Math. Z. 241, 597–611 (2002) **Mathematische Zeitschrift**
DOI: 10.1007/s00209-002-0434-6

Growth tightness for word hyperbolic groups

G. N. Arzhantseva[1], **I. G. Lysenok**[2]

[1] Section de Mathématiques, Université de Genève, CP 240, 1211 Genève 24, Switzerland (e-mail: Goulnara.Arjantseva@math.unige.ch
[2] Steklov Institute of Mathematics, Gubkina str. 8, 117966 Moscow, Russia (e-mail: lysionok@mi.ras.ru)

Received: 20 September 2001; in final form: 24 January 2002/
Published online: 5 September 2002 –

Abstract. We show that non-elementary word hyperbolic groups are growth tight. This means that, given such a group G and a finite set A of its generators, for any infinite normal subgroup N of G, the exponential growth rate of G/N with respect to the natural image of A is strictly less than the exponential growth rate of G with respect to A.

1. Introduction

Let G be a finitely generated group and A a finite set of generators for G. By $|x|$ we denote the *geodesic length* of an element $x \in G$ in the generators A, i.e. the length of a shortest word in the alphabet $A^{\pm 1}$ representing x. Let $B(n)$ denote the ball $\{g \in G \mid |g| \leq n\}$ of radius n in G.

The *exponential growth rate* of the pair (G, A) is the limit

$$\lambda(G, A) = \lim_{n\to\infty} \sqrt[n]{\#B(n)}$$

where $\#X$ denotes the number of elements of a finite set X. The existence of the limit follows from the submultiplicativity property of the function $\#B(n)$: $\#B(m+n) \leq \#B(n)\#B(m)$ for any $n, m \geq 0$, see for example [8, VI.C]. The *uniform exponential growth rate* $\lambda(G)$ of G is the infinum $\inf_A \lambda(G, A)$ over all finite generating sets A of G.

We say that a pair (G, A) is *growth tight* if $\lambda(G, A) > \lambda(G/N, \underline{A})$ for all infinite normal subgroups N of G, with $\underline{A}$ denoting the canonical image of A in G/N. (This is a modification of the definition in [5].) Observe that, for

The work has been supported by the Swiss National Science Foundation. The second author has been supported also by INTAS, grant No. 97–1259

a torsion-free group G, "infinite" means "any nontrivial" in this definition and it coincides then with one given in [5]. On the other hand, if N is finite it is not hard to see that we always have $\lambda(G, A) = \lambda(G/N, \underline{A})$.

It is known that for a free group F_k of rank $k \geq 2$ and for a free basis X_k of F_k, the pair (F_k, X_k) is growth tight. The direct product $F_k \times F_k$ generated by $X = (X_k \times \{1\}) \cup (\{1\} \times X_k)$ is an example of a pair which is not growth tight [5].

For results and applications related to the exponential growth rate and the growth tightness property, see [5,8,9]. We mention here that the exponential growth rate of the pair (G, A) and its logarithm, sometimes called the entropy of the pair (G, A), give rise to bounds of the growth of the volume of balls in a Riemannian manifold M with fundamental group G as well as to bounds of the topological entropy of the geodesic flow on M, see [6, Sect. 5.B].

Recall that a group G is called *Hopfian* if every epimorphism $G \to G$ is an isomorphism, i.e. there is no a proper quotient G/N of G isomorphic to G. As an immediate consequence of the above definitions we have the following

Observation. Let G be a finitely generated group. If there exists a finite generating set A of G such that $\lambda(G, A) = \lambda(G)$ and if (G, A) is growth tight, then G/N cannot be isomorphic to G for any infinite normal subgroup N of G.

In this paper, we restrict our attention to the class of word hyperbolic groups G in the sense of M.Gromov [7]. Note that any word hyperbolic group is finitely generated and even finitely presented [7, 2.1A, 2.2A], [4, Proposition 4.17]. It is known also that there are only finitely many conjugacy classes of finite subgroups in a word hyperbolic group G, see for example [2, Ch.III.Γ, Theorem 3.2]. This implies in particular that G/N is not isomorphic to G for any nontrivial finite normal subgroup N of G. Indeed, if such an isomorphism $\phi : G \to G/N$ would exist then we could define an infinite strictly increasing sequence N_i $(i \geq 0)$ of finite subgroups of G by $N_0 = N$ and $N_{i+1} = \psi^{-1}\phi(N_i)$ where $\psi : G \to G/N$ is the canonical epimorphism.

Thus, the observation provides a possible way to prove the Hopf property for word hyperbolic groups. Note that torsion-free word hyperbolic groups are already known to be Hopfian [10]. For word hyperbolic groups with torsion, the question whether or not they are Hopfian is still open.

Our main result concerns the second condition in the observation above and gives an affirmative answer to the question about growth tightness of word hyperbolic groups, posed by R. Grigorchuk and P. de la Harpe [5]. Recall that a word hyperbolic group is called *elementary* if it is either finite or a finite extension of the infinite cyclic group.

Theorem 1. *Let G be a non-elementary word hyperbolic group and A any finite set of generators for G. Then, for any infinite normal subgroup N of G,*

$$\lambda(G, A) > \lambda(G/N, \underline{A}),$$

where $\underline{A}$ is the canonical image of A in G/N.

To prove the theorem we first observe that the growth tightness of G is related to the growth tightness of certain languages over $A^{\pm 1}$ and we obtain a stronger result (Theorem 2 below) which implies Theorem 1.

We fix an arbitrary linear ordering on the alphabet $A^{\pm 1}$. A word w in the alphabet $A^{\pm 1}$ is called *geodesic* if w has the minimal possible length among all the words representing the same element of G. If w is geodesic and is minimal in the induced lexical ordering, among all geodesic words representing the same element of G, it is called *shortlex geodesic*. By $\mathcal{L}$ we denote the set of all shortlex geodesic words in $A^{\pm 1}$. Clearly, every element $x \in G$ has a unique representative in $\mathcal{L}$ which we denote $\overline{x}$. It is obvious also that any subword of a shortlex geodesic word is itself shortlex geodesic.

Definition 1. *Let $C > 0$. We say that two elements $x, y \in G$ are C-*close *if there are $u, v \in G$ such that $x = uyv$ and $|u|, |v| \leq C$. We say that x C-*contains *y if $\overline{x}$ has a subword representing an element which is C-close to y. For a given $w \in G$, we denote*

$$Z_{w,C} = \{x \in G \mid x \ C\text{-contains } w\}.$$

We define the exponential growth rate of the set $G \setminus Z_{w,C}$ by

$$\lambda_{w,C} = \limsup_{n\to\infty} \sqrt[n]{\#(B(n) \setminus Z_{w,C})} \tag{1}$$

Theorem 2. *Let G be a non-elementary word hyperbolic group and A a finite set of generators for G. Let $\lambda = \lambda(G, A)$ be the exponential growth rate of G with respect to A. Then there is a number $C = C(G, A)$ such that $\lambda_{w,C} < \lambda$ for any $w \in G$.*

Note that Theorem 2 is of somewhat technical nature because the definition of $\lambda_{w,C}$ uses the language of shortlex geodesic words in G. But it easily implies a stronger result which does not make use of any specific language. To formulate this result, we use, in an informal sense, "the strongest" form of the relation "x C-contains y" which leads to the "smallest possible" appropriate set $Z_{w,C}$ and, consequently, to the "largest possible" appropriate parameter $\lambda_{w,C}$.

We say that x *strongly C-contains* y if any geodesic word representing x contains a subword which is C-close to y. Let

$$\overline{Z}_{w,C} = \{x \in G \mid x \text{ strongly } C\text{-contains } w\}$$

and let $\overline{\lambda}_{w,C}$ be the exponential growth rate of the set $G \setminus \overline{Z}_{w,C}$, defined as in (1).

Theorem 3. *Let G be a non-elementary word hyperbolic group and A a finite set of generators for G. Let $\lambda = \lambda(G, A)$ be the exponential growth rate of G with respect to A. Then there is a number $C = C(G, A)$ such that $\overline{\lambda}_{w,C} < \lambda$ for any $w \in G$.*

Acknowledgments. This work has been done during the second author's visit to the University of Geneva in May 2001. He thanks P. de la Harpe for hospitality. We thank also the referee for useful comments.

2. Preliminaries

2.1. Languages

Let L be a language over a finite alphabet X, i.e. L is any set of words in the alphabet X. We denote

$$B_L(n) = \{x \in L \mid |x| \leq n\}, \quad \lambda_L = \limsup_{n\to\infty} \sqrt[n]{\#B_L(n)}$$

and, for any word $w \in L$, we define

$$L_w = \{x \in L \mid x \text{ contains } w \text{ as a subword}\}.$$

We say that L is *growth tight* if $\lambda_{L \setminus L_w} < \lambda_L$ for any $w \in L$.

Now let G be a finitely generated group, A a finite set of its generators, and $\mathcal{L}$ a language of shortlex geodesic words in the alphabet $A^{\pm 1}$. Assume that $\mathcal{L}$ is growth tight. It is then easy to see that the pair (G, A) is growth tight. Indeed, let N be a nontrivial normal subgroup of G. Choosing a word $w \in \mathcal{L}$ representing a nontrivial element of N we see that $\mathcal{L}_{G/N} \subseteq \mathcal{L} \setminus \mathcal{L}_w$ where $\mathcal{L}_{G/N}$ is the language of shortlex geodesic words in G/N. This obviously implies

$$\lambda(G/N, \underline{A}) = \lambda_{\mathcal{L}_{G/N}} \leq \lambda_{\mathcal{L}\setminus\mathcal{L}_w} < \lambda_{\mathcal{L}} = \lambda(G, A).$$

(Here $\underline{A}$, as above, denotes the canonical image of A in G/N.)

Thus to prove growth tightness of G it suffices to prove growth tightness of the language $\mathcal{L}$. This approach works well for example for proving growth tightness of finitely generated free groups. But in general, we observe that growth tightness of $\mathcal{L}$ implies that $\lambda(G/N, \underline{A}) < \lambda(G, A)$ for *any* nontrivial normal subgroup N of G; and as we have mentioned above, for finite N we have the equality. This means in particular that the language $\mathcal{L}$ is not growth tight if G has a nontrivial finite normal subgroup.

We use a slight modification of this approach for word hyperbolic groups. Namely, for $C > 0$ we define

$$L_{w,C} = \{x \in L \mid x \text{ has a subword representing an element of } G \\ \text{which is } C\text{-close to } w\}.$$

We say that L is *growth quasitight* if there exists $C = C(G, A) > 0$ such that for any $w \in L$ we have $\lambda_L > \lambda_{L \setminus L_{w,C}}$.

Now observe that growth quasitightness of $\mathcal{L}$ implies growth tightness of G. Let N be an infinite normal subgroup of G. We take a word $w \in \mathcal{L}$ representing an element of N and of length greater than $4C$. Suppose that a word $x \in \mathcal{L}$ has a subword y representing an element which is C-close to w, that is, $y = uwv$ in G for some u, v with $|u|, |v| \leq C$. We have $|y| \geq |w| - |u| - |v| > 2C$. This implies that x is not geodesic in G/N as its subword y is equal in G/N to a shorter word uv. Hence $\mathcal{L}_{G/N} \subseteq \mathcal{L} \setminus \mathcal{L}_{w,C}$ and again we have $\lambda(G/N, \underline{A}) = \lambda_{\mathcal{L}_{G/N}} \leq \lambda_{\mathcal{L} \setminus \mathcal{L}_{w,C}} < \lambda_{\mathcal{L}} = \lambda(G, A)$.

As a consequence, we obtain that Theorem 2 implies Theorem 1.

2.2. Auxiliary results on word hyperbolic groups

We fix a group G and a finite set A of generators for G for the rest of the paper. All words are assumed to be in the alphabet $A^{\pm 1}$. We shall make no essential distinction between words and elements of G. If w is a word then the notation $w \equiv xy$ means that w can be decomposed, as a word, into a product of two words which represent elements $x, y \in G$.

There are known several properties of a group G which are equivalent to its word hyperbolicity. For practical purposes, one of the most useful is the property of geodesic triangles in the Cayley graph of G being δ-thin, see [7, 6.3], [4, Proposition 2.21]. In the combinatorial language, it can be formulated as follows. Recall that the *Gromov inner product* of two elements $x, y \in G$ with respect to the identity element of G is defined to be

$$(x|y) = \frac{1}{2}(|x| + |y| - |x^{-1}y|).$$

Then G is word hyperbolic if and only if there is a number $\delta \geq 0$ such that the following is true:

(H1) *for any two geodesic words u and v, if $u \equiv u_1u_2$, $v \equiv v_1v_2$ and $|u_1| = |v_1| \leq (u|v)$ then $|u_1^{-1}v_1| \leq \delta$.*

From now on, we assume G to be word hyperbolic and fix a number $\delta \geq 0$ such that (H1) holds.

As a consequence of (H1), we have another property of G which is in fact original Gromov's definition, see [7, 6.3B], [4, 2.21]:

(H2) $\quad (x|y) \geq \min\{(x|z),(y|z)\} - 2\delta$ for any $x, y, z \in G$.

Lemma 1. *For any $x, y, z \in G$ the following assertions hold.*

(a) *If $(x^{-1}|y) \leq r$, $(y^{-1}|z) \leq t$ and $|y| > r + t + 2\delta$ then $((xy)^{-1}|z) \leq t + 2\delta$ and $(x^{-1}|yz) \leq r + 2\delta$. We have also*

$$|xyz| \geq |x| + |y| + |z| - 2(r + t + 2\delta).$$

(b) *If x is a shortest representative in its conjugacy class and $|x| \geq 2\delta + 2$ then $(x^{-1}|x) < \delta + 1$.*

Proof. (a): Let $(x^{-1}|y) \leq r$, $(y^{-1}|z) \leq t$ and $|y| > r + t + 2\delta$. We have

$$(y^{-1}|(xy)^{-1}) = \frac{1}{2}(|y| + |xy| - |x|) = |y| - (x^{-1}|y) > t + 2\delta.$$

Assuming $((xy)^{-1}|z) > t + 2\delta$, by (H2) we get $(y^{-1}|z) > t$ contrary to the hypothesis. Hence $((xy)^{-1}|z) \leq t + 2\delta$. Similarly, observing that $(y|yz) = |y| - (y^{-1}|z) > r + 2\delta$ from $(x^{-1}|y) \leq r$ we deduce by (H2) that $(x^{-1}|yz) \leq r + 2\delta$.

Now using $((xy)^{-1}|z) \leq t + 2\delta$ and $(x^{-1}|y) \leq r$, we obtain

$$|xyz| \geq |xy| + |z| - 2(t + 2\delta) \geq |x| + |y| + |z| - 2(r + t + 2\delta).$$

(b): Let x be a shortest representative in its conjugacy class and $|x| \geq 2\delta + 2$. Assume that $(x^{-1}|x) \geq \delta + 1$. Let $\tilde{x}$ be a geodesic word representing x and let $\tilde{x} \equiv x_1 x_2 x_3$ where $|x_1| = |x_3| = \delta + 1$. By (H1), $|x_3 x_1| \leq \delta$. Then $|x_3 x x_3^{-1}| = |x_3 x_1 x_2| \leq |x_2| + \delta < |x|$ contrary to the choice of x. $\square$

Definition 2 (a refinement of Definition 1). Let $x, y \in G$ and $0 \leq t_1 \leq t_2 \leq |x|$. We say that x (r, t_1, t_2)-*contains* y if $\overline{x} \equiv x' z x''$ where z is r-close to y and $|x'| \geq t_1$ and $|x'z| \leq t_2$. (This is equivalent to the condition that if we divide $\overline{x}$ into segments $\overline{x} \equiv x_1 w x_2$ with $|x_1| = t_1$, $|x_1 w| = t_2$ then w r-contains y.)

Lemma 2.

(a) *If x is r-close to y and y is t-close to z then x is $(r+t)$-close to z.*

(b) *Let $|x^{-1}y| \leq r$ and $\tilde{x}$, $\tilde{y}$ be geodesic words representing x and y, respectively. If $\tilde{x} \equiv x_1 x_2$ then $\tilde{y} \equiv y_1 y_2$ for some y_1, y_2 where $|x_1^{-1} y_1| \leq r + \delta$ and x_2 is $(r+\delta)$-close to y_2.*

(c) *Let x and y be r-close and let $\tilde{x}$ and $\tilde{y}$ be any geodesic words representing x and y, respectively. Then the following assertions are true. If $\tilde{x} \equiv x_1x_2$ then $\tilde{y} \equiv y_1y_2$ for some y_1 and y_2 where x_i is $(r+2\delta)$-close to y_i $(i=1,2)$. Any subword of $\tilde{x}$ is $(r+2\delta)$-close to some subword of $\tilde{y}$ and, in particular, is $(r+2\delta)$-contained in y.*

(d) *If $(x^{-1}|y) \le r$ then both x and y are $(r+\delta+1)$-contained in xy. Moreover, if w is a geodesic word representing xy then $w \equiv x_1y_1$ where $|x^{-1}x_1| \le r+\delta+1$ and $|yy_1^{-1}| \le r+\delta+1$.*

(e) *If x is (r,t_1,t_2)-contained in y, $t_1' \le t_2' \le |y|$ and $|t_i - t_i'| \le s$ $(i=1,2)$ then x is $(r+s,t_1',t_2')$-contained in y.*

(f) *If x is (r,t_1,t_2)-contained in y, and y is s-close to z then x is $(r+s+2\delta,t_1',t_2')$-contained in z for some t_1', t_2' with $|t_i - t_i'| \le 2s+2\delta$. In particular, by (e), if $t_2 \le |z|$ then x is $(r+3s+4\delta,t_1,t_2)$-contained in z.*

Proof. Assertion (a) easily follows from Definition 1.

(b): Let $\tilde{x} \equiv x_1x_2$ and $\tilde{y} \equiv y_1y_2$. Notice that $x_2 = x_1^{-1}y_1 \cdot y_2 \cdot (x^{-1}y)^{-1}$. Hence if $|x_1^{-1}y_1| \le r+\delta$ then by Definition 1, x_2 and y_2 are $(r+\delta)$-close. So it suffices to find a factorization $\tilde{y} \equiv y_1y_2$ with $|x_1^{-1}y_1| \le r+\delta$.

If $|x_1| \le (x|y)$ then by (H1), for $\tilde{y} \equiv y_1y_2$ with $|y_1| = |x_1|$ we have $|x_1^{-1}y_1| \le \delta$. Let $|x_1| > (x|y)$. Since $(x|y) + (x^{-1}|x^{-1}y) = |x|$ we have $|x_2| < (x^{-1}|x^{-1}y)$. Let w be a geodesic word representing $x^{-1}y$ and let $w \equiv w_1w_2$ where $|w_1| = |x_2|$. Then by (H1), $|x_2w_1| \le \delta$ and hence $|x_1^{-1}y| = |x_2x^{-1}y| \le |x_2w_1| + |w_2| \le \delta + r$. In this case, we can take $y_1 = y$ and $y_2 = 1$.

(c): Let $x = uyv$ where $|u|,|v| \le r$. Denoting $z = xv^{-1}$ we have $|x^{-1}z| = |v| \le r$ and $|zy^{-1}| = |u| \le r$.

We prove the first assertion. Let $\tilde{x} \equiv x_1x_2$. Using the arguments from the proof of (b) we find z_1 and z_2 such that $\tilde{z} \equiv z_1z_2$ and either $|x_1^{-1}z_1| \le \delta$ or $z_1 = z$ and $|x_1^{-1}z_1| \le r+\delta$. If $z_1 = z$ then $x_1 = uy(x_1^{-1}z_1)^{-1}$ and we can take $y_1 = y$. Otherwise, the application of (b) with $x := z^{-1}, y := y^{-1}$ gives a factorization $\tilde{y} \equiv y_1y_2$ with $|z_2y_2^{-1}| \le r+\delta$. Then $x_1 = uy_1w$ where

$$w = y_1^{-1}u^{-1}x_1 = y_1^{-1}u^{-1}z \cdot z^{-1} \cdot x_1 = y_2 \cdot z_2^{-1}z_1^{-1} \cdot x_1 = (z_2y_2^{-1})^{-1}(x_1^{-1}z_1)^{-1}$$

and hence $|w| \le r+2\delta$.

To prove the second assertion, we observe that the factorization $\tilde{y} \equiv y_1y_2$ given in the proof of (b), and hence in the previous argument, is monotone with respect to the factorization $\tilde{x} \equiv x_1x_2$; i.e., if $\tilde{x} \equiv x_1'x_2'$ and $|x_1'| \ge |x_1|$ then, for the corresponding y_1', we have $|y_1'| \ge |y_1|$. This and the previous argument imply that if $\tilde{x} \equiv x_1x_2x_3$ then there are y_1, y_2, y_3 such that $\tilde{y} \equiv y_1y_2y_3$ and $x_1 = uy_1w$, $x_1x_2 = uy_1y_2w'$ for some w, w' with $|w|,|w'| \le r+2\delta$. But then $x_2 = w^{-1}y_2w'$ thus finishing the proof of (c).

(d): If $|u^{-1}v| \le t$ or $|uv^{-1}| \le t$ then, obviously, u is t-close to v. So the first assertion follows from the second.

Suppose that $(x^{-1}|y) \le r$ and w is a geodesic representative for xy. We choose a geodesic representative $\tilde{x}$ for x and let $w \equiv x_1 y_1$, $\tilde{x} \equiv x_2 x_3$ where $(x|w) - 1 < |x_1| = |x_2| \le (x|w)$. By (H1), $|x_2^{-1} x_1| \le \delta$. Since $(x|w) + (x^{-1}|y) = |x|$ we have $|x_3| < (x^{-1}|y) + 1$ and hence

$$|x^{-1}x_1| = |x_3^{-1} \cdot x_2^{-1} x_1| < r + \delta + 1.$$

It remains to notice that $yy_1^{-1} = x^{-1}x_1$.

(e): easily follows from Definition 2.

(f): Let x be (r, t_1, t_2)-contained in y, and y be s-close to z. By Definition 2, $\overline{y} \equiv y_1 x_1 y_2$ where x_1 is r-close to x, $|y_1| \ge t_1$ and $|y_1 x_1| \le t_2$. Let $y = uzv$ where $|u|, |v| \le s$. The proof of (c) shows that $\overline{z}$ may be decomposed as $\overline{z} \equiv z_1 x_2 z_2$ where $y_1 = u z_1 w$, $y_1 x_1 = u z_1 x_2 w'$ and $x_1 = w^{-1} x_2 w'$ for some w and w' with $|w|, |w'| \le s + 2\delta$. By (a), x_2 is $(r + s + 2\delta)$-close to x. We have $||y_1| - |z_1|| \le 2s + 2\delta$ and $||y_1 x_1| - |z_1 x_2|| \le 2s + 2\delta$. Hence x is $(r + s + 2\delta, t_1', t_2')$-contained in z where $t_1' = \max\{0,\ t_1 - 2s - 2\delta\}$ and $t_2' = \min\{|z|,\ t_2 + 2s + 2\delta\}$. □

Lemma 3. *There is a number $D = D(G, A)$ with the following property. Given any $u, w \in G$ there exists a word $z \in \mathcal{L}$ such that $z \equiv z_1 z_2$ for some z_1, z_2 with $|u^{-1} z_1| \le D$ and $|w z_2^{-1}| \le D$.*

Proof. By [7, 5.1B], there are infinitely many conjugacy classes of elements in a non-elementary word hyperbolic group. We choose any $v_0 \in G$ such that $|v_0| \ge 14\delta + 2$ and v_0 is a shortest representative in its conjugacy class. Let $H = \langle v_0 \rangle$. Observe that H has infinite index in G since G is non-elementary. Then there are infinitely many double cosets HxH in G, as H is a quasiconvex subgroup of infinite index in a word hyperbolic group, see for example [1, 2.3 and Proposition 1]. We choose $x \in G$ such that $|x| \ge |v_0| + 2\delta + 1$ and x is a shortest representative in the double coset HxH.

By the choice of x, we have $|v_0^{\pm 1} x^{\pm 1}| \ge |x|$. This implies that $(v_0^{\pm 1}|x^{\pm 1}) \le \frac{1}{2}|v_0|$. Since $|x| > |v_0| + 2\delta$, by Lemma 1(a) we get $(v_0^{\pm 1}|x^{\pm 1} v_0^{\pm 1}) \le \frac{1}{2}|v_0| + 2\delta$ and

$$|v_0^{\pm 1} x v_0^{\pm 1}| \ge |x| - 4\delta \ge 12\delta + 2.$$

Hence for $\varepsilon = \pm 1$ we obtain

$$(v_0^{\varepsilon} | v_0^{\varepsilon} x^{\pm 1} v_0^{\pm 1}) = |v_0| - (v_0^{-\varepsilon} | x^{\pm 1} v_0^{\pm 1}) \ge \frac{1}{2}|v_0| - 2\delta.$$

Now by Lemma 1(b) and (H2), we have either $(u^{-1}|v_0) < 3\delta + 1$ or $(u^{-1}|v_0^{-1}) < 3\delta + 1$ and either $(w|v_0) < 3\delta + 1$ or $(w|v_0^{-1}) < 3\delta + 1$. Let $(u^{-1}|v_0^{\xi}) < 3\delta + 1$ and $(w|v_0^{\nu}) < 3\delta + 1$.

If $(u^{-1}|v_0^\varepsilon x v_0^\nu) \geq 5\delta + 1$, by (H2) with $x := v_0^\varepsilon$, $y := u^{-1}$ and $z := v_0^\varepsilon x v_0^\nu$ we would have

$$(u^{-1}|v_0^\varepsilon) \geq \min\left\{\frac{1}{2}|v_0| - 2\delta, 5\delta + 1\right\} - 2\delta \geq 3\delta + 1,$$

obtaining a contradiction. Hence

$$(u^{-1}|v_0^\varepsilon x v_0^\nu) < 5\delta + 1.$$

Similarly, we see that $((v_0^\varepsilon x v_0^\nu)^{-1}|w) < 5\delta + 1$. Since $|v_0^\varepsilon x v_0^\nu| \geq 12\delta + 2$, by Lemma 1(a),

$$((uv_0^\varepsilon x v_0^\nu)^{-1}|w) < 7\delta + 1.$$

By Lemma 2(d), $\overline{uv_0^\varepsilon x v_0^\nu} \equiv u_1 y$ where $|u^{-1}u_1| \leq 6\delta + 2$ and $|v_0^\varepsilon x v_0^\nu y^{-1}| \leq 6\delta + 2$. The last inequality implies

$$|y| \leq |v_0^\varepsilon x v_0^\nu| + 6\delta + 2 \leq |x| + 2|v_0| + 6\delta + 2.$$

Again by Lemma 2(d), $\overline{uv_0^\varepsilon x v_0^\nu w} \equiv u_2 w_1$ where $|(uv_0^\varepsilon x v_0^\nu)^{-1}u_2| \leq 8\delta + 2$ and $|ww_1^{-1}| \leq 8\delta + 2$. By Lemma 2(b), $u_2 \equiv u_3 y_1$ where $|u_1^{-1}u_3| \leq 9\delta + 2$ and y_1 is $(9\delta + 2)$-close to y. We have $\overline{uv_0^\varepsilon x v_0^\nu w} \equiv u_3 y_1 w_1$ where

$$|u^{-1}u_3| \leq |u^{-1}u_1| + |u_1^{-1}u_3| \leq 15\delta + 4$$

and

$$|w(y_1w_1)^{-1}| \leq |ww_1^{-1}| + |y_1| \leq 8\delta + 2 + |y| + 2(9\delta + 2) \leq |x| + 2|v_0| + 32\delta + 8.$$

Therefore, we can take $z = \overline{uv_0^\varepsilon x v_0^\nu w}$, $z_1 = u_3$, $z_2 = y_1 w_1$ and $D = |x| + 2|v_0| + 32\delta + 8$. ☐

Lemma 4. *There are numbers $E = E(G, A)$ and $\kappa = \kappa(G, A) > 0$ with the following property. Let $w \in G$ be any element. Let*

$$X_w = \{x \in G \mid |w^{-1}x_1| \leq E \text{ for some initial segment } x_1 \text{ of } \overline{x}\}.$$

Then, for any $n \geq |w|$,

$$\#(X_w \cap B(n)) \geq \kappa \, \#B(n - |w|).$$

Proof. We take $E = D$ and $\kappa = \frac{1}{(\#B(2D))^2}$ where D is as in Lemma 3. If $|w| \leq n < |w| + 2D$ then $\kappa \, \#B(n - |w|) < 1$ and $\#(X_w \cap B(n)) \geq 1$ as $w \in X_w$; so the required inequality holds. Let $n \geq |w| + 2D$. For a given w and any $u \in B(n - |w| - 2D)$ we take $z_u \equiv z_1 z_2 \in \mathcal{L}$ by Lemma 3 where $|w^{-1}z_1| \leq D$ and $|uz_2^{-1}| \leq D$. Clearly we have $z_u = w y_u u$ for some $y_u \in G$ with $|y_u| \leq 2D$. In particular, $z_u \in X_w \cap B(n)$.

By the pigeon hole principle, for some subset $U \subseteq B(n-|w|-2D)$ with $\#U \geq \frac{1}{\#\{y_u\}}\#B(n-|w|-2D)$ we have $y_{u_1}=y_{u_2}$ for all $u_1,u_2 \in U$. It follows that $z_{u_1} \neq z_{u_2}$ for $u_1,u_2 \in U$, $u_1 \neq u_2$. Hence

$$\#(X_w \cap B(n)) \geq \frac{1}{\#\{y_u\}}\#B(n-|w|-2D) \geq \frac{1}{\#B(2D)}\#B(n-|w|-2D).$$

By submultiplicativity of $\#B(n)$,

$$\#B(n-|w|-2D) \geq \frac{1}{\#B(2D)}\#B(n-|w|)$$

and we finally have

$$\#(X_w \cap B(n)) \geq \frac{1}{(\#B(2D))^2}\#B(n-|w|).$$

□

Lemma 5. *Let $r \geq 0$. There exists a number $F = F(G,A,r) > 0$ such that for any finite set X of elements of G there is a subset $U \subseteq X$ such that $\#U \geq \frac{1}{F}\#X$ and $|x^{-1}y| > r$ for any distinct $x,y \in U$.*

Proof. We take $F = \#B(r)$. Denote $B_x(r) = \{y \in G|\ |x^{-1}y| \leq r\}$. Clearly, $\#B_x(r) = \#B(r)$ for any $x \in G$. We choose subsequently arbitrary elements $x_1, x_2, \dots \in G$ such that

$$\begin{aligned} x_1 &\in X_0 = X, \\ x_2 &\in X_1 = X \setminus B_{x_1}(r), \\ &\dots \\ x_{k+1} &\in X_k = X \setminus \cup_{i=1}^k B_{x_i}(r), \\ &\dots \end{aligned}$$

We have $\#X_k \geq \#X - kF$, so there are at least $\frac{1}{F}\#X$ such x_i's. We set $U = \{x_1, x_2, \dots\}$. □

3. Proof of Theorems 2 and 3

By Lemma 2(a), (c), for any $x,y \in G$, if x is C-contained in y then x is strongly $(C+2\delta)$-contained in y. As an immediate consequence, we obtain that Theorem 2 implies Theorem 3. The rest of the section is devoted to proving Theorem 2.

By a result of Coornaert [3], there exists $\alpha = \alpha(G,A)$ such that $\#B(n) \leq \alpha\lambda^n$ for all n. We introduce the following constants where

D and E are as in Lemmas 3 and 4, respectively:

$$\begin{aligned} N_0 &= \log_\lambda 2\alpha + 2D + E, \\ C_1 &= N_0 + 2D + 3E + 4\delta, \\ C &= C_1 + 9D + 5E + 8\delta, \\ R &= N_0 + 4D + 4E + 4\delta. \end{aligned}$$

We fix an arbitrary $w \in G$ and assume that $\lambda_{w,C} = \lambda$. Our aim is to deduce a contradiction from this assumption.

Observe that any subword of a word in $\mathcal{L} \setminus \mathcal{L}_{w,C}$ lies in $\mathcal{L} \setminus \mathcal{L}_{w,C}$ as well. Since $\mathcal{L} \setminus \mathcal{L}_{w,C}$ forms a set of unique representatives for the elements in $G \setminus Z_{w,C}$, this easily implies that $\#(B(n) \setminus Z_{w,C})$ is a submultiplicative function on n. Then by [8, VI.C, Proposition 56], $\lambda_{w,C} = \inf_{n\geq 1} \#(B(n)\setminus Z_{w,C})^{1/n}$ and we consequently have $\#(B(n) \setminus Z_{w,C}) \geq \lambda^n$ for all n. Thus, with the assumption above, we have lower and upper bounds for $\#(B(n) \setminus Z_{w,C})$ and $\#B(n)$:

$$\lambda^n \leq \#(B(n) \setminus Z_{w,C}) \leq \#B(n) \leq \alpha\lambda^n. \tag{2}$$

Let N be any number such that

$$N \geq |w| + 2N_0 + 3D + 4E + 4\delta.$$

We reach a contradiction with (2) by proving that

$$\lim_{k\to\infty} \frac{\#(B(kN) \setminus Z_{w,C})}{\#B(kN)} = 0. \tag{3}$$

To do this, we introduce two series Y_k and Y_k^* $(k \geq 1)$ of subsets of G:

$$\begin{aligned} Y_k = \{x \in G \mid\ & x \text{ does not } (C, (i-1)N, iN)\text{-contain } w \text{ for } 1\leq i\leq k-1, \\ & x\ (C, (k-1)N, t)\text{-contains } w \text{ for some } t \leq kN\}, \\ Y_k^* = \{x \in G |\ & x \text{ does not } (C_1, (i-1)N, iN)\text{-contain } w \text{ for } 1\leq i\leq k-1, \\ & x\ (C_1, (k-1)N, t)\text{-contains } w \text{ for some } t \leq kN\}. \end{aligned}$$

It immediately follows from the definition that Y_i are pairwise disjoint and so are Y_i^*. We set

$$Z_k = \bigcup_{i=1}^{k} Y_i.$$

Clearly, $Z_k \subseteq Z_{w,C}$ for any k.

Lemma 6 (main lemma). *There is a number $\beta > 0$ such that for any $k \geq 1$ and $0 \leq i \leq k-1$,*

$$\#(Y_{i+1}^* \cap B(kN)) \geq \beta\, \#(B(iN) \setminus Z_i)\, \#B((k-i)N). \tag{4}$$

Before proving the lemma, we show how it implies (3) and hence proves Theorem 2. Since Y_i^* are pairwise disjoint we have

$$\sum_{i=0}^{k-1} \#(Y_{i+1}^* \cap B(kN)) \le \#B(kN).$$

With the inequality in Lemma 6 this gives

$$\beta \sum_{i=0}^{k-1} \#(B(iN) \setminus Z_i) \,\#B((k-i)N) \le \#B(kN).$$

By submultiplicativity of $\#B(n)$,

$$\frac{\#B(kN)}{\#B(iN)} \le \#B((k-i)N)$$

and hence, after dividing by $\#B(kN)$ we obtain

$$\beta \sum_{i=0}^{k-1} \frac{\#(B(iN) \setminus Z_i)}{\#B(iN)} \le 1.$$

But since this holds for any k and $Z_i \subseteq Z_{w,C}$, it follows that the series $\sum_{i=1}^{\infty} \frac{\#(B(iN)\setminus Z_{w,C})}{\#B(iN)}$ converges thus implying (3) as required.

Proof of Lemma 6. Let $k \ge 1$ and $0 \le i \le k-1$. For $m < n$ we denote

$$S(m,n) = B(n) \setminus B(m).$$

By the definition of N_0, the inclusion $Z_i \subseteq Z_{w,C}$ and (2),

$$\begin{aligned}
&\#(S(iN+2D+E,\ iN+N_0)\ \setminus\ Z_i) \\
&\qquad = \#(B(iN+N_0) \setminus Z_i) - \#(B(iN+2D+E) \setminus Z_i) \\
&\qquad \ge \lambda^{iN+N_0} - \alpha\lambda^{iN+2D+E} \ge \alpha\lambda^{iN+2D+E} \ge \#(B(iN) \setminus Z_i).
\end{aligned}$$

By Lemma 5, there is a number $F = F(G,A,R) > 0$ and a subset $U \subseteq S(iN+2D+E, iN+N_0) \setminus Z_i$ such that $|x^{-1}y| > R$ for any distinct $x, y \in U$ and

$$\#U \ge \frac{1}{F}\#(B(iN) \setminus Z_i). \tag{5}$$

For any $x \in U$ we define a set $V_x \subset G$. By Lemma 3, there is a word $\hat{x} \in \mathcal{L}$ such that

$$\hat{x} \equiv x_1 x_2 \quad \text{where} \quad |x^{-1}x_1| \le D,\ |wx_2^{-1}| \le D. \tag{6}$$

We define, as in Lemma 4,

$$V_x = \{y \in G \mid |\hat{x}^{-1}y_1| \leq E \text{ for some initial segment } y_1 \text{ of } \overline{y}\}.$$

We have $iN + 2D + E \leq |x| \leq iN + N_0$, $|w| + |x| - 2D \leq |\hat{x}| \leq |w| + |x| + 2D$ and hence

$$iN + |w| + E \leq |\hat{x}| \leq iN + |w| + N_0 + 2D. \tag{7}$$

Since $kN \geq iN + N \geq iN + |w| + N_0 + 2D$, by Lemma 4

$$\#(V_x \cap B(kN)) \geq \kappa\,\#(B(kN - |\hat{x}|)).$$

Using the upper bound for $|\hat{x}|$ and submultiplicativity of $\#B(n)$ we get

$$\begin{aligned}\#(B(kN - |\hat{x}|)) &\geq \#(B((k-i)N - |w| - N_0 - 2D)) \\ &\geq \frac{\#(B((k-i)N))}{\#(B(|w| + N_0 + 2D))}.\end{aligned}$$

This gives a lower bound for $\#(V_x \cap B(kN))$:

$$\#(V_x \cap B(kN)) \geq \frac{\kappa}{\#(B(|w| + N_0 + 2D))}\,\#(B((k-i)N)).$$

We prove the following two assertions about the sets V_x:

(i) $V_x \subseteq Y^*_{i+1}$ for any $x \in U$;
(ii) $V_x \cap V_{x'} = \emptyset$ for $x \neq x'$.

Observe that (i), (ii) and the lower bounds for $\#(V_x \cap B(kN))$ and (5) for $\#U$ imply the desired inequality (4) for $\beta = \frac{\kappa}{F\,\#(B(|w|+N_0+2D))}$ in an obvious way. To prove the lemma, it remains to prove (i) and (ii).

Proof of (i). Let $x \in U$ and $y \in V_x$. We check for y the conditions in the definition of Y^*_{i+1}. First we show that y (C_1, iN, t)-contains w for some $t \leq (i+1)N$. Since x_2 in (6) is D-close to w, by Definition 2 $\hat{x}$ $(D, |x_1|, |\hat{x}|)$-contains w. Let y_1 be the initial segment of $\overline{y}$ with $|\hat{x}^{-1}y_1| \leq E$. By Lemma 2(f), w is $(D+E+2\delta, t_1, t_2)$-contained in y_1, and therefore in y, where $|t_1 - |x_1|| \leq 2E + 2\delta$ and $|t_2 - |\hat{x}|| \leq 2E + 2\delta$. Now

$$|t_1 - iN| \leq ||x_1| - iN| + 2E + 2\delta \leq ||x| - iN| + D + 2E + 2\delta \leq N_0 + D + 2E + 2\delta$$

and by the choice of N,

$$\begin{aligned}(i+1)N - t_2 &\geq (i+1)N - 2E - 2\delta - |\hat{x}| \\ &\geq (i+1)N - 2E - 2\delta - iN - |w| - N_0 - 2D \\ &\geq N_0 + D + 2E + 2\delta.\end{aligned}$$

Observing that $|y| \geq |\hat{x}| - E \geq iN$ by the definition of V_x and (7), we conclude by Lemma 2(e) that w is $(N_0 + 2D + 3E + 4\delta, iN, t)$-contained in y for some $t \leq (i+1)N$ as required.

Now we prove that y does not $(C_1, (j-1)N, jN)$-contain w for all $1 \leq j \leq i$. Assume that y does $(C_1, (j-1)N, jN)$-contain w for some $1 \leq j \leq i$. Since

$$jN - |y_1| \leq jN - |\hat{x}| + E \leq (j-i)N - |w| + 2D + E \leq 2D + E,$$

by Lemma 2(e) w is $(C_1 + 2D + E, (j-1)N, t)$-contained in y_1 for some t with $jN - 2D - E \leq t \leq jN$. By Lemma 2(f), w is $(C_1 + 2D + 2E + 2\delta, t_1, t_2)$-contained in $\hat{x}$ where $|t_1 - (j-1)N| \leq 2E + 2\delta$ and $|t_2 - t| \leq 2E + 2\delta$. We have

$$|t_2 - jN| \leq |t_2 - t| + 2D + E \leq 2D + 3E + 2\delta.$$

Now $jN - |x_1| \leq jN - |x| + D \leq D$ and using Lemma 2(e) again we find that w is $(C_1 + 3D + 2E + 2\delta, t_1, t_2')$-contained in x_1 with $|t_2' - jN| \leq 3D + 3E + 2\delta$. Another application of Lemma 2(f) gives that w is $(C_1 + 4D + 2E + 4\delta, t_1', t_2'')$-contained in x where $|t_1' - (j-1)N| \leq 2D + 2E + 4\delta$ and $|t_2'' - jN| \leq 5D + 3E + 4\delta$. Again by Lemma 2(e), w is $(C_1 + 9D + 5E + 8\delta, (j-1)N, s)$-contained in x for some $s \leq jN$. Due to the choice of C and the definition of Y_k this means that $x \in Y_l$ for some $l \leq j$. But this is a contradiction with $x \notin Z_i$. This finishes the proof of (i).

Proof of (ii). Let $y \in V_x$. By Lemma 2(b), $|x_1^{-1} y'| \leq E + \delta$ for some initial segment y' of $\overline{y}$ and hence we have

$$|x^{-1} y'| \leq |x^{-1} x_1| + |x_1^{-1} y'| \leq E + D + \delta.$$

Assume that $y \in V_{x'}$ for some $x' \neq x$. Then there is another initial segment y'' of $\overline{y}$ such that $|(x')^{-1} y''| \leq E + D + \delta$. Since $x, x' \in S(iN, iN + N_0)$ we have $||x| - |x'|| \leq N_0$ and therefore $||y'| - |y''|| \leq N_0 + 2E + 2D + 2\delta$. Since both y' and y'' are initial segments of $\overline{y}$ this implies $|(y')^{-1} y''| \leq N_0 + 2E + 2D + 2\delta$ and finally we obtain $|x^{-1} x'| \leq N_0 + 4D + 4E + 4\delta$. But this contradicts the choice of R and the definition of U. Lemma 6 is proved. □

References

1. G. N. Arzhantseva, On quasiconvex subgroups of word hyperbolic groups, Geometriae Dedicata **85** (2001), 191–208.
2. M. R. Bridson, A. Haefliger, Metric spaces of non-positive curvature, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 319, Springer-Verlag, Berlin, 1999.

3. M. Coornaert, Mesures de Patterson–Sullivan sur le bord d'un espace hyperbolique au sens de Gromov, Pacific J. Math. **159** (1993), 241–270.
4. E. Ghys, P. de la Harpe (editors), Sur les groupes hyperboliques d'après Mikhael Gromov, Progress in Math. **83**, Birkhäuser, 1990.
5. R. Grigorchuk, P. de la Harpe, On problems related to growth, entropy, and spectrum in group theory, J. Dynam. Control Systems **3** (1997) 1, 51–89.
6. M. Gromov, Metric Structures for Riemannian and Non-Riemannian Spaces, edited by J. Lafontaine, P. Pansu, Progress in Math. **152**, Birkhäuser, 1999.
7. M. Gromov, Hyperbolic groups, in: Essays in group theory, ed. S.M. Gersten, M.S.R.I. Publ. **8**, Springer, 1987, 75–263.
8. P. de la Harpe, Topics in geometric group theory, Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 2000.
9. A. Sambusetti, Growth tightness of free and amalgamated products, preprint. Universitá degli Studi di Roma "La Sapienza", 2001.
10. Z. Sela, Endomorphisms of hyperbolic groups. I. The Hopf property, Topology **38** (1999) 2, 301–321.