-complex X is called *Cockcroft*). Then the projection map $p : \tilde{X} \to X$ kills spherical cycles. Thus if $c(f) \in C_2(X)$ is the chain of the van Kampen f for the null homotopic edge circuit w, we have $|c(f)|_1 \leq |w|_1$. Hence for a Cockcroft 2-complex X one can estimate the l_1-norm $|w|_1$ from below in X without appealing to $\tilde{X}$. The same remark applies to tuples $\mathbf{w}$ of null homotopic circuits. The next result gives us examples of Cockcroft 2-complexes, which we shall make use of in §5.

Proposition 3.7. *Suppose that M is a CW-complex which is a closed, orientable, aspherical 3-manifold and suppose that M possesses exactly one 3-cell. Then the 2-skeleton $M^{(2)}$ of M is a Cockcroft 2-complex.*

Proof. From the exact homology sequence for the pair $(M, M^{(2)})$, we see that the map $H_2(M^{(2)}) \to H_2(M)$ is injective. The composition of maps $\pi_2(M^{(2)}) \to \pi_2(M) \to H_2(M)$ is zero since M is aspherical. We deduce by a diagram chase that the Hurewicz map $\pi_2(M^{(2)}) \to H_2(M^{(2)})$ is zero. Hence $M^{(2)}$ is Cockcroft.

Proposition 3.8. *Suppose that $\mathcal{P}$ is a finite presentation such that there exists words w_n in the generators representing 1 in $G(\mathcal{P})$ with $\ell(w_n) \to \infty$ as $n \to \infty$, and such that for no positive constant A is it the case that $|w_n|_1 \leq A\ell(w_n)^2$ for all $n > 0$. Then $G(\mathcal{P})$ is not automatic.*

Proof. If $G(\mathcal{P})$ were automatic, then by Theorem 2.9 there exists $A > 0$ such that $\delta_{\mathcal{P}}(n) \leq An^2$ for all $n > 0$. But then we would have $|w_n|_1 \leq \lambda_{\mathcal{P}}(\ell(w_n)) \leq \delta_{\mathcal{P}}(\ell(w_n)) \leq A\ell(w_n)^2$ for all $n > 0$, which is contrary to hypothesis. □

§4. PROOF OF THEOREM B

In this section we fix $\mathcal{P} = \langle\, x, y \mid yx^k\bar{y} = x^l \,\rangle$ with $1 \leq k < l$. We shall construct words w_n in the generators of $\mathcal{P}$ such that $|w_n|_1$ increases more rapidly than any polynomial in $\ell(w_n)$; that is, for no constants $A, d > 0$ does one have $|w_n|_1 \leq A \cdot \ell(w_n)^d$ for all $n > 0$. This will establish Theorem B of §1. In addition it then follows from Proposition 3.8 that $G(\mathcal{P})$ is not automatic. The stronger assertion that $G(\mathcal{P})$ cannot even be imbedded as a subgroup of the fundamental group of a finite aspherical 2-complex follows from Theorem C, which will be proved in the next section. Similar considerations apply to the other Baumslag-Solitar presentations in Theorem B when $|k| \neq |l|$; the modifications of the arguments given here are straightforward and are left to the reader.

Observe that $X = K(\mathcal{P})$ is aspherical, by Lyndon's theorem [**LS**] Proposition 11.1, page 161. Hence by Lemma 3.3, $|w|_1 = |c(\tilde{f})|_1$ where $f : D \to X$ is any van Kampen diagram for w and $\tilde{f}$ is a lift to $\tilde{X}$.

We define inductively two sequences of non negative integers, a_i for $i \geq 1$ and b_j for $j \geq 0$, satisfying $a_1 = k$, $b_0 = 0$, and $k | a_i$, $0 \leq b_i \leq k-1$ for all i. If a_i and b_i have been defined with these properties for all $i \leq r$, we let b_{r+1} be the least natural number with $\frac{l}{k}a_r + b_{r+1}$ divisible by k, and we set $a_{r+1} = \frac{l}{k}a_r + b_{r+1}$. One sees immediately that the induction continues. The desired words w_n are defined by

$$w_n = [y^n x^k \bar{y} x^{b_2} \bar{y} x^{b_3} \ldots \bar{y} x^{b_n} \bar{y}, x]$$

where $[a, b] = ab\bar{a}\bar{b}$. Observe that since $0 \leq b_i < k$ we have $\ell(w_n) \leq 2(k+2)n + 2$, which is bounded by a linear function of n. In addition w_n represents 1 in $G(\mathcal{P})$. To see this observe that $u_n = y^n x^k \bar{y} x^{b_2} \ldots \bar{y} x^{b_n} \bar{y}$ represents $x^{\frac{l}{k}a_n}$ and a van Kampen diagram $\mathcal{D}_n$ for w_n is obtained by gluing two copies of a diagram for $v_n = u_n \cdot \bar{x}^{\frac{l}{k}a_n}$ together along a segment of the boundary labelled $x^{\frac{l}{k}a_n - 1}$. See Figure 2 below.

We show schematically a van Kampen diagram $\mathcal{E}_n$ for v_n below in Figure 3, drawn for $n = 4$.

The number of faces of $\mathcal{E}_n$ is $\frac{1}{k}(a_1 + a_2 + \cdots + a_n) \geq \frac{1}{k}(1 + \frac{l}{k} + \cdots + (\frac{l}{k})^{n-1}) = \frac{1}{k} \cdot \frac{l^n - k^n}{l-k}$. Thus if $\mathcal{D}_n$, the diagram constructed for w_n, has F_n faces, then $F_n \geq$

$\frac{2}{k}\frac{l^n-k^n}{l-k}$. Since this number increases exponentially with n whereas $\ell(w_n)$ increases at most linearly, it follows that F_n increases more rapidly than any polynomial in $\ell(w_n)$. If we can show that there is no cancellation in the chain associated to a lift of $\mathcal{D}_n$ to $\tilde{X}$ between any of the lifted faces, then it will follow that $|w_n|_1$ increases faster than any polynomial in $\ell(w_n)$. This will establish Theorem B.

We chose the base point in the 2-cell α of X as shown in Figure 4.

The base point in the domain of $\mathcal{D}_n$ is shown in Figure 2. If $\tilde{\alpha}$ is a lift of α to $\tilde{X}$, then $c(\tilde{\mathcal{D}}_n) = (1-x)\cdot c(\tilde{\mathcal{E}}_n)$ where the chains have coefficients in the integral group ring $\mathbb{Z}G$ with $G = \pi_1(X)$. We have

$$c(\tilde{\mathcal{E}}_n) = \{(1+x^l+x^{2l}+\cdots) + y(1+x^l+x^{2l}+\cdots)$$

$$+y^2(1+x^l+\cdots)+\cdots+y^{n-1}\}\cdot\tilde{\alpha}.$$

Thus we must show there is no cancellation among the terms of the following expression in $\mathbb{Z}G$:

$$(1-x)\{(1+x^l+x^{2l}+\cdots)+y(1+x^l+x^{2l}+\cdots)+\cdots+y^{n-1}\} \tag{4.1}$$

It is convenient to work in the group ring of a homomorphic image H of G. We map G into 2x2 upper triangular matrices over $\mathbb{Q}$ by

$$x \mapsto \xi = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$y \mapsto \eta = \begin{pmatrix} \frac{l}{k} & 0 \\ 0 & 1 \end{pmatrix}.$$

One checks the relation $\eta\xi^k\bar{\eta} = \xi^l$ is satisfied. In the homomorphic image fractional powers of ξ make sense; for example, $\xi^{\frac{l}{k}}$ means the matrix $\begin{pmatrix} 1 & \frac{l}{k} \\ 0 & 1 \end{pmatrix}$. These fractional powers of ξ represent distinct elements of H, so are independent over $\mathbb{Z}$. The image of the expression (4.1) in $\mathbb{Z}H$ is hence

$$(1-\xi)\{(1+\xi^{l}+\xi^{2l}+\cdots)+(1+\xi^{\frac{l}{k}l}+\xi^{2\frac{l}{k}l}+\cdots)\eta \\ +(1+\xi^{(\frac{l}{k})^2 l}+\cdots)\eta^2+\cdots+\eta^{n-1}\}. \tag{4.2}$$

Since terms with different powers of η in the last expression cannot cancel, the only potential cancellation is in the coefficient series of a single power of η. But the coefficients of η^r are of the form $\xi^{(\frac{l}{k})^r il}$ and $-\xi^{(\frac{l}{k})^r il+1}$, where $i \geq 0$. We need

Lemma 4.3. *For all $r, i, j \geq 0$ one has $(\frac{l}{k})^r il \neq (\frac{l}{k})^r jl + 1$.*

Proof. Otherwise we would have $l^{r+1}(i-j) = k^r$, which contradicts the initial assumption $l > k \geq 1$.

It follows that there is no cancellation in (4.2) and consequently none in (4.1). Thus $|w_n|_1 = F_n$, which we saw increases faster than any polynomial in $\ell(w_n)$. This completes the proof of Theorem B.

Remark 4.4. In the special case $\langle\, x, y \mid yx\bar{y} = x^2 \,\rangle$ all the numbers $b_i = 0$ and the diagram $\mathcal{D}_n$ has the form illustrated below in Figure 5 for $n = 3$.

We shall now indicate a geometrical proof of Theorem B based on a different principle. The starting point is the disc diagrams $\tilde{\mathcal{D}_n}$ in $\tilde{X}$. To show that the l_1-norm of the chain $c(\tilde{\mathcal{D}_n})$ in $C_2(\tilde{X})$ is equal to the number of faces in the domain, it suffices to show that the mapping $\tilde{f}_n$ from the domain of $\tilde{\mathcal{D}_n}$ to $\tilde{X}$ is injective, for if there is cancellation in the chain, then the corresponding faces must have the same image in $\tilde{X}$.

Definition 4.5. Let Y be a piecewise Euclidean 2 complex of non positive curvature. Recall that this means each cell of Y has the metric of an convex polygon in $\mathbb{R}^2$ and the metrics agree on overlapping edges (there may not exist any metric on Y inducing the given metric on the cells). The condition of non positive curvature means that every non trivial circuit without backtracking in the link of each vertex of Y has total angular measure at least 2π. Let $f : D \to Y$ be a reduced disc diagram in Y (so D is a topological disc) and give D the pull back piecwise Euclidean structure; thus D is of non positive curvature. It follows from results of Aleksandrov [A] that (1) the path metric on D agrees with the given metric on each cell of D and (2) D possesses unique geodesic segments. The geodesics here are defined in a purely local manner, the condition being that all paths in the link connecting the incoming ray to the outgoing ray have angular measure at least π.

Lemma 4.6. *If $f : D \to Y$ is a reduced disc diagram and Y is a piecewise Euclidean 2-complex of non positive curvature and if the angle sum at each interior vertex of D is precisely 2π, then f maps geodesic segments γ of D, which meet ∂D only at end points of γ, to geodesic segments of Y.*

Proof. We need only check the angle condition at a vertex $f(P)$ of $f(D)$, where P is a vertex lying on the interior of γ. But if some path in $\mathrm{Link}_Y(f(P))$ connecting the incoming ray to the outgoing ray had angular measure less than π, then combining that path with the image under f of one in the link of P in D connecting incoming to outgoing rays of γ (note that such a path in $\mathrm{Link}_D(P)$ has angular measure precisely π) would yield a nontrivial circuit in the link of $f(P)$ in Y of total angular measure $< 2\pi$. This is contrary to the assumption of non positive curvature on Y.

Definition 4.7. Let $f : D \to Y$ satisfy the hypotheses of Lemma 4.7. A subpolyhedral disc $E \subset D$ is called *convex* if the angle sum at each boundary point of E is at most π. Note that we do not require E to be a subcomplex for given cell structure on D.

Proposition 4.8. *Let $f : D \to Y$ satisfy the hypotheses of Lemma 4.7, where Y in addition is simply connected, and let E be a convex subpolyhedron of D. Then f maps E injectively onto its image.*

Proof. Suppose $P, Q \in E$ are distinct points such that $f(P) = f(Q)$. Let γ be the geodesic segment connecting P to Q. Observe that by convexity γ lies entirely in E. By Lemma 4.6, the image of γ under f satisfies the geodesic condition except at the endpoints. Since Y is simply connected, this means that $f \circ \gamma$ bounds a geodesic monogon. But this cannot exist in non positive curvature, by the piecewise

Euclidean analog of the Gauss-Bonnet theorem [GS1]. It follows that f is injective on E.

Now we return to the 2-complex $X = K(\mathcal{P})$, where $\mathcal{P} = \langle x, y \mid yx^py^{-1} = x^q\rangle$ for positive p, q. Let Y be the $\mathbb{Z}$-cover of X. We observe that Y has a piecewise Euclidean structure of non positive curvature (although X itself does not unless $p = q$). In this structure all angles are either $\frac{\pi}{2}$ or π. A picture of this piecewise Euclidean structure can be gotten by examining Figure 5, for the special case $p = 1, q = 2$. Here the covering transformation corresponding to y is a homothety on the edges covering x, and multiplies lengths of such edges by 2. In the general case, y is a homothety on edges covering x with a scale factor of $\frac{q}{p}$. This piecewise Euclidean structure lifts to one on $\tilde{X}$ so $\tilde{X}$ becomes a piecewise Euclidean 2-complex of non positive curvature.

Observe now that each diagram $\tilde{\mathcal{D}_n}$ can be imbedded in some $\tilde{\mathcal{D}_N}$ for large enough N in such a way that the smaller diagram is contained in a convex subpolyhedron of the larger one. Observe further that all interior angle sums in all of these diagrams is exactly π (see Figure 5 for the special case $p = 1, q = 2$). It follows from Proposition 4.8 that each map $\tilde{f_n} : \tilde{\mathcal{D}_n} \to \tilde{X}$ is injective. We have already observed that this implies there is no cancellation in the chain $c(\tilde{\mathcal{D}_n})$. Since the lengths of the boundary labels of these diagrams increase linearly with n, whereas their l_1-norms, and hence their areas, increase exponentially, this completes the geometric proof of Theorem B.

§5. Proof of Theorem C

We are given $f : X \to Y$, a π_1-injective map of finite aspherical 2-complexes. We first treat the special case when $X^{(1)}$ is a subcomplex of $Y^{(1)}$ and $f|_{X^{(1)}}$ is the inclusion $X^{(1)} \subseteq Y^{(1)}$. Choose the base point in X and let $G = \pi_1(X)$, $H = \pi_1(Y)$. We note that f lifts to a map $\tilde{f} : \tilde{X} \to \tilde{Y}$ of universal covers. In addition each 2-cell of $\tilde{X}$ determines upon applying $\tilde{f}$ a van Kampen diagram in $\tilde{Y}$ whose associated chain in $C_2(\tilde{Y})$ is well defined, since Y is aspherical. We have in fact

Lemma 5.1. *The induced map $\tilde{f}_2 : C_2(\tilde{X}) \to C_2(\tilde{Y})$ is an injective homomorphism of $\mathbb{Z}G$ modules.*

Proof. Consider the commutative diagram in Figure 6 below.

Observe that $\tilde{f}_1$ is a split injection since $X^{(1)} \subseteq Y^{(1)}$ and the vertical arrows are injective since X and Y are aspherical. The result follows by a diagram chase.

Hence we can consider the short exact sequence of chain complexes

$$0 \to C_\star(\tilde{X}) \xrightarrow{\tilde{f}_\star} C_\star(\tilde{Y}) \to Q_\star \to 0 \tag{5.2}$$

where $Q_\star$ is the quotient complex. Note that $C_\star(\tilde{X})$ and $C_\star(\tilde{Y})$ possess homology only in degree 0 where $H_0 = \mathbf{Z}$ is mapped isomorphically by $\tilde{f}_0$. It follows that $H_i(Q_\star) = 0$ for all $i \geq 0$. Thus we have the short exact sequence of $\mathbb{Z}G$ modules

$$0 \to Q_2 \to Q_1 \to Q_0 \to 0. \tag{5.3}$$

But we assumed f to be an inclusion on the 1-skeleta. This implies that Q_0 and Q_1 are free modules over $\mathbb{Z}G$, since $\mathbb{Z}H$ is free over $\mathbb{Z}G$. Thus (5.3) splits and Q_2

is projective over $\mathbb{Z}G$. Hence $\tilde{f}_2 : C_2(\tilde{X}) \to C_2(\tilde{Y})$ is a split injection of free $\mathbb{Z}G$ modules. Let $\rho' : C_2(\tilde{Y}) \to C_2(\tilde{X})$ be a splitting, so $\rho' \cdot \tilde{f}_2 = 1$. Note that $C_2(\tilde{X})$ is finitely generated and free over $\mathbb{Z}G$, although $C_2(\tilde{Y})$ need not be finitely generated.

Lemma 5.4. *Let $f : A \to B$ be a split monomorphism of free modules over a ring R where A is finitely generated. Let B have the free basis $\{ b_i \mid i \in I \}$. Then there exists a splitting $\rho : B \to A$ for f such that $\rho(b_i) = 0$ for all but a finite number of indices $i \in I$.*

Proof. Let $a_1, a_2, \ldots, a_n$ be a finite free basis for A and let $\rho' : B \to A$ be a splitting, so that $\rho' \cdot f = 1_A$. Since $f(a_j)$ is a finite linear combination of the b_i's, the set I_0 of indices i for which b_i has a non zero coefficient in some $f(a_j)$ is finite. We define $\rho : B \to A$ by $\rho(b_i) = \rho'(b_i)$ if $i \in I_0$ and $\rho(b_i) = 0$ if $i \notin I_0$. Then one checks that $\rho \cdot f = 1_A$.

We apply the lemma with $A = C_2(\tilde{X})$ and $B = C_2(\tilde{Y})$. The basis for $C_2(\tilde{Y})$ consists of multiples $t \cdot \tilde{\beta}$ where $\tilde{\beta}$ is a chosen lift to $\tilde{Y}$ of the 2-cell β of Y and where t describes a transversal of coset representatives for the image of G in H. We obtain a retraction $\rho : C_2(\tilde{Y}) \to C_2(\tilde{X})$, $\rho \cdot \tilde{f}_2 = 1$, with $\rho(t \cdot \tilde{\beta}) = 0$ for all but a finite number of coset representatives t. But this means that the matrix M for ρ in terms of the free bases for $C_2(\tilde{X})$ and $C_2(\tilde{Y})$ as free $\mathbb{Z}G$ modules has only a finite number of non zero entries. Thus M is bounded in l_1-norm and if we set A equal to the supremum of the l_1-norms of the matrix entries (in $\mathbb{Z}G$) of M, then we have

$$|\rho(u)|_1 \leq A \; |u|_1 \tag{5.5}$$

for all $u \in C_2(\tilde{Y})$. If we let $c \in C_2(\tilde{X})$ and calculate l_1-norms, we get $|c|_1 = |\rho \cdot \tilde{f}_2(c)|_1 \leq A \cdot |\tilde{f}_2(c)|_1$, where we have used (5.5) in the last inequality. In particular this holds if $c = c(\tilde{g})$, where $c(\tilde{g})$ is the 2-chain determined by the lift $\tilde{g}$ of a van Kampen diagram g for a circuit w in $X^{(1)}$; here w is assumed null homotopic in X. But this means that

$$|w|_{1,X} \leq A \; |w|_{1,Y} \tag{5.6}$$

for all such circuits w, where the subscripts X and Y refer to the l_1-norms for these respective spaces. Since (5.6) holds for all such circuits w, we deduce that

$$\lambda_X(n) \leq A \; \lambda_Y(n) \tag{5.7}$$

for all $n > 0$. This completes the proof of Theorem C in the special case.

In general, given $f : X \to Y$ a π_1-injective map of finite aspherical 2-complexes where Y has at least one 2-cell, we can modify Y by elementary 2-expansions to get the finite 2-complex Y' containing Y and a copy of $X^{(1)}$ as subcomplexes so that the map f is homotopic to a map $f' : X \to Y'$ where f' restricted to $X^{(1)}$ is the inclusion of $X^{(1)}$ in $Y'^{(1)}$. The result follows from Theorem 3.5 and the special case just considered. This completes the proof of Theorem C.

Remark 5.8. The hypotheses of asphericity of the 2-complexes in Theorem C cannot be dropped as the following example shows. Let $G = G(\mathcal{P})$ where $\mathcal{P} = \langle\, x, y, z, \mid \; [x, y] = z, [z, x] = 1, [z, y] = 1\,\rangle$; G is the 3-dimensional integral Heisenberg group. Let H be the 5-dimensional Heisenberg group, $H = G\langle\, x_1, x_2, y_1, y_2, z \mid [x_1, x_2] = z = [y_1, y_2], [x_i, y_j] = 1 \,\text{for all}\, i, j;\, z\,\text{central}\,\rangle$, and imbed G in H by $x \mapsto x_1$, $y \mapsto y_1$, and $z \mapsto z$. According to Thurston [**Th**], H satisfies the quadratic isoperimetric inequality. However G does not, as the argument in the next paragraph shows.

Remark 5.9. Another presentation for the 3-dimensional Heisenberg group G is $\mathcal{Q} = \langle x, y, t \mid [x, y] = 1,\; txt^{-1} = xy,\; tyt^{-1} = y\rangle$, as one can easily check (in this presentation, y represents the central commutator). This presentation is Cockcroft (that is, its associated 2-complex $K(\mathcal{Q})$ is Cockcroft). To see this, note that the 2-complex $K(\mathcal{Q})$ is the 2-skeleton of a cell structure of a closed, orientable, aspherical 3-manifold M with precisely one 3-cell. The manifold M is in fact a torus bundle over the circle with monodromy the automorphism ϕ of $\mathbb{Z}^2$ given by $\phi(x) = xy$, $\phi(y) = y$. It follows from Proposition 3.7 that $\mathcal{Q}$ is Cockcroft.

Now let $w_n = [x^n, t^n x^n t^{-n}]$, where $[u, v] = uvu^{-1}v^{-1}$. In G we have that $t^n x^n t^{-n} = \phi^n(x^n) = (xy^n)^n$ commutes with x^n, so w_n is null homotopic in $X := K(\mathcal{Q})$. We can visualize a van Kampen diagram f_n for w_n as follows. The top third is a trapezoid that gives a null homotopy for $u_n := t^n x^n t^{-n} \phi^n(x^n)^{-1}$. The bottom third is the reflection of the top third through a horizontal line. The middle third represents a van Kampen diagram for the relation $[x^n, \phi^n(x^n)] = 1$ (the middle third is glued to the top and bottom thirds along arcs whose label is $\phi^n(x^n)$). Now we can calculate the chain $c(f_n)$ for f_n in $C_2(X)$ and hence its l_1-norm as follows. The top and bottom thirds of f_n are oppositely oriented, so their chains exactly cancel. It remains to calculate the chain for the middle third. There is a short cut for doing this. Commutation of x with itself doesn't count (it represents a one dimensional subset of the diagram). All we need to count are the number of commutations between x^n and the y's occurring in $\phi^n(x^n)$. The easy way to do this is to use the Grassman product, $\Lambda^2(\mathbb{Z}^2)$ equipped with its basis $x \wedge y$ (and the l_1-norm with respect to this basis for this rank 1 lattice). Calculating $[x^n] \wedge [(xy^n)^n]$ and using additive notation instead of multipicative notation, we get $nx \wedge (nx + n^2 y) = n^3 x \wedge y$. It follows from Remark 3.6 that the l_1-norm of w_n is estimated below by $|c(f_n)|_1 = n^3$.

Since $\ell(w_n) = 8n$ increases linearly with n, whereas its l_1-norm, calculated in X, increases like a cubic polynomial in n, it follows from Lemma 3.4 that the Dehn function for G is at least cubic, and hence that G is not automatic. In fact one can show easily that there is a constant $B > 0$ such that $\delta_X(n) \le Bn^3$ for all $n > 0$. Hence G satisfies a cubic isoperimetric inequality and this cubic inequality is optimal.

Summarizing, we have the result that *the 3-dimensional Heisenberg group does not satisfy the quadratic isoperimetric inequality.*

Remark 5.10. We shall show more generally *the split extension* $G = \mathbb{Z}^2 \rtimes_\phi \mathbb{Z}$*, where*

$\phi \in \mathrm{Aut}(\mathbb{Z}^2) \cong \mathrm{Gl}_2(\mathbb{Z})$, *is automatic iff* ϕ *is of finite order*. The 'if' direction is easy, for if ϕ is of finite order, then G contains a free abelian group of rank 3 as a subgroup of finite index. It follows then from results of [CEHPT] that G is automatic.

In the converse direction, we shall show the stronger result that *if* ϕ *is not of finite order, then* G *does not satisfy the quadratic isoperimetric inequality.* Up to replacing G by a subgroup of index 2 (this does not change the growth of the Dehn function, by Proposition 2.11), we may assume that the determinant of ϕ is 1. In this case, if $\{x, y\}$ is a basis for the normal $\mathbb{Z}^2$ subgroup, the presentation $\mathcal{P} = \langle x, y, t \mid [x, y] = 1, txt^{-1} = \phi(x),\ tyt^{-1} = \phi(y)\rangle$ is Cockcroft, since its associated 2-complex is the 2-skeleton of a cell structure on a closed, orientable, aspherical 3-manifold with exactly one 3-cell.

In this case there are two subcases, depending on whether ϕ has 1 as an eigenvalue or not. If 1 is an eigenvalue, then the basis $\{x, y\}$ may be chosen so that the matrix for ϕ with respect to this basis has the form $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$, where $a \neq 0$. Then the argument given in Remark 5.9 applies unchanged to show that the G satisfies the cubic isoperimetric inequality and this is optimal (the sequence of words $w_n = [x^n, t^n x^n t^{-n}]$ has length linear in n but area cubic in n; the key point as before is the fact that the presentation $\mathcal{P}$ is Cockcroft).

Suppose now that $\det(\phi) = 1$, ϕ is of infinite order, and no eigenvalue of ϕ is 1. In this case it follows from the classification of elements of $\mathrm{Sl}_2(\mathbb{Z})$ that ϕ is a hyperbolic automorphism. It has two real eigenvalues $\lambda, \frac{1}{\lambda}$ and we may assume that $|\lambda| > 1, |\frac{1}{\lambda}| < 1$. In addition we may assume, after replacing ϕ by ϕ^2 if necessary, that $\lambda > 1$. Furthermore, an elementary number theoretic argument based on the quadratic formula for λ shows that λ is irrational. Let ξ, η be (real) eigenvectors for ϕ for the eigenvalues $\lambda, \frac{1}{\lambda}$, respectively; here ϕ is considered to be in $\mathrm{Sl}_2(\mathbb{R})$. We have $x = \alpha\xi + \gamma\eta$, $y = \beta\xi + \delta\eta$ for suitable real numbers α, β, γ and δ. This takes place in $\mathbb{R}^2 = \mathbb{R} \otimes \mathbb{Z}^2$. A calculation shows that

$$\phi^n(x) = (\alpha\delta\lambda^n - \frac{\beta\gamma}{\lambda^n})x - \alpha\gamma(\lambda^n - \frac{1}{\lambda^n})y.$$

Observe that, in this last expression, $\alpha\gamma \neq 0$, for otherwise x would be an eigenvector of ϕ, contradicting the irrationality of λ. It follows that the coefficient of y in $\phi^n(x)$, which is of course an integer, increases like λ^n; that is, it increases exponentially with n. From this point on, the argument is the same as in Remark 5.9 using the same family of words w_n. In this case one gets the stronger result that the area of w_n increases exponentially with n although the length of w_n increases linearly. Since we have analyzed all cases, it follows that G does not satisfy the quadratic isoperimetric inequalty if ϕ is of infinite order.

Remark 5.11. By virtue of Theorem C, there is interest in producing examples of groups G of geometric dimension 2 whose Dehn function is equivalent to its l_1-norm function; here G is said to be of geometric dimension 2 if $G = \pi_1(X)$, where X is a finite aspherical 2-complex. For example, suppose $G < H$, where G and H

are both groups of geometric dimension 2 and suppose that H is hyperbolic. If the Dehn function for G is equivalent to its l_1- norm function, then G is also hyperbolic. To see this, observe that by Theorem C, λ_G is dominated by λ_H, which is in turn dominated by δ_H. By assumption, λ_G is equivalent to δ_G. Since H is hyperbolic, it follows that δ_H is linearly bounded. Consequently δ_G is also linearly bounded, and hence G is hyperbolic. In the next paragraph, we shall give one example where λ_G is equivalent to δ_G.

Example 5.12. We shall show that if $X = K\langle x, y \mid [x, y] = 1\rangle$, then $\lambda_X = \delta_X$. Observe first that $\tilde{X} = \mathbb{R}^2$, where $\tilde{X}$ is the universal cover of X. If w is an edge circuit in $X^{(1)}$ which is null homotopic in X and such that no subpath of w is null homotopic, then w lifts to a simple closed circuit $\tilde{w}$ in $\mathbb{R}^2$. By the Jordan-Schoenfliess theorem, $\tilde{w}$ bounds a disc, which gives an *injective* disc diagram $f : D^2 \to \tilde{X}$ bounded by $\tilde{w}$. Hence $|c(\tilde{f})|_1 = \mathrm{Area}_X(w)$. From this observation, it is not difficult to complete the proof that $\lambda_X = \delta_X$.

Question. It seems reasonable to ask whether $\lambda_X = \delta_X$ for every finite 2-complex X admitting a piecewise Euclidean structure of non positive curvature (see 4.5).

§6. Examples

Early on in the discussions leading to [**BGSS**] the authors asked whether an amalgam of finitely generated free groups over a subgroup of finite index in each was automatic. We observed

Proposition 6.1. *Let F be a finitely generated free group and let N be a normal subgroup of finite index. Then the double $D = F \underset{N}{\star} F$ of F along N is automatic.*

Proof. If $G = F/N$ one has a homomorphism $\phi : G \to \operatorname{Out}(N)$ with $\phi(aN)$ induced by conjugation by a on N for $a \in F$. If we consider the short exact sequence of groups

$$1 \to N \to D \to G \star G \to 1, \tag{6.2}$$

then we get an induced homomorphism $\phi_1 : G \star G \to \operatorname{Out}(N)$ with $\operatorname{Im}\phi_1 = \operatorname{Im}\phi$. In particular $\operatorname{Im}\phi_1$ is a finite subgroup of $\operatorname{Out}(N)$, so the kernel K_1 of ϕ_1 is a normal subgroup of finite index in $G \star G$. But the kernel K_2 of the canonical homomorphism $G \star G \to G \times G$ is of finite index and free. Let $K_3 = K_1 \cap K_2$. It follows that K_3 is finitely generated free and of finite index in $G \star G$ and $\phi_1 | K_3$ is trivial. This implies that the extension (6.2) pulled back to K_3 is the direct product of K_3 with N. That is, D possesses a subgroup of finite index which is isomorphic to $N \times K_3$, the direct product of two finitely generated free groups. Thus $N \times K_3$ is automatic [**CEHPT**]. Since D has a subgroup of finite index which is automatic, D is also automatic.

However in general an amalgam of two finitely generated free groups over a subgroup of finite index in each is not automatic. The following construction is due to Stallings. Let F_1 be freely generated by elements a and s and let F_2 be freely generated by b and t. Let $A_1 = \mathrm{gp}\langle\, a, sa\bar{s}, s^2as^{-2}, s^3 \,\rangle < F_1$ and let $A_2 = \mathrm{gp}\langle\, b, tb^2t^{-1}, t^2bt^{-1}, t^3 \,\rangle < F_2$. Observe that $(F_i : A_i) = 3$ for $i = 1, 2$. Thus we can form $H = F_1 \underset{A_1 \cong A_2}{\star} F_2$ where we identify $a = b$, $sas^{-1} = tb^2t^{-1}$, $s^2as^{-2} = t^2bt^{-1}$, and $s^3 = t^3$. Then $t^{-1}sas^{-1}t = b^2 = a^2$, so $a^{(t^{-1}s)} = a^2$. One sees easily that the subgroup $\mathrm{gp}\langle\, a, t^{-1}s \,\rangle$ of H is in fact isomorphic to $G\langle\, x, y \mid yx\bar{y} = x^2 \,\rangle$. But $H = \pi_1(Y)$ with Y a finite aspherical 2-complex. It follows from Theorem A that H is not automatic.

Example 6.3. Let $\mathcal{P}_r = \langle\, x_0, x_1, \ldots, x_r \mid x_i^{x_{i+1}} = x_i^2;\ 0 \le i \le r-1 \,\rangle$. The Dehn function for $\mathcal{P}_r$ grows no slower than the function E_r; here $E_0(n) = n$ and $E_{r+1}(n) = 2^{E_r(n)}$ for all $n > 0, r > 0$. Let us sketch why this is so (the case $r = 1$ is contained in the proof of Theorem A). We consider first $\mathcal{P}_1$ and the words $w_{1,n} = [x_0^{x_1^n}, x_0]$. The methods of §4 show that the minimal van Kampen diagram for $w_{1,n}$ has $2(2^n - 1) \ge 2^n$ faces (see Figure 5) and that this number of faces is the same as the l_1-norm. Now specialize to $w_{1,2^n} = [x_0^{x_1^{2^n}}, x_0]$. We replace $x_1^{2^n}$ by $x_1^{(x_2^n)}$ to get $w_{2,n} = [x_0^{(x_1^{(x_2^n)})}, x_0]$. One has $l(w_{2,n}) = 2^3(n+1)$ and the minimal van Kampen diagram for it has at least $E_2(n) = 2^{2^n}$ faces and the number of faces is the l_1-norm. At the next stage we specialize to $w_{2,2^n}$ and replace $x_2^{2^n}$ by $x_2^{(x_3^n)}$. In general we have words $w_{k,n}$ with $l(w_{k,n}) = 2^{k+1}(n+1)$ but the minimal van Kampen diagram for $w_{k,n}$ has at least $E_k(n)$ faces and the number of faces is the l_1-norm. If we attempt to draw these minimal van Kampen diagrams the appearance is similar to the approximations to the Koch snowflake curve.

Next consider the 1-relator presentation $\mathcal{P} = \langle\, x, y \mid x^{x^y} = x^2 \,\rangle$. The group $G(\mathcal{P})$ maps onto $\mathbb{Z}$ by $x \mapsto 0$, $y \mapsto 1$. The kernel K can be calculated by the Reidemeister-Schreier method. It has the presentation

$\mathcal{Q} = \langle\, x_i;\, i \in \mathbb{Z} \mid x_i^{x_{i+1}} = x_i^2$ for all $i \in \mathbb{Z}\,\rangle$. In particular $\mathcal{P}_r$ is a subpresentation of $\mathcal{Q}$ for all $r \geq 1$ and the group imbeds as well.

By Theorem C, since $K(\mathcal{P})$ is aspherical, the l_1-norm function for $\mathcal{P}$ grows at least as fast as that for any of the presentations $\mathcal{P}_r$. Thus the Dehn function for $\mathcal{P}$ grows faster than the function E_r for all r. With more care one can actually show that $\delta_{\mathcal{P}}$ grows at least as fast as Δ, where $\Delta(n) = E_n(n)$. This is indeed a fast rate of growth and is our record so far for 1-relator groups.

Proposition 6.4. *Let $\mathcal{R} = \langle\, x, y, z, \mid x^{x^y} = x^2\,\rangle$. Then $G = G(\mathcal{R})$ is not an amalgam of finitely generated free groups amalgamating a finitely generated subgroup.*

Proof. The group $G(\mathcal{R})$ admits the group $G(\mathcal{P})$ above as a retract, so the Dehn function for $\mathcal{R}$ grows at least as fast as that of $\mathcal{P}$. In particular, since $\delta_{\mathcal{P}}$ grows faster that E_2, where $E_2(n) = 2^{2^n}$, it follows that $\delta_{\mathcal{P}}$ grows faster than any simple exponential. But an amalgam of finitely generated free groups amalgamating a finitely generated subgroup is asynchronously automatic by [**BGSS**], so its Dehn function is bounded by a simple exponential. Hence $\mathcal{R}$ is not an amalgam of such groups.

Remark 6.5. This result answers negatively a question raised by G. Baumslag. Baumslag and Shalen [**BS**] proved that a finitely presented group admitting a finite presentation of deficiency at least two is a proper amalgam of finitely generated subgroups over a finitely generated amalgamated subgroup. Baumslag asked whether in the case of a 1-relator group of deficiency at least two these subgroups could all be taken to be finitely generated and free. The example $\mathcal{R}$ above shows that this is not the case.

Remark 6.6. It can be shown that the group $G(\mathcal{P}_2)$ above admits a finite complete rewriting system [**Br**]. Thus a finitely presented group with a presentation that admits a finite complete rewriting system need not be asynchronously automatic. The group $G(\mathcal{P}_1)$ provides a simpler example of one which admits a finite complete rewriting system but is not synchronously automatic.

We finish with some open problems. We have not yet been able to decide whether the Baumslag-Solitar group $B_{1,2}$ is isomorphic to a subgroup of an automatic group. Indeed, in an earlier version of this article, we stated that we knew no restriction on the finitely presented subgroups H of general automatic groups other than that H must have a solvable word problem. Since then, Paul Schupp informed me of the fact that the theory of time-complexity [**HU**] puts restriction on such subgroups, based on the result that the word problem for automatic groups is solvable in $O(n^2)$-time [**CEHPT**].

In general, the diagrammatic algorithm we have considered in this paper is very inefficient. Its advantage is its universality, since a finitely presented group has a solvable word problem if and only if its Dehn function is recursive (Proposition 2.2 above). For example, $\mathbb{Z}^n$ satisfies the quadratic isoperimetric inequality but not the linear isoperimetric inequality if $n \geq 2$. However there is a linear time algorithm for the word problem of the presentation $\langle\, x_i;\ 1 \le i \le n \mid [x_i, x_j];\ 1 \le i, j \le n \,\rangle$. Namely, a word represents 1 if and only if its exponent sum in each variable is zero.

More generally we do not know whether $B_{1,2}$ can be isomorphic to a subgroup of a finitely presented group satisfying the quadratic isoperimetric inequality. By Thurston's results [**Th**] this last class of groups is very large, including $\mathrm{Sl}_n(\mathbb{Z})$ for all $n \ge 4$ and many nilpotent groups, like the $(2n+1)$-dimensional integral Heisenberg group for all $n \ge 2$. A characterization of finitely presented groups satisfying the quadratic isoperimetric inequality is lacking unlike the situation for the linear isoperimetric inequality: a finite presentation $\mathcal{P}$ satisfies $\delta_{\mathcal{P}}(n) \le An$ for all $n > 0$ iff $G(\mathcal{P})$ is hyperbolic [**Gr**].

Finally note the following immediate consequence of Theorem C: *Let* $X \overset{j}{\subseteq} Y \subseteq Z$ *be inclusions of finite connected subcomplexes* X *and* Y *of the aspherical 2-complex* Z*. Assume (i)* j *is* π_1*-injective, (ii)* λ_X *grows faster than any polynomial, and (iii)* δ_Y *is polynomially bounded. Then* Y *is not aspherical.* If the hypotheses (i)–(iii) could be realized in an example, this would give a negative answer to a question of J.H.C. Whitehead's, whether a subcomplex of an aspherical 2-complex is aspherical.

In fact, we remark that there are natural candidates among which to look for examples; namely, one should look among the labelled oriented trees, or LOT's, of Howie's [**Ho**]. A LOT is a pair (T, λ) where T is a tree (in Serre's sense [**Se**], so that geometric edges correspond to pairs of oriented edges $e, \bar{e}$ of T) together with a function $\lambda : E \to F(V)$ so that $\lambda(\bar{e}) = \lambda(e)^{-1}$. Here E is the set of edges and $F(V)$ is the free group freely generated by the set V of vertices of T. The presentation $\mathcal{P}(T, \lambda)$ associated to the LOT (T, λ) is given by $\langle V \mid (\partial_1 e)^{\lambda(e)} = \partial_0 e;\ e \in \mathcal{O} \rangle$, where $\mathcal{O}$ is an *orientation* on E, so $\#(\mathcal{O} \cap \{e, \bar{e}\}) = 1$ for all $e \in E$.

One verifies that the LOT (T', λ') illustrated below in Figure 8 has associated presentation $\mathcal{P}' = \mathcal{P}(T', \lambda')$ with group $G(\mathcal{P}') \cong G\langle x, y \mid yx\bar{y} = x^2 \rangle$.

Thus, if this LOT (T', λ') is a sub-LOT of the LOT (T, λ), then $X := K(\mathcal{P}(T', \lambda'))$ satisfies hypothesis (ii) above. In addition, it is easy to see, by an Euler characteristic argument, that if $\mathcal{Q}$ is obtained from $\mathcal{P}(T, \lambda)$ by adjoining any one of the vertices of T as a new relator, then $Z := K(\mathcal{Q})$ is contractible, since $\pi_1(Z)$ is trivial. One takes $Y := K(\mathcal{P}(T, \lambda))$. We do not know whether (i) and (iii) can both be satisfied in an example.

References

[[**A**]] A. D. Aleksandrov, *Odna teorema o treugolnikah v metricheskom prostranstve i nekotorye ee prilozheniya*, Trudy Mat. Inst. Im. V. A. Steklova **38** (1951), 5–23.

[[**BGSS**]] G. Baumslag, S. M. Gersten, M. Shapiro, and H. Short, *Automatic groups and amalgams*, these proceedings.

[[**Br**]] K. S. Brown, *The geometry of rewriting systems: a proof of the Anick-Groves-Squier theorem*, preprint.

[[**BS**]] G. Baumslag and P. B. Shalen, *Amalgamated products and finitely presented groups*, preprint.

[[**CDP**]] M. Coornaert, T. Delzant, and A. Papadopoulos, *Notes sur les groupes hyperboliques de Gromov*, Publication de I.R.M.A., Strasbourg.

[[**CEHPT**]] J. W. Cannon, D. B. A. Epstein, D. F. Holt, M. S. Paterson, and W. P. Thurston, *Word processing and group theory*, preprint.

[[**G**]] S. M. Gersten, *Reducible diagrams and equations over groups*, Essays in Group Theory, M.S.R.I. series Vol.8, S. M. Gersten, editor, Springer-Verlag, 1987.

[[**GH**]] W. Ballmann, E. Ghys, A. Haefliger, P. de la Harpe, E. Salem, R. Strebel, et M. Troyanov, *Sur les groupes hyperboliques d'après Mikhael Gromov.*

[[**Gr**]] M. Gromov, *Hyperbolic groups*, Essays in Group Theory, M.S.R.I. series Vol. 8, S. M. Gersten, editor, Springer-Verlag, 1987.

[[**GS1**]] S. M. Gersten and H. Short, *Small cancellation theory and automatic groups*, MSRI preprint 1989.

[[**GS2**]] S. M. Gersten and H. Short, *Small cancellation theory and automatic groups, part II*, MSRI preprint 1989.

[[**Ho**]] J. Howie, *On the asphericity of ribbon disc complements*, Trans. Amer. Math. Soc. **289** (1985), 281-302.

[[**HU**]] J. Hopcroft and J. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, 1979.

[[**LS**]] R. C. Lyndon and P. E. Schupp, *Combinatorial Group Theory*, Springer-Verlag, 1977.

[[**Re**]] Yu. G. Reshetnyak, *Ob odnom special'nom otobrazhenii konusa na mnogogrannik*, in Russian, Mat. Sbornik **53(95)** (1961).

[[**Se**]] J.-P. Serre, *Trees*, Springer-Verlag, 1980.

[[**Th**]] W. P. Thurston, oral communication.