The generalised word problem for subgroups of hyperbolic groups

Derek F. Holt and Sarah Rees

D. F. Holt, Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

E-mail address: D.F.Holt@warwick.ac.uk

S. Rees, School of Mathematics and Statistics, University of Newcastle, Newcastle NE1 7RU, UK

E-mail address: Sarah.Rees@newcastle.ac.uk

Abstract

We prove that the generalised word problem of a finitely generated subgroup of a finitely generated virtually free group is context-free, that a hyperbolic group must be virtually free if it has a torsion-free quasiconvex subgroup of infinite index with context-free generalised word problem, and that, for any hyperbolic group, the generalised word problem of a torsion-free quasiconvex subgroup is recognised by a real-time Turing machine.

2010 Mathematics Subject Classification: 20F10
Key words: generalised word problem, hyperbolic group, context-free language, real-time Turing machine

1 Introduction

Let $G = \langle X \rangle$ with $|X| < \infty$ be a group and $H \leq G$. The *word problem* $\mathrm{WP}(G, X)$ and *generalised word problem* $\mathrm{GWP}(G, H, X)$ are defined to be the preimages $\phi^{-1}(\{1_G\})$ and $\phi^{-1}(H)$ respectively, where ϕ is the natural map from the set of words over X to G. We are interested in the relationship between the algebraic properties of G (and H) and the formal language classes containing $\mathrm{WP}(G, X)$ and $\mathrm{GWP}(G, H, X)$. These questions have already been well studied for the word problem, but relatively little for the generalised word problem. Since, as is well known for $\mathrm{WP}(G, X)$, the question of membership of $\mathrm{WP}(G, X)$

and $\mathrm{GWP}(G,H,X)$ in a formal language family $\mathcal{F}$ is typically independent of the choice of X, we shall usually use the simpler notations $\mathrm{WP}(G)$ and $\mathrm{GWP}(G,H)$.

It is elementary to prove that $\mathrm{WP}(G)$ is regular if and only if G is finite, and well known that $\mathrm{WP}(G)$ is context-free if and only if G is virtually free [10]. It is similarly easy to show that $\mathrm{GWP}(G,H)$ is regular if and only if $|G:H|$ is finite, but conditions for $\mathrm{GWP}(G,H)$ to be context-free have not been much studied. Our first two results address this.

Theorem 1. *Let G be finitely generated and virtually free, and let H be a finitely generated subgroup of G. Then $\mathrm{GWP}(G,H)$ is deterministic context-free*

The conclusion of Theorem 1 may or may not hold if we drop the condition that H is finitely generated. Suppose that G is the free group on two generators a,b. For $H_1 = [G,G]$, the set $\mathrm{GWP}(G,H_1)$ consists of all words whose exponent sums in both a and b are zero, and is not context-free, but for the subgroup H_2 of words whose exponent sum in a is zero, the set $\mathrm{GWP}(G,H_2)$ is context-free.

We would like to know to what extent this result is best possible when H is finitely generated. We observe that, where $H_G := \cap_{g\in G} H^g$ is the *core* of H in G, the set $\mathrm{GWP}(G,H,X)$ is the same set of words as $\mathrm{GWP}(G/H_G, H/H_G, X)$. We know of no examples for which H is finitely generated with trivial core, and $\mathrm{GWP}(G,H)$ is context-free but G is not virtually free.

The following theorem is an attempt at a converse to Theorem 1.

Theorem 2. *Let G be a hyperbolic group, and let H be a quasiconvex subgroup of infinite index in G such that $|C_G(h) : C_H(h)|$ is finite for all $1 \neq h \in H$. If $\mathrm{GWP}(G,H)$ is context-free then G is virtually free.*

Our proof is dependent on the quasiconvexity of H (which implies its finite generation), and on the condition on centralisers, but we conjecture that neither condition is necessary for the result to hold. Note that the centraliser condition holds whenever H is torsion-free.

It is proved in [6] that the word problem of a hyperbolic group is the language of a real-time Turing machine. We prove an analogous result for the generalised word problem.

Theorem 3. *Let G be a hyperbolic group, and let H be a quasiconvex subgroup such that $|C_G(h) : C_H(h)|$ is finite for all $1 \neq h \in H$. Then $\mathrm{GWP}(G,H)$ is the language of a real-time Turing machine.*

In this case too we conjecture that the condition on centralisers is not necessary for the conclusion of the theorem. However the necessity of quasiconvexity is demonstrated by a construction found in [11], as follows. Let Q be a finitely presented group, and choose $\lambda > 0$. We can define a finitely presented group G and a normal 2-generated subgroup H of G, such that $G/H \cong Q$, and G satisfies the small cancellation condition $C'(\lambda)$; in particular we can choose Q with insoluble word problem, and choose $\lambda \leq 1/6$ to ensure G hyperbolic, and in that case $\mathrm{GWP}(G,H)$ is not even recursive.

2 Proof of Theorem 1

Let F be a free subgroup of G with $|G : F|$ finite, and let Y be a free generating set for F. Let $A = X \cup X^{-1}$, $B = Y \cup Y^{-1}$, and let $K = F \cap H$. The subgroup K, like H, is finitely generated. Let $\{t_1, \ldots, t_m\}$ be a right transversal of K in H and extend this to a right transversal $T = \{t_1, \ldots, t_n\}$ of F in G. Then any word w in A^* can be expressed (in G) as a word in C^*T, where C is the set of Schreier generators $u(i, a)$ for F defined by the equations $t_i a = u(i, a)t_j$, and hence as a word vt in B^*T.

The first step to recognise whether $w \in H$ is to rewrite it to the form vt, as above, using a transducer. Then $w \in H$ if and only if $t \in \{t_1, \ldots, t_m\}$ and $v \in K$. It remains for us to describe the operation of a deterministic pushdown automaton (pda) N to recognise those words over Y^* that lie in the subgroup K of the free group F. Note that this machine operates simultaneously, rather than sequentially, with the transducer, and it follows from the fact that context-free languages are closed under inverse gsms [8, Example 11.1, Theorem 11.2] that the combination of the two machines is a pda.

By [1] or [4, Proposition 4.1], any finitely generated subgroup K of a free group is L-*rational*, where L is the set of freely reduced words over a free generating set; that is, the set $L \cap K$ is a regular language.

We shall build our pda N out of a finite state automaton (fsa) M for which $L(M) \cap L = K \cap L$, which is (in effect) described in the proof of [4, Theorem 2.2] that K is L-quasiconvex.

The fsa M has the following properties.

(i) The states of M are denoted by $\sigma_1, \ldots, \sigma_m, \hat{\sigma}$.

(ii) $\sigma_1, \ldots, \sigma_m$ are the elements of a finite subset S of the set of right cosets Kg_i of K in F, including $\sigma_1 = K = Kg_1$ with $g_1 = 1$, and σ_1 is the start state and the single accepting state.

(iii) For $1 \leq i, j \leq m$ and $a \in A$, there is a transition $\sigma_i^a = \sigma_j$ if and only if $Kg_i a = Kg_j$. It follows from this that $\sigma_i^a = \sigma_j$ if and only if $\sigma_j^{a^{-1}} = \sigma_i$.

(iv) $\hat{\sigma}$ is a failure state, and is the target of all transitions that are not defined in (iii), including those from $\hat{\sigma}$.

(v) A word $w \in L$ lies in K if and only if it is in $L(M)$.

So $M \setminus \{\hat{\sigma}\}$ is a finite subgraph of the Schreier graph of K in F. The L-quasiconvexity condition is equivalent to the property that all prefixes of reduced words that represent elements of K lie within a bounded distance of K in the Schreier graph. It is a consequence of this that a word in L is accepted by M if and only if it represents an element of K. Note that $L(M)$ may contain some words not in L, but only contains words that represent elements of K.

Our pda N has the same states $\sigma_1, \ldots, \sigma_n, \hat{\sigma}$ as M, with σ_1 as the start state and sole accepting state. The transitions from the states σ_i are as in M. We need however to describe the operation of the stack, and transitions from the state $\hat{\sigma}$, which is non-accepting, but no longer a failure state.

The stack alphabet is the set $A \cup (A \times S)$. The second component of an element of $A \times S$ is used to record the state M is in immediately before it enters the state $\hat{\sigma}$. In addition, we use the stack to store the free reduction of the prefix of w that has been read so far.

The operations of the pda N that correspond to the various transitions of M are described in the following table. The absence of an entry in the 'push' column indicates that nothing is pushed.

Transition of M	Input state	Input symbol	Pop	Push	Output state
$\sigma_i^a = \sigma_j$	σ_i	a	a^{-1}		σ_j
	σ_i	a	$b \neq a^{-1}$	ba	σ_j
$\sigma_i^a = \hat{\sigma}$	σ_i	a	b	$b(a, \sigma_i)$	$\hat{\sigma}$
$\hat{\sigma}^a = \hat{\sigma}$	$\hat{\sigma}$	a	(a^{-1}, σ_i)		σ_i
	$\hat{\sigma}$	a	$(b, \sigma_i), \quad b \neq a^{-1}$	$(b, \sigma_i)a$	$\hat{\sigma}$
	$\hat{\sigma}$	a	a^{-1}		$\hat{\sigma}$
	$\hat{\sigma}$	a	$b \neq a^{-1}$	ba	$\hat{\sigma}$

Note that we have not specified that $b \neq a^{-1}$ in line 3, but in fact the condition $b = a^{-1}$ does not arise in this situation. Since there can be a symbol in $A \times S$ on the stack only when N is in state $\hat{\sigma}$, it is not possible to pop such a symbol when N is in state σ_i, so there are no such entries in the table.

The fact that $L(N) = \mathrm{GWP}(F, K)$ follows from the fact that $w \in L(N)$ if and only if $\bar{w} \in L(M)$, where $\bar{w}$ is the free reduction in L of w. We prove this by induction on the number k of reductions of the form $w_1 a a^{-1} w_2 \to w_1 w_2$ with $a \in A$ that we need to apply to reduce w to $\bar{w}$.

The case $k = 0$ of our induction follows from Property (v) of M and the observation that, if w is freely reduced, then w leads to the same state of M as it does of N. For in that case the only possible transitions as we read w are of the types described in lines 2,3,5,7 of the table.

For $k > 0$ it is enough to prove the statement

> $(*)$: if $wa \in L$, then the configuration of N after reading waa^{-1} is identical to the configuration after reading w.

It follows from $(*)$ that a word $w_1 a a^{-1} w_2$ in which aa^{-1} is the leftmost cancelling pair is in $L(N)$ if and only if $w_1 w_2$ is in $L(N)$, and hence we have the inductive step we need.

We can check the statement $(*)$ with reference to the table. There are up to seven possibilities for the type of transition of N as the final symbol a of wa is read.

For the first two of these, N is in state σ_i after reading w, and moves to a state σ_j. Since wa is in L, the top stack symbol after reading w is not a^{-1}. Hence the transition must be of the type described in line 2 of the table, and not as in line

1. That is, $b \neq a^{-1}$ is popped, and then ba is pushed. Recalling that $\sigma_i^a = \sigma_j$ in M if and only if $\sigma_j^{a^{-1}} = \sigma_i$, we see that the next transition of N, from wa on a^{-1}, is of the type described in line 1. Then the symbol a is popped, the symbol b is again on the top of the stack. and N returns to the state σ_i

We consider similarly the remaining five possibilities for the transition from w on a, and the subsequent transitions on a^{-1}, and verify $(*)$ for each of those configurations. This completes the proof of Theorem 1.

3 Proof of Theorem 2

For a finite generating set X of a group G, we define an *X-graph* to be a graph with directed edges labelled by elements of $A := X \cup X^{-1}$, in which, for each vertex p and each $a \in A$, there is a single edge labelled a with source p and, if this edge has target q, then there is an edge labelled a^{-1} from q to p. So Cayley and Schreier graphs $\mathcal{G}(G,X)$ and $\mathcal{S}(G,H,X)$ are examples of X-graphs. We shall denote the base points of Cayley and Schreier graphs by $1_{\mathcal{C}}$ and $1_{\mathcal{S}}$ respectively.

Following [3, Chapter 4], for $k \in \mathbb{N}$, We define the condition GIB(k) for $\mathcal{S} := \mathcal{S}(G,H,X)$ as follows.

> GIB(k): there exists $K \in \mathbb{N}$ such that, for any vertex p of $\mathcal{S}$ with $d(1_{\mathcal{S}}, p) \geq K$, the closed k-ball $B_k(p)$ of $\mathcal{S}$ is X-graph isomorphic to the k-ball $B_k(1_{\mathcal{C}})$ of $\mathcal{G}(G,X)$.

We say that $\mathcal{S}$ has GIB(∞) if it has GIB(k) for all $k \geq 0$. The following result is proved in [3, Theorem 4.3.1.1].

Proposition 3.1. *Let H be a quasiconvex subgroup of the hyperbolic group G. Then $\mathcal{S}(G,H,X)$ has* GIB(∞) *if and only if, for all $1 \neq h \in H$, the index $|C_G(h) : C_H(h)|$ is finite.*

This applies to groups G, H satisfying the hypotheses of Theorem 2. Since its proof may not be readily available, we shall sketch its proof in the final section of the paper.

Our proof of Theorem 2 has the same structure as the proof in [10] that groups with WP(G) context-free are virtually free, and it would be helpful for the reader to be familiar with that proof.

The first step is to prove that G has more than one end. Assuming that, we use Stalling's theorem [12] to conclude that G has a decomposition as an amalgamated free product $G = G_1 *_K G_2$, or as an HNN-extension $G = G_1 *_{K,t}$, over a finite subgroup K. Since G_1 (and G_2) are easily seen to be quasiconvex subgroups of G, it is not hard to show that the hypotheses of the theorem are inherited by the subgroup $H \cap G_1$ of G_1 (and $H \cap G_2$ of G_2), and so they too have more than one end, and we can apply the Dunwoody accessibility result to conclude that G is virtually free.

So we just need to prove that G has more than one end. Fix a finite generating set X of G. Then, as in [10], we consider a grammar in Chomsky normal form with no useless variables that derives $\mathrm{WP}(G, H, X)$. That is, each rule has the form $S \to \varepsilon$ (where S is the start symbol), $z \to z'z''$ or $z \to a$, where z, z', z'' are variables, and a is terminal, and we assume that S does not occur on the right hand side of any derivation. When a word w' can be derived from a word w by application of a single grammatical rule we write $w \Rightarrow w'$, and when a sequence of such rules is needed we write $w \Rightarrow^* w'$.

Let $z_1, \ldots, z_n$ be the variables of the grammar other than S, and for each z_i let u_i be a shortest word in A^* with $z_i \Rightarrow^* u_i$. Let L be the maximum length of the words u_i.

Let $w \in \mathrm{WP}(G, H, X)$ with $|w| > 3$, and fix a derivation of w in the grammar. We shall define a planar X-graph Δ with an associated X-graph homomorphism $\phi : \Delta \to \mathcal{S}(G, H, X)$. We start with a simple plane polygon with a base point, and edges labelled by the letters of w, and with ϕ mapping the base-point of Δ to $1_{\mathcal{S}}$. Note that ϕ is not necessarily injective.

If z_i occurs in the chosen derivation of w, then we have $w = vv_iv'$ with $z_i \Rightarrow^* v_i$; two such words v_i and v_j are either disjoint as subwords of w or related by containment. Since $z_i \Rightarrow^* u_i$, we also have $vu_iv' \in \mathrm{WP}(G, H, X)$. So we can draw a chord labelled u_i in the interior of Δ between the two ends of the subpath labelled v_i, and ϕ extends to this extension of Δ. If we do this for each such z_i for which $1 < |v_i| < |w| - 1$ then, as in [10, Theorem 1], we get a 'diagonal triangulation' of Δ, in which the sides are either boundary edges of Δ or internal chords of length at most L. (But note that, for the first derivation $S \to z_1z_2$, say, if $|v_1| > 1$ and $|v_2| > 1$ then, to avoid an internal bigon, we omit the chord labelled u_2.)

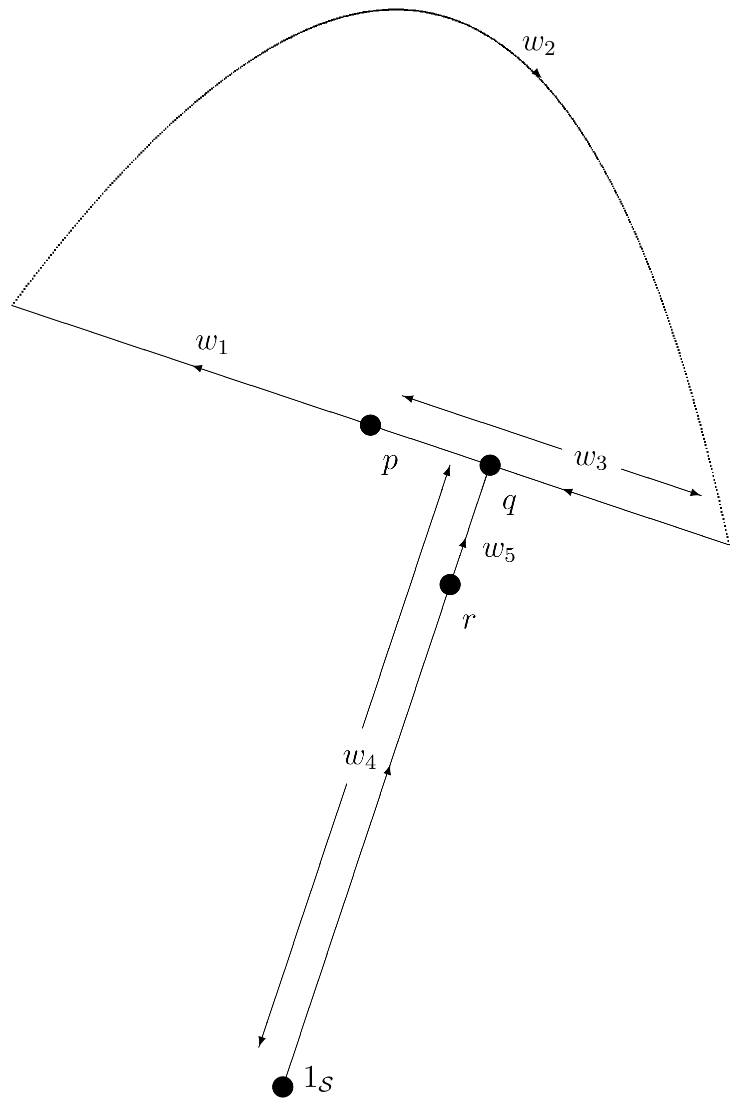
Suppose, for a contradiction, that G has just one end; that is, for any R, the complement in $\mathcal{G}(G, X)$ of any ball of radius R is connected. Then, for any R, we can find a word $w_1w_2w_3$ over X with $w_1w_2w_3 =_G 1$ which, starting at $1_{\mathcal{C}}$, labels a simple closed path in $\mathcal{G}(G, X)$, where $|w_1| = |w_3| = R$, w_3w_1 is geodesic, and no vertex in the path labelled w_2 is at distance less than R from $1_{\mathcal{C}}$.

Choose such a path with $R = 3L+1$. Choose k' such that the whole of the path lies in the ball $B_{k'}(1_{\mathcal{C}})$ of $\mathcal{G}(G, X)$, and let $k = k' + L$. Then, since $\mathcal{S}(G, H, X)$ satisfies $\mathrm{GIB}(k)$, there exists K such that, for any vertex p of $\mathcal{S}(G, H, X)$ with $d(1_{\mathcal{S}}, p) \geq K$, the ball $B_k(p)$ of $\mathcal{S}(G, H, X)$ is X-graph isomorphic to the ball $B_k(1_{\mathcal{C}})$ of $\mathcal{G}(G, X)$. Choose such a vertex p, and consider the path labelled $w_1w_2w_3$ of $\mathcal{S}(G, H, X)$ that is based at p.

Choose a vertex q on the path labelled $w_1w_2w_3$ with $d(1_{\mathcal{S}}, q)$ minimal, and let w_4 be the label of a geodesic path in $\mathcal{S}(G, H, X)$ from $1_{\mathcal{S}}$ to q. Then, for some cyclic permutation w' of $w_1w_2w_3$, we have a closed path in $\mathcal{S}(G, H, X)$ based at $1_{\mathcal{S}}$ and labelled $w_4w'w_4^{-1}$

We apply the above triangulation process to a planar X-graph Δ for $w_4w'w_4^{-1}$. Since q is the closest vertex to $1_{\mathcal{S}}$ on the loop labelled by w', and the path from $1_{\mathcal{S}}$ to q in the Cayley graph labelled by w_4 is geodesic, every vertex on that path is as close to q as to any other vertex of w', and so any vertex of w_4 that can be connected by the image of a chord of Δ' to a vertex of w' must be within distance at most L of q. Let r be the first such vertex on w_4 (as we move from $1_{\mathcal{S}}$ to q), and let w_5 be the suffix of w_4 that labels the path along w_4 from r to q. Then $|w_5| \leq L$.

So we can derive from our triangulation of Δ a triangulation of a planar diagram Δ' for the word $w_5 w' w_5^{-1}$, and there is an associated X-graph homomorphism ϕ' that maps this to the corresponding subpath in $\mathcal{S}(G, H, X)$. (Note that the images of w_5 and w_5^{-1} under ϕ' are equal, but that ϕ' is injective when restricted to w'.) By our choice of $k = k' + L$, the image of ϕ' lies entirely within $B_k(p)$, which is X-graph isomorphic to $B_k(1_{\mathcal{C}})$. So the distances in $\mathcal{S}(G, H, X)$ between vertices in this image are the same as in any path with the same label in $\mathcal{G}(G, X)$.



As in [10], we colour, using three colours, the vertices of the boundary paths of Δ' that are labelled w_1, w_2, w_3 (where vertices on two of these subwords get both associated colours), and we colour the vertices on w_5 and w_5^{-1} with the same colour (or colours) as q. As in [10, Lemma 5], we conclude that there is a triangle in the triangulation whose vertices use all three colours between them. One (or even two) of these vertices could be on the subpath labelled w_5, and two 2-coloured vertices in the triangle might coincide, but, since any vertex on w_5 is within distance L of q, replacing vertices on w_5 by q as necessary, we end up with a triangle of three (not necessarily distinct) vertices p_1, p_2, p_3 with p_i on w_i, and with $d(p_i, p_j) \leq 2L$ for each i, j. At least one of p_1, p_3 must be within distance L of p. But then $d(p, p_2) \leq 3L$, contradicting our assumption that w_2 is outside $B_{3L}(p)$. This completes the proof of Theorem 2.

4 Proof of Theorem 3

We prove Theorem 3 by constructing a generalised Dehn algorithm over the alphabet $A \cup \{H\}$ to solve $\mathrm{GWP}(G, H, X)$, where $A = X \cup X^{-1}$. We define such an algorithm to be a system of length reducing rewrite rules for strings over that alphabet which, for any word w over A, rewrites the string Hw to H precisely when w represents an element of H. We verify that the conditions of [7, Theorem 4.1] hold, and hence that the algorithm can be programmed on a real-time Turing machine.

We observe that both [7, Theorem 4.1] and [6, Proposition 2.1] (which we shall use to deduce the conditions we need) were originally stated and proved for word problems of groups, but in fact their proofs remain valid for generalised word problems. Similarly, generalised Dehn algorithms, also known as Cannon's algorithms, were defined in [5] in relation to word problems, but here we are using the concept a little more generally, so that it can be applied to the generalised word problem.

Suppose that G, H satisfy the hypotheses of Theorem 3. By [3, Theorem 4.1.3.3] or [9], the Schreier graph $\mathcal{S} := \mathcal{S}(G, H, X)$ is δ-hyperbolic for some $\delta > 0$ (that is, geodesic triangles in $\mathcal{S}$ are δ-thin). Let k be an integer with $k \geq 4\delta$. By Proposition 3.1, $\mathcal{S}$ satisfies $\mathrm{GIB}(k)$. Let K be an integer that satisfies the condition in the definition of $\mathrm{GIB}(k)$, and let $R = 2K$. We can assume that $K \geq \max(k, 2)$. We define our rewrite system to consist of all rules of the following two forms:

$$\begin{aligned} Hv_1 &\to Hv_2, \qquad |v_2| < |v_1| \leq R, \\ u_1 &\to u_2, \qquad |u_2| < |u_1| \leq k, \end{aligned}$$

where $v_1, v_2, u_1, u_2 \in A^*, v_1 v_2^{-1} \in H, u_1 =_G u_2$.

In order to apply [7, Theorem 4.1] we need to verify that, whenever $w \in A^*$ and Hw is reduced according to the above algorithm, the length of the shortest string v over A with $Hv = Hw$ is bounded below by a linear function of $|w|$.

We shall use [6, Proposition 2.1]: if u (of length > 1) is a k-local geodesic in a

δ-hyperbolic graph, with $k \geq 4\delta$, then the distance between the endpoints of u is at least $|u|/2 + 1$.

So suppose that Hw is reduced according to the above algorithm. If w has length at most R, then Hw is geodesic, and so there is nothing to prove. So suppose that $|w| > R$, and let w_1 be the prefix of length R of w. We aim to show that every vertex of $\mathcal{S}$ that comes after w_1 on the path from $1_{\mathcal{S}}$ labelled w lies outside of $B_K(1_{\mathcal{S}})$. Choose w_2 so that w_1w_2 is maximal as a prefix of w subject to all vertices of w_2 lying outside of $B_K(1_{\mathcal{S}})$. Then, since w_1 is geodesic of length $R = 2K$, we have $|w_2| \geq K - 1$, and so $|w_1w_2| \geq 3K - 1$. Since w_1 is geodesic in $\mathcal{S}$, and that part of the path labelled w_1w_2 that lies outside of the K-ball is a k-local geodesic in $\mathcal{S}$ (because it is isometric to part of the Cayley graph), we see that the whole of the path labelled w_1w_2 is a k-local geodesic in $\mathcal{S}$. So we can apply [6, Proposition 2.1] to deduce from the δ-hyperbolicity of $\mathcal{S}$ that

$$d_{\mathcal{S}}(1_{\mathcal{S}}, Hw_1w_2) \geq (|w_1w_2| + 1)/2 \geq (3K + 1)/2 > K + 1.$$

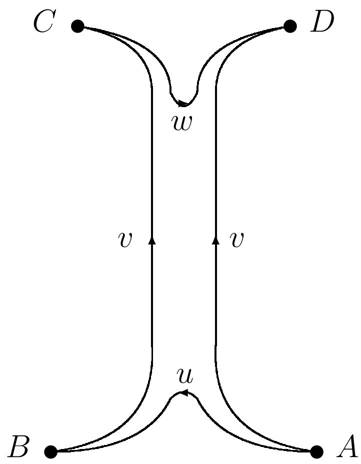
It follows that, if w_1w_2 were not already equal to w, then it would be extendible to a longer prefix of w. So $w_1w_2 = w$, and the above inequality gives us the linear lower bound $d_{\mathcal{S}}(1_{\mathcal{S}}, Hw) \geq (|w| + 1)/2$ on $d_{\mathcal{S}}(1_{\mathcal{S}}, Hw)$. Since, for any $w \in H$, the coset Hw is equal to $1_{\mathcal{S}}$, the existence of this lower bound ensures that, for such w, the string Hw is reduced by the algorithm to H, and so the algorithm that we have described is correct. And it allows us to apply [7, Theorem 4.1] and deduce that the algorithm can be programmed on a real-time Turing machine. This completes the proof of Theorem 3.

5 Sketch of proof of Proposition 3.1

Suppose first that $|C_G(h) : C_H(h)|$ is infinite for some $1 \neq h \in H$, and let w be a word representing h. Then, for any $K > 0$, there exists a word $v \in C_G(h)$ labelling a path in $\mathcal{S} := \mathcal{S}(G, H, X)$ from $1_{\mathcal{S}}$ to a vertex p with $d(1_{\mathcal{S}}, p) > K$, and there is a loop labelled w based at p in $\mathcal{S}$, but no such loop based at $1_{\mathcal{C}}$ in $\mathcal{G} := \mathcal{G}(G, X)$. So GIB$(|w|)$ fails in $\mathcal{S}$.

Suppose conversely that GIB(k) fails in $\mathcal{S}$ for some k. Then there are vertices p of $\mathcal{S}$ at arbitrarily large distance from $1_{\mathcal{S}}$ such that $B_k(p)$ is not X-graph isomorphic to the ball $B_k(1_{\mathcal{C}})$ in $\mathcal{G}$; at any such vertex p there is at least one loop, based at p, within $B_k(p)$, and labelled by a word which is not equal to the identity of G, and so cannot label a loop in $\mathcal{G}$. Since the number of words that can label loops in a ball of radius k in $\mathcal{S}$ is finite, some such word w appears infinitely often in this situation, and we can choose it to be geodesic over X. Now for any integer N, there is a word v of length greater that N, labelling a geodesic in $\mathcal{S}$ from $1_{\mathcal{S}}$ to a vertex p, from which there is a loop in $\mathcal{S}$ labelled by w.

For such a word v, we have $hvw = v$ for some $h \in H$. Let u be a geodesic word labelling h. Then we have a geodesic quadrilateral vertices $A = 1_{\mathcal{C}}, B, C, D$ in $\mathcal{G}(G, X)$ with sides AB, BC, CD, AD labelled u, v, w, v, respectively, as shown in the figure.



By the hyperbolicity of G, each vertex of AB lies within a distance 2δ of some vertex on BC, CD or DA, where δ is the constant of hyperbolicity. Furthermore, since H is quasiconvex in G, there is a constant λ, such that each vertex of AB is within a distance λ of a vertex of $\mathcal{G}$ representing an element of H. Since each vertex of w lies at distance at least $|v| - k$ from any vertex in H, by choosing $|v| > k + 2\delta + \lambda$ we can ensure that none of the vertices of AB is 2δ-close to any vertex of CD. So the vertices of AB must all be 2δ-close to vertices in BC or DA. But, since v labels a geodesic path from $1_{\mathcal{S}}$ in $\mathcal{S}$, at most $2\delta + \gamma$ vertices on BC or on DA can be within $2\delta + \gamma$ of a vertex in H. So each vertex of AB is at distance at most 2δ from one of at most $4\delta + 2\gamma$ vertices and, since the total number of vertices in $\mathcal{G}$ with that property is bounded, we see that $|AB| = |u|$ is bounded by some expression in $|X|$, δ and λ.

By hyperbolicity of G, the two paths BC and AD labelled v must synchronously L-fellow travel for some L (which depends on the upper bounds on $|w|$ and $|u|$). Let $m > 0$. Then, by choosing v sufficiently long, we can ensure that some word u' appears as a word-difference between BC and AD at least m times. That is, v has consecutive subwords $v_0, \ldots, v_m, v'$, such that $v = v_0 v_1 v_2 \cdots v_m v'$, and $h v_0 v_1 v_2 \cdots v_i u' =_G v_0 v_1 v_2 \cdots v_i$ for each i with $0 \leq i \leq m$. The case $i = 0$ gives $u' = v_0^{-1} h^{-1} v_0$, and it follows from this that $g_i := v_0 (v_1 v_2 \cdots v_i) v_0^{-1} \in C_G(h)$ for $1 \leq i \leq m$. Also, since v labels a geodesic in $\mathcal{S}$, the elements $v_0 v_1, v_0 v_1 v_2, \ldots v_0 v_1 v_2 \cdots v_m$ lie in distinct cosets of H and hence so do the g_i. Since we can choose m arbitrarily large, this contradicts the finiteness of $|C_G(h) : C_H(h)|$.

References

[1] A.V. Anisimov and F.D. Seifert. Zur algebraischen Charakteristik der durch kontext-freie Sprachen definierten Gruppen. *Elektron. Informationsverarb. Kybernet.*, 11:695–702, 1975.

[2] M.J. Dunwoody. The accessibility of finitely presented groups. *Invent. Math.*, 81:449–457, 1985.

[3] R. Foord. Automaticity and growth in certain classes of groups and monoids. PhD thesis, University of Warwick, 2000.

[4] S.M. Gersten and H.B. Short. Rational subgroups of biautomatic groups. *Ann. of Math.*, 134(1):125–158, 1991.

[5] O. Goodman and M. Shapiro. On a generalization of Dehn's algorithm. *Internat. J. Algebra Comput.*, 18:1137–1177,2008.

[6] D.F. Holt. Word-hyperbolic groups have real-time word problem, *Internat. J. Algebra Comput.*, 10:221–227, 2000.

[7] D.F. Holt and S. Rees, Solving the word problem in real time, *J. London Math Soc.* 63 (2001), 623–639.

[8] John E. Hopcroft and Jeffrey D. Ullman, *Introduction to automata theory, languages and computation*, Addison-Wesley, 1979.

[9] I. Kapovich. The geometry of relative Cayley graphs for subgroups of hyperbolic groups, `http://arxiv.org/abs/math/0201045v2`.

[10] D.E. Muller and P.E. Schupp. Groups, the theory of ends, and context-free languages. *J. Comp. System Sci.*, 26:295–310, 1983.

[11] E. Rips. Subgroups of small cancellation groups, *Bulletin London Math. Soc.*, 14(1):45–47, 1982.

[12] J. Stallings. *Group Theory and Three-Dimensional Manifolds*, volume 4 of *Yale Mathematical Monographs*. Yale University Press, 1971.