Formal language theory and the geometry of 3-manifolds.

Comment. Math. Helvetici 71 (1996) 525–555 0010-2571/96/040525-31$1.50 + 0.20/0

Formal language theory and the geometry of 3-manifolds

MARTIN R. BRIDSON AND ROBERT H. GILMAN

Abstract. Automatic groups were introduced in connection with geometric problems, in particular with the study of fundamental groups of 3-manifolds. In this article the class of automatic groups is extended to include the fundamental group of every compact 3-manifold which satisfies Thurston's geometrization conjecture. Toward this end, the class $\mathscr{C}_{\mathscr{A}}$ of asynchronously $\mathscr{A}$-combable groups is introduced and studied, where $\mathscr{A}$ is an arbitrary full abstract family of languages. For example $\mathscr{A}$ may be the family of regular languages *Reg*, context-free languages *CF*, or indexed languages *Ind*. The class $\mathscr{C}_{Reg}$ consists of precisely those groups which are asynchronously automatic. It is proved that $\mathscr{C}_{Ind}$ contains all of the above fundamental groups, but that $\mathscr{C}_{CF}$ does not. Indeed a virtually nilpotent group belongs to $\mathscr{C}_{CF}$ if and only if it is virtually abelian.

Introduction

In the last several years there has been a remarkable interplay between topology, geometry, group theory, and the theory of formal languages. Thurston's work on the classification problem has shown that the topology of 3-dimensional manifolds is inextricably linked with the study of rigid geometries; in a series of essays Gromov introduced powerful geometric techniques into combinatorial group theory; and the work of Epstein *et al.* [E+], and earlier that of Muller and Schupp [MS] showed that the geometry of the Cayley graph of a group is intimately connected with the properties of associated formal languages. In this article we focus on a point of intersection of these three bodies of knowledge: normal forms for 3-manifold groups. We shall give upper and lower bounds on the logical complexity of the languages associated to geometrically efficient normal forms for elements in the fundamental group of any compact 3-manifold that satisfies Thurston's geometrization conjecture.

Before stating our results we describe something of the context in which they are set. A choice of generators for a finitely generated group G may be thought of as a map $\mu: \Sigma \to G$ which extends to a surjective monoid homomorphism $\mu: \Sigma^* \twoheadrightarrow G$,

The first author was partially supported by NSF grant DMS-9203500 and FNRS (Suisse). He also wishes to thank the University of Geneva for its hospitality while this paper was being written. The second author thanks the Institute for Advanced Study for its hospitality while this paper was being written.

where Σ is a finite set and Σ^* is the free monoid on Σ. Let Γ denote the corresponding Cayley graph. Γ can be viewed as a metric space with each edge isometric to the unit interval, and words in Σ^* can be thought of as edge-paths in Γ beginning at the identity. A *combing* of G is a selection for each $g \in G$ of one such path ending at g. Thus a combing is just a geometric way of looking at a set of normal forms for G. A reasonable measure of our understanding of the geometry of G is the extent to which we are able to construct efficient combings of it. Ideally we would like to construct a combing which has a simple description as a formal language (i.e., subset of Σ^*) and is such that the routes to nearby points are uniformly close (the *fellow-traveller property*).

Much of the motivation for approaching combinatorial group theory via combings comes from the desire to understand the fundamental groups of compact 3-manifolds. The fundamental groups of 2-manifolds are completely understood, and requiring a finitely presented group to be the fundamental group of a compact 4-manifold is a vacuous condition. A natural challenge arises in dimension three. The work of Thurston shows that it would be rather misleading to describe the class of 3-manifold groups as being given by a purely topological condition. Indeed, the existence of natural geometric structures for 3-manifolds makes the task of understanding the geometry and complexity of the optimal normal forms for all 3-manifold groups seem enticingly approachable. On the other hand, because of the dominant role which the fundamental group plays in 3-manifold topology (see [H]), seeking to understand all 3-manifold groups is much like trying to classify all compact 3-manifolds. Such a classification is clearly beyond the reach of present knowledge, but the situation is by no means hopeless and conjecturally the structure of the solution is beautifully simple.

Thurston's geometrization conjecture asserts that the interior of each component in the prime/toral decomposition of a compact 3-manifold ([K], [M], [JS], [J]) can be modelled on one of the eight 3-dimensional geometries ([S], [T2]). This conjecture has been proved in many cases, for example in the case where the fundamental group of the manifold contains $\mathbb{Z}^2$ (see [CJ], [G]), and in the case where the given manifold contains an incompressible surface (see [Mo], [T1]). We shall continue the present discussion under the hypothesis that Thurston's conjecture is true and use the adjective 'geometrizable' to emphasize this assumption. However, we should emphasize that the validity of our results does not depend upon this conjecture, although it certainly serves to motivate them.

Recall that a group is automatic [E+] if it admits a combing which satisfies the fellow-traveller property and forms a regular language – the simplest type of language in the Chomsky hierarchy (see [HU]). The drive to understand 3-manifold groups motivated much of the early development of automatic group theory, beginning with the seminal work of Cannon on the algorithmic structure of

Kleinian groups [C] and culminating in the Epstein–Thurston classification of those compact geometrizable 3-manifolds whose fundamental groups are automatic [E+]. The upshot of this classification, much of which was worked out independently by Shapiro [Sha], is that the fundamental group of a compact geometrizable 3-manifold M is automatic if and only if none of the factors in the prime decomposition of M is a closed manifold modelled on one of the geometries *Nil* or *Sol*. Furthermore, Epstein and Holt [E+] showed that the fundamental group of a closed *Nil* manifold is not even *asynchronously* automatic. (This term refers to a generalization of an automatic group in which one allows monotone reparametrization of paths when defining the fellow-traveller property – see Section 2.) Brady has shown that the fundamental group of a closed *Sol* manifold is not asynchronously automatic [Br].

These results lead us naturally to the task of expanding the theory of automatic groups to include the fundamental groups of closed *Nil* and *Sol* manifolds. The first step in this direction was taken in [B2], where it was shown that the fundamental group of every compact geometrizable 3-manifold M admits a combing that satisfies the asynchronous fellow-traveller property. This result shows that there exist normal forms which allow one to navigate effectively in $\pi_1 M$ (or, equivalently, in $\tilde{M}$) but leaves unanswered the deeper question of how difficult it is to generate a list of the preferred routes.

In the present article we answer this last question by obtaining upper and lower bounds on the logical complexity of the geometrically efficient combings admitted by the fundamental groups of all compact geometrizable 3-manifolds. In order to obtain these bounds we consider two generalizations of automatic group theory. On the one hand, we prove that the widely-suggested strategy of enlarging the class of languages considered to include all context-free languages does not give rise to a theory rich enough to embrace the fundamental groups of closed *Nil* manifolds. On the other hand, we prove that allowing a slightly larger class of languages, the so-called *indexed languages* (see Section 1) does give rise to a class of groups that includes the fundamental group of every compact geometrizable 3-manifold (or, more generally, 3-orbifold.)

THEOREM A. *Let G be a virtually nilpotent group and let $\mu\colon \Sigma^* \twoheadrightarrow \Gamma$ be a choice of finite generating set for G. Suppose that $L \subseteq \Sigma^*$ is a combing of G that satisfies the asynchronous fellow-traveller property. If L is a context-free language then G is virtually abelian.*

COROLLARY. *Let M be a closed 3-manifold modelled on the geometry Nil and let L be a combing of $\pi_1 M$. If L satisfies the asynchronous fellow-traveller property then L is not context-free.*

THEOREM B. *Let M be a compact 3-manifold which satisfies the geometrization conjecture, and let $\mu: \Sigma^* \twoheadrightarrow \pi_1 M$ be a choice of generators. Then, there exists a combing $L \subseteq \Sigma^*$ which satisfies the asynchronous fellow-traveller property and is an indexed language.*

This paper is organized as follows. In Section 1 we review elements of the theory of formal languages and prove some structural results for context-free languages of polynomial growth. In Section 2 we define the notion of an asynchronous $\mathscr{A}$-combing for an arbitrary full abstract family of languages $\mathscr{A}$. This definition generalizes the notion of an asynchronous automatic structure. After establishing some basic properties of such combings we concentrate on the case where $\mathscr{A}$ is the class of context-free languages. In this setting, a study of the structure of quasiconvex subgroups (Definition 2.11) enables us to prove that the combing which the fundamental group of a compact nonpositively curved 3-manifold M inherits from the geometry of the universal cover can be uniformly approximated by a context-free combing if and only if $\pi_1 M$ is hyperbolic in the sense of Gromov [G].

It is in Section 3 that we prove Theorem B. We first show that the combings of torus bundle groups constructed in [B2] are not context-free; our proof highlights precisely the subtlety which renders context-free languages inadequate in this setting. Having isolated this difficulty we then show that it can be resolved by the use of indexed languages.

In Section 4 we prove Theorem A. The proof begins with the observation that since virtually nilpotent groups have polynomial growth, so do all of their combings; hence we are able to apply the structural results of Section 1. The geometry which these results force on combings can be detected in abelian quotients of nilpotent groups, and we prove Theorem A by exploiting the resulting tension between the geometry of these quotients and the behaviour of commutator relations in nilpotent groups.

Section 1: Formal languages

In this section we develop some results about formal languages, i.e., subsets of finitely generated free monoids, which we will require later. See Hopcroft and Ullman [HU] for an introduction to formal language theory.

For any set Σ denote by Σ^* the free monoid over Σ. For our purposes Σ will always be finite. The elements of Σ^* are words in the letters Σ; the length of a word w is denoted $|w|$. The identity element of Σ^* is the empty word ε, and multiplication of words is by concatenation. Given $L \subseteq \Sigma^*$, we let L^* denote the submonoid of Σ^* generated by L. In this case of singleton sets we usually omit the brackets: a^*

instead of $\{a\}^*$. If $L, M \subseteq \Sigma^*$, then $LM = \{w \mid w = xy, x \in L, y \in M\}$ is the product of L and M. If $L \subseteq \Sigma^*$, then Σ is called an alphabet for L, and L is called a language over Σ.

We will use regular, context-free, indexed, and bounded languages. A bounded language is by definition a subset of $w_1^* w_2^* \cdots w_n^*$ for some choice of words w_i. It is clear that bounded languages have polynomial growth; i.e., the number of words of length at most k in a bounded language is bounded above by a polynomial in k.

LEMMA 1.1. [GS] *The class of bounded languages contains all finite languages and is closed under union and product.*

Proof. Clearly every singleton $\{w\}$ is bounded, so the second assertion implies the first. If $L_1 \subseteq w_1^* w_2^* \cdots w_n^*$ and $L_2 \subseteq v_1^* v_2^* \cdots v_m^*$, then $L_1 L_2$ and $L_1 \cup L_2$ are both sublanguages of $w_1^* w_2^* \cdots w_n^* v_1^* v_2^* \cdots v_m^*$. □

The three other language classes form an increasing hierarchy. We consider these classes simultaneously as instances of a full abstract family of languages. A full AFL is a class of languages which contains a nonempty language and is closed under the operations listed in Lemma 1.2(i) below (see [HU, Chapter 11] for more details).

LEMMA 1.2. *Let $\mathcal{A}$ be a full AFL.*

(i) *$\mathcal{A}$ is closed under homomorphism (i.e., image under homomorphism of the ambient free monoid to another finitely generated free monoid), inverse homomorphism (likewise), intersection with regular sets, union, product, and generation of submonoid.*

(ii) *$\mathcal{A}$ contains all regular languages.*

(iii) *If L_1 and L_2 are both in $\mathcal{A}$, then the language L consisting of ε and all words $w = u_1 \cdots u_n$ with successive u_i chosen alternately from L_1 and L_2 is also in $\mathcal{A}$.*

(iv) *The classes of regular, context-free, and indexed languages are each a full AFL.*

Proof. Part (i) is true by definition; for (ii) see [HU] and for (iv) [HU] and [A]. Finally, the language L in (iii) can be constructed from L_1 and L_2 by the operations given in (i): $L = (L_1L_2)^* \cup (L_1L_2)^*L_1 \cup (L_2L_1)^* \cup (L_2L_1)^*L_2$. □

Regular, context-free and indexed languages are generated by corresponding types of grammars (see [HU]). Briefly, an indexed grammar has disjoint finite sets N, Σ, F of nonterminals, terminals, and indices respectively. The language (over Σ) generated by the grammar is obtained by beginning with a designated start symbol

$S \in N$, and performing substitutions. The allowed substitutions are determined by a finite set of productions, and the grammar is completely described by N, Σ, F, S, and the set of productions. Productions of three types are allowed:

(1) $A \to \alpha$ (2) $A \to Bf$ (3) $Af \to \alpha$, where A and B are nonterminals, $\alpha \in (N \cup \Sigma)^*$, and f is an index. Roughly speaking, the corresponding substitutions consist of replacing the lefthand side of a production by the righthand side. This procedure generates words in $(N \cup \Sigma \cup F)^*$. The words in Σ^* which can be derived in this way are the language generated by the grammar.

Words in $(N \cup \Sigma \cup F)^*$ are called sentential forms. To apply a production of type (1) to a sentential form β one finds an occurrence of A in β and substitutes α for A. The string of indices δ (possibly the empty string) following the occurrence of A is repeated after each nonterminal in α. In case (2) one simply substitutes Bf for A. Finally in case (3) one finds an occurrence of Af followed by a string of indices δ, substitutes α for Af, and repeats δ after each nonterminal in α. An example of an indexed grammar occurs after Lemma 3.15.

A context-free grammar is the same as an indexed grammar except that there are no indices, and a regular grammar is a context-free grammar in which all productions have a righthand side consisting of an element of $\Sigma \cup \{\varepsilon\}$ possibly followed by a single nonterminal.

The next result gives a property of context free languages which does not hold for indexed languages (as can be deduced from Corollary 3.4 and the proof of Theorem 2.17). This property is crucial for the proof of Theorem A in Section 4. We remark in passing that the following proof actually shows that a context free language is bounded if it has subexponential growth.

PROPOSITION 1.3. *If $L \subseteq \Sigma^*$ is a context free language of polynomial growth, then L is bounded.*

Proof. The proof is by induction on the number of nonterminals in a context-free grammar generating L. For any derivation of $w \in L$ from S there will be a last occurrence of S, say: $S \xrightarrow{*} \alpha S \beta \xrightarrow{*} w$. That is, S does not appear in the derivation $\alpha S \beta \xrightarrow{*} w$ except for the initial term. It follows that $w = xyz$ with

$$\alpha \xrightarrow{*} x \qquad S \xrightarrow{*} y \qquad \beta \xrightarrow{*} z.$$

We claim that there are words t and u such that if any $w \in L$ is decomposed as above, then $x \in t^*$ and $z \in u^*$. Consider any two instances

(1) $S \xrightarrow{*} \alpha_1 S \beta_1 \xrightarrow{*} x_1 y_1 z_1$

(2) $S \xrightarrow{*} \alpha_2 S \beta_2 \xrightarrow{*} x_2 y_2 z_2$.

The first part of (2), $S \xrightarrow{*} \alpha_2 S \beta_2$, may be used to expand the S in the middle of (1) and vice versa. By iterating this process, one sees that every word in the submonoid

$\{x_1, x_2\}^*$ occurs as a prefix of some word in L, and consequently that the number of words of length k in $\{x_1, x_2\}^*$ is bounded above by a polynomial in k. In particular for some m the 2^m words of length m over $\{r, s\}$ cannot all remain distinct when the substitutions $r = x_1x_2$ and $s = x_2x_1$ are made. It follows that $x_1x_2 = x_2x_1$. By [Lo, Section 1.3], every word x' in a free monoid determines a unique word $t = t(x')$ such that x' is a power of t and t is not a proper power. Furthermore, if $x'x'' = x''x'$ then $t(x') = t(x'')$. Consequently, in the present setting we deduce that x_1 and x_2 are both powers of $t = t(x_1) = t(x_2)$, whence t is a word with the desired properties. A similar argument on suffixes instead of prefixes yields the desired word u.

Now let L' be the sublanguage of L consisting of words y with derivations $S \xrightarrow{*} y$ such that the only occurrence of S in the derivation is the visible one. By the preceding argument $L \subseteq t^*L'u^*$, and so it suffices to show that L' is bounded. If S is the only nonterminal in the grammar for L, then L' is finite hence bounded by Lemma 1.1. Otherwise for each nonterminal $A \neq S$ let L_A be the language whose grammar is obtained by deleting S from the grammar for L and taking A as the start symbol. By the induction hypothesis L_A is bounded. A derivation of $y \in L'$ must begin with a production $S \to \alpha$ such that $\alpha = Z_1 \cdots Z_n$ where each Z_i is either a terminal or a nonterminal not equal to S. It follows that $y = y_1 \cdots y_n$ with $y_i \in L_{Z_i}$ if Z_i is a nonterminal and $y_i = Z_i$ if Z_i is a terminal. Thus y lies in a product of n bounded languages determined by $S \to \alpha$, and L' lies in the union of the products determined by all the productions $S \to \alpha$ with S not occurring in α. L' is bounded by Lemma 1.1. □

LEMMA 1.4. *Any bounded language $L \subseteq \Sigma^*$ is a finite union of bounded languages L_λ such that for each λ there exists an integer r (depending on λ) and a choice of words $v_{\lambda,0}, \ldots, v_{\lambda,r} \in \Sigma^*$ and $u_{\lambda,1}, \ldots, u_{\lambda,r} \in \Sigma^*$ with*

$$L_\lambda = \{v_{\lambda,0}u_{\lambda,1}^{n_1}v_{\lambda,1} \cdots u_{\lambda,r}^{n_r}v_{\lambda,r} \mid (n_1, \ldots, n_r) \in S_\lambda\},$$

where $S_\lambda \subseteq \mathbb{N}^r$ is empty if $r = 0$ (in which case $L_\lambda = \{v_{\lambda,0}\}$) and otherwise there exist r-tuples in S_λ whose smallest entry is arbitrarily large.

Proof. By hypothesis $L \subseteq w_1^* \cdots w_n^*$ for some words $w_j \in \Sigma^*$. Consider subsets $J \subseteq \{1, 2, \ldots, n\}$ such that the exponents of all w_j with $j \in J$ can be made simultaneously arbitrarily large by elements of L. The proof is by induction on the number of such subsets.

If there are no such subsets, then L is finite and is the union of a finite number of languages L_λ each consisting of a single word $v_{\lambda,0}$. Otherwise, pick a subset J maximal with respect to inclusion. By maximality, there exists an integer M such that for any $w = w_1^{e_1} \cdots w_n^{e_n} \in L$, if the exponents e_j with $j \in J$ are all greater than M, then all of the other exponents are less than or equal to M.

Let L' be the set of all $w \in L$ which can be written as

$$w = w_1^{e_1} \cdots w_n^{e_n}$$

with $e_j > M$ for all $j \in J$ and $e_j \leq M$ for all $j \notin J$. By induction $L - L'$ has the required decomposition, so it only remains to show that L' can be written as a union of finitely many sublanguages of the desired form. Clearly L' is a union of languages L'' such that the exponents e_j, $j \notin J$ are fixed for each L''. If some L'' is such that all of the exponents e_j, $j \in J$ can be made simultaneously arbitrarily large by choosing suitable $w \in L''$, then L'' is of the desired form with the $\{w_j \mid j \in J\}$ playing the role of the $u_{\lambda,p}$ in the statement of the lemma, and the remaining w_j's being multiplied together to form the $v_{i,p}$'s. Otherwise the induction hypothesis applies to L''. □

Section 2: $\mathscr{A}$-combings of groups

In this section we review and extend some basic results concerning combings of groups, with particular emphasis on the preservation of linguistic and geometric structure by such operations as change of generators, passing to a commensurable group, and the formation of free products, all of which are operations which we shall need to invoke in our study of normal forms for 3-manifold groups. We shall also consider subgroups which are quasiconvex with respect to a given combing (see Definition 2.11). Our main results in this direction (Theorems 2.15 and 2.17) allow us to extend the results of [B3] and thus show that the relationship of context-free combings to the study of spaces of nonpositive curvature is much the same as that of regular combings.

DEFINITION 2.1. A choice of generators for a group G is a surjective monoid homomorphism $\mu: \Sigma^* \twoheadrightarrow G$, where Σ^* is the free monoid over the alphabet (set) Σ. We assume that Σ is equipped with an involutory permutation written $a \to a^{-1}$ such that $\mu(a^{-1}) = (\mu(a))^{-1}$. Elements of Σ are called letters and elements of Σ^* words.

Convention. All groups are understood to be finitely generated and all alphabets finite unless otherwise noted.

We denote letters in Σ by $a, b, \ldots$, words in Σ^* by $u, v, w, \ldots$, and elements of G by $g, h, \ldots$. We shall also use the symbol ε to denote the empty word in Σ^*. The involutory permutation of Definition 2.1 extends uniquely to a permutation $w \to w^{-1}$ of Σ^* satisfying the condition $(wv)^{-1} = v^{-1}w^{-1}$, and $\mu(w^{-1}) = (\mu(w))^{-1}$ for all $w \in \Sigma^*$. When there is no risk of ambiguity, we shall write $\bar{w}$ rather than $\mu(w)$ for the image in G of a word w, and more generally $\bar{X}$ for the image of a set of words X. The length of w will be denoted $|w|$.

Any choice of generators determines a left invariant word metric on G.

DEFINITION 2.2. $d_\mu(g, h) = \min\{|w| \mid w \in \Sigma^*, \mu(w) = g^{-1}h\}$.

According to the context, we may write d_Σ, d_G, or just d rather than d_μ. If $H \subseteq G$ is a subgroup of finite index in G, then for any word metrics d_G on G and d_H on H there exists a positive constant c such that for all $h, h' \in H$

$$(1/c)d_H(h', h) \le d_G(h', h) \le c d_H(h', h). \tag{2.1}$$

When discussing geometric properties of G it is natural to view words in Σ^* as paths in G in the following sense.

DEFINITION 2.3. A discrete path in a set X is an eventually constant map $p\colon \mathbb{N} \to X$. The collection of all discrete paths in X is denoted $\mathcal{P}(X)$.

If $\mu\colon \Sigma^* \twoheadrightarrow G$ is a choice of generators, then any word $w = a_1 \cdots a_n \in \Sigma^*$ determines a discrete path in G, beginning at $1 \in G$, obtained by defining $p_w(t)$ to be the image in G of the prefix of length t in w if $t \le |w|$, and to be constant at $\bar{w}$ thereafter. Henceforth we shall identify w with p_w and often talk of words as discrete paths in G. For any metric space (X, d) one natural notion of the distance between two discrete paths p and q is $\max_i\{d(p(i), q(i))\}$, which records the maximum separation of points ('fellow-travellers') traversing the two paths at unit speed. However, for our purposes it is more useful to consider the asynchronous analogue of this notion. In other words when measuring the distance between points traversing the paths under consideration we allow these points to stop for a while whenever this helps them to stay close to their fellow-traveller. The appropriate technical device to encode this idea is the set of monotone reparametrizations

$$\mathcal{R} = \{\rho\colon \mathbb{N} \to \mathbb{N} \mid \rho(0) = 0;\ \rho(n+1) \in \{\rho(n), \rho(n)+1\} \text{ for all } n;\ \rho \text{ unbounded}\}.$$

DEFINITION 2.4. Given a metric space (X, d), the distance between two discrete paths $p, q \in \mathcal{P}(X)$ is defined to be:

$$D_X(p, q) = \min_{\rho, \rho' \in R} \{\max_{n \ge 0} \{d(p(\rho(n)), q(\rho'(n)))\}\}.$$

In the case of the metric space (G, d_μ) we may write D_μ or D_Σ, or even just D, for D_G.

LEMMA 2.5. *The following conditions hold.*

(i) *For any metric space (X, d), D_X is a pseudometric on $\mathscr{P}(X)$; and if d and d' are Lipschitz equivalent metrics on X, then the induced pseudometrics D_X and D'_X on $\mathscr{P}(X)$ are Lipschitz equivalent.*

(ii) *Consider $\mu\colon \Sigma^* \to G$, and suppose that $\nu\colon \Lambda^* \twoheadrightarrow G$ is another choice of generators. If $\tau\colon \Sigma^* \to \Lambda^*$ is a monoid homomorphism with $\mu = \nu \circ \tau$, then there is a constant k such that $D_\Lambda(\tau(w), \tau(v)) \leq k D_\Sigma(w, v)$ for all $v, w \in \Sigma^*$.*

(iii) *If H is a subgroup of finite index in G then for any choice of word metrics on H and G there exists a constant C such that*

$$(1/C) D_G(p, q) \leq D_H(p, q) \leq C D_G(p, q)$$

for all $p, q \in \mathscr{P}(H)$.

Proof. For the triangle inequality in part (i) one needs to observe first that if $\rho_1, \rho_2 \in \mathscr{R}$, then there are $\rho, \rho' \in \mathscr{R}$ such that $\rho_1 \circ \rho = \rho_2 \circ \rho'$, and second that if $D_X(p, q)$ is realized by reparametrizations $p \circ \rho_1$ and $q \circ \rho_2$, then for any $\rho \in \mathscr{R}$ it is also realized by $p \circ \rho_1 \circ \rho$ and $q \circ \rho_2 \circ \rho$.

Part (iii) follows from (i) and equation (2.1). For part (ii), let k' bound the lengths $|\tau(a)|$, $a \in \Sigma$. The path $p_{\tau(w)}$ is obtained from p_w by inserting subpaths of d_Λ-length at most k'. □

DEFINITION 2.6. Let $\mu\colon \Sigma^* \twoheadrightarrow G$ be a choice of generators, d the induced metric on G, and D the induced pseudometric on $\Sigma^* \subseteq \mathscr{P}(G)$. Recall that a full AFL is a class of languages satisfying certain conditions discussed in Section 1. Let L be a language over Σ.

(i) If $D(w, v) \leq k$, then w and v are called asynchronous k-fellow-travellers.

(ii) L is said to satisfy the asynchronous fellow-traveller property if there exists a constant k such that $D(w, v) \leq k$ for all $w, v \in L$ with $d(\bar{w}, \bar{v}) \leq 1$.

(iii) L is called a combing if it projects bijectively to G.

(iv) L is said to be an asynchronous $\mathscr{A}$-combing if it satisfies conditions (ii) and (iii) and belongs to the full *AFL* $\mathscr{A}$.

Example. In the above terminology, the asynchronous automatic groups of [E+] are precisely those which admit an asynchronous regular combing [BGSS, Theorem 7.3 and Lemma 7.2(1)].

LEMMA 2.7. *If G admits an asynchronous $\mathscr{A}$-combing with respect to one choice of generators, then it admits such a combing with respect to any choice of generators.*

More precisely, if L is an asynchronous $\mathscr{A}$-combing with respect to the choice of generators $\mu: \Sigma^* \twoheadrightarrow G$, and $\nu: \Lambda^* \twoheadrightarrow G$ is another choice of generators, then for any monoid homomorphism $\tau: \Sigma^* \to \Lambda^*$ such that $\mu = \nu \circ \tau$, $\tau(L)$ is an asynchronous $\mathscr{A}$-combing.

Proof. Let $L_1 = \tau(L)$. It is clear that L_1 is a combing of G. Since full AFL's are closed under homomorphism, L_1 is in $\mathscr{A}$. Finally, L_1 is asynchronously bounded by Lemma 2.5(ii). □

LEMMA 2.8. *If $L \subseteq \Sigma^*$ is an asynchronous $\mathscr{A}$-combing of G and $L_1 \subseteq \Sigma^*$ is obtained from L by replacing one word $w \in L$ with any other $v \in \Sigma^*$ such that $\bar{v} = \bar{w}$, then L_1 is an asynchronous $\mathscr{A}$-combing of G.*

Proof. It is clear that L_1 is a combing of G. Since the transition from L to L_1 can be accomplished by intersection and union with regular sets, Lemma 1.2 ensures that L_1 is in $\mathscr{A}$. As L_1 is a combing, there are only finitely many $u \in L_1$ with $d(\bar{u}, \bar{v}) \leq 1$. Thus it suffices to increase the constant k in the fellow-traveller property for L by adding $\max\{D(u, v) \mid d(\bar{u}, \bar{v}) \leq 1\}$ to it. If k' is the resulting constant then L_1 satisfies the asynchronous fellow-traveller property with constant k'. □

PROPOSITION 2.9. *If G_1 and G_2 both have an asynchronous $\mathscr{A}$-combing then so too does the free product $G_1 * G_2$.*

Proof. Recall that every non-identity element $g \in G_1 * G_2$ can be expressed uniquely as a product $g_1 \cdots g_n$ of non-identity elements chosen alternately from G_1 and G_2. Thus the language L consisting of all words $w = u_1 \cdots u_n$ with successive u_j chosen alternately from L_1 and L_2 is a combing of $G_1 * G_2$. Lemma 1.2 implies that L is in $\mathscr{A}$. The proof that L has the asynchronous fellow-traveller property is straightforward and standard, *cf.*, [E+], [AB, Sec. 10], [B2]. □

Our next objective is to establish the relationship between the existence of asynchronous $\mathscr{A}$-combings for subgroups and their existence for the ambient group.

PROPOSITION 2.10. *Let H be a subgroup of finite index in G. If H admits an asynchronous $\mathscr{A}$-combing, then so too does G.*

Proof. Let $L_H \subseteq \Sigma^*$ be an asynchronous $\mathscr{A}$-combing of H. Fix coset representatives $1 = g_1, \ldots, g_m$ with $G = \bigcup Hg_i$ and append corresponding new symbols $b_1, \ldots, b_m$ (and their formal inverses) to Σ_H to obtain Σ_G. Extend $\Sigma_H^* \twoheadrightarrow H$ to Σ_G^* by $b_i \mapsto g_i$. By Lemma 1.2, the language $L := \bigcup_i \{wb_i \mid w \in L_H\}$ lies in $\mathscr{A}$. L is obviously a combing of G, so it only remains to establish the asynchronous fellow-traveller property.

Let K be the maximum of the constants appearing in equation (2.1) and Lemma 2.5(iii). By hypothesis there is a constant K' such that $D_H(u_1, u_2) \leq K'$ for all $u_1, u_2 \in L_H$ with $d_H(\overline{u_1}, \overline{u_2}) \leq 1$. If $w_1 b_i, w_2 b_j \in L$ with $d_G(\overline{w_1 b_i}, \overline{w_2 b_j}) \leq 1$, then $d_G(\overline{w_1}, \overline{w_2}) \leq 3$. Thus $d_H(\overline{w_1}, \overline{w_2}) \leq 3K$, and the triangle inequality for D_H yields $D_H(w_1, w_2) \leq 3KK'$. Hence $D_G(w_1, w_2) \leq 3K^2K'$ and $D_G(w_1 b_i, w_2 b_j) \leq 3K^2K' + 2$. $\square$

We shall see that in the framework of $\mathscr{A}$-combings the passage to subgroups of finite index is a more delicate matter than the preceding argument for finite extensions. However, the arguments involved are more potent in that they apply not only to subgroups of finite index but also to a rich variety of subgroups that satisfy the following geometric condition introduced by Gromov [G, Sec. 7.3].

DEFINITION 2.11. If X is a metric space and S is a subset of $\mathscr{P}(X)$, then a subset $Y \subseteq X$ is said to be S-quasiconvex if there exists a constant k such that the image of every path $p \in S$ whose endpoints lie in Y is entirely contained in the k-neighborhood of Y. In particular, if L is a combing of a finitely generated group G and $H \subseteq G$ is a subgroup, then H is L-quasiconvex if there exists a constant k such that for every $w \in L$ with $\bar{w} \in H$ and every prefix u of w there is an element $h \in H$ with $d(\bar{u}, h) \leq k$.

Example. If H is a subgroup of finite index in G, then H is quasiconvex with respect to any subset $L \subseteq \Sigma^*$, since every point of G lies within a uniformly bounded distance of H.

We shall need the following notion of *compatibility* for sets of paths in $\mathscr{P}(G)$, particularly as applied to languages $L \subseteq \Sigma^* \subseteq \mathscr{P}(G)$.

DEFINITION 2.12. Let X be a metric space. We define a relation among subsets $S, S' \in \mathscr{P}(X)$ by saying that $S \preccurlyeq S'$ if there is a constant K such that for every $p \in S$ one can find $p' \in S'$ with the same endpoints as p and $D_X(p, p') \leq K$. Two subsets $S, S' \in \mathscr{P}(X)$ are said to be compatible if $S \preccurlyeq S'$ and $S' \preccurlyeq S$.

In the case that the metric space is a group with a word metric, Lemma 2.5(ii) guarantees that the validity of $S \preccurlyeq S'$ does not depend on the particular word metric chosen. Thus we may speak of compatible subsets of $\mathscr{P}(G)$ without reference to a particular word metric. Compatibility has the following properties, whose proofs are trivial.

LEMMA 2.13.

(i) *If $S, S' \subseteq \mathscr{P}(X)$ are compatible, then a subset $Y \subseteq X$ is S-quasiconvex if and only if it is S'-quasiconvex.*

(ii) *Suppose $\mu\colon \Sigma \twoheadrightarrow G$ and $\mu_1\colon \Sigma_1 \twoheadrightarrow G$ are choices of generators. If $L \subseteq \Sigma^* \subseteq \mathscr{P}(G)$ and $L_1 \subseteq \Sigma_1^* \subseteq \mathscr{P}(G)$ are compatible subsets of $\mathscr{P}(G)$, then L satisfies the asynchronous fellow-traveller property if and only if L_1 does.*

The usefulness of the notion of quasiconvexity to the study of $\mathscr{A}$-combings is largely a consequence of the following result, which has been widely used in the case where $\mathscr{A}$ is the class of regular languages.

LEMMA 2.14. *Let $\mu\colon \Sigma^* \to G$ be a choice of generators for G and let $\mathscr{A}$ be a full AFL. Suppose that $L \subseteq \Sigma^*$ belongs to $\mathscr{A}$ and that the subgroup $H \subseteq \mu(L)$ is L-quasiconvex. Then there exist a finite subset $\Sigma_H \subseteq \Sigma^*$, a language $L_H \subseteq \Sigma_H^*$ (where Σ_H^* is the submonoid of Σ^* generated by Σ_H), a constant $k > 0$, and a bijection $\gamma\colon L \cap \mu^{-1}(H) \to L_H$ such that*

(i) *Σ_H freely generates the submonoid Σ_H^*.*

(ii) *The restriction of μ to Σ_H^* is a choice of generators for H.*

(iii) *Considered as a language over Σ_H, L_H belongs to $\mathscr{A}$.*

(iv) *For all $w \in L \cap \mu^{-1}(H)$, $\mu(\gamma(w)) = \mu(w)$ and $D_G(\gamma(w), w) \leq k$.*

In particular $L \cap \mu^{-1}(H)$ and L_H are compatible subsets of $\mathscr{P}(G)$.

From Lemma 2.14 and Lemma 2.13 we obtain:

THEOREM 2.15. *If G admits an $\mathscr{A}$-combing L, and $H \subseteq G$ is an L-quasiconvex subgroup, then H admits an $\mathscr{A}$-combing L_H with the property that a subgroup $K \subseteq H$ is L-quasiconvex if and only if it is L_H-quasiconvex. Further, if L satisfies the asynchronous fellow-traveller property, then so too does L_H.*

Combining Lemma 2.14 with Proposition 2.10 we have:

THEOREM 2.16. *Let G be a finitely generated group, let $H \subseteq G$ be a subgroup of finite index, and let $\mathscr{A}$ be a full AFL. Then, G admits an asynchronous $\mathscr{A}$-combing if and only if H admits an asynchronous $\mathscr{A}$-combing.*

Proof of Lemma 2.14. The key idea of the proof is to use the quasiconvexity of H to construct a certain finite subgraph Γ_0 of the Schreier diagram $\Gamma_{G/H}$ for G/H. Recall that $\Gamma_{G/H}$ is the directed labelled graph with vertex set G/H and edges $Hx \xrightarrow{a} Hy$ whenever $a \in \Sigma$ with $Hx\bar{a} = Hy$. In particular for each such edge there is an inverse edge $Hy \xrightarrow{a^{-1}} Hx$. It follows from this definition that for every $w \in \Sigma^*$ and vertex Hx there is exactly one path in $\Gamma_{G/H}$ with label w that begins at Hx. Further the labels of paths which both begin and end at H are precisely those words in Σ^* which represent elements of H. Here path means edge-path in a graph as opposed to discrete path in a metric space.

We first construct a map $\gamma: \Sigma^* \to \Sigma^*$ whose restriction to $L_0 := L \cap \mu^{-1}(H)$ will have the properties stated in the lemma. For each vertex Hx fix a path in $\Gamma_{G/H}$ of minimum length from H to Hx. Let $\rho(Hx)$ be the label of this path, and notice that $\rho(H) = \varepsilon$, the empty word. Notice also that all of these paths are simple, in the sense that they do not pass through any vertex twice. Any word $w = a_1 \cdots a_n \in \Sigma^*$ is the label of a unique path from H. In this path each a_i is the label of an edge from some Hx_{i-1} to Hx_i. Let $\gamma: \Sigma^* \to \Sigma^*$ be the map that sends w to the word obtained by replacing each of its letters a_i by the word $\rho(Hx_{i-1})a_i\rho(Hx_i)^{-1}$. If $w \in \mu^{-1}(H)$, then $Hx_0 = Hx_n = H$. Consequently $\gamma(w)$ is obtained from w by interpolating words $\rho(Hx_i)\rho(Hx_i)^{-1}$ whence $\mu(\gamma(w)) = \mu(w)$. In particular the first assertion of (iv) holds.

Our next goal is to construct the set Σ_H. Σ_H will be a subset of Δ, the collection of all words $\rho(Hx)a\rho(Hx\bar{a})^{-1}$. Notice that Δ is closed under the taking of inverses. We claim that Δ *freely* generates a submonoid of Σ^*. If not, then two distinct products of elements from Δ would be equal in Σ^*; and consequently one of the elements would be a proper prefix in Σ^* of a second. However, it follows from the choice of the words $\rho(Hx)$ that this cannot happen.

We now give each edge $Hx \xrightarrow{a} Hy$ of $\Gamma_{G/H}$ the additional label $\rho(Hx)a\rho(Hy)^{-1} \in \Delta$. Thus a directed edge-path which was originally labelled by $w \in \Sigma^*$ now has $\gamma(w) \in \Delta^*$ as a second label, and the image of γ in Δ^* is just the set of Δ-labels of paths beginning at H. Furthermore, the monoid homomorphism $\pi: \Delta^* \to \Sigma^*$ defined by $\pi(\rho(Hx)a\rho(Hy)^{-1}) = a$ satisfies $\pi(\gamma(w)) = w$. Thus $\gamma: \Sigma^* \to \Delta^*$ is injective.

Because H is L-quasiconvex there is a constant k such that for every $w \in L_0$ and every prefix v of w there exists $h \in H$ with $d(\bar{v}, h) \leq k$. Consequently $\rho(Hv)$ has length at most k. Thus if we define Γ_0 to be the finite subgraph of $\Gamma_{G/H}$ determined by all vertices Hv with $\rho(Hv)$ of length at most k, then every $w \in L_0$ is the Σ-label of a path in Γ_0 that begins and ends at H. Let Σ_H be the subset of Δ consisting of those elements which occur as Δ-labels of edges in Γ_0. Since Γ_0 is finite, Σ_H is too; and because Δ is a set of free generators, Σ_H is too. Thus (i) holds. By construction, Σ_H is closed under inverse; and for each $h \in H$ there exists $w \in L_0$ with $h = \mu(w) = \mu(\gamma(w))$. It follows that μ maps Σ_H^* onto H and (ii) holds.

Let R be the set of all Σ-labels of paths in Γ_0 from H to H, and let R_H be the corresponding set of Δ-labels. Because Γ_0 is finite, R and R_H are regular subsets of Σ^* and Σ_H^* respectively. It follows from the definitions of γ and π that the restrictions of γ to R and π to R_H are inverse maps. Further we know by the construction of Γ_0 that $L_0 = R \cap L$. Hence $L_0 \in \mathscr{A}$ by Lemma 1.2.

Define $L_H = \gamma(L_0)$. From the preceding paragraph γ maps L_0 bijectively to L_H and $L_H = R_H \cap \pi_H^{-1}(L_0)$ where π_H is the restriction of π to Σ_H^*. Lemma 1.2 yields

$L_H \in \mathscr{A}$. Thus (iii) holds. The last part of (iv) follows from the fact that $\gamma(w)$ is obtained from w by interpolating words $\rho(Hx)\rho(Hx)^{-1}$ for some vertices Hx in Γ_0. $\square$

Next we apply the preceding results to the study of context free combings of virtually abelian groups, thus generalizing [AB, 9.1].

Two groups are said to be commensurable if they contain isomorphic subgroups of finite index. In the statement of the next theorem, two subgroups of a group G are said to be *strictly commensurable* if they contain a common subgroup H of finite index. Notice that in the following statement we place no restriction on the geometry of the language L, in particular it is not required to satisfy any fellow-traveller property.

THEOREM 2.17. *Let G be a virtually abelian group. For every context free combing L of G there exist among the strict commensurability classes of subgroups of G only finitely many which contain an L-quasiconvex subgroup.*

Proof. For the duration of this proof 'commensurability' will mean strict commensurability. Let A be an abelian subgroup of finite index in G. It suffices to show that there exist among the commensurability classes of subgroups of A only finitely many which contain an L-quasiconvex subgroup. Furthermore, by Theorem 2.15, we may assume that $G = A$, i.e., that G is abelian.

Fix a choice of generators $\mu: \Sigma^* \twoheadrightarrow G$ and a context-free combing $L \subset \Sigma^*$. Because it maps bijectively onto an abelian group, L must have polynomial growth, whence by Proposition 1.3 it is a bounded language. That is, there exists an integer m and words $w_1, \ldots, w_m \in \Sigma^*$, such that

$$L \subseteq w_1^* w_2^* \cdots w_m^*.$$

We shall show that every L-quasiconvex subgroup of G is commensurable with the subgroup generated by a subset of $\{\overline{w_i} \mid 1 \leq i \leq m\}$. Suppose that $H \subseteq G$ is L-quasiconvex, and let k be as in Definition 2.11. Let c be the number of words in Σ^* of length at most k. Suppose that $h \in H$ is represented by $w_1^{i_1} \cdots w_m^{i_m} \in L$. Given $i_j > c$, we let $v = w_1^{i_1} \cdots w_{j-1}^{i_{j-1}}$ and consider the sequence of prefixes vw_j^r, $0 \leq r \leq i_j$. By quasiconvexity there is for each r an element $h_r \in H$ and a word $u_r \in \Sigma^*$ of length at most k such that $\mu(vw_j^r u_r) = h_r$. But $i_j > c$, so $u_r = u_s$ for some r, s with $0 \leq r < s \leq c$. Since G is abelian, it follows that w_j^{s-r} represents an element of H.

Let K be the subgroup of H generated by its intersections with the finitely many cyclic subgroups $\langle \overline{w_i} \rangle$. The word w_j^{s-r} of the last paragraph represents an element of K. It follows that because G is abelian, the image in G of $\{w_1^{i_1} \cdots w_m^{i_m} \mid i_j \leq c$ for all $1 \leq j \leq m\}$ contains a set of coset representatives for K in H. Thus K has finite index in H. On the other hand, K is of finite index in the subgroup generated by those $\overline{w_i}$ for which $\langle \overline{w_i} \rangle \cap H \neq 1$. □

The results of [B3] demonstrate the incompatibility of regular combings with many combings that arise in the study of manifolds of nonpositive curvature. We recall briefly how the latter combings are defined (see [B3] or [AB] for more details). If a group G acts properly and cocompactly by isometries on a 1-connected geodesic metric space X of nonpositive curvature, then one can map G to X by fixing a basepoint $x_0 \in X$ and making the identification $g \mapsto g \cdot x_0$. Every pair of points in X is joined by a unique geodesic, and one can uniformly approximate these geodesics by discrete paths in G to obtain a combing $\sigma(X)$ of G. It is not difficult to show that, up to compatibility, $\sigma(X)$ is independent of the choice of x_0 and the chosen approximation of geodesics. One can also show that $\sigma(X)$ satisfies the asynchronous fellow-traveller property; this is a reflection of the convexity of the metric in spaces of nonpositive curvature.

[B3] contains several results comparing the behaviour of $\sigma(X)$ to that of regular combings. The arguments presented there can be extended directly to the case of context-free combings by casting Theorems 2.17 and 2.15 above in the roles of [AB, 9.1] and [B3, 2.2] respectively. We record two such results.

THEOREM 2.18. *Let M be a closed real-analytic manifold of nonpositive curvature. The associated combing $\sigma(\tilde{M})$ of $\pi_1 M$ is compatible with a context-free combing if and only if $\pi_1 M$ is word-hyperbolic* (*in the sense of Gromov* [G]).

For the following result we do not assume that M satisfies Thurston's geometrization conjecture.

THEOREM 2.19. *Let M be a closed 3-manifold of nonpositive curvature. The associated combing $\sigma(\tilde{M})$ of $\pi_1 M$ is compatible with a context-free combing if and only if $\pi_1 M$ is word-hyperbolic.*

Section 3: Indexed combings of 3-manifold groups

In this section we shall prove Theorem B by showing that if a group occurs as the fundamental group of a compact geometrizable 3-dimensional manifold or orbifold, then it admits an asynchronous indexed combing. More generally, we show that any group which is commensurable with a free product of automatic groups and finite extensions of semidirect products of the form $\mathbb{Z}^2 \rtimes \mathbb{Z}$ admits an asynchronous indexed combing. The need to introduce indexed combings arises from an inadequacy of context free combings as demonstrated by the results in the next section. The root of this inadequacy can be traced to the fact that the normal forms for groups $G = \mathbb{Z}^n \rtimes \mathbb{Z}$ constructed in [B2] are never context free.

First we recall the definition of these combings. Every element $g \in G$ can be written uniquely as $g = t^m x$, where t is a generator for the righthand factor in the given semidirect product decomposition and x lies in the lefthand factor. We identify the lefthand factor with the integer lattice in Euclidean n-space and choose generators $\mu: \Sigma_n \to \mathbb{Z}^n$ with $\Sigma_n = \{e_i, e_i^{-1} \mid i = 1, \dots, n\}$ and $\{\mu(e_1), \dots, \mu(e_n)\}$ the standard basis (rectangular coordinates) for $\mathbb{Z}^n \subseteq \mathbb{E}^n$. For each $x \in \mathbb{Z}^n$ we let ℓ_x denote the Euclidean straight line segment joining the origin to x, and we choose any $\gamma_x \in \Sigma_n^*$ such that when viewed as a path in $\mathbb{Z}^n \subseteq \mathbb{E}^n$ this word satisfies $D_{\mathbb{E}^n}(\hat{\ell}_x, \gamma_x) \leq \sqrt{n}$, where $\hat{\ell}_x$ is the restriction of the path ℓ_x (parametrized by arc length) to the integer points of its domain. The following result is proved in [B2, Sections 2 and 3].

THEOREM 3.1. *$L_0 := \{t^m \gamma_x \mid m \in \mathbb{Z}, x \in \mathbb{Z}^n\} \subseteq (\Sigma_n \cup \{t, t^{-1}\})^*$ is a combing with the asynchronous fellow-traveller property.*

COROLLARY 3.2. *L_0 is not a context free language.*

Proof. If L_0 were a context free language, then its image under the monoid projection $(\Sigma_n \cup \{t, t^{-1}\})^* \twoheadrightarrow \Sigma_n^*$ would also be context free (Lemma 1.2). But as this image is precisely $L_1 := \{\gamma_x \mid x \in \mathbb{Z}^n\}$, a combing of $\mathbb{Z}^n$ with respect to which every cyclic subgroup is quasiconvex, Theorem 2.17 implies that L_1 is not context free. □

We now turn our attention to the task of obtaining an upper bound on the complexity of the language L_0 in the case of relevance to 3-manifolds, $n = 2$. The following result contrasts sharply with Theorem 2.15.

THEOREM 3.3. *View $\mathbb{Z}^2$ as the integer lattice in the Euclidean plane $\mathbb{E}^2$, and let $\mu: \Sigma_2 \to \mathbb{Z}^2$ be the choice of generators defined above. There exists an asynchronous indexed combing $L \subseteq \Sigma_2^*$ such that for all $w \in L$, $D_{\mathbb{E}^2}(\hat{\ell}_{\bar{w}}, p_w) \leq \sqrt{2}$ where p_w is the discrete path in $\mathbb{Z}^2$ corresponding to w.*

The proof of this theorem is deferred until the end of the section.

COROLLARY 3.4. *For every choice of generators $\mu: \Sigma^* \twoheadrightarrow \mathbb{Z}^2$ there exists an asynchronous indexed combing $L \subseteq \Sigma^*$ with respect to which every cyclic subgroup of $\mathbb{Z}^2$ is quasiconvex.*

Proof. It is immediate that every cyclic subgroup is quasiconvex with respect to the combing L of Theorem 3.3. If $\nu: A \to \mathbb{Z}^2$ is another choice of generators, let $\tau: \Sigma_2^* \to A^*$ be a monoid homomorphism such that $\mu = \nu \circ \tau$. By Lemma 2.7, $\tau(L)$ is an asynchronous indexed combing of $\mathbb{Z}^2$. It is straightforward to check that L and $\tau(L)$ are compatible subsets of $\mathscr{P}(\mathbb{Z}^2)$, in the sense of Definition 2.11. Hence by Lemma 2.13 we are done. □

COROLLARY 3.5. *Every semidirect product of the form $G = \mathbb{Z}^2 \rtimes \mathbb{Z}$ admits an asynchronous indexed combing.*

Proof. Let L be the indexed combing yielded by Theorem 3.3. By Theorem 3.1 $L_0 := \{t^* \cup (t^{-1})^*\}L$ is an asynchronous combing of G, and by Lemma 1.2 L_0 is an indexed language over $\{t, t^{-1}\} \cup \Sigma_2$. □

COROLLARY 3.6. *Every group G which is commensurable with a free product of asynchronously automatic groups and finite extensions of semidirect products of the form $\mathbb{Z}^2 \rtimes \mathbb{Z}$ admits an asynchronous indexed combing.*

Proof. Asynchronously automatic groups are precisely those which admit asynchronous regular combings [E+], [BGSS, Theorem 7.3 and Lemma 7.2(1)], so the preceding corollary together with Proposition 2.9 and Corollary 2.16 yields the desired conclusion. □

The results of Epstein and Thurston [E+, Chapter 12] imply that the fundamental group of any compact orientable 3-manifold that satisfies Thurston's geometrization conjecture is the free product of an automatic group and (possibly) finite extensions of semidirect products of the form $\mathbb{Z}^2 \rtimes \mathbb{Z}$ (the latter being the fundamental groups of closed *Nil* and *Sol* manifolds). One can remove the hypothesis of orientability at the expense of passing to a subgroup of index at most two. Indeed, by passing to subgroups of higher index one can generalize this result to include orbifolds. Thus Corollary 3.6 completes the proof of Theorem B.

Proof of Theorem 3.3. We begin by reducing to the first quadrant $\mathscr{Q} = \{(m, n) \mid m > 0, n \geq 0\}$. Let $\Sigma = \{a, b\}$, and define $\mu: \Sigma^* \to \mathbb{Z}^2$ by $\mu(a) = (1, 0)$, $\mu(b) = (0, 1)$.

LEMMA 3.7. *It suffices to find $L_1 \subseteq \{a, b\}^*$ such that:*

(i) *μ maps L_1 bijectively to $\mathscr{Q}$;*

(ii) *if $wv \in L_1$, then the distance in the plane from $\bar{w}$ to the line segment from the origin to $\overline{wv}$ is less than* 1;

(iii) *L_1 is an indexed language.*

Proof. If $wv \in L_1$ and $\bar{w} = (x, y)$, then $x + y = |w|$. In particular the Euclidean distance from $(0, 0)$ to $\bar{w}$ is a monotonically increasing function of $|w|$. It follows from this observation together with (ii) that $D_{\mathbb{E}^2}(\hat{\ell}_{\bar{w}}, p_w) \leq \sqrt{2}$ for all $w \in L_1$. One can rotate $\mathscr{Q}$ into the other quadrants to obtain languages L_2, L_3, L_4 which satisfy corresponding conditions. For example, L_2 is obtained by replacing b by a^{-1} and a by b in every word in L_1. The language $L = L_1 \cup L_2 \cup L_3 \cup L_4 \cup \{\varepsilon\}$ is indexed by Lemma 1.2. □

It remains to find L_1 satisfying conditions (i) to (iii). We make $\mathscr{Q}$ into a graph by adding horizontal and vertical edges of unit length directed from initial vertices (i, j) to terminal ones $(i + 1, j)$ or $(i, j + 1)$. Fix $(m, n) \in \mathscr{Q}$ and consider the closed line segment $\ell(m, n^+)$ in the plane from $(0, 0)$ to (m, n^+) where n^+ is chosen slightly larger than n but small enough to ensure that

$$\begin{aligned}&\ell(m, n^+) \text{ does not contain any lattice points except } (0,0), \text{ and}\\&\text{there are no lattice points in the interior of the triangle}\\&\text{with vertices } (0,0), (m,n), (m,n^+).\end{aligned} \tag{3.1}$$

Notice that condition (3.1) guarantees that the sequence of edges which intersect $\ell(m, n^+)$ depends only on (m, n) and not on the particular choice of n^+. In particular the sequence of edges intersected is unchanged if $\ell(m, n^+)$ is replaced by $\ell(m, n_1)$ for any n_1 strictly between n and n^+.

DEFINITION 3.8. For $(m, n) \in \mathscr{Q}$ and $\ell(m, n^+)$ chosen as above, the sequence formed by recording (in order) an h for each horizontal edge intersected by $\ell(m, n^+)$ and a v for each vertical edge is called the crossing sequence $\kappa(m, n)$.

Example. $\kappa(2, 3) = hvhhv$.

LEMMA 3.9. *The following conditions hold for all* $(m, n) \in \mathscr{Q}$:

(i) *The first edge intersected by* $\ell(m, n^+)$ *will be vertical with initial vertex* $(1, 0)$ *if* $m > n$ *and will be horizontal with initial point* $(0, 1)$ *if* $m \leq n$.

(ii) *The last edge intersected by* $\ell(m, n^+)$ *will always be vertical with initial point* (m, n), *and if* $n > 0$, *the penultimate edge will always be horizontal.*

(iii) *If* $\ell(m, n^+)$ *intersects two edges in succession and the second edge is vertical, then the initial point of the second edge is obtained from the initial point of the first edge by increasing the* x*-coordinate by* 1. *If the second edge is horizontal, the* y*-coordinate is increased by* 1.

Proof. The first two assertions are clear. A pair of successive edges intersected by $\ell(m, n^+)$ forms two sides of a square. The first edge is one of the two (closed) sides of the square that contain the vertex closest to the origin, and the second edge contains the vertex farthest from the origin. It follows by consideration of the posible cases that initial points of successive edges are related as in (iii). □

For each $(m, n) \in \mathcal{Q}$ consider the path $\gamma(m, n)$ in $\mathbb{Z}^2$ starting at $(0, 0)$ and proceeding through the initial points of all edges intersected by $\ell(m, n^+)$. Let L_1 be the language over Σ consisting of all the paths $\gamma(m, n)$ for $(m, n) \in \mathcal{Q}$. Since the last initial point is always (m, n), condition (i) of Lemma 3.7 holds. By condition (3.1) each (closed) edge which intersects $\ell(m, n^+)$ also intersects the line from $(0, 0)$ to (m, n). Consequently condition (ii) of Lemma 3.7 is valid also. It follows in a straightforward way from Lemma 3.9 that $\gamma(m, n)$ is obtained by replacing v by a and h by b in $\kappa(m, n)$. Thus the next theorem completes the proof of Theorem 3.3.

THEOREM 3.10. *The set* $\{\kappa(m, n) \mid (m, n) \in \mathcal{Q}\}$ *is an indexed language.*

The proof of Theorem 3.10 is divided into a number of lemmas. Fix $(m, n) \in \mathcal{Q}$ and choose n^+ and $\ell(m, n^+)$ as above. Let s be the slope of $\ell(m, n^+)$. Note that $s > 0$. The points on $\ell(m, n^+)$ whose x-coordinate is a positive integer are

$$(1, s), (2, 2s), \ldots, (m, ms).$$

Thus the vertical edges intersected by $\ell(m, n^+)$ have initial points

$$(1, \lfloor s \rfloor), (2, \lfloor 2s \rfloor), \ldots, (m, \lfloor ms \rfloor)$$

where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. All the remaining edges are horizontal. It follows from Lemma 3.9 that the sequence of initial points of edges intersected by $\ell(m, n^+)$ is

$$(0, 1), (0, 2), \ldots, (0, \lfloor s \rfloor), (1, \lfloor s \rfloor), (1, \lfloor s \rfloor + 1), \ldots, (1, \lfloor 2s \rfloor)$$
$$(2, \lfloor 2s \rfloor), (2, \lfloor 2s \rfloor + 1), \ldots, (m, \lfloor ms \rfloor)$$

where it is understood that the sum of the x and y coordinates increases by 1 at each step. In particular if $\lfloor s \rfloor = 0$, then the sequence begins with $(1, \lfloor s \rfloor)$; and if $\lfloor s \rfloor = \lfloor 2s \rfloor$ then the sequence is $\ldots, (1, \lfloor s \rfloor), (2, \lfloor 2s \rfloor), \ldots$.

LEMMA 3.11. *For any s greater than but sufficiently close to* n/m:
(i) $\kappa(m, n) = h^{\lfloor s \rfloor} v h^{\lfloor 2s \rfloor - \lfloor s \rfloor} v h^{\lfloor 3s \rfloor - \lfloor 2s \rfloor} v \cdots h^{\lfloor ms \rfloor - \lfloor (m-1)s \rfloor} v.$
(ii) *If* $n > 0$, *then* $\kappa(m, n) = v^{\lfloor 1/s \rfloor} h v^{\lfloor 2/s \rfloor - \lfloor 1/s \rfloor} h v^{\lfloor 3/s \rfloor - \lfloor 2/s \rfloor} h \cdots v^{\lfloor n/s - (n-1)/s \rfloor} h v.$
(iii) *If* $n = 0$, *then* $\kappa(m, n) = v^m$.

Proof. The first assertion follows from the preceding analysis of the vertical edges intersected by $\ell(m, n^+)$. For the second part one uses Lemma 3.9 to analyze the horizontal edges intersected by $\ell(m, n^+)$. The last assertion is clear from the definition of κ. □

First we apply Lemma 3.11 to the case in which $\kappa(m, n)$ begins with h. In this case $n \geq m$, and $\kappa(m, n) = h^{\lfloor s \rfloor} v \cdots$ with $\lfloor s \rfloor \geq 1$. The other blocks of h's have lengths $\lfloor ks \rfloor - \lfloor (k-1)s \rfloor$ for $1 \leq k \leq m$. Application of the estimate $x - 1 < \lfloor x \rfloor \leq x$ yields $s - 1 < \lfloor ks \rfloor - \lfloor (k-1)s \rfloor < s + 1$. Thus the '$h$-blocks' have length $\lfloor s \rfloor$ or $\lfloor s \rfloor + 1$, where by 'h-block' we mean a maximal subword of the form h^i. Consequently the following definition makes sense.

DEFINITION 3.12. If $\kappa(m, n) = h^i v \cdots$ with $i > 0$, define $H(\kappa(m, n))$ to be the sequence obtained by removing i h's from each h-block of $\kappa(m, n)$.

LEMMA 3.13.
(i) *If $\kappa(m, n)$ begins with h, then $H(\kappa(m, n))$ begins with v and $H(\kappa(m, n)) = \kappa(m, n - m\lfloor n/m \rfloor)$.*
(ii) *If $\kappa(m, n)$ begins with v, then $H^{-1}(\kappa(m, n))$ consists of all sequences obtained by choosing a positive integer k and replacing each v in $\kappa(m, n)$ by $h^k v$.*

Proof. The first part of (i) is clear. For the second pick s greater than but close enough to n/m to compute $\kappa(m, n)$ by Lemma 3.11. We may assume that s is close enough so that $\lfloor s \rfloor = \lfloor n/m \rfloor$. Decrease s if necessary to make $t = s - \lfloor s \rfloor$ close enough to $n/m - \lfloor n/m \rfloor$ to compute $\kappa(m, n - m\lfloor n/m \rfloor)$. As $\lfloor kt \rfloor - \lfloor (k-1)t \rfloor = \lfloor ks \rfloor - \lfloor (k-1)s \rfloor - \lfloor s \rfloor$, Lemma 3.11(i) yields the rest of (i).

Consider $\kappa(m, n)$ as in (ii); in particular $n < m$ lest $\kappa(m, n)$ begin with h. By Definition 3.12 every member of $H^{-1}(\kappa(m, n))$ has the desired form. To show that $H^{-1}(\kappa(m, n))$ contains every sequence promised by (ii), observe that by Lemma 3.11(i) $\kappa(m, n + km) = h^k v \cdots$ if $k \geq 1$. By (i) $H(\kappa(m, n + km)) = \kappa(m, n)$. □

Now we apply Lemma 3.11 to the case in which $\kappa(m, n)$ begins with v and contains an h; that is, the case $1 \leq n < m$.

DEFINITION 3.14. *If $\kappa(m, n) = v^i h \cdots$ with $i > 0$, define $V(\kappa(m, n))$ be the sequence obtained by removing i v's from each v-block of $\kappa(m, n)$ except the last (which is of length 1).*

The proof of the following lemma is entirely similar to that of Lemma 3.13.

LEMMA 3.15.

(i) *If* $\kappa(m, n) = v^i h \cdots$ *with* $i > 0$, *then* $V(\kappa(m, n))$ *begins with* h *and* $V(\kappa(m, n)) = \kappa(m - n\lfloor m/n \rfloor, n)$ *unless* m *is a multiple of* n *in which case* $V(\kappa(m, n)) = \kappa(n, n)$.

(ii) *If* $\kappa(m, n) = h \cdots$, *then* $V^{-1}(\kappa(m, n))$ *consists of all sequences obtained by choosing a positive integer* k *and replacing each* h *in* $\kappa(m, n)$ *by* $v^k h$.

We can combine H and V to form a reduction operator which by repeated application reduces any crossing sequence $\kappa(m, n)$ containing h's and v's to a power of v (by a computation equivalent to the Euclidean algorithm). The inverse of this reduction operator generates all crossing sequences by starting with the sequences v^k, $k > 0$ and alternately replacing all v's by $h^i v$ and all h's by $v^j h$. The generation procedure can be described by an indexed grammar. Consider the indexed grammar with terminals h, v, nonterminals S, T, U, V, H, index symbols p, q, r, and start symbol S. The productions are

$$S \to Tp \text{ or } Up \qquad T \to Tq \text{ or } Tr \text{ or } Uq \qquad U \to VU \text{ or } V$$

$$Hp \to h \qquad Hq \to H \qquad Hr \to VH$$

$$Vp \to v \qquad Vp \to HV \qquad Vr \to V.$$

For example, a derivation of the crossing sequence $\kappa(2, 3) = hvhhv$ is

$$\begin{aligned} &S \to Tp \to Tqp \to Trqp \to Uqrqp \to Vqrqp \to HrqpVrqp \to \\ &\quad VqpHqpVrqp \to VqpHqpVqp \to HpVpHqpVqp \to HpVpHpVqp \to \\ &\quad HpVpHpHpVp \to hVpHpHpVp \to hvHpHpVp \to hvhHpVp \to hvhhVp \to hvhhv. \end{aligned}$$

First we will show that for any derivation $S \xrightarrow{*} w \in \{h, v\}^*$, w is a crossing sequence. The derivation must begin $S \xrightarrow{*} Uxp$ where x is some word over $\{q, r\}$ and either $x = \varepsilon$ or x begins with q. The word x encodes the sequence of replacements mentioned above. The next step in the derivation will be $Uxp \to Vxp$ or $Uxp \to VxpUxp$. Since every U must be replaced by a V in order to reach w, there is no loss of generality in assuming that the derivation begins $S \xrightarrow{*} (Vxp)^m$ for some $m \geq 1$. Now the two lemmas below complete the proof of Theorem 3.10.

DEFINITION 3.16. $\mathscr{F}$ is the set of sentential forms α such that:

(1) There are no terminals in α, and each nonterminal is either a V or an H.

(2) The nonterminals spell out a crossing sequence (if each V is replaced by v and H by h).

(3) Every nonterminal is followed by the same index word xp with $x \in \{q, r\}^*$; and if $|x| > 0$, then x begins with q or r according to whether α begins with V or H.

LEMMA 3.17. *If $\alpha \xrightarrow{*} w$ and α is a sentential form in the set $\mathcal{F}$ defined above, then w is a crossing sequence.*

Proof. We argue by induction on $|x|$. If $|x| = 0$, then the only productions applicable are $Vp \to v$ and $Hp \to h$, and Definition 3.16(2) implies that w is a crossing sequence. Otherwise either

(1) α begins with V and $x = q^i x'$ where $i \neq 0$ and $x' = \varepsilon$ or x' begins with r; or

(2) α begins with H and $x = r^i x'$ where $i \neq 0$ and $x' = \varepsilon$ or x' begins with q.

These cases are similar, and we consider just the first. The only productions applicable are $Vq \to HV$, $Hq \to H$. As these productions must be applied to each factor Vxp or Hxp respectively at some point in the derivation, we may as well apply them to all factors at the same time to obtain $\alpha \xrightarrow{*} \alpha' \xrightarrow{*} w$ where α' is obtained from α by replacing each V by $H^i V$ and each x by x'. By Lemma 3.13 (Lemma 3.15 in the other case), $\alpha' \in \mathcal{F}$, and w is a crossing sequence by induction. □

LEMMA 3.18. *Every crossing sequence is derivable from S.*

Proof. Clearly every crossing sequence of the form v^i is derivable from S. By Lemmas 3.13 and 3.15 it suffices to show that

(1) If $S \xrightarrow{*} w = v \cdots$ and w' is obtained from w by replacing each v in w with $h^j v$, then $S \xrightarrow{*} w'$; and

(2) If $S \xrightarrow{*} w = h \cdots$ and w' is obtained from w by replacing each h in w with $v^j h$, then $S \xrightarrow{*} w'$.

We will prove the first assertion; the second is proved similarly. First, if $w = v^i$, then $S \xrightarrow{*} (Vq^j p)^i \xrightarrow{*} ((Hp)^j Vp)^i \xrightarrow{*} (h^j v)^i = w'$. Otherwise $w = v^i h \cdots$, and by the argument used in the proof of Lemma 3.17

$$S \xrightarrow{*} (Vxp)^m \xrightarrow{*} \alpha = Hr^i p \cdots \xrightarrow{*} \alpha' = V^i p H p \cdots \xrightarrow{*} w.$$

In particular $x = \cdots r^i p$. It follows that

$$S \xrightarrow{*} (Vxq^j p)^m \xrightarrow{*} Hr^i q^j p \cdots \xrightarrow{*} V^i q^j p H q^j p \cdots \xrightarrow{*} ((Hp^j) Vp)^i Hp \cdots \xrightarrow{*} w'. \qquad \square$$

Section 4: Bounded combings of nilpotent groups

This section is entirely devoted to the proof of the following theorem.

THEOREM A. *Let G be a virtually nilpotent group and let $\mu\colon \Sigma^* \twoheadrightarrow G$ be a choice of finite generating set for G. Suppose that the language $L \subseteq \Sigma^*$ is a combing of G that satisfies the asynchronous fellow-traveller property. If L is a context-free language, then G is virtually abelian.*

By Theorem 2.16 we may assume that G is nilpotent. Since finitely generated nilpotent groups have polynomial growth [W] and L is a combing, L must also have polynomial growth. In light of Proposition 1.3 we see that Theorem A is a special case of the following result, the geometry of which is explained in [BG].

THEOREM 4.1. *Let G be a nilpotent group and let $\mu: \Sigma^* \twoheadrightarrow G$ be a choice of generators for G. Suppose that $L \subseteq \Sigma^*$ satisfies the asynchronous fellow-traveller property and that $\mu|_L$ is surjective. If L is a bounded language then G is virtually abelian.*

LEMMA 4.2. *Let G be a finitely generated nilpotent group.*
(1) *Every subgroup of G is finitely generated and nilpotent.*
(2) *The set of torsion elements of G is a finite normal subgroup.*
(3) *For any integer $n \geq 1$, the subgroup of G generated by all n-th powers of elements of G is of finite index in G.*
(4) *If G is of class two (i.e., if the commutator subgroup lies in the center), then for any $g \in G$ the map $h \rightarrow [g, h] := g^{-1}h^{-1}gh$ is a group homomorphism from G to the commutator subgroup.*

Proof. These properties are well known and can be proved by induction on the nilpotence class [Ha]. □

LEMMA 4.3. *Let G be a finitely generated torsion free nilpotent group.*
(1) *The central quotient $G/Z(G)$ is torsion free.*
(2) *For any $g \in G$ and integer $m \neq 0$ the centralizers $C(g)$ and $C(g^m)$ are equal.*
(3) *If $g \in G - \{1\}$ then the set of all roots of g in G generates a cyclic subgroup.*
(4) *Suppose $g_1, g_2 \in G$ and each g_i has no proper roots in G. If $hg_1^m h^{-1} = g_2^n$ for some positive integers m, n, then $hg_1h^{-1} = g_2$.*
(5) *If G is virtually abelian, then it is abelian.*

Proof. Part (1) is [Ba, Corollary 2.11]. For (2) one simply casts $C(g^m)$ in the role of G in part (1) and observes that if the inclusion $C(g) \subseteq C(g^m)$ were strict then g would not be in the centre of $C(g^m)$, whereas g^m is. (3) follows easily, because (2) implies that all of the roots of g lie in the centre of $C(g)$, which is a finitely generated abelian group. (4) follows easily from (3). Finally, if G were virtually abelian, it would have a normal abelian subgroup A of finite index n. Thus $g^n \in A$ for all $g \in G$ whence $A \subseteq Z(G)$ by (2). Now (1) yields $A = G$. □

LEMMA 4.4. *A finitely generated nilpotent group H is either virtually abelian or has a quotient G such that:*
(1) *G is torsion free.*
(2) *$G/Z(G)$ is free abelian of rank at least two.*
(3) *If $h \in G - Z(G)$ then $C(h)$ is a normal subgroup of finite index in G.*

(4) *Suppose* $g_1, g_2 \in G$ *and each* g_i *has no proper roots in* G. *If* $hg_1^m h^{-1} = g_2^n$ *for some positive integers* m, n *with* $m \neq 0 \neq n$, *then* $C(g_1) = C(g_2)$.

(5) *For any choice of generators* $\mu\colon \Sigma^* \twoheadrightarrow G$ *there is a constant* P *such that for all* $w \in \Sigma^*$ *if* $\bar{w} \notin Z(G)$ *then* $d(1, \bar{w}^n) \geq n/P$.

Proof. Assume H is nilpotent but not virtually abelian. By standard arguments one can show that H modulo its torsion subgroup T is not virtually abelian. Thus we assume H that is torsion free. By Lemma 4.3 $H/Z(H)$ is torsion free, and if it is virtually abelian then it is abelian. This observation applies equally to the successive quotients of H by terms in its ascending central series. Let Z_n be the last term in the ascending central series such that H/Z_n is not abelian (i.e., $H = Z_{n+2}$) and define $G = H/Z_n$. Notice that G is not virtually abelian but $G/Z(G)$ is; in other words G has class 2. $G/Z(G)$ is torsion free and hence a free abelian group. If $G/Z(G)$ were cyclic, then G would be abelian. Thus (1) and (2) hold. It follows from (2) that all centralizers in G are normal subgroups. If $g \in G - Z(G)$ then its centralizer could not have finite index n in G or else $h^n \in C(g)$ for all $h \in G$ would imply $g \in Z(G)$ by Lemma 4.3(2). (3) implies (4) because $C(g_1)$ and $C(g_2)$ are conjugate (Lemma 4.3(4)) and normal.

Finally, if $v\colon \Lambda^* \twoheadrightarrow A$ is any set of free generators for the free abelian group A, then d_v is the restriction of the corresponding ℓ_1-norm on the real vector space with basis Λ, and hence $d_A(1, v(w^n)) \geq n$ for any $w \in \Lambda^*$ with $v(w) \neq 1$. Thus, by (2.1), any choice of generators affords a constant P such that $d_A(1, v(w^n)) \geq n/P$. Setting $A = G/Z(G)$ now yields (5). □

Proof of Theorem 4.1. Suppose G and L are as in the statement of Theorem 4.1, and assume that G is not virtually abelian. We will argue by contradiction. Observe that the hypothesis of Theorem 4.1 is valid in all quotients, thus we may assume that G satisfies the conclusions of Lemma 4.4. Because L is a bounded language, Lemma 1.4 allows us to write it as the union of finitely many sublanguages L_λ of the form

$$L_\lambda = \{w \mid w = v_{\lambda,0} u_{\lambda,1}^{n_1} v_{\lambda,1} \cdots u_{\lambda,r}^{n_r} v_{\lambda,r} \,|(n_1, \ldots, n_r) \in S_\lambda\}, \tag{4.1}$$

where r is an integer depending on λ and either $r = 0$ and S_λ is empty or $S_\lambda \subseteq \mathbb{N}^r$ contains r-tuples whose smallest entry is arbitrarily large.

Now we make some changes in L_λ which do not alter the image of L_λ in G. Nor do these changes alter the fact that L_λ is bounded, or the fact that S_λ contains r-tuples with all entries arbitrarily large (we say nothing about whether L_λ remains context free). If a word $u_{\lambda,j}$ has a trivial image in G, we delete it from (4.1); and otherwise if $\overline{u_{\lambda,j}}$ has proper roots in G, we invoke Lemma 4.3 to replace $u_{\lambda,j}$ by a word whose image in G is a root of $\overline{u_{\lambda,j}}$ but has no proper roots itself. In other words, without loss of generality we may assume that every $u_{\lambda,j}$ occurring in (4.1) has nontrivial image in G and that each $u_{\lambda,j}$ projects to an element of G with no proper roots in G.

In what follows, when referring to $w \in L_\lambda$ we shall assume that it is decomposed as in formula (4.1). Two words $u_{\lambda,j}, u_{\lambda,j+j'}$ in that decomposition may have the same or different images in $G/Z(G)$. We shall focus our attention on some $L_\lambda \subseteq L$ with the maximal number of sequences $u_{\lambda,j}, u_{\lambda,j+1} \cdots u_{\lambda,j+j'}$ such that $u_{\lambda,j}$ and $u_{\lambda,j+j'}$ have distinct and nontrivial images in $G/Z(G)$ and all intervening u's lie in $Z(G)$. There must be such sequences because otherwise (4.1) and the fact that L projects onto G would force $G/Z(G)$ to be a union of a finite number of cyclic subgroups contrary to Lemma 4.4(2).

The intuition behind the somewhat strange sequences considered above is as follows. Given a word w for which all of the exponents n_j in (4.1) are large, one can imagine this as a path in the Cayley graph of $G/Z(G)$ consisting of a sequence of long straight segments (corresponding to iterates of those $u_{\lambda,j}$ which project non-trivially) joined by short segments (corresponding to the $v_{\lambda,j}$). If two successive $u_{\lambda,j}$ were to have the same non-trivial image in $G/Z(G)$ then a distant observer would see no 'corner' between the corresponding straight segments, and the above notion of maximality among the sublanguages L_λ is based (roughly speaking, but not exactly) on counting the number of perceivable corners in this sense.

For convenience we assume that the sublanguage L_λ which is maximal with respect to the above count of special subsequences $u_{\lambda,j}, u_{\lambda,j+1} \cdots u_{\lambda,j+j'}$ is given by $\lambda = 0$, and we fix r to be the corresponding integer from (4.1). Let r' denote the maximum index for which $u_{0,r'}$ has a nontrivial image in $G/Z(G)$.

Before proceeding to the proof of Theorem 4.1 we pause to outline our strategy. Our argument shall proceed by judicious choices of λ,

$$w = v_0 u_1^{n_1} v_1 \cdots u_r^{n_r} v_r \in L_0 \quad \text{and} \quad x = y_0 z_1^{m_1} y_1 \cdots z_s^{m_s} y_s \in L_\lambda \tag{4.2}$$

(where the factorisations are as in (4.1) but with the notation simplified slightly), the exponents n_j will be large and the images of w and x will be a controlled distance Q apart in G. The idea of the proof is to make the paths in G determined by w and x close over a distance which is very large compared to Q; so close in fact that certain subwords are forced to be conjugate. We then show that these conjugacy relations force the images of these paths in $\hat{G} := G/C(\overline{u_r})$ to be uniformly close (independent of Q) and taking Q to be large yields a contradiction.

More precisely, the hypothesis that L satisfies the asynchronous fellow-traveller property will yield a bound on $D(w, x)$ (a bound which depends on Q), and we shall use this bound to show that subwords $u_j^{n_j}$ of w with $\overline{u_j} \notin Z(G)$ have a substantial overlap with subwords $z_k^{m_k}$ of x. By this we mean that large segments of the subwords $u_j^{n_j}$ and $z_k^{m_k}$ correspond under the distance minimizing monotone reparametrizations used in the definition of $D(w, x)$. In other words, if w is written as $w = \underline{w_1 u_j^{n_j} w_2}$ and x as $x = x_1 z_k^{m_k} x_2$, then for integers i lying in a large interval we have $d(\overline{w_1 u_j^i}, \overline{x_1 z_k^i}) \le D(w, x) + K'$, where the constant K' depends only on L. We shall choose w so as to ensure that this interval is long enough to allow a word-difference argument of the type found in [E+], which forces a power of $\overline{u_j}$ to be conjugate in G to a power of $\overline{z_k}$. We shall then use the control over conjugation relations yielded by Lemma 4.2(4) and Lemma 4.4(4) to obtain an estimate on the distance between the images of w and x in $\hat{G}$, and this estimate will contradict the choice of Q.

In order to implement this line of attack on Theorem 4.1 we must first dispense with some preparatory technicalities. Consider $\hat{G} = G/C(\overline{u_{r'}})$ with the generators obtained by projecting the generators Σ of G, and let $\hat{d}$ denote the corresponding word metric. It is important to note that $\hat{d}$ is invariant with respect to the natural left action of G on $\hat{G}$. We will abuse notation to the extent of allowing $\hat{d}(g, h)$ to stand for the distance between the images of $g, h \in G$ in $\hat{G}$. Clearly $\hat{d}(g, h) \le d(h, g)$, and for any positive integer Q the ball of radius Q around $1 \in G$ projects onto the ball of the same radius around $1 \in \hat{G}$.

The following is an immediate consequence of Lemma 4.3(4):

LEMMA 4.5. *There exists a constant A such that if $h^{-1}\overline{u_{r'}}^m h = \overline{u_{\lambda,j}}^n$ for any positive integers m, n, any λ, and any $u_{\lambda,j}$ occurring in* (4.1), *then* $\hat{d}(1, h) \le A$.

In what follows we shall need the following additional constants B, K, P, Q, S, which we record here for the convenience of the reader.

$B := \max_\lambda |v_{\lambda,0} u_{\lambda,1} v_{\lambda,1} \cdots u_{\lambda,r} v_{\lambda,r}|$.

If $w_1, w_2 \in L$ with $d(\overline{w_1}, \overline{w_2}) \le 1$ then $D(w_1, w_2) \le K$.

If $w \in \Sigma^*$ with $\bar{w} \notin Z(G)$, then $d(1, \overline{w^n}) \ge n/P$.

$Q := A + 2B + 1$.

$S := |\Sigma|$.

The constant K exists because L satisfies the asynchronous fellow-traveller property, and P comes from Lemma 4.4.

Proceeding with the proof of Theorem 4.1, we now fix $w \in L_0$ with all exponents n_j satisfying

$$n_j \ge P(B^2 S^{QK+B} + 2B + 2QK). \tag{4.3}$$

Because $\hat{G}$ is infinite (Lemma 4.4(3)), we can choose $x \in L$ with $d(\bar{w}, \bar{x}) = \hat{d}(\bar{w}, \bar{x}) = Q$. We assume that w and x are decomposed as in equation (4.2) and henceforth use the notation established there.

LEMMA 4.6. *There is a partial map $\sigma\colon \{1, \ldots, r\} \to \{1, \ldots, s\}$ such that*
(1) *the domain of σ is all j with $\overline{u_j} \notin Z(G)$;*
(2) *σ is nondecreasing;*
(3) *$\overline{u_j}$ is conjugate in G to $\overline{z_{\sigma(j)}}$ for all j in the domain of σ.*

Proof. By the triangle inequality for D we have that $D_G(w, x) \leq QK$. Consider a subword $u_j^{n_j}$ of w from the decomposition (4.2) such that $u_j \notin Z(G)$, and write $w = w_1 w_2 w_3$ with $w_2 = u_j^{n_j}$. It follows from the definition of the pseudometric D that x factors as $x = x_1 x_2 x_3$ with $d(\overline{w_1}, \overline{x_1}) \leq D(w, x) < KQ$ and $d(\overline{w_1 w_2}, \overline{x_1 x_2}) \leq QK$.

The last two inequalities imply that $d(1, \overline{x_2}) \geq d(1, \overline{w_2}) - 2KQ \geq n_j/P - 2KQ \geq B^2 S^{QK+B} + 2B$. As no subword y_k or z_k is longer than B, it must be that x_2 involves at least $BS^{QK+B} + 2$ subwords y_k and z_k from (4.2). It follows that x_2 itself contains a subword consisting of a product of at least BS^{QK+B} of the y_k and z_k. Since the definition of B implies that the parameter s of (4.2) is no greater than B, one such z_k must occur to a power at least S^{QK+B}. In other words $x_2 = x_4 z_k^{m_k} x_5$ with $m_k \geq S^{QK+B}$ for some k. We define $\sigma(j) = k$. This definition yields the first two assertions of Lemma 4.5.

Now, it follows from the monotonicity of the reparametrizations in the definition of D that the image in G of each word $x_1 x_4 z_{\sigma(j)}^m$, $0 \leq m \leq S^{QK+B}$ is a distance at most QK from the image of some word of the form $w_1 u_j^n u'$, where u' is an initial segment of u; and n depends in a monotonically nondecreasing fashion on m. Consequently the image in G of $x_1 x_4 z_{\sigma(j)}^m$ is a distance at most $QK + B$ from the image of $w_1 u_j^n$. For each $x_1 x_4 z_{\sigma(j)}^m$ we fix a shortest path in the Cayley graph realizing this distance. The word labelling this path has length at most $QK + B$, and there are fewer than S^{QK+B} possibilities for such a label. Thus two of these paths have the same label.

Let us suppose that the label ℓ_j is shared by the paths connecting the images of $x_1 x_4 z_{\sigma(j)}^m$ and $x_1 x_4 z_{\sigma(j)}^{m+m'}$ with those of $w_1 u_j^n$ and $w_1 u_j^{n+n'}$ respectively. Since these paths are distinct, $m' > 0$, whence $n' \geq 0$. In fact, $n' > 0$ lest $z_{\sigma(j)}$ have trivial image in G. Thus in G we have the equation $\overline{\ell_j u_j^{n'} \ell_j^{-1}} = \overline{z_{\sigma(j)}^{m'}}$. It follows from Lemma 4.3(4) that $\overline{u_j}$ is conjugate to $\overline{z_{\sigma(j)}}$. □

COROLLARY 4.7. *If $k > \sigma(r')$, then either $z_k \in Z(G)$, or else z_k has the same image as $z_{\sigma(r')}$ in $G/Z(G)$.*

Proof. With the notation of Lemma 4.6, $\overline{u_j}$ and $\overline{z_{\sigma(j)}}$ have the same image in $G/Z(G)$. It follows that the number of subsequences $u_{i,j}, u_{i,j+1} \cdots u_{i,j+j'}$ of the type governing the choice of L_0 which occur in w is no less than the number of such subsequences which occur in x up to and including $z_{\sigma(r')}^{n_{\sigma(r')}}$. By the choice of L_0 there are no more such subsequences in x. □

For the final stage of the proof we fix $s' = \sigma(r')$. Recall that $\overline{u_{0,r'}} \notin Z(G)$ but $\overline{u_{0,j}} \in Z(G)$ for all $j > r'$. The paths in the Cayley graph of G which occurred with label $\ell_{r'}$ in the proof of Lemma 4.6, connect the images in G of $x_1 x_4 z_{s'}^m$ and $x_1 x_4 z_{s'}^{m+m'}$ with those of $w_1 u_{r'}^n$ and $w_1 u_{r'}^{n+n'}$ respectively. Because $\overline{u_j} \in Z(G) \subseteq C(\overline{u_{r'}})$ for all $j > r'$, we have:

$$\begin{aligned}\hat{d}(\overline{w_1 u_{r'}^n}, \bar{w}) &= \hat{d}(\overline{w_1 u_{r'}^n}, \overline{w_1 u_{r'}^{n_{r'}} v_{r'} u_{r'+1}^{n_{r'+1}} \cdots v_r})\\ &= \hat{d}(1, \overline{u_{r'}^{n_{r'}-n} v_{r'} u_{r'+1}^{n_{r'+1}} \cdots v_r})\\ &\le \hat{d}(1, \overline{v_{r'} v_{r'+1} \cdots v_r})\\ &\le d(1, \overline{v_{r'} v_{r'+1} \cdots v_r}) \le B. \end{aligned} \tag{4.4}$$

Here we have used the fact that $\hat{d}$ is invariant under left translation by G.

By Lemma 4.6, $\overline{z_{s'}}$ is conjugate to $\overline{u_{r'}}$, and hence by Lemma 4.4(4), $\overline{z_{s'}} \in C(\overline{u(r')})$. It then follows from Corollary 4.7 that $\overline{z_k}$ is also in $C(\overline{u(r')})$ for $k > s'$. Thus

$$\begin{aligned}\hat{d}(\overline{x_1 x_4 z_{s'}^m}, \bar{x}) &= \hat{d}(\overline{x_1 x_4 z_{s'}^m}, \overline{x_1 x_4 z_{s'}^{m_{s'}} y_{s'} z_{s'+1}^{m_{s'+1}} \cdots y_s})\\ &= \hat{d}(1, \overline{z_{s'}^{m_{s'}-m} y_{s'} z_{s'+1}^{m_{s'+1}} \cdots y_s})\\ &= \hat{d}(1, \overline{y_{s'} \cdots y_s})\\ &\le \hat{d}(1, \overline{y_{s'} \cdots y_s}) \le B. \end{aligned} \tag{4.5}$$

Finally, as there is a path in the Cayley graph of G from $\overline{x_1 x_4 z_{s'}^m}$ to $\overline{w_1 u_{r'}^n}$ with label $\alpha = (z_1 x_4 z_{s'}^m)^{-1} w_1 u_{r'}^n$ (note that $\bar{\alpha} = \ell_{r'}$) it follows that $\bar{\alpha}$ conjugates $\overline{z_{s'}}$ to $\overline{u_{r'}}$. Lemma 4.5 yields:

$$\hat{d}(\overline{x_1 x_4 z_k^t}, \overline{w_1 u_r^s}) = \hat{d}(1, \bar{\alpha}) \le A.$$

Combining inequalities (4.4), (4.5) and (4.6) with the triangle inequality for $\hat{d}$ we see that $Q = \hat{d}(\bar{w}, \bar{x}) \le A + 2B$, which contradicts our choice of Q. Thus we have successfully argued *ad absurdum*, and we conclude that if G satisfies the hypotheses of Theorem 4.1 then it must be virtually abelian. □

REFERENCES

[A] A. V. AHO, *Indexed grammars – an extension of context-free grammars*, J. Assoc. Comp. Mach. *15* (1968), 647–671.

[AB] J. M. ALONSO and M. R. BRIDSON, *Semihyperbolic groups*, Proc. London Math. Soc. *70* (1993), 56–114.

[B1] M. R. BRIDSON, *On the geometry of normal forms in discrete groups*, Proc. London Math. Soc. (3) *67* (1993), 596–616.

[B2] M. R. BRIDSON, *Combings of semidirect products and 3-manifold groups*, J. Geom. and Funct. Anal. (GAFA) *3* (1993), 263–278.

[B3] M. R. BRIDSON, *Regular combings, nonpositive curvature and the quasiconvexity of abelian subgroups*, J. Pure and Appl. Algebra *88* (1993), 23–35.

[Ba] G. BAUMSLAG, *Lecture Notes on Nilpotent Groups*, Conf. Board of the Math. Sci. Regional Conf. Series in Math., vol. 2, Amer. Math. Soc., Providence, RI, 1971.

[BGSS] G. BAUMSLAG, S. GERSTEN, M. SHAPIRO and H. SHORT, *Automatic groups and amalgams*, J. Pure and Appl. Algebra *76* (1991), 229–316.

[BG] M. R. BRIDSON and R. H. GILMAN, *On bounded languages and the geometry of nilpotent groups*, in *Combinatorial and Geometric Group Theory (Edinburgh, 1993)*, LMS Lecture Notes (204) Cambr. Univ. Pr. 1995, 1–15.

[Br] N. BRADY, *The geometry of asynchronous automatic structures*, Ph.D. Thesis, UC Berkeley, 1993.

[C] J. CANNON, *The combinatorial structure of cocompact discrete hyperbolic groups*, Geom. Dedicata *16* (1984), 123–148.

[CJ] A. CASSON and D. JUNGREIS, *Convergence groups and Seifert fibered 3-manifolds*, Invent. Math. *118* (1994), 441–456.

[E+] 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, Boston, MA, 1992.

[G] M. GROMOV, *Hyperbolic Groups*, in *Essays in group theory*, S. M. Gersten (ed.), MSRI Publ. vol. 8, Springer Verlag, New York, 1987, 75–263.

[Ga] DAVID GABAI, *Convergence groups are Fuchsian groups*, Annals of Mathematics *136* (1992), 447–519.

[GS] S. GINSBURG and E. H. SPANIER, *Bounded Algol-like languages*, Trans. A.M.S. *113* (1964), 333–368.

[GSh] S. M. GERSTEN and H. B. SHORT, *Rational subgroups of biautomatic groups*, Ann. of Math. (2) *134* (1991), 125–158.

[H] J. HEMPEL, *3-manifolds*, Annals of Math. Studies 86, Princ. Univ. Press, Princeton, NJ, 1976.

[Ha] MARSHALL HALL, JR., *The Theory of Groups*, Macmillan, New York, 1959.

[HU] J. E. HOPCROFT and J. D. ULLMAN, *Introduction to Automata Theory, Languages and Computation*, Addison-Wesley, Reading, MA, 1979.

[JS] W. JACO and P. SHALEN, *Seifert fibred spaces in 3-manifolds*, Memoirs of the AMS 220, AMS, Providence, RI, 1979.

[J] K. JOHANNSON, *Homotopy equivalences of 3-manifolds with boundary*, Springer Lecture Notes in Math. *761*, Springer Verlag, Berlin, 1979.

[K] H. KNESER, *Geschlossen Flächen in dreidimensionalen Mannigfaltigkeiten*, Jahr. der Deustcher math. Ver. *38* (1929), 248–260.

[Lo] M. LOTHAIRE, *Combinatorics on Words*, Addison-Wesley, Reading, MA, 1983.

[Mi] J. W. MILNOR, *A unique factorization theorem for 3-manifolds*, Amer. J. Math. *84* (1962), 1–7.

[Mo] J. MORGAN, *On Thurston's uniformization conjecture for three dimensional manifolds*, in *The Smith conjecture*, H. Bass and J. Morgan (eds.), Academic Press, Orlando, Fla., 1984.

[MS] D. E. MULLER and P. E. SCHUPP, *Groups, the theory of ends and context-free languages*, J. Computer and System Sciences *26* (1983), 295–310.

[S] G. P. SCOTT, *The geometry of 3-manifolds*, Bull. LMS *15* (1983), 401–487.

[Sha] M. SHAPIRO, *Automatic structure and graphs of groups*, in *Topology '90*, Proceedings of the research semester in low-dimensional topology at Ohio State, de Gruyter Verlag, 1992.

[T1] W. P. THURSTON, *The geometry and topology of 3-manifolds*, preprint, Princeton University, 1991.

[T2] W. P. THURSTON, *Three dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. AMS *6* (1982), 357–381.

[W] J. A. WOLF, *Growth of finitely generated solvable groups and curvature of Riemannian manifolds*, J. Diff. Geom. *2* (1968), 421–446.

Martin R. Bridson
Department of Mathematics, Princeton University, Princeton, NJ 08544
Current address: Mathematical Institute, 24–29 St. Giles, Oxford, OX1 3LB
E-mail address: bridson@math.princeton.edu

Robert H. Gilman
Department of Mathematics, Stevens Institute of Technology, Hoboken, NJ 07030
E-mail address: rgilman@stevens-tech.edu
Homepage: http://attila.stevens-tech.edu/~rgilman

Received April 17, 1994