

# Subgroups generated by two pseudo-Anosov elements in a mapping class group. II. Uniform bound on exponents

Koji Fujiwara  
Graduate School of Information Science  
Tohoku University  
Sendai, 980-8579, Japan  
fujiwara@math.is.tohoku.ac.jp

June 11, 2013

## Abstract

Let  $S$  be a compact orientable surface, and  $\text{Mod}(S)$  its mapping class group. Then there exists a constant  $M(S)$ , which depends on  $S$ , with the following property. Suppose  $a, b \in \text{Mod}(S)$  are independent (i.e.,  $[a^n, b^m] \neq 1$  for any  $n, m \neq 0$ ) pseudo-Anosov elements. Then for any  $n, m \geq M$ , the subgroup  $\langle a^n, b^m \rangle$  is free of rank two, and convex-cocompact in the sense of Farb-Mosher. In particular all non-trivial elements in  $\langle a^n, b^m \rangle$  are pseudo-Anosov. We also show that there exists a constant  $N$ , which depends on  $a, b$ , such that  $\langle a^n, b^m \rangle$  is free of rank two and convex-cocompact if  $|n| + |m| \geq N$  and  $nm \neq 0$ .

## 1 Introduction

This is the second half of our study on subgroups generated by two pseudo-Anosov elements in a mapping class group. We will improve the results we obtained in the first half of the study [4], but one can read this paper independently. We explain the improvement after we state the main results in this section.

## 1.1 Hyperbolic isometry and (quasi-)axis

A geodesic space is called  $\delta$ -hyperbolic for  $\delta \geq 0$  if for any geodesics  $\alpha, \beta, \gamma$  which form a triangle,  $\alpha$  is contained in the  $\delta$ -neighborhood of  $\beta \cup \gamma$  ([5]). Let  $\Gamma$  be a  $\delta$ -hyperbolic graph. Let  $a$  be an isometry of  $\Gamma$ . If there exist a point  $x \in \Gamma$  and a constant  $C > 0$  such that  $d(x, a^n(x)) \geq Cn$  for any  $n > 0$ , then  $a$  is called *hyperbolic*.

Suppose  $a$  is a hyperbolic isometry. If there exists a geodesic  $\alpha$  such that  $a(\alpha)$  is contained in the  $C$ -neighborhood of  $\alpha$  for some  $C \geq 0$ ,  $\alpha$  is called a *quasi-axis* of  $a$ . By  $\delta$ -hyperbolicity of  $\Gamma$ , it then follows that  $a(\alpha)$  is in the  $2\delta$ -neighborhood of  $\alpha$ . If  $\alpha$  and  $\beta$  are quasi-axes of  $a$  (they are geodesics by definition), then they are contained in the  $2\delta$ -neighborhood of each other. If  $C = 0$  we say  $\alpha$  is an *axis*. We remark that an axis and even a quasi-axis may not exist for  $a$ , but there is always a *quasi-geodesic* which is invariant by  $a$ . That will be good enough for our argument (See section 2.6). In the literature, quasi-axis sometimes means a quasi-geodesic which is invariant by  $a$ . Our definition is different.

For two points  $x, y \in \Gamma$ , we may denote a geodesic joining them by  $[x, y]$  although they are not unique. We may write the distance between the two points by  $|x - y|$ .

For an isometry  $a$ , we define its *translation length*,  $\text{tr}(a)$ , by

$$\text{tr}(a) = \lim_{n \rightarrow \infty} \frac{|x - a^n(x)|}{n} \geq 0$$

for a point  $x$ . It is easy to see  $\text{tr}(a)$  does not depend on the choice of  $x$ . The isometry  $a$  is hyperbolic iff  $\text{tr}(a) > 0$ .

It is known (§7,8 [5]) that  $a$  is hyperbolic if there is a point  $p \in \Gamma$  such that the following is satisfied. In particular, the element  $a$  has infinite order. We remark that if  $\delta > 0$  we can replace  $(\delta + 1)$  by  $\delta$ .

$$|p - a(p)| \leq |a(p) - a^{-1}(p)| - 100(\delta + 1). \quad (1)$$

The argument is geometric. Consider the following infinite path, which is  $a$ -invariant.

$$\beta = \cdots [a^{-1}(p), p] \cup [p, a(p)] \cup [a(p), a^2(p)] \cup [a^2(p), a^3(p)] \cup \cdots$$

For each  $n > 0$ , by  $\delta$ -hyperbolicity,  $[p, a^n(p)]$  is contained in the  $n\delta$ -neighborhood of  $[p, a(p)] \cup \cdots \cup [a^n(p), a^{n+1}(p)]$ . Using (1), one can show it is indeed contained in the  $3\delta$ -neighborhood. It then follows

that  $\beta$  is a quasi-geodesic. If  $a$  has a (geodesic) quasi-axis, then it is contained in the  $4\delta$ -neighborhood of  $\beta$ .

In section 2 we use a similar geometric idea (Proposition 1) to give a sufficient condition for two isometries to generate a free group.

## 1.2 Convex co-compact subgroup in $\text{Mod}(S)$

Suppose that a finitely generated group  $G$  is acting on  $\Gamma$  by isometries. Fix a finite generating set and let  $|a|$  be the word metric of  $a \in G$ . Let  $x \in \Gamma$  be a point and consider the map from  $G$  to  $\Gamma$  defined by sending  $a \in G$  to  $a(x) \in \Gamma$ . We call this map as *the embedding by an orbit* of the action by  $G$ . If there exist constants  $L, C > 0$  such that for all  $a \in G$

$$|a|/L - C \leq d(x, a(x)) \leq L|a| + C,$$

then we say the map is *quasi-isometric*.

Our main application is regarding subgroups of mapping class groups. Let  $S$  be a compact orientable surface, and  $\text{Mod}(S)$  its mapping class group. Let  $\mathcal{C}(S)$  be the curve graph of  $S$ , on which  $\text{Mod}(S)$  acts by isometries (see for example [10], [16] for the definition).

Masur-Minsky [16] showed that  $\mathcal{C}(S)$  is  $\delta$ -hyperbolic and an element  $a \in \text{Mod}(S)$  is pseudo-Anosov if and only if it acts as a hyperbolic isometry on  $\mathcal{C}(S)$ , and moreover that there always exists a quasi-axis.

For a subgroup  $G < \text{Mod}(S)$ , Farb-Mosher [3] introduced the notion of *convex-cocompact* in terms of the action on Teichmuller space. It has been shown ([6], [12]) that  $G$  is convex-cocompact iff for a point  $c \in \mathcal{C}(S)$ , the map from  $G$  to  $\mathcal{C}(S)$  sending  $g$  to  $g(c)$  is quasi-isometric. Note that the choice of the generating set and the point  $c$  is not important.

## 1.3 Main results

In Section 2 we discuss subgroups generated by powers of two hyperbolic isometries on a hyperbolic graph, and obtain sufficient conditions for them to be free. That section is the main technical part of the paper. We put an overview of the argument in Section 2.1. The final result is the following. The point is that although the constant  $M$  depends on  $\Gamma$  and the action of  $G$  on  $\Gamma$ , it does not depend on  $a$  and  $b$ . It will become clear how the constant  $M$  depends on the action.

**Theorem 14.** *Suppose  $G$  acts acylindrically on a  $\delta$ -hyperbolic graph  $\Gamma$ . Then there exists a constant  $M$  with the following property.*

*Suppose  $a, b \in G$  act hyperbolically. Assume for any  $p, q \neq 0$ ,  $[a^p, b^q] \neq 1$  in  $G$ . Then for any  $n, m \geq M$ ,  $\langle a^n, b^m \rangle$  is free of rank two. Moreover, the embedding of  $\langle a^n, b^m \rangle$  by an orbit in  $\Gamma$  is quasi-isometric. In particular, all non-trivial elements in  $\langle a^n, b^m \rangle$  are hyperbolic on  $\Gamma$ .*

The *acylindricity* of an action (see section 2.3 for the definition) is a weak assumption on properness, and in particular, the result applies to a word-hyperbolic group and its action on a Cayley graph, therefore, if  $G$  is a word-hyperbolic group, then there exists  $M$  such that for any two elements  $a, b \in G$  of infinite order, either the subgroup  $\langle a, b \rangle$  is elementary or else for any  $n, m \geq M$ ,  $\langle a^n, b^m \rangle$  is free and quasi-convex in  $G$ . It seems this claim is new (see [5, 8.2 E] for the statement without a bound on  $n, m$ ). The result also immediately applies to mapping class groups. That is our motivation and we show the following in Section 3.

**Theorem 16.** *Let  $S$  be a compact orientable surface, and  $\text{Mod}(S)$  its mapping class group. Then there exists a constant  $M(S)$  with the following property. Suppose  $a, b \in \text{Mod}(S)$  are pseudo-Anosov elements such that  $[a^n, b^m] \neq 1$  for any  $n, m \neq 0$ . ( $a, b$  are called independent.) Then for any  $n, m \geq M$ , the subgroup  $\langle a^n, b^m \rangle$  is free of rank two, and convex-cocompact in the sense of Farb-Mosher. In particular all non-trivial elements in  $\langle a^n, b^m \rangle$  are pseudo-Anosov.*

It was known ([10], [17]) that  $\langle a^n, b^m \rangle$  is free for sufficiently large  $n, m$ . A uniform bound on  $n, m$  is new. In our previous study [4], a uniform bound on one of  $n, m$  was shown. Namely, there exists a constant  $N(S)$  such that  $\langle a^n, b^m \rangle$  is free and convex-cocompact if one of  $n, m$  is at least  $N$  and the other one is sufficiently large. The previous study has also been used to show the *uniform exponential growth* of a mapping class group by Mangahas [15].

The theorem concerns only pseudo-Anosov elements  $a, b$ . It is unknown if a uniform bound such as  $M(S)$  exists for two elements  $a, b$  of infinite order in general such that the subgroup  $\langle a^n, b^m \rangle$  is free if  $n, m \geq M$ . Note that the subgroup is never convex-cocompact unless both  $a$  and  $b$  are pseudo-Anosov. In the case that both  $a, b$  are Dehn twists ([8]), and more generally, positive multi-twists ([7, Theorem 3.2]), it is known that  $\langle a^n, b^m \rangle$  is free or abelian if  $n, m \geq 2$ . Recently,

Leininger and Margalit [14] have shown that if  $S$  is the  $n$ -times punctured sphere, then for any two elements  $a, b \in \text{Mod}(S)$ ,  $\langle a^N, b^N \rangle$  is either free or abelian for  $N = n!$ .

It would be interesting to know for which  $(n, m)$ , the subgroup  $\langle a^n, b^m \rangle$  is free for given  $a, b$  in the above theorem. The following theorem says that for given  $a, b$ , the subgroup  $\langle a^n, b^m \rangle$  is free except for finitely many pairs  $(n, m)$ . It is not clear if the number of those exceptional pairs  $(n, m)$  is bounded, but we know that the constant  $N$  depends on  $a, b$  in the following theorem (see Example 18).

**Theorem 17.** *Let  $S$  be a compact orientable surface and  $a, b$  two independent pseudo-Anosov elements. Then there exists  $N$  such that for any  $n \geq N$ , both  $\langle a, b^n \rangle$  and  $\langle b, a^n \rangle$  are free of rank two, and convex-cocompact. In particular,  $\langle a^n, b^m \rangle$  is free of rank two, and convex-cocompact if  $|n| + |m| \geq 2N$  and  $nm \neq 0$ .*

The author would like to thank Z. Sela for many insightful suggestions, and L. Mosher for his interest and comments. He is grateful to M. Bestvina and T. Delzant. The work was partly done during his stay at MSRI in Fall 2007. He appreciates their hospitality. He wishes to acknowledge the support of JSPS-KAKENHI-19340013.

## 2 Free subgroups

### 2.1 Overview

In this section, we will find sufficient conditions for certain powers of two *independent* hyperbolic isometries  $a, b$  to generate a free group. The final results are Theorem 9 and Theorem 14.

It is well-known that for sufficiently large  $n, m > 0$ ,  $a^n, b^m$  generate a free group (Proposition 2). The argument is an application of a geometric fact on a  $\delta$ -hyperbolic space (Proposition 1). The goal of this section is to give an upper bound on  $n, m$  which does not depend on  $a$  and  $b$ .

In Section 2.3, by analyzing the argument for Proposition 2 carefully, we first show that there is an upper bound on  $n$  and  $m$  if the translation length of  $a$  and  $b$  are comparable (Proposition 6). A more difficult case is that one of the translation length, say for  $b$ , is much smaller than the translation length of  $a$ . In this case, if we use the same argument, we need to take the exponent  $m$  for  $b$  very large so that  $a^n, b^m$  generate a free group. In Section 2.4, we use a different

idea to deal with this case and show there is an upper bound on  $n$  and  $m$  if the translation of one of  $a$  and  $b$  is much smaller than the translation length of the other (Proposition 7). This part is tedious (the idea is elementary, but we put many details), but we think this is a main technical achievement of the paper.

As we explained, the two propositions are complimentary to each other, and combining them, we obtain an upper bound on the both exponents which does not depend on  $a$  and  $b$  in Section 2.5 (Theorem 9).

We first prove those results under the assumption that  $a$  and  $b$  have *quasi-axes*. Then in section 2.6 we explain that the assumption is indeed redundant, which gives Theorem 14. For our application in this paper, we only need Theorem 9, but we prove Theorem 14 for potential application in the future.

Another technical issue is the properness of an action. We argue under the assumption of *acylindricity*, which is weaker than the action being properly discontinuous (see Section 2.3). We will need that when we discuss application in Section 3.

## 2.2 Nielsen condition

In this section, we review a well-known fact (Proposition 2) and its proof. We start with a fundamental result from [5] (see §7 and 8).

**Proposition 1** (7.2C [5], Three points condition). *Let  $\Gamma$  be a  $\delta$ -hyperbolic graph. Let  $\varepsilon > 100\delta$  be a constant. Let  $p_i \in \Gamma (i \geq 1)$  be points such that for all  $i \geq 1$*

$$|p_i - p_{i+2}| \geq \max(|p_i - p_{i+1}|, |p_{i+1} - p_{i+2}|) + \varepsilon. \quad (2)$$

Then, for each  $i \geq 3$ ,

$$\lambda |p_1 - p_i| \geq \sum_{j=1}^{i-1} |p_j - p_{j+1}|, \quad (3)$$

where  $\lambda = (\frac{\varepsilon}{100} - \delta)^{-1} \max_{1 \leq j \leq i-1} |p_j - p_{j+1}|$ . In particular,  $|p_1 - p_i| > 0$  for all  $i \geq 3$ .

We call the inequality (2) *the three points condition* (for  $p_i, p_{i+1}, p_{i+2}$  and a constant  $\varepsilon$ ). If the three points condition is satisfied for any three consecutive points in a sequence, then we say the sequence satisfies the three point condition.

Proposition 1 has been used to derive a condition for sufficiently large powers of hyperbolic isometries  $a, b$  of  $\Gamma$  with quasi-axes  $\alpha, \beta$  to generate a free group in terms of  $\alpha, \beta$  (Proposition 2). In that argument, it will be important how much of  $\beta$  is contained in the  $10\delta$ -neighborhood of  $\alpha$  and vice-versa. We thus define *the  $10\delta$ -overlap* of  $\alpha$  and  $\beta$ , denoted by  $\alpha \cap_{10\delta} \beta$ , as follows.

$$\alpha \cap_{10\delta} \beta = (\alpha \cap N_{10\delta}(\beta)) \cup (\beta \cap N_{10\delta}(\alpha)).$$

Let  $|\alpha \cap_{10\delta} \beta|$  denote the diameter of this set.  $|\alpha \cap_{10\delta} \beta|$  can be  $\infty$ . If it is finite, by the  $\delta$ -hyperbolicity, the longest segment of  $\alpha$ , the longest segment of  $\beta$  and the longest geodesics which are contained in  $\alpha \cap_{10\delta} \beta$  all have length between  $|\alpha \cap_{10\delta} \beta| - 20\delta$  and  $|\alpha \cap_{10\delta} \beta| + 20\delta$ , and those segments are in the  $20\delta$ -neighborhood of each other.

The following fact is well-known ([5], see also [13]). Notice that the exponents  $n$  and  $m$  which satisfy the inequalities depend on  $a$  and  $b$ . For readers who have not seen a proof, we give details of the argument, since we will generalize the statement using the same idea.

**Proposition 2** (Nielsen condition). *Suppose isometries  $a, b$  act hyperbolically on a  $\delta$ -hyperbolic graph  $\Gamma$  with quasi-axes  $\alpha, \beta$ , respectively. Suppose  $|\alpha \cap_{10\delta} \beta| < \infty$ . If  $1 \leq n, m \in \mathbb{Z}$  are such that*

$$\mathrm{tr}(a^n) \geq |\alpha \cap_{10\delta} \beta| + 100(\delta + 1), \quad \mathrm{tr}(b^m) \geq |\alpha \cap_{10\delta} \beta| + 100(\delta + 1)$$

*then  $\langle a^n, b^m \rangle < \mathrm{Isom}(\Gamma)$  is free of rank two. Moreover, for a point  $x \in \Gamma$ , the embedding of the subgroup  $\langle a^n, b^m \rangle$  to  $\Gamma$  by sending  $w$  to  $w(x)$  is quasi-isometric.*

*Proof.* We use Proposition 1 with  $\varepsilon = 100(\delta + 1)$ . We first show that  $\langle a^n, b^m \rangle$  is free, then argue independently the moreover part. It turns out that for a certain choice of  $x$ , the embedding is not only quasi-isometric, but also bi-Lipschitz, which implies that  $\langle a^n, b^m \rangle$  is free. In that sense, the first part is not necessary, but we hope it will make the whole argument more transparent in this way.

Set

$$A = a^n, B = b^m.$$

We remark that  $\alpha, \beta$  are quasi-axes of  $A, B$ , respectively. Let  $w$  be a non-empty reduced word on  $A, B$  and we prove that the action of  $w$  on  $\Gamma$  is non-trivial, therefore,  $w$  is a non-trivial element in  $\mathrm{Isom}(\Gamma)$ . It suffices to find a point  $p \in \Gamma$  with  $w(p) \neq p$ . (In this proof the point  $p$  does not depend on the word  $w$ ).

Suppose  $|\alpha \cap_{10\delta} \beta| = 0$ . Let  $\ell$  be a geodesic segment which realizes the distance between  $\alpha$  and  $\beta$ , and  $p \in \ell$  the mid point. (See Figure 1.) We claim  $w(p) \neq p$ . To see it, let

$$w = A^{n_1} B^{m_1} \dots A^{n_i} B^{m_i},$$

where  $n_1, m_i$  are possibly 0. We discuss the case such that both  $n_1, m_i$  are not 0. Set

$$p_0 = p, p_1 = A^{n_1}(p), p_2 = A^{n_1} B^{m_1}(p), p_3 = A^{n_1} B^{m_1} A^{n_2}(p), \dots, \\ p_{2i-1} = A^{n_1} B^{m_1} \dots A^{n_i}(p), p_{2i} = A^{n_1} B^{m_1} \dots A^{n_i} B^{m_i}(p).$$

Then the the sequence of points  $p_j$  satisfies the inequalities (2) in Proposition 1. Indeed, for example, for  $p_1, p_2, p_3$ , if we apply the element  $B^{-m_1} A^{-n_1}$ , then we get  $B^{-m_1}(p), p, A^{n_2}(p)$ . Since  $m_1, n_2 \neq 0$ , those three points satisfy the three point condition by the  $\delta$ -hyperbolicity of  $\Gamma$ . Now, by Proposition 1, we get  $|p_0 - p_{2i}| > 0$ , namely  $|p - w(p)| > 0$ . We can argue similarly if  $n_1$  or  $m_i$  is 0, and omit the details.

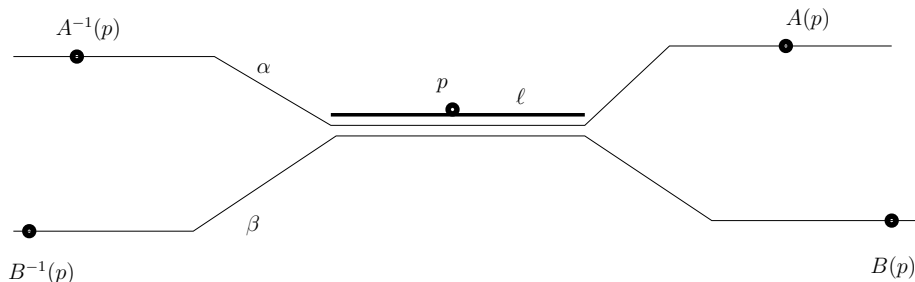


Figure 1: Nielsen condition

Suppose  $|\alpha \cap_{10\delta} \beta| = D > 0$ . Let  $\ell$  be the longest geodesic which is contained in  $\alpha \cap_{10\delta} \beta$ . Then  $D - 20\delta \leq |\ell| \leq D + 20\delta$ . Let  $p$  be the mid point of  $\ell$ . Define the points  $p_j$  in the same way as before, then by using the assumption on  $\text{tr}(a^n)$  and  $\text{tr}(b^m)$ , we get  $|p - w(p)| > 0$  by Proposition 1.

Now we argue that the embedding by the orbit of a point  $x$  is quasi-isometric with respect to the word metric for  $a^n, b^m$ . We remind that the choice of the point  $x$  is not important. Since  $\langle a^n, b^m \rangle$  is finitely generated, the embedding is always Lipschitz with respect to the constant  $\max(|a^n(x) - x|, |b^m(x) - x|)$ . We will show that there



exist a point  $x$ , constants  $L > 0, C \geq 0$  such that for any non-trivial reduced word  $w$  on  $a^n, b^m$ ,

$$\frac{|w|}{L} - C \leq |w(x) - x|,$$

where  $|w|$  is the word metric with respect to  $a^n, b^m$ . Indeed we will have  $C = 0$  if  $x = p$ .

The argument is a modification of the previous one, so that we use Proposition 1. For each  $j$  set

$$B_j = [p_{2j}, p_{2j-1}], \quad A_{j+1} = [p_{2j}, p_{2j+1}].$$

They are geodesic segments. Note that

$$|p_{2j} - p_{2j-1}| = |p - B^{m_j}(p)| = |m_j| \text{tr}(B), \quad |p_{2j} - p_{2j+1}| = |p - A^{n_{j+1}}(p)| = |n_{j+1}| \text{tr}(A).$$

Thus the right hand side of the inequality (3) is proportional to the word length of  $w$  when we vary  $w$ . The left hand side of the inequality is  $\lambda|p_0 - p_{2i}| = \lambda|p - w(p)|$ , so that if there is an upper bound on  $\lambda > 0$  when we vary the word  $w$ , then we would be done. But since  $\lambda = \max_j |p_j - p_{j+1}|$ ,  $\lambda$  can be arbitrarily large when we vary  $w$ .

As a remedy, we will divide each of geodesic segments  $A_j, B_j$ , namely, introduce certain points on them such that there is an upper bound on the distance between any two consecutive points, then apply Proposition 1 to this new sequence of points. The point is that introducing new points does not change the property that the right hand side of the inequality (3) is proportional to  $|w|$ , while the constant  $\lambda$  for the new sequence will have an upper bound.

We start the argument. Again, we discuss the case that both  $n_1, m_i$  are not 0. First, between  $p_0$  and  $p_1$ , if  $n_1 > 0$ , define points by

$$p_{0,0} = p_0, p_{0,1} = A(p), p_{0,2} = A^2(p), \dots, p_{0,n_1} = A^{n_1}(p) = p_1,$$

and if  $n_1 < 0$ , define the points  $p_{0,k}$  for  $0 \leq k \leq -n_1$  similarly using actions by  $A^{-1}, A^{-2}, \dots, A^{-n_1}$  on  $p$ . Next, between  $p_1$  and  $p_2$ , if  $m_1 > 0$ , define points by

$$p_{1,0} = p_1, p_{1,1} = A^{n_1}B(p), p_{1,2} = A^{n_1}B^2(p), \dots, p_{1,m_1} = A^{n_1}B^{m_1}(p) = p_2,$$

and if  $m_1 < 0$ , define the points similarly as follows.

$$p_{1,0} = p_1, p_{1,1} = A^{n_1}B(p), p_{1,2} = A^{n_1}B^2(p), \dots, p_{1,|m_1|} = A^{n_1}B^{m_1}(p) = p_2.$$

We define points similarly between  $p_{2j}$  and  $p_{2j+1}$ , and also between  $p_{2j}$  and  $p_{2j-1}$  for all  $j$ . We obtain a sequence of points  $p_{j,k}$  with the canonical order (the lexicographical order on  $(j, k)$ ). By definition, the distance between any two consecutive points is either  $\text{tr}(A)$  or  $\text{tr}(B)$ . Also, by our assumption, the sequence of points satisfies the three points condition of Proposition 1 such that  $\lambda = \max(\text{tr}(A), \text{tr}(B))$ , which no more depends on  $w$ . Here, we regard, for example,  $p_{0,n_1}$  and  $p_{1,0}$  are the same point, which is  $p_1$ . By the proposition (for the first inequality), we get

$$\begin{aligned} |p_0 - p_{2i}| &\geq \frac{1}{\lambda} \sum_{j=0}^{i-1} \left( \sum_{k=1}^{|n_{j+1}|} |p_{2j,k-1} - p_{2j,k}| + \sum_{k=1}^{|m_{j+1}|} |p_{2j+1,k-1} - p_{2j+1,k}| \right) \\ &= \sum_{j=0}^{2i-1} |p_j - p_{j+1}| = \frac{1}{\lambda} \left( \text{tr}(A) \sum_{j=1}^i |n_j| + \text{tr}(B) \sum_{j=1}^i |m_j| \right) \\ &\geq \frac{\min(\text{tr}(A), \text{tr}(B))}{\lambda} |w|. \end{aligned}$$

Set  $L' = \frac{\min(\text{tr}(A), \text{tr}(B))}{\lambda} > 0$ . We get  $|p - w(p)| \geq L'|w|$ . The same bound holds if  $n_1$  or  $m_i$  is 0 as well. We have shown that the embedding is bi-Lipschitz, for this particular choice of a base point and a generating set.  $\square$

We are interested in finding an upper bound on  $n, m > 0$  such that  $\langle a^n, b^m \rangle$  is not free under some condition on the action of  $G$  on  $\Gamma$ , provided that  $D = |\alpha \cap_{10\delta} \beta| < \infty$ . Before we discuss that, we analyze the case when  $D = \infty$ , namely, hyperbolic isometries  $a$  and  $b$  have a common quasi-axis. For example, suppose  $G$  is a word-hyperbolic group and  $\Gamma$  is a Cayley graph, which is  $\delta$ -hyperbolic. Then  $D = \infty$  implies that  $[a, b^k] = 1$  for some  $k > 0$  and  $[b, a^l] = 1$  for some  $l > 0$ , where the commutator of two elements is defined by

$$[f, g] = f^{-1}g^{-1}fg.$$

To see the first claim, take a point  $x \in \alpha$ , the common quasi-axis, and look at the set of points

$$[a, b](x), [a, b^2](x), [a, b^3](x), \dots$$

They are all in the  $20\delta$ -neighborhood of  $x$ . Since the action of  $G$  is proper, there are only finitely many elements  $g \in G$  with  $|x - g(x)| \leq$

$20\delta$ , therefore there must be distinct integers  $n, m > 0$  such that  $[a, b^n] = [a, b^m]$ . This implies  $[a, b^{n-m}] = 1$ . The second claim is similar.

Notice that in the previous argument, we do not need that  $D$  is infinite, but it is enough if  $D$  is sufficiently large. To formulate a precise statement (Lemma 5), we consider a certain condition, *acylindricity*, on the action in the next section.

## 2.3 Acylindrical action

In this section, we assume certain properness of an action, acylindricity, and improve Proposition 2 to Proposition 6.

Let  $\Gamma$  be a  $\delta$ -hyperbolic graph, and  $G$  a group acting on  $\Gamma$  by isometries. Bowditch [1] defined that the action is *acylindrical* if for any  $R > 0$ , there exist  $K(R), L(R) \geq 1$  such that for any vertices  $x, y \in \Gamma$  with  $d(x, y) \geq L(R)$ , the following set has at most  $K(R)$  elements:

$$\{g \in G \mid d(x, g(x)) \leq R, d(y, g(y)) \leq R\}.$$

**Lemma 3.** *Suppose  $G$  acts on a  $\delta$ -hyperbolic graph  $\Gamma$ . If the action is acylindrical with constants  $K(R), L(R)$ , then there exists an integer  $P \geq 1$  such that for any element  $a \in G$  which acts hyperbolically on  $\Gamma$  with a quasi-axis, we have  $\text{tr}(a^P) \geq 1$ . The constant  $P$  depends only on  $\delta$  and  $K(200\delta)$ .*

*Convention 4* (Subscript of a constant). To keep track of constants, we may number a constant by the number of the claim which the constant first appears, for example, the constant  $P$  in Lemma 3 will be  $P_3$ . We may omit the subscript if there is no confusion.

*Proof.* If  $\delta = 0$ ,  $\Gamma$  is a tree. Then  $\text{tr}(a) \geq 1$ . Set  $P = 1$ . Suppose  $\delta > 0$ . Set  $R = 100\delta$ . Let  $\alpha$  be a (geodesic) quasi-axis of  $a$ . Take a point  $x \in \alpha$ . Let  $y \in \alpha$  be a point with  $|x - y| \geq L(2R)$ . If  $|a^i(x) - x| \leq R$  for some  $i$ , then  $|a^i(y) - y| \leq 2R$ . This is because  $a^i(\alpha)$  is in the  $2\delta$ -neighborhood of  $\alpha$ . By the acylindricity, this implies that there is  $I$  with  $1 \leq I \leq K(2R) = K(200\delta)$  such that  $|a^I(x) - x| > R$ . It then follows that for any  $n \geq 1$ ,  $|a^{In}(x) - x| > n(R - 10\delta)$ . Thus

$$\text{tr}(a) \geq \frac{R - 10\delta}{I} \geq \frac{90\delta}{K(200\delta)}.$$

Choose an integer  $P$  such that  $P \geq \frac{K(200\delta)}{90\delta}$ . □

**Lemma 5.** *Let  $\Gamma$  be a  $\delta$ -hyperbolic graph, and  $G$  a group acting on  $\Gamma$  acylindrically with constants  $K(R), L(R)$ . Suppose  $a, b \in G$  act hyperbolically with quasi-axes  $\alpha, \beta \subset \Gamma$ , respectively.*

*If  $a^n b \neq ba^n$  for all  $n \neq 0$  or  $b^n a \neq ab^n$  for all  $n \neq 0$ , then*

$$|\alpha \cap_{10\delta} \beta| < 4P_3K(20\delta)L(20\delta) \max(\text{tr}(a), \text{tr}(b)) + 100\delta.$$

By Convention 4,  $P_3$  is the constant from Lemma 3.

*Proof.* To argue by contradiction, suppose that the inequality was false. Set  $K = K(20\delta), L = L(20\delta)$ . For concreteness, suppose  $\text{tr}(b) \leq \text{tr}(a)$ . By our assumption, since  $|\alpha \cap_{10\delta} \beta|$  is much larger than  $2\delta$ , the set  $\alpha \cap_{10\delta} \beta$  looks like a narrow tube.

Let  $\ell \subset \alpha$  be the longest segment which is contained in  $\alpha \cap_{10\delta} \beta$ . Then, by our assumption,  $|\ell| \geq 4PKL\text{tr}(a) + 80\delta$ . Take a point  $p \in \ell$  such that the following points are in  $\ell$  (See Figure 2.)

$$p, a(p), a^2(p), \dots, a^{4PKL}(p).$$

Set

$$x = a^{PKL}(p), y = a^{2PKL}(p) \in \ell.$$

Since  $y = a^{PKL}(x)$ , and by Lemma 3,  $d(x, y) \geq PKL\text{tr}(a) \geq KL \geq L$ .

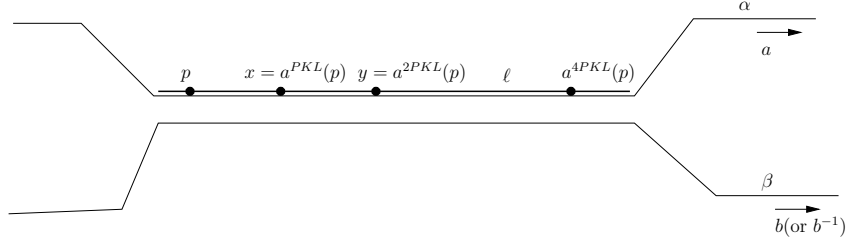


Figure 2: Apply the acylindricity to the pair  $x, y$ .

*Claim.* For each  $i(1 \leq i \leq PKL)$ ,

$$d(x, [b, a^i](x)) \leq 20\delta, d(y, [b, a^i](y)) \leq 20\delta.$$

We first consider the special case that  $\delta = 0$ , namely,  $\Gamma$  is a tree. Then,  $\alpha \cap_{10\delta} \beta$  coincides the segment  $\alpha \cap \beta$ , and also the segment  $\ell$ , therefore, all above points  $a^n(p), 1 \leq n \leq 4PKL$ , are in  $\alpha \cap \beta$ . We

want to show  $x = [b, a^i](x)$ , but this is obvious since when we apply  $a^i, b, a^{-i}$ , then  $b^{-1}$  to  $x$ , the point moves within  $\ell$ . Thus,  $[b, a^i](x) = x$ . If  $\delta > 0$ , we can show that the point moves in the  $10\delta$ -neighborhood of  $\ell$  when we apply  $a^i, b, a^{-i}$  followed by  $b^{-1}$  to  $x$ . Therefore, we get  $d(x, [b, a^i](x)) \leq 20\delta$  by estimating the error terms from the tree case using triangle inequality. We leave the details to readers. (See Figure 3.) We can show  $d(y, [b, a^i](y)) \leq 20\delta$  in the same way. We got the claim.

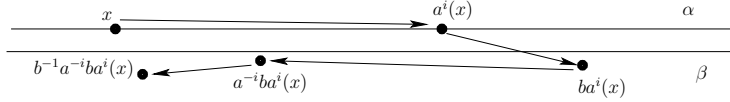


Figure 3: How commutators  $[b, a^i]$  act near  $\alpha \cap_{10\delta} \beta$ .

Since  $|x - y| \geq L = L(20\delta)$ , by the acylindricity of the action, it follows from the claim that there are at most  $K$  distinct elements in the set  $[b, a^i](1 \leq i \leq PKL)$ . By the pigeon-hall principle,  $[b, a^i] = [b, a^j]$  for some  $i \neq j, (1 \leq i, j \leq PKL)$ . It follows that  $a^i b^{-1} a^{-i} = a^j b^{-1} a^{-j}$ , therefore,  $a^{i-j} b^{-1} = b^{-1} a^{i-j}$ . We get  $[b, a^n] = 1$  for some  $n \neq 0$ . The same argument applies to the elements  $[a, b^i]$  since  $\text{tr}(b) \leq \text{tr}(a)$ , therefore we also get  $[a, b^n] = 1$  for some  $n \neq 0$  as well. This is a contradiction.  $\square$

Combining Proposition 2 and Lemma 5, we obtain the following. This says that we can find a global bound on one of the exponents, but the other bound depends on the ratio  $\text{tr}(a)/\text{tr}(b)$ .

**Proposition 6.** *Let  $G$  be a group which acts on a  $\delta$ -hyperbolic graph  $\Gamma$  acylindrically with constants  $K(R), L(R)$ . Then, there exists a constant  $N = N_{\mathcal{G}} \geq 1$  with the following property.  $N$  depends only on  $\delta, K(20\delta), L(20\delta)$  and  $K(200\delta)$ .*

*Suppose  $a, b \in G$  act hyperbolically with quasi-axes. Assume  $[a^n, b^m] \neq 1$  for all  $n, m \neq 0$ . Suppose there exists a number  $q \geq 1$  such that*

$$\text{tr}(a)/q \leq \text{tr}(b) \leq \text{tr}(a).$$

*Then, for any  $n \geq N$  and  $m \geq qN$ ,  $\langle a^n, b^m \rangle$  is free. Moreover, the embedding of  $\langle a^n, b^m \rangle$  in  $\Gamma$  is quasi-isometric.*

*Proof.* Put  $K = K(20\delta), L = L(20\delta)$ . It then suffices to set

$$N = 4P_3KL + 200(\delta + 1)P_3.$$

Remember that  $P_3$  depends only on  $\delta$  and  $K(200\delta)$ . By Lemma 5, since  $\text{tr}(b) \leq \text{tr}(a)$ , we have  $|\alpha \cap_{10\delta} \beta| < 4PKL\text{tr}(a) + 100\delta$ . Since  $\text{tr}(a) \leq q\text{tr}(b)$ , we have  $|\alpha \cap_{10\delta} \beta| < 4qPKL\text{tr}(b) + 100\delta$ . Therefore, if  $n \geq N$  and  $m \geq qN$ , then the conditions of Proposition 2 are satisfied for  $a^n, b^m$ , so that we conclude  $\langle a^n, b^m \rangle$  is free, and the embedding is quasi-isometric.  $\square$

## 2.4 Another condition for freeness

In this section we discuss the case when  $\text{tr}(b)/\text{tr}(a)$  is small as opposed to Proposition 6. We also apply Proposition 1 to a certain sequence of points, but we need a slightly different idea to construct the sequence from a given word  $w$ . The argument is elementary but lengthy, and takes most part of this section.

**Proposition 7.** *Suppose  $G$  acts on a  $\delta$ -hyperbolic graph  $\Gamma$  acylindrically with constants  $K(R), L(R)$ . Assume that  $f, g \in G$  act hyperbolically with quasi-axes  $\alpha, \beta$ . Assume that  $[f^s, g^t] \neq 1$  for any  $s, t \neq 0$ . Let  $D = |\alpha \cap_{10\delta} \beta|$ . Suppose that the following conditions are satisfied.*

1.  $\text{tr}(g) \leq \text{tr}(f)$ .
2.  $D \leq 2\text{tr}(f)$ .

*Then there exists a constant  $N = N\gamma > 0$  with the following property.  $N$  depends only on  $\delta, K(20\delta), L(20\delta)$  and  $L(200\delta)$ .*

*If  $n \geq N$ , then  $\langle g, f^n \rangle$  is free of rank two. Moreover, the embedding of  $\langle g, f^n \rangle$  in  $\Gamma$  by an orbit is quasi-isometric.*

A few remarks are in order before the lengthy proof. Unlike Proposition 6, the roles of two elements  $f$  and  $g$  are not symmetric. For example, in the conclusion we take powers only for  $f$ . Also, when we construct sequence of points  $p_i$  in the argument, we detect the action of  $a = f^n$  more closely than the action of  $b = g$ , since  $a$  moves a base point much more than  $b$  does. That is summarized as the condition (\*) in the claim. As usual, in Part 1, we first show the subgroup  $\langle g, f^n \rangle$  is free, by constructing a certain sequence of points  $p_i$  which satisfies the three points condition because of the condition (\*). Then in Part 2, we show that the embedding of the subgroup by an orbit is quasi-isometric by interpolating the points in the sequence  $p_i$  as before. Part 2 is most complicated in the paper.

*Proof.* First of all,  $D < \infty$  by our assumption. As in the proof for Proposition 2, the argument is slightly different if  $D = 0$  than the case

$D > 0$ , and from the view point of Proposition 1, it is easier than the case  $D > 0$ . So, we discuss the case  $D > 0$  in detail, and then discuss the case  $D = 0$  briefly.

Assume  $D > 0$ .

**Part 1.** As usual, we first prove that  $\langle g, f^n \rangle$  is free. Set  $K = K(20\delta), L = L(20\delta)$ . This is the only place where constants  $K(R), L(R)$  are used. Set

$$E = D + 100(\delta + 1) + 10P_3KL\text{tr}(f).$$

There exists a constant  $N$ , which depends only on  $K, L, P, \delta$ , such that if  $n \geq N$ , then  $\text{tr}(f^n) \geq 1000E$ . We used  $\text{tr}(f) \geq 1/P$  and  $2\text{tr}(f) \geq D$ . This is the only place the condition  $2\text{tr}(f) \geq D$  is used. Fix such constant  $N$ . Note that  $N \geq 10000$  by the definition of  $E$ .

For  $n \geq N$ , set  $a = f^n, b = g$ . We will show  $\langle a, b \rangle$  is free. Let  $w$  be a non-empty reduced word on  $a, b$ , and we show  $w$  is non-trivial in  $\text{Isom}(\Gamma)$ . There are three cases according to the form of  $w$ . In the case (O) it is clear that  $w \neq 1$  in  $\text{Isom}(\Gamma)$ .

(O)  $w = b^m (m \neq 0)$

(I)  $w = a^{n_1} b^{m_1} \dots a^{n_i} b^{m_i} (i \geq 1)$  such that  $n_1 \neq 0, n_i \neq 0$  and  $m_i$  is possibly 0,

(II)  $w = b^{m_0} a^{n_1} b^{m_1} \dots a^{n_i} b^{m_i} (i \geq 1)$  such that  $m_0 \neq 0, n_i \neq 0$  and  $m_i$  is possibly 0.

For (I) and (II), we can find a point  $p \in \Gamma$  such that  $|p - w(p)| > 0$ , therefore  $w$  is not 1 in  $\text{Isom}(\Gamma)$ . The argument is very similar to each other, so we only discuss the case (I) in detail.

Assume we are in the case (I). Let  $m$  be the mid point of a longest segment  $\ell$  which is contained in  $\alpha \cap_{10\delta} \beta$ . We can take a point  $x \in \alpha$  and a point  $y \in \beta$  such that  $|x - y| \leq 2\delta$  and  $|m - x|, |m - y| \leq 4\delta$ . It would be easier to follow the discussion if we imagine that  $\alpha$  and  $\beta$  coincide in  $\ell$  and that  $m = x = y$ , although there are actually errors of the order of  $\delta$ .

To show  $|x - w(x)| > 0$ , we interpolate  $x$  and  $w(x)$  by the following points, which gives a sequence satisfying the three points condition.

$$\begin{aligned}
s_0 &= y \\
p_1 &= x, q_1 = a^{n_1}(x), r_1 = a^{n_1}(y), s_1 = a^{n_1}b^{m_1}(y), \\
p_2 &= a^{n_1}b^{m_1}(x), q_2 = a^{n_1}b^{m_1}a^{n_2}(x), r_2 = a^{n_1}b^{m_1}a^{n_2}(y), s_2 = a^{n_1}b^{m_1}a^{n_2}b^{m_2}(y) \\
&\dots \\
p_i &= a^{n_1}b^{m_1} \dots a^{n_{i-1}}b^{m_{i-1}}(x), q_i = a^{n_1}b^{m_1} \dots a^{n_{i-1}}b^{m_{i-1}}a^{n_i}(x), \\
r_i &= a^{n_1}b^{m_1} \dots a^{n_{i-1}}b^{m_{i-1}}a^{n_i}(y), s_i = a^{n_1}b^{m_1} \dots a^{n_{i-1}}b^{m_{i-1}}a^{n_i}b^{m_i}(y), \\
p_{i+1} &= a^{n_{i-1}}b^{m_{i-1}}a^{n_i}b^{m_i}(x).
\end{aligned}$$

Now set for each  $i \geq j \geq 1$

$$A_j = [p_j, q_j], B_j = [r_j, s_j].$$

**Claim.** The following conditions are satisfied for all  $1 \leq j$ .

$$\left\{ \begin{array}{l}
(1). |q_j - r_j| \leq 2\delta, |p_j - s_{j-1}| \leq 2\delta. \\
(2). |A_j| \geq 1000E. \\
(3). |A_j \cap_{10\delta} B_j| \leq E. \\
(4). |B_{j-1} \cap_{10\delta} A_j| \leq E. \\
(5). |A_j \cap_{10\delta} A_{j+1}| \leq E.
\end{array} \right. \quad (*)$$

We verify those later, and proceed to show the set of the conditions (1)–(5) implies that the sequence  $\{p_j\}$  satisfy the three points condition for  $990E$ , therefore we get  $|p_1 - p_{i+1}| > 0$ , namely,  $|w(x) - x| > 0$ .

For  $j \geq 1$ , we define

$$C_j = [p_j, q_{j+1}], D_j = [p_j, p_{j+1}].$$

The conditions (3),(4),(5) imply that  $d(C_j, B_j) \leq 2E$  for all  $j$ . It follows using (3) and (4) that  $d(q_j, C_j) \leq 4E, d(p_{j+1}, C_j) \leq 4E$ . This is because, for each  $j$ , the geodesic quadrilateral with the corners  $p_j, q_j, p_{j+1}, q_{j+1}$  is  $2\delta$ -thin (see Figure 4, 5).

Therefore, for all  $1 \leq j \leq i$ ,  $|D_j \cap_{10\delta} D_{j+1}| \leq 5E$ . It implies that for all  $1 \leq j \leq i - 1$ ,

$$|p_j - p_{j+2}| \geq \max(|p_j - p_{j+1}|, |p_{j+1} - p_{j+2}|) + 990E.$$

We checked the three points condition for the constant  $990E$  and the sequence  $\{p_j\}$ , therefore by Proposition 1, since  $E > 100\delta$ , we get that  $|p_1 - p_{i+1}| > 0$ .



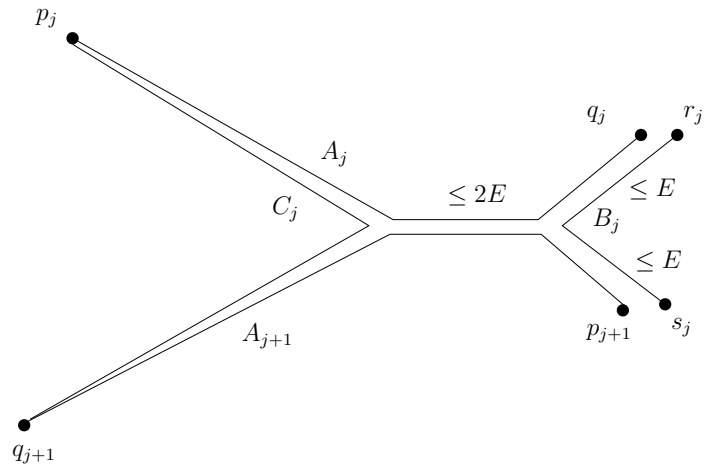


Figure 4:  $A_j \cap_{10\delta} A_{j+1} \neq \emptyset$

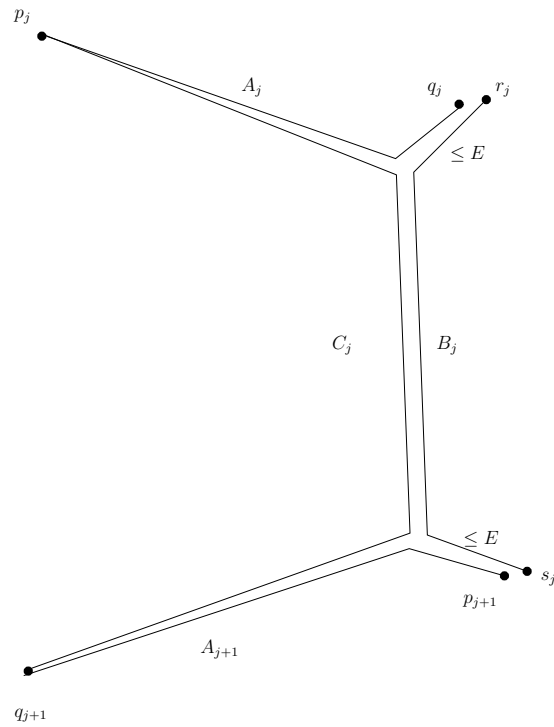


Figure 5:  $A_j \cap_{10\delta} A_{j+1} = \emptyset$

We are left to verify the condition (\*). (1) follows from  $|x-y| \leq 2\delta$ . Since  $n_j \neq 0$  and  $|A_j| = \text{tr}(a^{n_j})$ , we have  $|A_j| \geq \text{tr}(a) \geq 1000E$ . We get (2).

Both  $|A_j \cap_{10\delta} B_j|$  and  $|B_{j-1} \cap_{10\delta} A_j|$  are at most  $|\alpha \cap_{10\delta} \beta|$ , therefore  $\leq D \leq E$ . We get (3) and (4).

To show (5), suppose not, i.e.,  $|A_j \cap_{10\delta} A_{j+1}| > E$  for some  $j \geq 1$ . Set

$$w_{j-1} = a^{n_1} b^{m_1} \dots a^{n_{j-1}} b^{m_{j-1}}, \quad w_j = a^{n_1} b^{m_1} \dots a^{n_j} b^{m_j}$$

(If  $j = 1$ , then set  $w_0 = 1$ .) Then  $A_j$  is contained in  $w_{j-1}(\alpha)$ , which is a quasi-axis of  $w_{j-1} f w_{j-1}^{-1}$ , and also  $A_{j+1}$  is contained in  $w_j(\alpha)$ , which is a quasi-axis of  $w_j f w_j^{-1}$ . Let  $\ell = [u, v] \subset A_{j+1}$  be a geodesic of length  $E$  which is contained in  $A_j \cap_{10\delta} A_{j+1}$ . For concreteness, suppose the point  $u$  is mapped toward  $v$  by (sufficiently big) positive powers of  $w_{j-1} f w_{j-1}^{-1}$ , and the point  $v$  is mapped toward  $u$  by (sufficiently big) positive powers of  $w_j f w_j^{-1}$ . (Otherwise we take the inverse of the elements in the following.) The direction of the actions makes sense since the set  $A_j \cap_{10\delta} A_{j+1}$  looks like a long narrow tube, which contains  $\ell$ . Note that  $\text{tr}(w_j f w_j^{-1}) = \text{tr}(w_{j-1} f w_{j-1}^{-1}) = \text{tr}(f)$  and  $10\text{tr}(f)PKL \leq |\ell|$ . (See Figure 6.) We apply the following elements

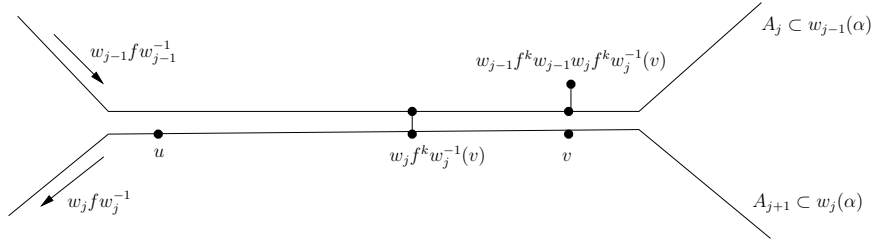


Figure 6:

to the point  $v$ .

$$w_{j-1} f w_{j-1}^{-1} w_j f w_j^{-1}, w_{j-1} f^2 w_{j-1}^{-1} w_j f^2 w_j^{-1}, \dots, w_{j-1} f^{PKL} w_{j-1}^{-1} w_j f^{PKL} w_j^{-1}.$$

Each of the elements moves  $v$  at most  $10\delta$ . Set  $v' = w_j f^{PKL} w_j^{-1}(v)$ . Then, each of the above elements also moves  $v'$  at most  $10\delta$ . On the other hand  $|v - v'| = PKL \text{tr}(f) \geq PKL/P \geq L$ . By the acylindricity, there must be  $1 \leq I < J \leq PKL$  such that

$$w_{j-1} f^I w_{j-1}^{-1} w_j f^I w_j^{-1} = w_{j-1} f^J w_{j-1}^{-1} w_j f^J w_j^{-1}$$

Since  $w_{j-1}^{-1}w_j = a^{n_j}b^{m_j}$ , this implies that  $[f^{I-J}, a^{n_j}b^{m_j}] = 1$ . Since  $a = f^n$ , it follows that  $[f^{I-J}, b^{m_j}] = 1$ . This is a contradiction (since  $b = g$ ). We showed (5). The case (I) is completed.

For the case (II), we show  $|y - w(y)| > 0$ . The argument is similar, and we omit the details. So far, we have shown that  $\langle g, f^n \rangle$  is free.

**Part 2.** Now, we show that the embedding of  $\langle g, f^n \rangle$  in  $\Gamma$  by an orbit is quasi-isometric. Our situation is same as in the proof of Proposition 2. Since the upper bound is trivial as before, we prove a lower bound. Note that it is enough to get a desired uniform lower bound for the case (O), (I) and (II) separately. The case (O) is trivial, and the argument is similar for (I) and (II). We only discuss the case (I) in detail.

As in the proof of Proposition 2, the reason why the above argument does not give a desired uniform lower bound in terms of  $|w|$  is that  $|p_{j+1} - p_j|$ , the length of  $D_j$ , is unbounded when we vary  $w$ . As before, we introduce interpolating points such that there is a uniform (for all  $w$ ) upper bound on the distance between two consecutive points, and the three points condition is satisfied for  $\varepsilon = E > 100\delta$ . Between  $p_j$  and  $q_j$  for each  $j$ , we define points using the action of  $f^n = a$  as follows:

if  $n_j > 0$ , define (see Figure 7)

$$\begin{aligned} p_j &= p_{j,0} = a^{n_1}b^{m_1} \dots a^{n_{j-1}}b^{m_{j-1}}(x), \\ p_{j,1} &= a^{n_1}b^{m_1} \dots a^{n_{j-1}}b^{m_{j-1}}a(x), \\ &\dots \\ p_{j,n_j} &= a^{n_1}b^{m_1} \dots a^{n_{j-1}}b^{m_{j-1}}a^{n_j}(x) = q_j. \end{aligned}$$

Note that for each  $j$ , the distance between any two consecutive points is  $\text{tr}(a)$ .

If  $n_j < 0$ , then define points similarly by  $p_{j,k} = a^{n_1}b^{m_1} \dots a^{n_{j-1}}b^{m_{j-1}}a^{-k}(x)$  for  $0 \leq k \leq -n_j$ .

Next, we also define points between  $r_j$  and  $s_j$ . In order to choose them with control on the distance between two consecutive new points, we fix an integer  $Q \geq 1$  such that

$$(10E \leq) \frac{\text{tr}(a)}{100} \leq \text{tr}(b^Q) \leq \frac{\text{tr}(a)}{50}.$$

Such  $Q$ , which depends on  $n$ , exists since  $\text{tr}(b) \leq \text{tr}(f) \leq \frac{\text{tr}(a)}{100}$ . This is the only place where the condition  $\text{tr}(g) \leq \text{tr}(f)$  is used. (Remember

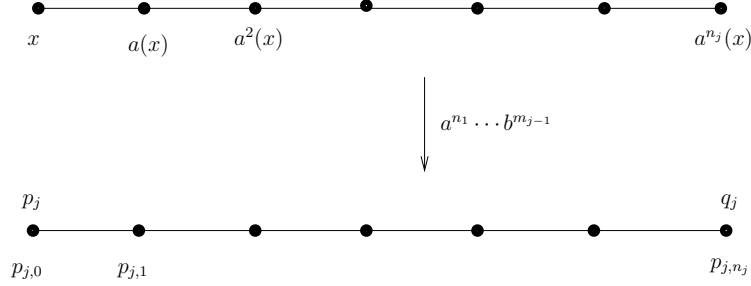


Figure 7: New points between  $p_j, q_j$  ( $n_j > 0$ )

$b = g$ .) Note that we use this  $Q$  for all  $j$ . To define points, we write (uniquely) for each  $j$ ,

$$m_j = o_j Q + l_j,$$

such that  $o_j \in \mathbb{Z}$  and that  $0 \leq l_j < Q$  if  $m_j \geq 0$  and  $-Q < l_j \leq 0$  if  $m_j < 0$ . For each  $j$ , define points as follows between  $r_j$  and  $s_j$ : if  $m_j \geq 0$  and  $o_j \geq 2$  then (see Figure 8)

$$\begin{aligned} r_j = r_{j,0} &= a^{n_1} b^{m_1} \dots a^{n_{j-1}} b^{m_{j-1}} a^{n_j}(y), \\ r_{j,1} &= a^{n_1} b^{m_1} \dots a^{n_{j-1}} b^{m_{j-1}} a^{n_j} b^Q(y), \\ &\dots \\ r_{j,o_j-1} &= a^{n_1} b^{m_1} \dots a^{n_{j-1}} b^{m_{j-1}} a^{n_j} b^{(o_j-1)Q}(y) \\ r_{j,o_j} &= a^{n_1} b^{m_1} \dots a^{n_{j-1}} b^{m_{j-1}} a^{n_j} b^{m_j}(y) = s_j. \end{aligned}$$

Here, if  $o_j \neq 0$ , then the distance between any two consecutive points is  $\text{tr}(b^Q)$  except for the last pair of points, and it is between  $\text{tr}(b^Q)$  and  $2\text{tr}(b^Q)$  for the last pair. If  $o_j = 1$ , we do not produce any new points. In this case,  $\text{tr}(b^Q) \leq |r_j - s_j| \leq 2\text{tr}(b^Q)$ . Also, if  $o_j = 0$ , then we do not produce any new points. In that case,  $|r_j - s_j| < \text{tr}(b^Q)$ .

If  $m_j < 0$ , we define points similarly, by  $r_{j,k} = a^{n_1} b^{m_1} \dots a^{n_{j-1}} b^{m_{j-1}} a^{n_j} b^{-kQ}(y)$  for  $0 \leq k \leq -o_j - 1$  if  $-o_j \geq 2$ . Otherwise we do not produce new points.

In this way, we obtain a new set of points  $\{p_{j,k}, r_{j,k}\}$  with the order which it naturally inherits from the order on the sequence

$$p_1, q_1, r_1, s_1, \dots, p_i, q_i, r_i, s_i.$$

The distance of any two consecutive points in this new sequence is at most  $\text{tr}(a)$ . But as a trade-off, the new sequence may not satisfy the

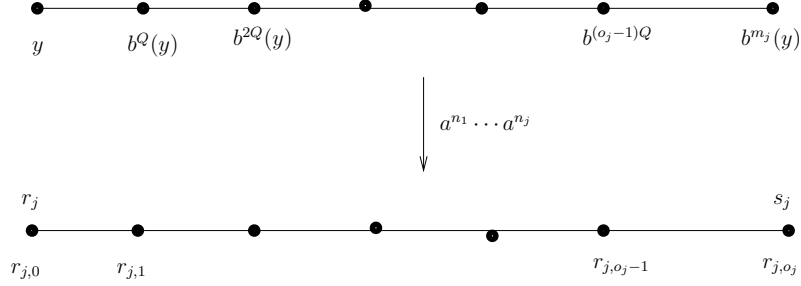


Figure 8: New points between  $r_j, s_j$  ( $m_j > 0$ )

three points condition any more. Therefore we modify the sequence by removing points. First, remove all  $r_j$ . ( $r_j$  is at most  $2\delta$ -close to  $q_j$ .) Next, remove all  $s_j$  except for the last point  $s_i$ . ( $s_j$  is at most  $2\delta$ -close to  $p_{j+1}$ .) Finally, if  $o_j = 0$ , then remove  $q_j$ . We get a subsequence of points with order, which we denote by  $\{u_k\}$ .

See Figure 9,10,11. In those figures, two consecutive points  $u_k, u_{k+1}$  are joined by a solid line, where an interval by a thin solid line appears after we remove points while an interval by a thick solid line exists before we remove points. The intervals of dashed line do not exist because we removed points. Removed points are described by white dots in the figures.

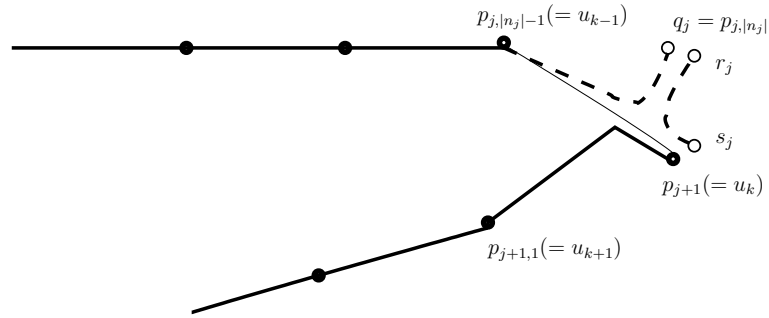


Figure 9:  $q_j, r_j, s_j$  are removed if  $o_j = 0$

Now, we argue that the sequence  $\{u_k\}$  satisfies the three points condition of Proposition 1, and that there is a uniform upper bound on the distance between any two consecutive points in the sequence. In the following argument, it would help to keep in mind the following

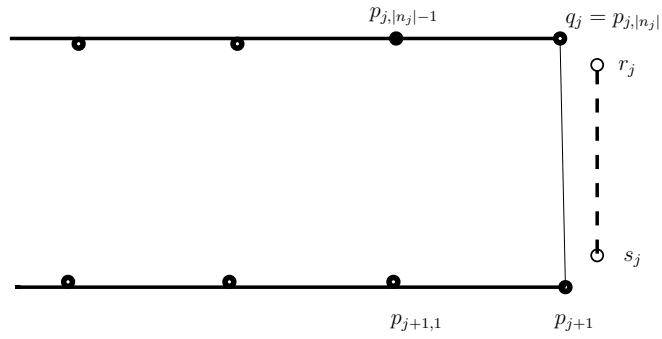


Figure 10:  $r_j, s_j$  are removed if  $o_j = 1$

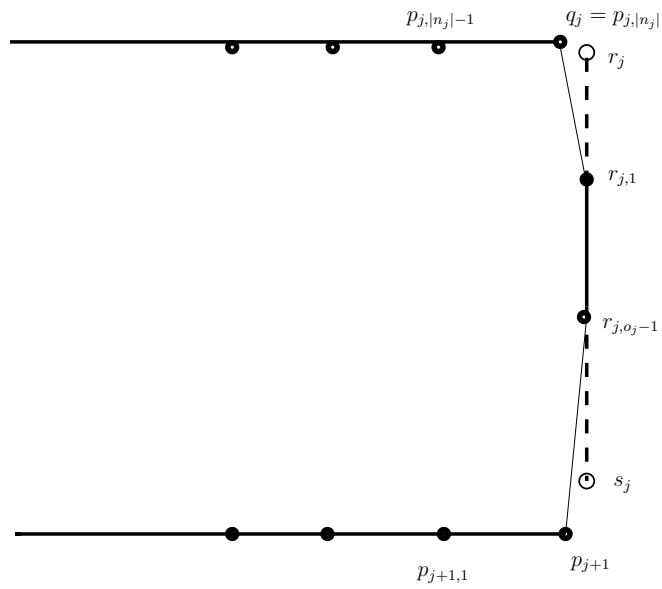


Figure 11:  $r_j, s_j$  are removed if  $o_j \geq 2$

rough estimate of constants.

$$\delta \ll E \ll \text{tr}(b^Q) \ll \text{tr}(a).$$

First, we remark that we removed points such that the distance of any two consecutive points in  $\{u_k\}$  is in one of the following intervals, in particular there is a uniform upper bound:

- (i) between  $\frac{4\text{tr}(a)}{5}$  and  $\frac{6\text{tr}(a)}{5}$
- (ii) between  $\frac{\text{tr}(a)}{100}$  and  $\frac{\text{tr}(a)}{10}$ .

Moreover,

- (iii) the sequence of intervals (between two consecutive points) starts with  $|n_1|$  intervals of length (i), followed by  $|o_1|$  intervals of length (ii), followed by  $|n_2|$  intervals of length (i), followed by  $|o_2|$  intervals of length (ii),  $\dots$ ,  $|n_i|$  intervals of length (i), then  $|o_i|$  intervals of length (ii) at the end. In particular, each interval has length  $\geq \frac{\text{tr}(a)}{100} \geq 10E \geq 1000$ .

This is because removing  $r_j$  and  $s_j$  changes the distance between two consecutive points at most  $4\delta$ . And removing  $q_j$  happens only when  $o_j = 0$ , therefore  $|r_j - s_j| \leq \text{tr}(b^Q) \leq \frac{\text{tr}(a)}{50}$ . In conclusion, the distance between any two consecutive points in the subsequence is at most  $\text{tr}(a) + 4\delta + \frac{\text{tr}(a)}{50} \leq \frac{6}{5}\text{tr}(a)$ .

Next, we argue that the sequence  $\{u_k\}$  satisfies the three points condition of Proposition 1 for the constant  $E$ . This is by  $\delta$ -hyperbolicity, and essentially by the same reason as the sequence  $D_j$  satisfies the condition in the previous discussion (see Figure 4, 5). If  $o_j \geq 1$  (Figure 10, 11), the three points (in other words, two consecutive intervals) condition is nearly obvious since an interval of thin line is  $10\delta$ -neighborhood of the corresponding thick dashed line interval. The less obvious case is when  $o_j = 0$ . See Figure 9. We explain why the three points condition is satisfied for  $u_{k-1} = p_{j,|n_j|-1}$ ,  $u_k = p_{j+1}$ ,  $u_{k+1} = p_{j+1,1}$ . It suffices to show that  $|[u_{k-1}, u_k] \cap_{10\delta} [u_k, u_{k+1}]|$  is at most  $\frac{1}{2} \min(|u_{k-1} - u_k|, |u_k - u_{k+1}|) - E$ , which is  $\geq \frac{1}{2} \frac{4}{5} \text{tr}(a) - E \geq \frac{39}{100} \text{tr}(a)$ . First,  $[u_{k-1}, u_k]$  is contained in the  $10\delta$ -neighborhood of  $[u_{k-1}, q_j] \cup [r_j, s_j]$  by  $\delta$ -hyperbolicity. Therefore,  $[u_{k-1}, u_k] \cap_{10\delta} [u_k, u_{k+1}]$  is contained in the  $10\delta$ -neighborhood of  $([u_{k-1}, q_j] \cup [r_j, s_j]) \cap_{10\delta} [u_k, u_{k+1}]$ , which is contained in  $20\delta$ -neighborhood of  $([u_{k-1}, q_j] \cap_{10\delta} [u_k, u_{k+1}]) \cup [r_j, s_j]$ . But we know that  $|[u_{k-1}, q_j] \cap_{10\delta} [u_k, u_{k+1}]| \leq E \leq \frac{\text{tr}(a)}{1000}$  and  $|r_j - s_j| \leq \frac{\text{tr}(a)}{50}$ . The first inequality is by one of the conditions we obtained, which is  $|A_j \cap_{10\delta} A_{j+1}| \leq E$  for all  $j$ . Therefore, since  $|r_j - q_j| \leq 2\delta$ , the diameter of the union of those two sets is smaller than  $\frac{39}{100} \text{tr}(a)$ , which

shows our claim. We verified the three points condition for  $E$  for the sequence  $\{u_k\}$ , which starts at  $x$  and ends at  $w(x)$ .

By Proposition 1 (for the first inequality below), we get

$$\begin{aligned} \lambda|w(x) - x| &\geq \sum_k |u_k - u_{k+1}| \geq 1000 \sum_{j=1}^i (|n_j| + |o_j|) \geq \sum_{j=1}^i 500(|n_j| + |o_j| + 1) \\ &\geq \frac{500}{Q} \sum_{j=1}^i (|n_j| + |m_j|) = \frac{500}{Q}|w|, \end{aligned}$$

where  $\lambda = (\frac{E}{100} - \delta)^{-1} \max_k |u_k - u_{k+1}|$ . The second inequality follows from the remark (iii) above. We have the third inequality because  $|n_j| \geq 1$  for all  $j$ , and the fourth one because  $Q(|o_j| + 1) \geq |m_j|$  for all  $j$ . Moreover, since  $|u_k - u_{k+1}|$  is at most  $2\text{tr}(a)$ , we have  $\lambda \leq (\frac{E}{100} - \delta)^{-1} 2\text{tr}(a) = \lambda_0$  for all  $w$ . Set  $L = \frac{Q\lambda_0}{500}$ , then we get  $L|w(x) - x| \geq |w|$  for all  $w$ .  $L$  depends on  $a, b$  but not on  $w$  since so does  $Q$ . We completed the case (I). Note that we showed that the embedding is bi-Lipschitz for the point  $x$  and the word metric  $|w|$  for the collection of words  $w$  of the case (I).

The case (II) is similar, and we omit the details. We finished the argument under the assumption that  $D > 0$ .

Finally, assume that  $D = 0$ . The argument will be essentially same as the case that  $D > 0$ , so we discuss only the part which is different (see the proof of Proposition 2. We discussed the case  $D = 0$  then the case  $D > 0$  as well). Take the constants  $K, L, P, E, N$  as before. Set  $a, b$  as before. Now let  $\ell$  be a geodesic which realizes the distance between  $\alpha, \beta$ . Let  $m, x, y$  be the midpoint of  $\ell$ . Then argue as before in the rest. We define segments  $A_j, B_j$  which satisfy the conditions (1)-(5). (In this case,  $q_j = r_j$  and  $p_j = s_{j-1}$ .) This case is easier in the sense that since  $\alpha \cap_{10\delta} \beta = \emptyset$ , we have  $A_j \cap_{10\delta} B_j = \emptyset, B_j \cap_{10\delta} A_{j+1} = \emptyset$ , therefore the union of those geodesics, which we need to analyze in terms of metric, looks nearly like a tree with those segments as edges. We omit the details.  $\square$

Before we state a main theorem, we state a proposition which can be shown similarly to Proposition 7. The conclusion is weaker since  $N_8$  depends on  $f, g$ , but we do not require the conditions 1 and 2 in Proposition 7 regarding  $\text{tr}(f), \text{tr}(g)$ .

**Proposition 8.** *Suppose  $G$  acts on a  $\delta$ -hyperbolic graph  $\Gamma$  acylindrically with constants  $K(R), L(R)$ . Assume that  $f, g \in G$  act hyperbolically with quasi-axes  $\alpha, \beta$ . Assume that  $[f^s, g^t] \neq 1$  for any  $s, t \neq 0$ .*



Then there exists a constant  $N_g = N > 0$ , which depends on  $f, g$ , such that if  $n \geq N$ , then  $\langle g, f^n \rangle$  is free of rank two. Moreover, the embedding of  $\langle g, f^n \rangle$  in  $\Gamma$  by an orbit is quasi-isometric.

*Proof.* The argument is very similar to the proof of Proposition 7, and easier since the difficult part was to obtain a uniform constant  $N_7$  for all  $f, g$ . We only indicate where we need to modify the argument.

Choose constants  $K, L, P$  in the same way. Let  $D = |\alpha \cap_{10\delta} \beta|$ . By Lemma 5,  $D < \infty$ . Define the constant  $E$  in the same way. Now fix a constant  $N > 0$  such that if  $n \geq N$ , then

$$\mathrm{tr}(f^n) \geq 1000E, \text{ and } \mathrm{tr}(g) \leq \frac{\mathrm{tr}(f^n)}{100}.$$

The constant  $N$  depends on  $f, g$  as opposed to Proposition 7, where we could choose  $N_7$  uniformly on  $f, g$  to have those two inequalities for all  $n \geq N_7$  because of the conditions 1 and 2. Note that the condition 2 (and condition 1) in Proposition 7 was used only to choose  $N_7$  uniformly on  $f, g$  such that we have the first (and the second, respectively) inequality in the above for all  $n \geq N_7$ .

Let  $n \geq N$  be a constant. Then we have the two inequalities in the above. Set  $a = f^n, b = g$  as before. Then, using the first inequality, we can show that  $\langle a, b \rangle$  is free exactly same as for Proposition 7. What is essential in the argument is that  $\mathrm{tr}(a)$  is much larger than  $D$ , the  $10\delta$ -overlap of the quasi-axes of  $a, b$ . (Remember that  $E \geq D$ .) We remark that the condition 1 was irrelevant until this part since it does not matter even if  $\mathrm{tr}(g) = \mathrm{tr}(b)$  is much larger than  $\mathrm{tr}(f)$  or  $\mathrm{tr}(a) = \mathrm{tr}(f^n)$ .

To show that the embedding is quasi-isometric for Proposition 7, we chose a constant  $Q$  such that  $\frac{\mathrm{tr}(a)}{100} \leq \mathrm{tr}(b^Q) \leq \frac{\mathrm{tr}(a)}{50}$ . This was possible since  $\mathrm{tr}(b) \leq \frac{\mathrm{tr}(a)}{100}$ . (For this we used the condition 1.) This is exactly the second inequality in the above, and we have chosen  $N$  to have this inequality if  $n \geq N$ . With this  $Q$  we apply the same argument for the rest. We omit details.  $\square$

## 2.5 Upper bound on both exponents

The following is the main theorem of Section 2.

**Theorem 9.** *Suppose  $G$  acts acylindrically for constants  $K(R), L(R)$  on a  $\delta$ -hyperbolic graph  $\Gamma$ . Then there exists a constant  $M_g$ , which depends only on  $\delta$  and  $K(20\delta), L(20\delta), L(200\delta)$  with the following property.*

Suppose  $a, b \in G$  act hyperbolically with quasi-axes  $\alpha, \beta$ . Assume for any  $p, q \neq 0$ ,  $[a^p, b^q] \neq 1$  in  $G$ . Then for any  $n, m \geq M$ ,  $\langle a^n, b^m \rangle$  is free of rank two. Moreover, the embedding of  $\langle a^n, b^m \rangle$  by an orbit in  $\Gamma$  is quasi-isometric. In particular, all non-trivial elements in  $\langle a^n, b^m \rangle$  are hyperbolic on  $\Gamma$ .

*Proof.* Set  $K = K(20\delta)$ ,  $L = L(20\delta)$ . Take constants  $P_3$ ,  $N_6$  and  $N_7$  by Lemma 3, Proposition 6 and Proposition 7. Set

$$M = 10KLP_3N_6 + 2000(\delta + 1)P_3 + N_7.$$

$M$  depends only on  $\delta$  and  $K(20\delta)$ ,  $L(20\delta)$ ,  $L(200\delta)$  and does not depend on  $a, b$ . It suffices to show

**Claim.**  $\langle a^n, b^m \rangle$  is free and the embedding to  $\Gamma$  is quasi-isometric if  $n, m \geq M$ .

To show this, set  $D = |\alpha \cap_{10\delta} \beta|$ . Without loss of generality, we may assume that  $\text{tr}(b) \leq \text{tr}(a)$ . By Lemma 5,  $D < 4PKL\text{tr}(a) + 100\delta < \infty$ .

*Step 1.* We may assume  $\frac{\text{tr}(b)}{\text{tr}(a)} \leq \frac{N_6}{M}$ .

This is because otherwise we can show Claim as follows. Assume  $\frac{\text{tr}(b)}{\text{tr}(a)} > \frac{N_6}{M}$ . Set  $q = \frac{M}{N_6}$ , and apply Proposition 6 to  $a, b$ . Then, if  $n \geq N_6$  and  $m \geq qN_6 = M$ , then  $\langle a^n, b^m \rangle$  is free and the embedding is quasi-isometric. Since  $M > N_6$ , we get Claim.

Note that it follows that  $10KLP\text{tr}(b) \leq \text{tr}(a)$  since  $\frac{N_6}{M} < \frac{1}{10KLP}$ .

*Step 2.* We may assume  $M\text{tr}(b) \leq D + 100(\delta + 1)$ .

This is because otherwise we get Claim as follows. Assume the contrary. It immediately follows that both  $\text{tr}(a^n)$ ,  $\text{tr}(b^m)$  are  $> D + 100(\delta + 1)$ . (We use  $\text{tr}(b) \leq \text{tr}(a)$ .) Then by Proposition 2, we get Claim.

Note that since  $10KL + 2000(\delta + 1) \leq M\text{tr}(b)$  by the way we chose  $M$ , it follows that  $10KL + 1000\delta \leq D$ .

*Step 3.* We have  $D \leq 2\text{tr}(a)$  (assuming the inequalities in Step 1 and 2).

To argue by contradiction, assume  $D > 2\text{tr}(a)$ . Then we get a contradiction using the same idea as for Lemma 5 concerning the action of commutators  $[b^i, a]$ . Since  $D \geq 1000\delta$  by Step 2, the set  $\alpha \cap_{10\delta} \beta$  looks like a narrow tube. Therefore, it makes sense to talk about the direction of the action by  $a$  and  $b$  along this tube, and furthermore, the direction of  $a$  coincides the direction of one of  $b$  or  $b^{-1}$ . In the following, we assume that the actions by  $a, b$  have the same direction along  $\alpha \cap_{10\delta} \beta$ , otherwise, we consider  $b^{-1}$  instead of  $b$ .

Let  $\ell = [p, p'] \subset \alpha$  be the longest segment contained in  $\alpha \cap N_{2\delta}(\beta)$  such that  $a$  moves  $p$  toward  $p'$ , i.e.,  $a(p) \in [p, p']$ . We know  $|\ell| \geq D - 100\delta$ . Since  $D \geq 1000\delta$  by Step 2, it follows  $|\ell| \geq \frac{9}{10}D$ .

We claim that for all  $1 \leq i \leq PKL$ , we have  $d(p, [b^i, a](p)) \leq 20\delta$ . If  $\delta = 0$  then this is obvious. Indeed, first of all,  $\alpha$  and  $\beta$  coincide in  $\ell$  in this case. Also, if we apply  $a, b^i, a^{-1}$  then  $b^{-i}$  in this order to  $p$ , the point moves in  $\ell$ . This is because since  $10KLP\text{tr}(b) \leq \text{tr}(a)$  by Step 1 and  $D > 2\text{tr}(a)$ , we have  $\text{tr}(a) + \text{tr}(b^i) \leq \frac{11}{10}\text{tr}(a) \leq \frac{11}{20}D \leq |\ell|$  for all  $1 \leq i \leq PKL$ . Now it is trivial that  $p = [b^i, a](p)$ . If  $\delta > 0$ , when we apply  $a, b^i, a^{-1}$  then  $b^{-i}$  to  $p$ , the point moves in the  $10\delta$ -neighborhood of  $\ell$ , and we get  $d(p, [b^i, a](p)) \leq 20\delta$  for all  $1 \leq i \leq PKL$ . (see Fig. 3 with the roles of  $a$  and  $b$  exchanged.)

Let  $q \in \ell$  be the point with  $L = d(p, q)$ . (Note that  $L \leq \frac{D}{10}$  by Step 2.) By the same reason as for  $p$ , we have  $d(q, [b^i, a](q)) \leq 20\delta$  for all  $1 \leq i \leq PKL$ . Now by the acylindricity, since  $d(p, q) = L = L(20\delta)$ , there must be  $1 \leq I < J \leq PKL$  such that  $[b^I, a] = [b^J, a]$ , therefore  $[b^{I-J}, a] = 1$ , which gives a contradiction. This is the end of Step 3.

But if  $D \leq 2\text{tr}(a)$ , we can apply Proposition 7 to  $a, b$ . Therefore, if  $n \geq N_7$ , then  $\langle b, a^n \rangle$  is free and the embedding is quasi-isometric. Since  $M > N_7$ , we have shown the Claim. Note that  $\langle b^m, a^n \rangle$  is a subgroup of  $\langle b, a^n \rangle$ .  $\square$

*Remark 10.* It is more difficult to deal with the normal subgroup generated by  $a^n, b^m$ , or even just by  $a^n$  (see Question 11 [11]). See the work of Delzant [2].

*Remark 11.* Theorem 9 is regarding two elements, but one can ask if there exists a constant  $M$  such that if  $a, b, c \in G$  are hyperbolic elements with certain condition (for example, pairwise independence), then  $\langle a^\ell, b^m, c^n \rangle$  is free for any  $\ell, m, n \geq M$ . We remark that the rank of the free subgroup may not be three. Take two hyperbolic elements  $a, b \in G$  which satisfy the commutator assumption in Theorem 9 (i.e., independent). For any  $M > 0$ , set  $c = a^M b a^{-M}$ . Then, the pairs  $a, c$  and  $b, c$  are also independent, but  $\langle a^M, b^M, c^M \rangle$  is equal to  $\langle a^M, b^M \rangle$ .

## 2.6 Hyperbolic isometries without quasi-axes

So far, we have been discussing hyperbolic isometries with quasi-axes. The existence of quasi-axes, which are geodesics by our definition, is a restriction but indeed not really necessary for our arguments since we can use certain quasi-geodesics and modify the original arguments. We discuss this issue in this section. Readers may skip this section

since we do not use this for our main application to mapping class groups in Section 3.

A path  $\alpha$  parametrized by the arc-length is called a  $(K, \varepsilon)$ -quasi-geodesic for  $0 < K \leq 1$  and  $0 \leq \varepsilon$  if for all  $t, s$  we have  $K|t - s| - \varepsilon \leq d(\alpha(t), \alpha(s))$ .

The following fact is elementary (see [4] for details).

*Fact 12* (quasi-geodesic axis). If  $a$  is a hyperbolic isometry of a  $\delta$ -hyperbolic graph  $\Gamma$ , there exists a  $(K, \varepsilon)$ -quasi-geodesic  $\alpha$  for some  $K, \varepsilon$  such that

1.  $a^n(\alpha)$  and  $\alpha$  are in the  $30\delta$ -neighborhood of each other for any  $n$ . (Namely,  $\alpha$  is almost invariant by  $a$ .)
2. Let  $p, q \in \alpha$ . Then the subpath of  $\alpha$  between  $p, q$  and a geodesic  $[p, q]$  are in the  $10\delta$ -neighborhood of each other.

We call such path  $\alpha$  as a *quasi-geodesic axis* of  $a$  in this paper. To be precise, we should use the term *quasi-geodesic quasi-axis*, but we make it shorter.

One can easily show from (1) and (2) that any two quasi-geodesic axes of  $a$  are in the  $30\delta$ -neighborhood of each other. Note that (2) is concerning only the path, but not the element  $a$ . Also, the quasi-geodesic constants of  $\alpha$  are not important for our purpose. What is useful for us is (2).

For example, Lemma 3 gives a uniform positive lower bound of  $\text{tr}(a)$  for all hyperbolic isometry  $a \in G$  with a quasi-axis if the action of  $G$  is acylindrical, but, indeed the assumption on the existence of quasi-axes is redundant. The proof of Lemma 3 easily generalizes by using quasi-geodesic axes instead of quasi-axes (see [4] for the precise argument), and we obtain the following.

**Lemma 13.** *Suppose  $G$  acts on a  $\delta$ -hyperbolic graph  $\Gamma$ . If the action is acylindrical with constants  $K(R)$  and  $L(R)$ , then there exists an integer  $P \geq 1$  such that for any element  $a \in G$  which acts hyperbolically on  $\Gamma$ , we have  $\text{tr}(a^P) \geq 1$ . The constant  $P$  depends only on  $\delta$  and  $K(200\delta)$ .*

Of course, this constant  $P$  is maybe larger than  $P_3$ . The existence of such  $P$ , but not the actual number, is essential for our argument.

We restate Theorem 9 in the following form for a potential application. The only difference is that we do not assume that there are quasi-axes for  $a$  and  $b$ .

**Theorem 14.** *Suppose  $G$  acts acylindrically for constants  $K(R), L(R)$  on a  $\delta$ -hyperbolic graph  $\Gamma$ . Then there exists a constant  $M'_{14}$ , which depends only on  $\delta$  and  $K(20\delta), L(20\delta), L(200\delta)$ , with the following property.*

*Suppose  $a, b \in G$  act hyperbolically. Assume for any  $p, q \neq 0$ ,  $[a^p, b^q] \neq 1$  in  $G$ . Then for any  $n, m \geq M'$ ,  $\langle a^n, b^m \rangle$  is free of rank two. Moreover, the embedding of  $\langle a^n, b^m \rangle$  by an orbit in  $\Gamma$  is quasi-isometric. In particular, all non-trivial elements in  $\langle a^n, b^m \rangle$  are hyperbolic on  $\Gamma$ .*

*Proof.* The proof is very similar to the one for Theorem 9. Basically, we use quasi-geodesic axes instead of quasi-axes for hyperbolic isometries. The proof of Theorem 9 relies on Proposition 2, Lemma 3, Lemma 5, Proposition 6 and Proposition 7. We have already generalized Lemma 3 to Lemma 13. We modify the statement and the proof of each of the other ones, which we only outline here.

As for Proposition 2, replace quasi-axes  $\alpha, \beta$  by quasi-geodesic axes  $\alpha, \beta$ . Accordingly, replace all  $\alpha \cap_{10\delta} \beta$  by  $\alpha \cap_{1000\delta} \beta$  in the statement. Then the original proof works with minor modification using the properties (1) and (2) in Fact 12 of the quasi-geodesic axes.

In Lemma 5, replace quasi-axes  $\alpha, \beta$  for  $a, b$  by quasi-geodesic axes in the assumption. We also replace the inequality in the conclusion by

$$|\alpha \cap_{1000\delta} \beta| < 4P_{13}K(20\delta)L(20\delta) \max(\text{tr}(a), \text{tr}(b)) + 10000\delta.$$

Then, the proof is same after an appropriate modification regarding constants.

As for Proposition 6, Proposition 7, replace quasi-axes  $\alpha, \beta$  by quasi-geodesic axes. Use  $\alpha \cap_{1000\delta} \beta$  instead of  $\alpha \cap_{10\delta} \beta$  in the statement. Then the original proof works with minor modification. Having done them all, we modify the proof of Theorem 9 to fit our setting. Of course, we always use the constant  $P_{13}$  instead of  $P_3$ . The constant  $M'_{14}$  is maybe larger than  $M_9$ . We omit details.  $\square$

*Remark 15.* Proposition 8 also holds if we drop the assumption on the existence of quasi-axes  $\alpha, \beta$ . The argument is also very similar to the original one.

### 3 Application to mapping class group

We discuss mapping class groups in this section. We apply results from Section 2 to pseudo-Anosov elements. Theorem 16 and 17 are main results of the paper.

#### 3.1 Uniform estimate

We apply Theorem 9 to the mapping class group,  $\text{Mod}(S)$ , of a compact orientable surface  $S$ . Let  $\mathcal{C}(S)$  be the curve graph of  $S$  (see for example [10], [16] for the definition). Masur-Minsky [16] showed that  $\mathcal{C}(S)$  is  $\delta$ -hyperbolic and an element  $a \in \text{Mod}(S)$  is pseudo-Anosov if and only if it acts as a hyperbolic isometry on  $\mathcal{C}(S)$ , and moreover ([1]) there always exists a quasi-axis. Bowditch [1] showed that the action is acylindrical.

For a subgroup  $G < \text{Mod}(S)$ , Farb-Mosher [3] introduced the notion of *convex-cocompact*. It has been shown ([6], [12]) that  $G$  is convex-cocompact iff for a point  $c \in \mathcal{C}(S)$ , the map from  $G$  to  $\mathcal{C}(S)$  sending  $g$  to  $g(c)$ , namely the embedding by an orbit, is quasi-isometric.

The following is an immediate consequence of Theorem 9. Apply it to the action of  $\text{Mod}(S)$  to  $\mathcal{C}(S)$ . Two pseudo-Anosov elements  $a, b$  are called *independent* if  $[a^n, b^m] \neq 1$  for any  $n, m \neq 0$  (cf.[10]).

**Theorem 16.** *Let  $S$  be a compact orientable surface, and  $\text{Mod}(S)$  its mapping class group. Then there exists a constant  $M(S)$  with the following property. Suppose  $a, b \in \text{Mod}(S)$  are pseudo-Anosov elements such that  $[a^n, b^m] \neq 1$  for any  $n, m \neq 0$ . Then for any  $n, m \geq M$ ,  $\langle a^n, b^m \rangle$  is free of rank two, and convex-cocompact. In particular all non-trivial elements in  $\langle a^n, b^m \rangle$  are pseudo-Anosov.*

#### 3.2 Non-uniform estimate and example

Let  $a, b \in \text{Mod}(S)$  be two independent pseudo-Anosov elements. It would be interesting to know for which  $(n, m)$ ,  $\langle a^n, b^m \rangle$  is free of rank two, and convex-cocompact. The following theorem says that it is the case except for finitely many  $(n, m)$ . We do not know if the number of the exceptional pairs is bounded.

**Theorem 17.** *Let  $S$  be a compact orientable surface and  $a, b$  two independent pseudo-Anosov elements. Then there exists  $N_{1\gamma} = N$ ,*

which depends on  $a, b$ , such that for any  $n \geq N$ , both  $\langle a, b^n \rangle$  and  $\langle b, a^n \rangle$  are free of rank two, and convex-cocompact. In particular,  $\langle a^n, b^m \rangle$  is free of rank two, and convex-cocompact if  $|n| + |m| \geq 2N$  and  $nm \neq 0$ .

*Proof.* Apply Proposition 8 to  $a, b$  for the action on  $\mathcal{C}(S)$ . The constant  $N_8$  will do.  $\square$

The constant  $N_{17}$  must depend on  $a, b$  as the following example shows.

*Example 18.* Let  $S$  be a compact orientable surface which is not a sphere with less than four punctures or a torus. If  $n > 0$  is sufficiently large, then there exist two independent pseudo-Anosov elements  $f, g \in \text{Mod}(S)$  such that  $\langle g, f^n \rangle$  is not free.

To see this, take  $f, a \in \text{Mod}(S)$  such that  $f$  is pseudo-Anosov,  $a$  is non-trivial torsion and that  $\langle f, a \rangle$  is not virtually cyclic. To find such  $a, f$ , first take a non-trivial torsion element  $a \in \text{Mod}(S)$  such that there is a non-trivial and non-peripheral simple closed curve  $\sigma$  on  $S$  which is not homotopic to  $a(\sigma)$ . One can find such  $a$  easily. Then one can find a desired  $f$ . For example, take any pseudo-Anosov element  $h$  on  $S$ . Let  $d$  be a Dehn-twist along  $\sigma$ . Set  $f = d^m h d^{-m}$ . We choose a sufficiently large  $m > 0$  later. It is clear that  $f$  is pseudo-Anosov, and the two laminations which are invariant by  $f$ , which we regard as a set of two points,  $\text{fix}(f)$ , in the boundary of the Teichmüller space of  $S$ , must be moved by  $a$  (i.e.  $\text{fix}(f) \cap a(\text{fix}(f)) = \emptyset$ ) if  $m$  is sufficiently large. For such  $m$ , it follows by a standard argument that  $\langle f, a \rangle$  is not virtually cyclic (cf. [9]).

Now, for sufficiently large  $n$ ,  $f^n a$  is pseudo-Anosov, and independent from  $f$ . One can show this using the curve graph of  $S$ ,  $\mathcal{C}(S)$ , which is  $\delta$ -hyperbolic. If necessary, replace  $f$  by some power of it in advance, and we may assume that  $f$  leaves a geodesic  $\gamma$  in  $\mathcal{C}(S)$  invariant. By our assumption  $\gamma \cap_{10\delta} a(\gamma)$  is bounded. For each  $n > 0$ , one can find a line which is invariant by  $f^n a$  using a piece of  $\gamma$ , a fundamental domain for the action of  $f^n$ , and the action of  $a$ . Then, for sufficiently large  $n$ , using  $\delta$ -hyperbolic geometry of  $\mathcal{C}(S)$ , one can show that the line is indeed a quasi-geodesic, therefore  $f^n a$  is pseudo-Anosov. Moreover, for sufficiently large  $n$ , the quasi-geodesic has two points at infinity of  $\mathcal{C}(S)$  which are disjoint from the two points for  $\gamma$ . It implies that  $f^n a$  and  $f$  are independent. Set  $g = f^n a$ . Then  $\langle g, f^n \rangle$  is not free since it contains the torsion element  $a$ .

## References

- [1] B. Bowditch, Tight geodesics in the curve complex. *Invent. Math.* 171 (2008), no. 2, 281–300.
- [2] T. Delzant, Sous-groupes distingués et quotients des groupes hyperboliques. *Duke Math. J.* 83 (1996), no. 3, 661–682.
- [3] B. Farb, L. Mosher, Convex cocompact subgroups of mapping class groups. *Geom. Topol.* 6 (2002), 91–152
- [4] K. Fujiwara, Subgroups generated by two pseudo-Anosov elements in a mapping class group. I. Uniform exponential growth, (previously “Subgroups generated by two elements in a mapping class group”). In *Groups of Diffeomorphisms*, 283–296, Advanced Studies in Pure Mathematics 52, Mathematical Society of Japan. 2008.
- [5] M. Gromov, Hyperbolic groups, in *Essays in group theory*, 75–263, Math. Sci. Res. Inst. Publ., 8, Springer, New York, 1987.
- [6] U. Hamenstädt, Word hyperbolic extensions of surface groups. eprint 2005, arXiv:math/0505244.
- [7] H. Hamidi-Tehrani, Groups generated by positive multi-twists and the fake lantern problem. *Algebr. Geom. Topol.* 2 (2002), 1155–1178
- [8] A. Ishida, The structure of subgroup of mapping class groups generated by two Dehn twists. *Proc. Japan Acad. Ser. A Math. Sci.* 72 (1996), no. 10, 240–241.
- [9] N. V. Ivanov, *Subgroups of Teichmüller modular groups*. Translations of Mathematical Monographs, 115. American Mathematical Society, 1992.
- [10] N. V. Ivanov, Mapping class groups, in *Handbook of Geometric Topology*, (R. Daverman and R. Sher, eds.), 523–633 North-Holland, 2002.
- [11] N. V. Ivanov, Fifteen problems about the mapping class groups. *Problems on mapping class groups and related topics*, 71–80, Proc. Sympos. Pure Math., 74, Amer. Math. Soc. 2006.
- [12] R. P. Kent IV, C. J. Leininger, Shadows of mapping class groups: capturing convex cocompactness. *Geom. Funct. Anal.* 18 (2008), no. 4, 1270–1325.



- [13] M. Koubi, Croissance uniforme dans les groupes hyperboliques. *Ann. Inst. Fourier* 48 (1998), no. 5, 1441–1453.
- [14] C. J. Leininger, D. Margalit, Two-generator subgroups of the pure braid group. eprint 2008, arXiv:0812.1783.
- [15] J. Mangahas, Uniform uniform exponential growth of subgroups of the mapping class group, eprint 2008, arXiv:0805.0133.
- [16] H. Masur and Y. Minsky, Geometry of the complex of curves. I. Hyperbolicity. *Invent. Math.* 138 (1999), no. 1, 103–149.
- [17] J. McCarthy, A “Tits-alternative” for subgroups of surface mapping class groups. *Trans. Amer. Math. Soc.* 291 (1985), no. 2, 583–612.