


# Cogrowth for group actions with strongly contracting elements

GOULNARA N. ARZHANTSEVA and CHRISTOPHER H. CASHEN 

*Universität Wien, Fakultät für Mathematik, Oskar-Morgenstern-Platz 1, 1090 Wien,  
Österreich*

*(e-mail: goulnara.arzhantseva@univie.ac.at, christopher.cashen@univie.ac.at)*

*(Received 15 March 2018 and accepted in revised form 15 October 2018)*

*Abstract.* Let  $G$  be a group acting properly by isometries and with a strongly contracting element on a geodesic metric space. Let  $N$  be an infinite normal subgroup of  $G$  and let  $\delta_N$  and  $\delta_G$  be the growth rates of  $N$  and  $G$  with respect to the pseudo-metric induced by the action. We prove that if  $G$  has purely exponential growth with respect to the pseudo-metric, then  $\delta_N/\delta_G > 1/2$ . Our result applies to suitable actions of hyperbolic groups, right-angled Artin groups and other CAT(0) groups, mapping class groups, snowflake groups, small cancellation groups, etc. This extends Grigorchuk's original result on free groups with respect to a word metric and a recent result of Matsuzaki, Yabuki and Jaerisch on groups acting on hyperbolic spaces to a much wider class of groups acting on spaces that are not necessarily hyperbolic.

Key words: cogrowth, exponential growth, contracting element, mapping class groups, CAT(0) groups

2010 Mathematics Subject Classification: 20F69, 37C35 (Primary); 20F65, 20F67, 20F36 (Secondary)

## 1. Introduction

We consider the exponential growth rate  $\delta_G$  of the orbit of a group  $G$  acting properly on a geodesic metric space  $X$ . In various notable contexts this asymptotic invariant is related to the Hausdorff dimension of the limit set of  $G$  in  $\partial X$  and to analytical and dynamical properties of  $G \backslash X$  such as the spectrum of the Laplacian, divergence rates of random walks, volume entropy, and ergodicity of the geodesic flow.

In some cases of special interest, the value of half the growth rate of the ambient space  $X$  is distinguished. For example, when  $X = \mathbb{H}^n$  and  $H$  is a torsion-free discrete group of isometries of  $X$ , the Elstrodt–Patterson–Sullivan formula [24] for the bottom of the spectrum of the Laplacian of  $H \backslash X$  has a phase change when the ratio of  $\delta_H$  to the volume

entropy of  $X$  is  $1/2$ . Similarly, if  $X$  is a Cayley tree of a finite-rank free group  $F_n$  and  $H$  is a subgroup, then the Grigorchuk cogrowth formula [13] for the spectral radius of  $H \setminus X$  has a phase change at  $\delta_H/\delta_{F_n} = 1/2$ . Our main result says that, in great generality, normal subgroups land decisively on one side of this distinguished value.

**THEOREM 1.1.** *Suppose that  $G$  is a group acting properly by isometries on a geodesic metric space  $X$  with a strongly contracting element and with purely exponential growth. If  $N$  is an infinite normal subgroup of  $G$ , then  $\delta_N/\delta_G > 1/2$ , where the growth rates  $\delta_G$  and  $\delta_N$  are computed with respect to  $G \curvearrowright X$ .*

The ratio  $\delta_N/\delta_G$  is known as the *cogrowth* of  $Q := G/N$ . The hypotheses will be explained in detail in the next section. Briefly, the existence of a strongly contracting element means that some element of  $G$  acts hyperbolically on  $X$ , though  $X$  itself need not be hyperbolic, and pure exponential growth is guaranteed if the action has a strongly contracting element and an orbit of  $G$  in  $X$  is not too badly distorted.

In negative curvature, the strict lower bound on cogrowth has been shown in various special cases [5, 16, 21, 23]. For  $X = G = F_n$ , the strict lower bound on cogrowth is due to Grigorchuk [13].

Grigorchuk and de la Harpe [14, p. 69] (see also [15, Problem 36]) asked whether the strict lower cogrowth bound also holds when  $F_n$  is replaced by a non-elementary Gromov hyperbolic group and  $X$  is one of its Cayley graphs. This long-open problem was recently answered affirmatively by Matsuzaki, Yabuki and Jaerisch [19] (see also a survey by Matsuzaki [18]). Their result applies more generally to groups of divergence type acting on hyperbolic spaces. Theorem 1.1 gives an alternative proof of the positive answer to Grigorchuk and de la Harpe's question and goes much beyond. In comparison, Matsuzaki, Yabuki and Jaerisch's result applies to more general actions if one restricts to actions on *hyperbolic spaces*, while Theorem 1.1 applies to many renowned non-hyperbolic examples.

**COROLLARY 1.2.** *For the following  $G \curvearrowright X$ , for every infinite normal subgroup  $N$  of  $G$  we have  $\delta_N/\delta_G > 1/2$ .*

- (1)  $G$  is a non-elementary hyperbolic group acting cocompactly on a hyperbolic space  $X$ .
- (2)  $G$  is a relatively hyperbolic group and  $X$  is hyperbolic such that  $G \curvearrowright X$  is cusp uniform and satisfies the parabolic gap condition.
- (3)  $G$  is a right-angled Artin group defined by a finite simple graph that is neither a single vertex nor a join and  $X$  is the universal cover of its Salvetti complex.
- (4)  $X$  is a CAT(0) space and  $G$  acts cocompactly with a rank-one isometry on  $X$ .
- (5)  $G$  is the mapping class group of a surface of genus  $g$  and  $p$  punctures, with  $6g - 6 + 2p \geq 2$ , and  $X$  is the Teichmüller space of the surface with the Teichmüller metric.

Results (3)–(5) are new, only known as consequences of Theorem 1.1. Further new examples include wide classes of snowflake groups [2] and of infinitely presented graphical and classical small cancellation groups [1] and hence many so-called infinite ‘monster’ groups.

The generality of Theorem 1.1 is striking. Previous successes in showing the strict lower bound on cogrowth have relied on fairly sophisticated results concerning Patterson–Sullivan measures on the boundary of a hyperbolic space or ergodicity of the geodesic flow on  $G \backslash X$ . These tools are not available in our general setting. Instead, we use the geometry of the group action directly to estimate orbit growth. The idea of our argument is as follows.

- (1) If  $G$  contains a strongly contracting element for  $G \curvearrowright X$ , then so does every infinite normal subgroup  $N$  of  $G$ . Let  $c \in N$  be such an element.
- (2) By passing to a high power of  $c$ , if necessary, we may assume that its translation length is much larger than the constants describing its strong contraction properties. In this case the growth  $\delta_{[c]}$  of the set  $[c]$  of conjugates of  $c$  is exactly  $\delta_G/2$ .
- (3) A ‘tree’s worth’ of copies of  $[c]$  injects into the normal closure  $\langle\langle c \rangle\rangle$  of  $c$ , which is a subgroup of  $N$ . It follows that the growth rate of  $\langle\langle c \rangle\rangle$ , and hence of  $N$ , is strictly greater than  $\delta_{[c]} = \delta_G/2$ . In this step we use the ‘hyperbolicity’ of the action of  $c$ , as quantified by strong contraction, to provide geometric separation between copies of  $[c]$ .

We used this strategy in our paper with Tao [2] (see also references therein) to prove *growth tightness* of  $G \curvearrowright X$  for actions having a strongly contracting element. The key point was to estimate the growth rate of the quotient of  $G$  by the normal closure of  $c$ . We chose a section  $A$  of the quotient map and built a tree’s worth of copies of it by translating by a high power of  $c$ . By construction, the set  $A$  did not contain words containing high powers of  $c$  as subwords, so translates of  $A$  by powers of  $c$  were geometrically separated. There is a serious difficulty in applying step (3) for cogrowth, because  $[c]$  *does* contain words with arbitrarily large powers of  $c$  as subwords. Indeed, any word of  $G$  can occur as a subword of an element of  $[c]$ , so we do not get the same nice geometric separation as hoped for in step (3) and consequently our abstract tree’s worth of copies of  $[c]$  does not inject into  $G$ . We overcome this difficulty by quantifying how this mapping fails to be an injection. We show that there is asymptotically at least half of  $[c]$  for which the map is an injection and we use this half of  $[c]$  to complete step (3).

For an example where the conclusion of the theorem does not hold, consider the group  $G = F_2 \times F_2$  acting on its Cayley graph  $X$  with respect to the generating set  $(S \cup 1) \times (S \cup 1)$ , where  $S$  is a free generating set of  $F_2$ . The  $F_2$  factors are normal and have growth rate exactly half the growth rate of  $G$ . The action  $G \curvearrowright X$  does not have a strongly contracting element.

## 2. Preliminaries

We write  $x \overset{*}{\prec} y$ ,  $x \overset{+}{\prec} y$ , or  $x < y$  if there is a universal constant  $C > 0$  such that  $x < Cy$ ,  $x < y + C$ , or  $x < Cy + C$ , respectively. We define  $\overset{*}{\succ}$ ,  $\overset{+}{\succ}$ ,  $\succ$ ,  $\overset{*}{\succ}$ ,  $\overset{+}{\succ}$ , and  $\succsim$  similarly.

Throughout, we let  $(X, d, o)$  be a based geodesic metric space and let  $G$  be a group acting isometrically on  $X$ . For  $Y \subset X$  and  $r \geq 0$ , let  $B_r(Y) := \{x \in X \mid \exists y \in Y, d(x, y) < r\}$  and  $\bar{B}_r(Y) := \{x \in X \mid \exists y \in Y, d(x, y) \leq r\}$ . Let  $B_r := B_r(o)$  and let  $S_r^\Delta := B_{r+\Delta} - B_r$ .

There are an induced pseudo-metric and a semi-norm on  $G$  given by  $d(g, h) := d(g.o, h.o)$  and  $|g| := d(o, g.o)$ .

2.1. *Growth.* The (exponential) growth rate of a subset  $Y \subset X$  is

$$\delta_Y := \limsup_{r \rightarrow \infty} \frac{\log \#Y \cap \bar{B}_r}{r}.$$

The Poincaré series of a countable subset  $Y$  of  $X$  is

$$\Theta_Y(s) := \sum_{y \in Y} \exp(-sd(y, o)).$$

For any  $\Delta > 0$ , we also consider the series

$$\Theta_Y^{S, \Delta}(s) := \sum_{i=0}^{\infty} (\#Y \cap S_{\Delta i}^{\Delta(i+1)}) \exp(-s\Delta i),$$

$$\Theta_Y^{B, \Delta}(s) := \sum_{i=0}^{\infty} (\#Y \cap \bar{B}_{\Delta i}) \exp(-s\Delta i).$$

The series  $\Theta_Y^{B, \Delta}(s)$  and  $\Theta_Y^{S, \Delta}(s)$  agree with  $\Theta_Y(s)$  up to multiplicative error depending on  $\Delta$  and  $s$ , so they all converge and diverge together. Now,  $\Theta_Y(s)$  converges for  $s > \delta_Y$  and diverges for  $s < \delta_Y$ . The set  $Y$  is said to be *divergent*, or *of divergent type*, if  $\Theta_Y(s)$  diverges at  $s = \delta_Y$ .

We say that  $Y \subset X$  has *purely exponential growth* if there exist  $\delta > 0$  and  $\Delta > 0$  such that  $\#Y \cap S_r^\Delta \asymp \exp(\delta r)$ . Recall that this means that there is a constant  $C > 0$ , independent of  $r$ , such that  $\exp(\delta r)/C \leq \#Y \cap S_r^\Delta \leq C \exp(\delta r)$ .

An action  $G \curvearrowright X$  is (metrically) *proper* if for all  $x \in X$  and  $r \geq 0$  the set  $\{g \in G \mid d(x, g.o) \leq r\}$  is finite. When  $G \curvearrowright X$  is proper, we extend all the preceding definitions to subsets  $H$  of  $G$  by taking  $Y = H.o$ ; for example,

$$\delta_H := \limsup_{r \rightarrow \infty} \frac{\log \#H.o \cap \bar{B}_r}{r} = \limsup_{r \rightarrow \infty} \frac{\log \#\{h \in H \mid |h| \leq r\}}{r}.$$

When  $G \curvearrowright X$  is cocompact or, more generally, has a quasi-convex orbit, the growth of  $\#S_r^\Delta \cap G.o$  is coarsely sub-multiplicative, which, when  $\delta_G > 0$ , implies an exponential lower bound on  $\#S_r^\Delta \cap G.o$ . Conversely, if  $G \curvearrowright X$  contains a strongly contracting element, then the growth of  $\#S_r^\Delta \cap G.o$  is coarsely super-multiplicative, which implies the corresponding exponential upper bound. For instance, Coornaert [9] proved that a quasi-convex-cocompact, exponentially growing subgroup of a hyperbolic group has purely exponential growth. More generally, in [2] we introduced the following condition that implies that the pseudo-metric induced by a group action behaves like a word metric for growth purposes: the *complementary growth* of  $G \curvearrowright X$  is the growth rate of the set of points of  $G.o$  that can be reached from  $o$  by a geodesic segment in  $X$  that stays completely outside of a neighborhood of  $G.o$ , except near its end points. We say that  $G \curvearrowright X$  has *complementary growth gap* if the complementary growth is strictly less than  $\delta_G$ . Yang [25] proved that if  $G$  acts properly with a strongly contracting element and  $0 < \delta_G < \infty$ , then complementary growth gap implies purely exponential growth.

For relatively hyperbolic groups the complementary growth gap specializes to the *parabolic growth gap* of [11], which requires that the growth of parabolic subgroups of

a relatively hyperbolic group is strictly less than the growth rate of the whole group. For another non-cocompact example, we showed in [2] that the action of the mapping class group of a hyperbolic surface on its Teichmüller space has complementary growth gap.

For a non-example, consider the integers  $\mathbb{Z}$  acting parabolically on the hyperbolic plane. Hyperbolic geodesics connecting  $o$  to  $n.o$  for large  $n$  travel deeply into a horoball at the fixed point of  $\mathbb{Z}$  on  $\partial\mathbb{H}^2$ , far from the orbit of  $\mathbb{Z}$ . Although  $\mathbb{Z}$  has 0 exponential growth in any word metric, in terms of this action on  $\mathbb{H}^2$  it has exponential growth due entirely to the distortion of the orbit.

2.2. *Contraction.* A subset  $Y$  of  $X$  is  $C$ -strongly contracting, for a ‘contraction constant’  $C \geq 0$ , if for all  $x, x' \in X$ , if  $d(x, x') \leq d(x, Y)$ , then the diameter of  $\pi_Y(x) \cup \pi_Y(x')$  is at most  $C$ , where  $\pi_Y(x) := \{y \in Y \mid d(x, y) = d(x, Y)\}$ . A set is called *strongly contracting* if there exists a  $C \geq 0$  such that it is  $C$ -strongly contracting. The *projection distance in  $Y$*  is  $d_Y^\pi(x, x') := \text{diam } \pi_Y(x) \cup \pi_Y(x')$ . We extend these definitions to sets  $Z \subset X$  by  $\pi_Y(Z) := \bigcup_{z \in Z} \pi_Y(z)$  and  $d_Y^\pi(Z, Z') := \text{diam } \pi_Y(Z) \cup \pi_Y(Z')$ .

Strong contraction of  $Y$  is equivalent [2, Lemma 2.4] to the *bounded geodesic image property*: for all  $C \geq 0$ , there exists  $C' \geq C$  such that if  $Y$  is  $C$ -strongly contracting, then, for every geodesic  $\gamma$  in  $X$ , if  $\gamma \cap B_{C'}(Y) = \emptyset$ , then  $\text{diam } \pi_Y(\gamma) \leq C'$ .

COROLLARY 2.1. *Suppose that  $Y$  is  $C$ -strongly contracting and  $C'$  is as above. Suppose that  $\gamma$  is a geodesic defined on an interval  $[a, b]$ , possibly infinite. Let  $t_0 := \inf\{t \mid d(\gamma(t), Y) < C'\}$  and let  $t_1 := \sup\{t \mid d(\gamma(t), Y) < C'\}$ . Then  $\text{diam } \pi_Y(\gamma([a, t_0])) \leq C'$  and  $\text{diam } \pi_Y(\gamma([t_1, b])) \leq C'$ , while  $\gamma([t_0, t_1]) \subset \bar{B}_{3C'}(Y)$ . If  $a$  and  $b$  are finite and  $\text{diam } \pi_Y(\gamma(a)) \cup \pi_Y(\gamma(b)) > C'$ , then  $\pi_Y(\gamma(a)) \subset \bar{B}_{2C'}(\gamma(t_0))$  and  $\pi_Y(\gamma(b)) \subset \bar{B}_{2C'}(\gamma(t_1))$ .*

An infinite-order element  $c \in G$  is said to be a *strongly contracting element* for  $G \curvearrowright X$  if the set  $\langle c \rangle.o$  is strongly contracting. In this case  $\mathbb{Z} \rightarrow X : i \mapsto c^i.o$  is a quasi-isometric embedding and  $c$  is contained in a maximal virtually cyclic subgroup  $E(c)$ . This subgroup, which is alternately known as the *elementarizer* or *elementary closure* of  $c$ , can also be characterized as the maximal subgroup consisting of elements  $g \in G$  such that  $g^{-1}\langle c \rangle g$  is at bounded Hausdorff distance from  $\langle c \rangle$ . Since  $E(c).o$  is coarsely equivalent to  $\langle c \rangle.o$ , the set  $E(c).o$  is also strongly contracting. Note that  $E(c) = E(c^n)$  for every  $n \neq 0$ . Thus, when considering  $E(c).o$ , we can pass to powers of  $c$  freely without changing the set  $E(c^n).o$  and in particular without changing its contraction constant.

For a strongly contracting element  $c$ , let  $\mathcal{E} := E(c).o$  and let  $\mathbf{Y}$  be the collection of distinct  $G$ -translates of  $\mathcal{E}$ . Bestvina, Bromberg and Fujiwara [3] axiomatized the geometry of projection distances in  $\mathbf{Y}$ . With Sisto [4], they showed that by a small change in the projections and projection distances, a cleaner set of axioms is satisfied—these will allow us to make an inductive argument in the next section. The following is [4, Theorem 4.1] applied to  $\mathbf{Y}$ . We list here only those axioms that we will make use of and that are not immediate from our particular definitions of  $\mathbf{Y}$ ,  $\pi_Y$ , and  $d_Y^\pi$ . A detailed verification that  $\mathbf{Y}$  satisfies the hypotheses of [4, Theorem 4.1] can be found in [2].

**THEOREM 2.2.** *There exists  $\theta \geq 0$  such that for each  $\mathcal{Y} \in \mathbf{Y}$  there is a projection  $\pi'_{\mathcal{Y}}$  taking elements of  $\mathbf{Y}$  to subsets of  $\mathcal{Y}$  such that for all  $\mathcal{X} \in \mathbf{Y}$  and  $g \in G$ , we have  $\pi'_{\mathcal{Y}}(\mathcal{X}) \subset B_{\theta}(\pi_{\mathcal{Y}}(\mathcal{X}))$  and  $\pi'_{g\mathcal{Y}}(g\mathcal{X}) = g\pi'_{\mathcal{Y}}(\mathcal{X})$ . Furthermore, there are distance maps  $d_{\mathcal{Y}}(\mathcal{X}, \mathcal{Z}) = \text{diam } \pi'_{\mathcal{Y}}(\mathcal{X}) \cup \pi'_{\mathcal{Y}}(\mathcal{Z})$  with  $|d_{\mathcal{Y}} - d'_{\mathcal{Y}}| \leq 2\theta$  such that, for  $\theta' := 11\theta$ , the following axioms are satisfied for all  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \in \mathbf{Y}$ :*

- (P 0)**  $d'_{\mathcal{Y}}(\mathcal{X}, \mathcal{X}) \leq \theta$  when  $\mathcal{X} \neq \mathcal{Y}$ ;
- (P 1)** if  $d'_{\mathcal{Y}}(\mathcal{X}, \mathcal{Z}) > \theta$ , then  $d'_{\mathcal{X}}(\mathcal{Y}, \mathcal{Z}) \leq \theta$  for all distinct  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ ;
- (SP 3)** if  $d_{\mathcal{Y}}(\mathcal{X}, \mathcal{Z}) > \theta'$ , then  $d_{\mathcal{Z}}(\mathcal{X}, \mathcal{W}) = d_{\mathcal{Z}}(\mathcal{Y}, \mathcal{W})$  for all  $\mathcal{W} \in \mathbf{Y} - \{\mathcal{Z}\}$ ;
- (SP 4)**  $d_{\mathcal{Y}}(\mathcal{X}, \mathcal{X}) \leq \theta'$  when  $\mathcal{X} \neq \mathcal{Y}$ .

For more details on strongly contracting elements and many examples, see [2].

**PROPOSITION 2.3.** (Lemma 2.2 and Proposition 2.3 of [4]) *With  $\theta'$  as in Theorem 2.2, for each  $\mathcal{X}$  and  $\mathcal{Z}$  in  $\mathbf{Y}$  define  $\mathbf{Y}(\mathcal{X}, \mathcal{Z}) := \{\mathcal{Y} \in \mathbf{Y} - \{\mathcal{X}, \mathcal{Z}\} \mid d_{\mathcal{Y}}(\mathcal{X}, \mathcal{Z}) > 2\theta'\}$  and  $\mathbf{Y}[\mathcal{X}, \mathcal{Z}] := \mathbf{Y}(\mathcal{X}, \mathcal{Z}) \cup \{\mathcal{X}, \mathcal{Z}\}$ . There is a total order  $\sqsubset$  on  $\mathbf{Y}[\mathcal{X}, \mathcal{Z}]$  such if  $\mathcal{Y}_0 \sqsubset \mathcal{Y}_1 \sqsubset \mathcal{Y}_2$ , then  $d_{\mathcal{Y}_1}(\mathcal{Y}_0, \mathcal{Y}_2) = d_{\mathcal{Y}_1}(\mathcal{X}, \mathcal{Z})$ . The relation  $\mathcal{Y}_0 \sqsubset \mathcal{Y}_1$  is defined by each of the following equivalent conditions:*

- $d_{\mathcal{Y}_0}(\mathcal{X}, \mathcal{Y}_1) > \theta'$ ;
- $d_{\mathcal{Y}_1}(\mathcal{X}, \mathcal{Y}_0) \leq \theta'$ ;
- $d_{\mathcal{Y}_1}(\mathcal{Y}_0, \mathcal{Z}) > \theta'$ ;
- $d_{\mathcal{Y}_0}(\mathcal{Y}_1, \mathcal{Z}) \leq \theta'$ .

3. *Embedding a tree’s worth of copies of  $[c]$*

For a subset  $H \subset G$ , let  $H^* := H - \{1\}$  and consider  $\hat{H} := \bigcup_{k=1}^{\infty} (H^*)^k$ . We consider  $\hat{H}$  to be a ‘tree’s worth of copies of  $H$ ’ in allusion to the case of the free product  $H * \mathbb{Z}/2\mathbb{Z}$  when  $H$  is a group. The group  $H * \mathbb{Z}/2\mathbb{Z}$  acts on a tree with vertex stabilizers conjugate to  $H$  and every element that is not equal to 1 or the generator  $z$  of  $\mathbb{Z}/2\mathbb{Z}$  has a unique expression as  $z^{\alpha} h_1 z h_2 z \cdots h_k z^{\beta}$  for some  $k \in \mathbb{N}$ ,  $\alpha, \beta \in \{0, 1\}$ , and  $h_i \in H^*$ .

The naive map  $\hat{H} \rightarrow X : (h_1, \dots, h_k) \mapsto h_1 c \cdots h_k c.o$ , where  $c$  is a strongly contracting element, is clearly not an injection for  $H = [c]$ , as it gives collisions  $(h^{-1}, h) \mapsto h^{-1} c h c.o \leftarrow (h^{-1} c h)$ . To avoid collisions we remove a fraction of  $[c]$  in four steps and use a slightly different map. The main technical result is as follows.

**PROPOSITION 3.1.** *Under the hypothesis of Theorem 1.1, let  $c$  be a strongly contracting element. After possibly passing to a power of  $c$ , there is a subset  $G_4 \subset [c]$  that is divergent, has  $\delta_{G_4} = \delta_G/2$ , and for which the map  $\hat{G}_4 \rightarrow X : (g_1, \dots, g_k) \mapsto (\prod_{i=1}^k g_i c^2).o$  is an injection.*

The main theorem follows by an argument analogous to the one we used in [2], which we reproduce for the reader’s convenience.

*Proof of Theorem 1.1.* Let  $c' \in G$  be a strongly contracting element for  $G \curvearrowright X$ . Suppose that  $N < E(c')$ . Since  $N$  is infinite, it has a finite-index subgroup in common with  $\langle c' \rangle$ . But conjugation by an element of  $G$  fixes  $N$ , so it moves  $\langle c' \rangle$  by a bounded Hausdorff distance, which means that  $G = E(c')$  is virtually cyclic and  $N$  is a finite-index subgroup of  $G$ . However,  $\langle c' \rangle$  has an undistorted orbit in  $X$ . Since this is a finite-index subgroup

of  $G$ , the growth of  $G$  is only linear, contradicting the exponential growth hypothesis. Thus, we may assume that  $G$  is not virtually cyclic and that  $N$  contains an element  $g$  that is not in  $E(c')$ . We showed in [2, Proposition 3.1] that for sufficiently large  $n$  the element  $c := g^{-1}(c')^{-n}g(c')^n$  is a strongly contracting element of  $N$ .

Consider  $G_4$  as provided by Proposition 3.1 with respect to  $c$ . Then  $\hat{G}_4$  injects into  $X$  and, moreover, the image is contained in  $\langle\langle c \rangle\rangle.o \subset N.o$ . Therefore, the growth rate of  $N$  is at least as large as the growth rate of the image of  $\hat{G}_4$ , which we estimate using its Poincaré series:

$$\begin{aligned} \Theta_{\hat{G}_4}(s) &= \sum_{k=1}^{\infty} \sum_{(g_1, \dots, g_k) \in (G_4^*)^k} \exp(-s|g_1c^2 \cdots g_kc^2|) \\ &\geq \sum_{k=1}^{\infty} \sum_{(g_1, \dots, g_k) \in (G_4^*)^k} \exp\left(-sk|c^2| - s \sum_{i=1}^k |g_i|\right) \\ &= \sum_{k=1}^{\infty} \exp(-sk|c^2|) \sum_{(g_1, \dots, g_k) \in (G_4^*)^k} \prod_{i=1}^k \exp(-s|g_i|) \\ &= \sum_{k=1}^{\infty} \exp(-sk|c^2|) \left( \sum_{g \in G_4^*} \exp(-s|g|) \right)^k \\ &= \sum_{k=1}^{\infty} (\exp(-s|c^2|)\Theta_{G_4^*}(s))^k. \end{aligned}$$

Since  $G_4$  is divergent, for sufficiently small positive  $\epsilon$  we have  $\Theta_{G_4^*}(\delta_{G_4} + \epsilon) \geq \exp((\delta_{G_4} + \epsilon)|c^2|)$ , so  $\Theta_{\hat{G}_4}(\delta_{G_4} + \epsilon)$  diverges, which implies that  $\delta_{\hat{G}_4} \geq \delta_{G_4} + \epsilon$ . Thus,  $\delta_N \geq \delta_{\hat{G}_4} \geq \delta_{G_4} + \epsilon > \delta_{G_4} = \delta_G/2$ .

The remainder of this section is devoted to the construction of the set  $G_4$  satisfying the conclusion of Proposition 3.1. Here is a brief overview. We need a subset of  $[c]$  such that the given map is an injection. It would be preferable if we could take conjugates of  $c$  by elements  $g$  that have no long projection to any element of  $\mathbf{Y}$ . It is easy to build an injection based on such elements, but, unfortunately, there are too few of them in our setting—the growth rate of the set of such elements is strictly smaller than  $\delta_G$ , so the growth rate of  $c$ -conjugates by such elements is strictly smaller than  $\delta_G/2$ . Instead, we consider elements  $g$  that do not have long projections to  $\mathcal{E}$  and  $g\mathcal{E}$ ; in a sense, these are elements ‘orthogonal to  $\mathbf{Y}$  at their end points’, rather than ‘orthogonal to  $\mathbf{Y}$ ’ throughout. The desired condition can be achieved with a small modification near the ends of  $g$ , so this does not change the growth rate. We call this set of elements  $G_1$  and the conjugates of (a power of)  $c$  by these elements  $G_2$ . We define  $G_3$  by passing to a maximal subset of  $G_2$  such that elements are sufficiently far apart. This does not change the set much; in particular, the growth rate is unchanged. However, it will be an important point for the injection argument, because we show in Lemma 3.5 that if  $g$  and  $h$  are in  $G_3$ , then  $g\mathcal{E} = h\mathcal{E}$  implies that  $g = h$ . The final refinement is to pass to the subset  $G_4$  of  $G_3$  of elements that are not ‘in the shadow’ of some other element of  $G_3$ , that is to say, elements  $g$  such that there does not exist  $h$  such



that a geodesic from  $o$  to  $g.o$  passes close to  $h.o$ . The crux of the argument, Lemma 3.6, is to show that at least half of  $G_3$  is unshadowed, so  $G_4$  is divergent with growth rate  $\delta_G/2$ . Finally, in Lemma 3.7, we check that  $G_4$  gives the desired injection.

Fix an element  $f_0 \in G$  such that  $f_0\mathcal{E}$  is disjoint from  $\mathcal{E}$ ,  $o \in \pi_{\mathcal{E}}(f_0.o)$ , and  $f_0.o \in \pi_{f_0\mathcal{E}}(o)$ . To see that such an element exists, first note that there exists  $g \in G - E(c)$ , for instance, as in the first paragraph of the proof of Theorem 1.1. If  $\mathcal{E}$  and  $g\mathcal{E}$  are disjoint, let  $f_1$  and  $f_2$  be elements of  $G$  such that  $f_1.o \in \mathcal{E}$  and  $f_2.o \in g\mathcal{E}$  realize the minimum distance between  $\mathcal{E}$  and  $g\mathcal{E}$ . Then the element  $f_0 := f_1^{-1}f_2$  satisfies our requirements. If  $g\mathcal{E}$  and  $\mathcal{E}$  are not disjoint, consider  $g\mathcal{E}$  and  $c^n g\mathcal{E}$  for some  $n$ . If they intersect, then, by (P 0),

$$2\theta \geq d_{\mathcal{E}}^{\pi}(g\mathcal{E}, g\mathcal{E}) + d_{\mathcal{E}}^{\pi}(c^n g\mathcal{E}, c^n g\mathcal{E}) \geq d_{\mathcal{E}}^{\pi}(g\mathcal{E}, c^n g\mathcal{E}) \geq |c^n|.$$

This is impossible once  $n$  is sufficiently large as  $c$  is strongly contracting. So,  $g\mathcal{E}$  and  $c^n g\mathcal{E}$  are disjoint for such  $n$  and we get  $f_0$  by the previous argument after replacing  $g$  with  $g^{-1}c^n g$ .

Since  $\mathcal{E}$  and  $f_0\mathcal{E}$  are disjoint and  $o$  and  $f_0.o$  are contained in one another's projections, strong contraction of  $c$ , and hence of  $\mathcal{E}$ , gives a constant  $C \geq 0$  such that

$$d_{f_0\mathcal{E}}^{\pi}(o, f_0.o) = \text{diam } \pi_{f_0\mathcal{E}}(o) \leq C \quad \text{and} \quad d_{\mathcal{E}}^{\pi}(o, f_0.o) = \text{diam } \pi_{\mathcal{E}}(f_0.o) \leq C. \quad (1)$$

In the following, we use the following notation:  $|f_0|$  is the length of the element  $f_0$  just defined;  $\Delta$  is as in the definition of purely exponential growth of  $G$ ;  $C$  is a contraction constant for  $\mathcal{E}$ ;  $C'$  is the corresponding constant from Corollary 2.1;  $\theta$  and  $\theta'$  are as in Theorem 2.2;  $K$  is a fixed constant strictly greater than  $\max\{C, \theta + \theta'/2\}$ . We call these, collectively, 'the constants'. The terms 'small' and 'close' mean bounded by some combination of the constants. When possible we decline to compute these explicitly, since only finitely many such combinations appear in the proof, except where noted. Furthermore,  $\Delta$  depends only on  $G$  and the others depend only on  $\mathcal{E} = E(c).o$ . Since  $E(c) = E(c^p)$  for all  $p \neq 0$ , we can, and will, pass to high powers of  $c$  to make  $|c^p|$  much larger than all of the constants and combinations of them that we encounter.

Set  $G_1 := \{g \in G \mid d_{\mathcal{E}}^{\pi}(o, g.o) \leq 2K \text{ and } d_{g\mathcal{E}}^{\pi}(o, g.o) \leq 2K \text{ and } g\mathcal{E} \neq \mathcal{E}\}$ . This is a subset of  $G$  that is closed under taking inverses.

LEMMA 3.2. *For every  $g \in G$ , at least one of the elements  $g, f_0g, gf_0$ , or  $f_0gf_0$  belongs to  $G_1$ .*

*Proof.* First, consider  $g \notin E(c)$  with  $|g| \leq K$ . Recall that  $g \in E(c)$  if and only if  $g\mathcal{E} = \mathcal{E}$ . By definition,  $\pi_{\mathcal{E}}(g.o)$  is the set of points of  $\mathcal{E}$  minimizing the distance to  $g.o$ . By hypothesis,  $o$  is a point of  $\mathcal{E}$  at distance at most  $K$  from  $g.o$ , so  $d(g.o, \pi_{\mathcal{E}}(g.o)) \leq K$ , and  $d_{\mathcal{E}}^{\pi}(o, g.o) = \text{diam } \{o\} \cup \pi_{\mathcal{E}}(g.o) \leq 2K$ . The same argument for  $o$  projecting to  $g\mathcal{E}$  gives  $d_{g\mathcal{E}}^{\pi}(o, g.o) \leq 2K$ . Thus, elements  $g$  of this form already belong to  $G_1$ .

Next, consider an element  $g \in E(c)$  such that  $|g| \leq K$ . Since  $g \in E(c)$ , we have  $f_0g\mathcal{E} = f_0\mathcal{E} \neq \mathcal{E}$  and  $\pi_{\mathcal{E}}(g.o) = g.o$ , so  $d_{\mathcal{E}}^{\pi}(o, g.o) = d(o, g.o) \leq K$ . Using this estimate and (1), we see that

$$\begin{aligned} d_{f_0g\mathcal{E}}^{\pi}(o, f_0g.o) &\leq d_{f_0g\mathcal{E}}^{\pi}(o, f_0.o) + d_{f_0g\mathcal{E}}^{\pi}(f_0.o, f_0g.o) \\ &= d_{f_0\mathcal{E}}^{\pi}(o, f_0.o) + d_{\mathcal{E}}^{\pi}(o, g.o) \leq C + K < 2K. \end{aligned}$$



In the other direction, using the fact that  $o \in \pi_{\mathcal{E}}(f_0.o) \subset \pi_{\mathcal{E}}(f_0\mathcal{E})$ , along with (P 0),

$$d_{\mathcal{E}}^{\pi}(o, f_0g.o) \leq d_{\mathcal{E}}^{\pi}(o, f_0\mathcal{E}) \leq d_{\mathcal{E}}^{\pi}(f_0\mathcal{E}, f_0\mathcal{E}) \leq \theta < K.$$

Note that we did not use  $d_{\mathcal{E}}^{\pi}(o, g.o) \leq K$  for this direction—the inequality is valid for any  $g \in E(c)$ .

Suppose that  $g \notin E(c)$  and  $d_{\mathcal{E}}^{\pi}(o, g.o) > K$ ; then

$$\theta < K < d_{\mathcal{E}}^{\pi}(o, g.o) = d_{f_0\mathcal{E}}^{\pi}(f_0.o, f_0g.o) \leq d_{f_0\mathcal{E}}^{\pi}(\mathcal{E}, f_0g\mathcal{E}).$$

This contradicts (P 0) if  $\mathcal{E} = f_0g\mathcal{E}$ , since, by hypothesis,  $f_0\mathcal{E} \neq \mathcal{E}$  and  $f_0g\mathcal{E} \neq f_0\mathcal{E}$ . Thus,  $\mathcal{E}$ ,  $f_0\mathcal{E}$ , and  $f_0g\mathcal{E}$  are distinct and we can apply (P 1) to get

$$d_{\mathcal{E}}^{\pi}(o, f_0g.o) \leq d_{\mathcal{E}}^{\pi}(f_0\mathcal{E}, f_0g\mathcal{E}) \leq \theta < K.$$

For  $|g| \leq K$  we are done, either  $g$  or  $f_0g$  is in  $G_1$ , and for  $|g| > K$  we have shown that there is at least one choice of  $g' \in \{g, f_0g\}$  such that  $g'\mathcal{E} \neq \mathcal{E}$  and  $d_{\mathcal{E}}^{\pi}(o, g'.o) \leq K$ . If  $d_{g'\mathcal{E}}^{\pi}(o, g'.o) \leq K$ , then we are done, so suppose not. Consider the possibility that  $g'f_0\mathcal{E} = \mathcal{E}$ . Then  $g'f_0.o \in \mathcal{E}$ , so  $o \in \pi_{\mathcal{E}}(f_0.o)$  implies that  $g'.o \in \pi_{g'\mathcal{E}}(g'f_0.o) \subset \pi_{g'\mathcal{E}}(\mathcal{E})$ . Since  $g'\mathcal{E} \neq \mathcal{E}$ , (P 0) says that  $d_{g'\mathcal{E}}^{\pi}(\mathcal{E}, \mathcal{E}) \leq \theta$ , so

$$K < d_{g'\mathcal{E}}^{\pi}(o, g'.o) \leq d_{g'\mathcal{E}}^{\pi}(\mathcal{E}, \mathcal{E}) \leq \theta < K.$$

This is a contradiction, so  $\mathcal{E}$ ,  $g'\mathcal{E}$ , and  $g'f_0\mathcal{E}$  are distinct. Observe that, since  $g'.o \in \pi_{g'\mathcal{E}}(g'f_0.o)$ ,

$$d_{g'\mathcal{E}}^{\pi}(\mathcal{E}, g'f_0\mathcal{E}) \geq d_{g'\mathcal{E}}^{\pi}(o, g'f_0.o) \geq d_{g'\mathcal{E}}^{\pi}(o, g'.o) > K > \theta.$$

Thus, by (P 1) and the fact that  $g'f_0.o \in \pi_{g'f_0\mathcal{E}}(g'.o)$ , we have  $d_{g'f_0\mathcal{E}}^{\pi}(o, g'f_0.o) \leq d_{g'f_0\mathcal{E}}^{\pi}(\mathcal{E}, g'\mathcal{E}) \leq \theta < K$ .

To check that the first inequality has not been spoiled, use the fact that  $d_{g'\mathcal{E}}^{\pi}(\mathcal{E}, g'f_0\mathcal{E}) > \theta$ , so (P 1) implies that  $d_{\mathcal{E}}^{\pi}(g'\mathcal{E}, g'f_0\mathcal{E}) \leq \theta$ , which gives

$$d_{\mathcal{E}}^{\pi}(o, g'f_0.o) \leq d_{\mathcal{E}}^{\pi}(o, g'.o) + d_{\mathcal{E}}^{\pi}(g'.o, g'f_0.o) \leq K + d_{\mathcal{E}}^{\pi}(g'\mathcal{E}, g'f_0\mathcal{E}) < K + \theta < 2K. \quad \square$$

Define  $\phi_0: G \rightarrow G_1$  by fixing  $G_1$  and sending an element  $g \in G - G_1$  to an arbitrary element of the non-empty set  $\{f_0g, gf_0, f_0gf_0\} \cap G_1$ . The map  $\phi_0$  is surjective, at most 4-to-1, and changes norm by at most  $2|f_0|$ .

For each  $p \in \mathbb{N}$ , define  $G_{2,p} := \{g^{-1}c^p g \mid g \in G_1\}$  and  $\phi_{1,p}: G_1 \rightarrow G_{2,p}: g \mapsto g^{-1}c^p g$ .

LEMMA 3.3. *If  $p$  is sufficiently large, then, for every  $g \in G_1$ , we have*

$$2|g| + |c^p| - 8C' - 8K \leq |\phi_{1,p}(g)| \leq 2|g| + |c^p|.$$

*Proof.* The upper bound is clear. We derive a lower bound from strong contraction. From the definition of  $G_1$ , it follows that  $\pi_{g^{-1}\mathcal{E}}(o) \subset \bar{B}_{2K}(g^{-1}.o)$  and  $\pi_{g^{-1}\mathcal{E}}(g^{-1}c^p g.o) \subset \bar{B}_{2K}(g^{-1}c^p.o)$ , so

$$|c^p| - 4K \leq d_{g^{-1}\mathcal{E}}^{\pi}(o, g^{-1}c^p g.o) \leq |c^p| + 4K. \tag{2}$$

Let  $\gamma$  be a geodesic from  $o$  to  $g^{-1}c^p g.o$ . Its end points have projection to  $g^{-1}\mathcal{E}$  at distance at least  $|c^p| - 4K \gg C'$  from one another, for  $p$  sufficiently large, as  $c$  is strongly contracting. Thus, for  $t_0$  and  $t_1$  as in Corollary 2.1, we have  $d(\gamma(t_0), \pi_{g^{-1}\mathcal{E}}(o)) \leq 2C'$ , so  $d(\gamma(t_0), g^{-1}.o) \leq 2C' + 2K$  and, similarly,  $d(\gamma(t_1), g^{-1}c^p.o) \leq 2C' + 2K$ . We have

$$\begin{aligned} |\phi_{1,p}(g)| &= |\gamma| = d(o, \gamma(t_0)) + d(\gamma(t_0), \gamma(t_1)) + d(\gamma(t_1), g^{-1}c^p g.o) \\ &\geq (d(o, g^{-1}.o) - (2C' + 2K)) + (d(g^{-1}.o, g^{-1}c^p.o) - 2(2C' + 2K)) \\ &\quad + (d(g^{-1}c^p.o, g^{-1}c^p g.o) - (2C' + 2K)) \\ &= 2|g| + |c^p| - 8C' - 8K. \end{aligned} \quad \square$$

The following lemma also follows from (2).

LEMMA 3.4. *Let  $g^{-1}c^p g = \phi_{1,p}(g) \in G_{2,p}$ . If  $p$  is sufficiently large, then  $g^{-1}\mathcal{E} \in \mathbf{Y}(\mathcal{E}, g^{-1}c^p g\mathcal{E})$ .*

We also claim that  $\phi_{1,p}$  is bounded-to-one, independent of  $p$ . To see this, fix  $g \in G_1$  and consider  $h \in G_1$  such that  $\phi_{1,p}(g) = \phi_{1,p}(h)$ . Then  $gh^{-1}$  commutes with  $c^p$ , so  $gh^{-1} \in E(c^p) = E(c)$ . Thus,

$$\begin{aligned} |gh^{-1}| &= d_{\mathcal{E}}^{\pi}(o, gh^{-1}.o) \\ &\leq d_{\mathcal{E}}^{\pi}(o, g.o) + d_{\mathcal{E}}^{\pi}(g.o, gh^{-1}.o) \\ &= d_{\mathcal{E}}^{\pi}(o, g.o) + d_{hg^{-1}\mathcal{E}}^{\pi}(h.o, o) \\ &= d_{\mathcal{E}}^{\pi}(o, g.o) + d_{\mathcal{E}}^{\pi}(h.o, o) \leq 4K. \end{aligned}$$

So,  $h$  satisfies  $h^{-1}.o \in \bar{B}_{4K}(g^{-1}.o)$ . By properness of  $G \curvearrowright X$ ,  $\#G.o \cap \bar{B}_{4K}(g^{-1}.o) = \#G.o \cap \bar{B}_{4K}(o)$  is finite.

Let  $G_{3,p}$  be a maximal  $(6K + 1)$ -separated subset of  $G_{2,p}$ , that is, a subset that is maximal for inclusion among those with the property that  $d(g.o, h.o) \geq 6K + 1$  for distinct elements  $g$  and  $h$ . Let  $\phi_{2,p}: G_{2,p} \rightarrow G_{3,p}$  be a choice of closest point. This map is surjective. By maximality,  $\phi_{2,p}$  moves points a distance less than  $6K + 1$ . Thus, by properness of  $G \curvearrowright X$ , the map  $\phi_{2,p}$  is bounded-to-one, independent of  $p$ .

LEMMA 3.5. *If  $p$  is sufficiently large, then  $g^{-1}c^p g\mathcal{E} = h^{-1}c^p h\mathcal{E}$  for  $g^{-1}c^p g$  and  $h^{-1}c^p h$  in  $G_{3,p}$  implies that  $g^{-1}c^p g = h^{-1}c^p h$ .*

*Proof.* Since  $g \in G_1$ ,  $d_{g\mathcal{E}}^{\pi}(o, g.o) \leq 2K$  and

$$\begin{aligned} d_{g^{-1}c^p g\mathcal{E}}^{\pi}(o, g^{-1}c^p g.o) &\leq d_{g^{-1}c^p g\mathcal{E}}^{\pi}(o, g^{-1}c^p.o) + d_{g^{-1}c^p g\mathcal{E}}^{\pi}(g^{-1}c^p.o, g^{-1}c^p g.o) \\ &\leq d_{g^{-1}c^p g\mathcal{E}}^{\pi}(\mathcal{E}, g^{-1}\mathcal{E}) + 2K. \end{aligned}$$

Furthermore,  $g \in G_1$  implies that  $\mathcal{E} \neq g^{-1}\mathcal{E} \neq g^{-1}c^p g\mathcal{E}$ . By (2),  $d_{g^{-1}\mathcal{E}}^{\pi}(\mathcal{E}, g^{-1}c^p g\mathcal{E}) \geq |c^p| - 4K \gg \theta$ , so, by (P 0),  $\mathcal{E} \neq g^{-1}c^p g\mathcal{E}$ . Thus,  $\mathcal{E}$ ,  $g^{-1}\mathcal{E}$ , and  $g^{-1}c^p g\mathcal{E}$  are distinct and we can apply (P 1) to see that  $d_{g^{-1}c^p g\mathcal{E}}^{\pi}(\mathcal{E}, g^{-1}\mathcal{E}) \leq \theta < K$ . Inserting this into the previous inequality gives

$$d_{g^{-1}c^p g\mathcal{E}}^{\pi}(o, g^{-1}c^p g.o) < 3K. \tag{3}$$

The same computation applies for  $h$ , so  $\pi_{g^{-1}c^p g \mathcal{E}}(o) \subset \bar{B}_{3K}(g^{-1}c^p g.o) \cap \bar{B}_{3K}(h^{-1}c^p h.o)$ . Thus,  $g^{-1}c^p g$  and  $h^{-1}c^p h$  are elements at distance at most  $6K$  in a  $(6K + 1)$ -separated set; hence, they are equal.  $\square$

For each  $D \geq 0$ , consider the set  $G'_{4,p,D}$  consisting of elements  $g^{-1}c^p g \in G_{3,p}$  such that there exists a different element  $h^{-1}c^p h \in G_{3,p}$  such that  $h^{-1}c^p h c^{2p}.o$  is within distance  $D$  of a geodesic  $\gamma$  from  $o$  to  $g^{-1}c^p g.o$ . Define  $G_{4,p,D} := G_{3,p} - G'_{4,p,D}$ .

LEMMA 3.6. *For all  $D \geq 0$ , for  $p$  sufficiently large,  $G_{4,p,D}$  is divergent and  $\delta_{G_{4,p,D}} = \delta_G/2$ .*

*Proof.* The maps  $\phi_{2,p}$ ,  $\phi_{1,p}$ , and  $\phi_0$  are surjective and bounded-to-one, with bound independent of  $p$ , so their composition is as well. Furthermore, we know how they change norm:  $\phi_0$  moves points at most  $2|f_0|$ ,  $\phi_{2,p}$  moves them less than  $6K + 1$ , and  $|\phi_{1,p}(g)|$  is estimated in Lemma 3.3. Putting these together, for any  $r \geq 0$  and  $g \in G \cap S_r^\Delta$ , we have

$$2r + |c^p| - 4|f_0| - 8C' - 14K - 1 \leq |\phi_{2,p} \circ \phi_{1,p} \circ \phi_0(g)| < 2r + |c^p| + 2\Delta + 4|f_0| + 6K + 1. \tag{4}$$

Let  $t := 2r + |c^p| - 4|f_0| - 8C' - 14K - 1$ ,  $E := 4|f_0| + 4C' + 10K + 1$ , and  $\Delta' := 2(\Delta + E)$ , so that (4) shows that

$$\phi_{2,p} \circ \phi_{1,p} \circ \phi_0(G \cap S_r^\Delta) \subset G_{3,p} \cap S_t^{\Delta'} \subset \phi_{2,p} \circ \phi_{1,p} \circ \phi_0(G \cap S_{r-E}^{\Delta+2E}).$$

This lets us compare the size of spherical shells in  $G_{3,p}$  and  $G$ :

$$\#G \cap S_{r-E}^{\Delta+2E} \geq \#G_{3,p} \cap S_t^{\Delta'} \stackrel{*}{\succ} \#G \cap S_r^\Delta. \tag{5}$$

Pure exponential growth of  $G$  says that  $\#G \cap S_r^\Delta \stackrel{*}{\asymp} \exp(r\delta_G)$ . Combining this with (5), we have

$$\#G_{3,p} \cap S_t^{\Delta'} \stackrel{*}{\asymp} \exp(\delta_G r) \stackrel{*}{\asymp} \exp(-\delta_G |c^p|/2) \exp(t\delta_G/2). \tag{6}$$

This tells us that  $\delta_{G_{3,p}} = \delta_G/2$  and  $G_{3,p}$  is divergent.

Now, we will estimate an upper bound for  $\#G'_{4,p,D} \cap S_r^{\Delta'}$  and see that for large  $p$  and  $r$  it is less than half of  $\#G_{3,p} \cap S_r^{\Delta'}$ . Thus, to get  $G_{4,p,D}$ , we threw away less than half of  $G_{3,p}$ , at least outside a sufficiently large radius. We conclude that  $\delta_{G_{4,p,D}} = \delta_G/2$  and  $G_{4,p,D}$  is divergent.

Consider  $g^{-1}c^p g \in G'_{4,p,D} \cap S_r^{\Delta'}$  for any  $r > 7|c^p|$ . By definition of  $G'_{4,p,D}$ , there exists  $h^{-1}c^p h \in G_{3,p}$  such that  $h^{-1}c^p h \neq g^{-1}c^p g$  and  $h^{-1}c^p h c^{2p}.o$  is close to a geodesic  $\gamma$  from  $o$  to  $g^{-1}c^p g.o$ .

Let  $\square$  be the order of Proposition 2.3 on  $\mathbf{Y}[\mathcal{E}, g^{-1}c^p g \mathcal{E}]$ . The first step of the proof is to show that  $\mathcal{E}$ ,  $g^{-1}\mathcal{E}$ ,  $g^{-1}c^p g \mathcal{E}$ ,  $h^{-1}\mathcal{E}$ , and  $h^{-1}c^p h \mathcal{E}$  are distinct elements of  $\mathbf{Y}[\mathcal{E}, g^{-1}c^p g \mathcal{E}]$  and that the ordering is one of the two possibilities shown in Figures 1 and 2.

By Lemma 3.4,  $\mathcal{E} \square g^{-1}\mathcal{E} \square g^{-1}c^p g \mathcal{E}$ , so these three are distinct. Similarly,  $\mathcal{E}$ ,  $h^{-1}\mathcal{E}$ , and  $h^{-1}c^p h \mathcal{E}$  are distinct. Lemma 3.5 implies that  $g^{-1}c^p g \mathcal{E} \neq h^{-1}c^p h \mathcal{E}$ .

We have  $|c^p| + 2|g| \geq |g^{-1}c^p g| \stackrel{+}{\succ} |h^{-1}c^p h c^{2p}|$ , since  $h^{-1}c^p h c^{2p}.o$  is close to a geodesic from  $o$  to  $g^{-1}c^p g.o$ . On the other hand, any geodesic from  $o$  to  $h^{-1}c^p h c^{2p}.o$  has projection to  $h^{-1}c^p h \mathcal{E}$  of diameter greater than  $|c^{2p}| - 3K$  by (3). This is much larger than

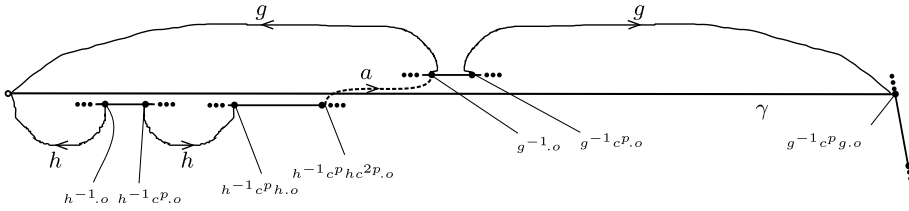


FIGURE 1.  $h^{-1}c^p h \mathcal{E}$  before  $g^{-1} \mathcal{E}$ , that is,  $h^{-1}c^p h \mathcal{E} \sqsubset g^{-1} \mathcal{E}$ .

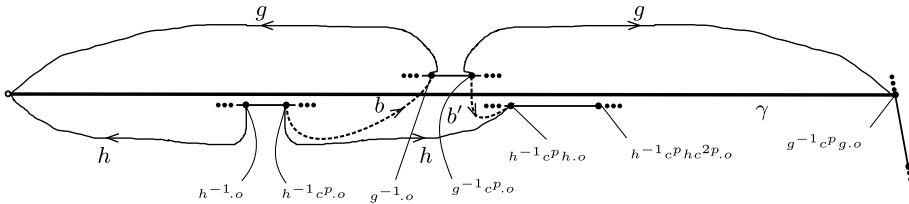


FIGURE 2.  $h^{-1}c^p h \mathcal{E}$  after  $g^{-1} \mathcal{E}$ , that is,  $g^{-1} \mathcal{E} \sqsubset h^{-1}c^p h \mathcal{E}$ .

$C'$  when  $p$  is large, so  $|h^{-1}c^p h c^{2p}| \stackrel{+}{\succ} |h^{-1}c^p h| + |c^{2p}| \stackrel{+}{\succ} 3|c^p| + 2|h|$  by Corollary 2.1 and Lemma 3.3. Thus,

$$|g| \stackrel{+}{\succ} |h| + |c^p|. \tag{7}$$

However, by definition of  $G_1$ , if  $h^{-1} \mathcal{E} = g^{-1} \mathcal{E}$ , then

$$4K \geq d_{g^{-1} \mathcal{E}}^\pi(o, g^{-1}.o) + d_{h^{-1} \mathcal{E}}^\pi(o, h^{-1}.o) \geq d(g^{-1}.o, h^{-1}.o) \geq |g| - |h| \stackrel{+}{\succ} |c^p|.$$

This is a contradiction for sufficiently large  $p$ . Similar considerations show that  $h^{-1} \mathcal{E} \neq g^{-1}c^p g \mathcal{E}$ , since  $o$  projects close to  $h^{-1}.o$  in  $h^{-1} \mathcal{E}$ , by definition of  $G_1$ , and close to  $g^{-1}c^p g.o$  in  $g^{-1}c^p g \mathcal{E}$ , by (3), but  $|h| \ll |g^{-1}c^p g|$ , by Lemma 3.3 and (7).

Next, we show that  $h^{-1} \mathcal{E}$  and  $h^{-1}c^p h \mathcal{E}$  belong to  $\mathbf{Y}[\mathcal{E}, g^{-1}c^p g \mathcal{E}]$  and in the course of the proof we will observe that  $g^{-1} \mathcal{E} \neq h^{-1}c^p h \mathcal{E}$ . By hypothesis, there exists  $t$  such that  $d(\gamma(t), h^{-1}c^p h c^{2p}.o) \leq D$ . This implies that  $d_{h^{-1}c^p h \mathcal{E}}^\pi(\gamma(t), h^{-1}c^p h c^{2p}.o) \leq 2D$ . Since  $d_{h^{-1}c^p h \mathcal{E}}^\pi(o, h^{-1}c^p h.o) < 3K$ , by (3), we have  $d_{h^{-1}c^p h \mathcal{E}}^\pi(o, \gamma(t)) \geq |c^{2p}| - 2D - 3K$ , which is large for  $p$  sufficiently large. Let  $t_0$  and  $t_1$  be the first and last times  $\gamma$  is distance  $C'$  from  $h^{-1}c^p h \mathcal{E}$ , as in Corollary 2.1 with respect to  $h^{-1}c^p h \mathcal{E}$ . We cannot have  $t \leq t_0$ , since then  $d_{h^{-1}c^p h \mathcal{E}}^\pi(o, \gamma(t)) \leq C'$ , which is a contradiction for large  $p$ .

If  $t \geq t_1$ , then  $d_{h^{-1}c^p h \mathcal{E}}^\pi(\gamma(t), g^{-1}c^p g.o) \leq C'$ , so

$$\begin{aligned} d_{h^{-1}c^p h \mathcal{E}}^\pi(o, g^{-1}c^p g.o) &\geq d_{h^{-1}c^p h \mathcal{E}}^\pi(o, \gamma(t)) - d_{h^{-1}c^p h \mathcal{E}}^\pi(\gamma(t), g^{-1}c^p g.o) \\ &\geq |c^{2p}| - 3K - 2D - C'. \end{aligned}$$

If  $t_0 < t < t_1$ , then we use Corollary 2.1 to say that  $d_{h^{-1}c^p h \mathcal{E}}^\pi(o, g^{-1}c^p g.o) \geq |\gamma(t_0, t_1)| - 4C'$  and then estimate

$$\begin{aligned}
 |\gamma(t_0, t_1)| &\geq d(\gamma(t_0), \gamma(t)) \\
 &\geq d(\pi_{h^{-1}c^ph\mathcal{E}}(\gamma(t_0)), \pi_{h^{-1}c^ph\mathcal{E}}(\gamma(t))) - C' - D \\
 &\geq d_{h^{-1}c^ph\mathcal{E}}^\pi(\gamma(t_0), \gamma(t)) - \text{diam } \pi_{h^{-1}c^ph\mathcal{E}}(\gamma(t_0)) \\
 &\quad - \text{diam } \pi_{h^{-1}c^ph\mathcal{E}}(\gamma(t)) - C' - D \\
 &\geq d_{h^{-1}c^ph\mathcal{E}}^\pi(\gamma(t_0), \gamma(t)) - 2C - C' - D \\
 &\geq d_{h^{-1}c^ph\mathcal{E}}^\pi(o, \gamma(t)) - d_{h^{-1}c^ph\mathcal{E}}^\pi(\gamma(t_0), o) - 2C - C' - D \\
 &\geq |c^{2p}| - 2D - 3K - C' - 2C - C' - D.
 \end{aligned}$$

Thus,  $h^{-1}c^ph\mathcal{E} \in \mathbf{Y}[\mathcal{E}, g^{-1}c^pg\mathcal{E}]$  once  $p$  is sufficiently large. Additionally, this shows that  $g^{-1}\mathcal{E} \neq h^{-1}c^ph\mathcal{E}$ , because, by (2) and (P 0), we have  $d_{g^{-1}\mathcal{E}}^\pi(\mathcal{E}, g^{-1}c^pg\mathcal{E}) \overset{+}{\asymp} |c^p|$ , while the estimates above show that  $d_{h^{-1}c^ph\mathcal{E}}^\pi(\mathcal{E}, g^{-1}c^pg\mathcal{E}) \overset{+}{\asymp} |c^{2p}|$ , and these are incompatible for sufficiently large  $p$ . Thus, the five axes are distinct.

From Corollary 2.1, we deduce that

$$d(h^{-1}c^phc^{2p}.o, h^{-1}\mathcal{E}) \overset{+}{\asymp} d_{h^{-1}c^ph\mathcal{E}}^\pi(h^{-1}c^phc^{2p}.o, h^{-1}\mathcal{E}) \overset{\pm}{\asymp} |c^{2p}|.$$

Thus, for large enough  $p$ , we have  $d(h^{-1}c^phc^{2p}.o, h^{-1}\mathcal{E}) \geq D \geq d(\gamma(t), h^{-1}c^phc^{2p}.o)$ , so strong contraction of  $h^{-1}\mathcal{E}$  implies that  $d_{h^{-1}\mathcal{E}}^\pi(\gamma(t), h^{-1}c^phc^{2p}.o) \leq C$ . Since  $o$  projects close to  $h^{-1}.o$  in  $h^{-1}\mathcal{E}$  and  $h^{-1}c^phc^{2p}.o \in h^{-1}c^ph\mathcal{E}$  projects close to  $h^{-1}c^p.o$ , Corollary 2.1 says that  $\gamma$  must pass close to  $h^{-1}c^p.o$ . Now, we can run the same argument as for  $h^{-1}c^ph\mathcal{E}$  to see that  $h^{-1}\mathcal{E} \in \mathbf{Y}[\mathcal{E}, g^{-1}c^pg\mathcal{E}]$  once  $p$  is sufficiently large.

The first step of the proof is completed by observing that  $g^{-1}\mathcal{E} \sqsubset h^{-1}\mathcal{E}$  implies that  $|h| \overset{+}{\asymp} |g| + |c^p|$ , which cannot be true when  $p$  is sufficiently large, by (7). Thus,  $h^{-1}\mathcal{E}$  comes before  $g^{-1}\mathcal{E}$  and  $h^{-1}c^ph\mathcal{E}$  under  $\sqsubset$ , and we are left with the possibilities that  $h^{-1}c^ph\mathcal{E} \sqsubset g^{-1}\mathcal{E}$ , as in Figure 1, or the converse, as in Figure 2.

In the case of Figure 1, we have  $h^{-1}c^ph\mathcal{E} \sqsubset g^{-1}\mathcal{E}$ , so the projection of  $h^{-1}c^phc^{2p}.o$  to  $g^{-1}\mathcal{E}$  is close to the projection of  $o$ , which we know to be close to  $g^{-1}.o$ . Write  $g^{-1}.o = h^{-1}c^phc^{2p}.a.o$  as in Figure 1 with  $|g| \overset{\pm}{\asymp} 2|h| + 3|c^p| + |a|$ .

In the case of Figure 2, we have  $h^{-1}\mathcal{E} \sqsubset g^{-1}\mathcal{E}$  and  $g^{-1}\mathcal{E} \sqsubset h^{-1}c^ph\mathcal{E}$ . The former implies that the projection of  $h^{-1}c^p.o$  to  $g^{-1}\mathcal{E}$  is close to the projection of  $o$ , which we know to be close to  $g^{-1}.o$ , while the latter implies that the projection of  $h^{-1}c^ph.o$  to  $g^{-1}\mathcal{E}$  is close to the projection of  $g^{-1}c^pg.o$ , which we know to be close to  $g^{-1}c^p.o$ . Write  $g^{-1}.o = h^{-1}c^pb.o$  with  $|g| \overset{\pm}{\asymp} |h| + |c^p| + |b|$  and write  $h.o = bc^pb'.o$  as in Figure 2 with  $|h| \overset{\pm}{\asymp} |b| + |c^p| + |b'|$ ; together these give  $|g| \overset{\pm}{\asymp} 2|b| + 2|c^p| + |b'|$ .

Suppose that we are in the case of Figure 2, so there are elements  $b$  and  $b'$  such that  $(r - |c^p|)/2 \overset{\pm}{\asymp} |g| \overset{\pm}{\asymp} 2|b| + 2|c^p| + |b'|$ . Since  $G$  has purely exponential growth, if  $i \leq |b| < i + 1$ , there are, up to a bounded multiplicative error independent of  $p, r$ , and  $i$ , at most  $\exp(\delta_G i)$  possible choices for  $b$  and at most  $\exp(\delta_G((r - 5|c^p|)/2 - 2i))$  choices of  $b'$ , so there is an upper bound for the number of possible elements  $g$  by a multiple of

$$\sum_{i=0}^{(r-5|c^p|)/4} \exp(\delta_G i) \exp\left(\delta_G\left(\frac{r - 5|c^p|}{2} - 2i\right)\right) < \frac{\exp(r\delta_G/2)}{\exp(5\delta_G|c^p|/2)(1 - \exp(-\delta_G))}. \tag{8}$$

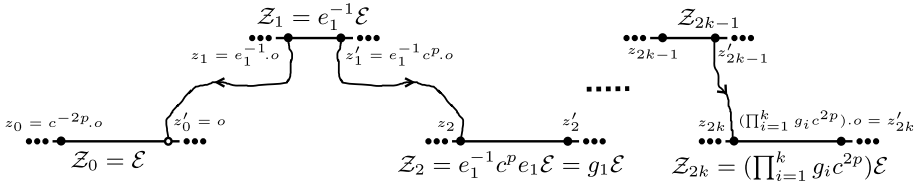


FIGURE 3.  $(\prod_{i=1}^k g_i c^{2p}).o$ .

The case of Figure 1 is similar, but gives an even smaller upper bound†. Thus, for all sufficiently large  $p$  and  $r$ ,

$$\#G'_{4,p,D} \cap S_r^{\Delta'} \lesssim \exp(-5\delta_G|c^p|/2) \exp(r\delta_G/2). \tag{9}$$

Combining (6) and (9) gives

$$\#G'_{4,p,D} \cap S_r^{\Delta'} \lesssim \exp(-2|c^p|\delta_G) \cdot \#G_{3,p} \cap S_r^{\Delta'}. \tag{10}$$

Crucially, the multiplicative constant in this asymptotic inequality does not depend on  $p$ , so, for  $p$  sufficiently large,  $\exp(2|c^p|\delta_G)$  is more than twice the multiplicative constant and (10) becomes a true inequality  $\#G'_{4,p,D} \cap S_r^{\Delta'} < \frac{1}{2}\#G_{3,p} \cap S_r^{\Delta'}$ . We conclude that to get  $G_{4,p,D}$  from  $G_{3,p}$  we threw away fewer than half of the points of  $G_{3,p}$  in each spherical shell  $S_r^{\Delta'}$  such that  $r > 7|c^p|$ . □

LEMMA 3.7. For all sufficiently large  $D$ , for all sufficiently large  $p$ , the map  $\hat{G}_{4,p,D} \rightarrow X : (g_1, \dots, g_k) \mapsto (\prod_{i=1}^k g_i c^{2p}).o$  is an injection.

Proof. Consider a point  $(\prod_{i=1}^k g_i c^{2p}).o$  in the image. Set  $g_0 := c^{-2p}$ . Suppose that for each  $i$ , we have  $g_i = e_i^{-1}c^pe_i$  for  $e_i \in G_1$ . For  $0 \leq i \leq k$ , set  $z'_{2i} := (\prod_{j=0}^i g_j c^{2p}).o$ ,  $z_{2i} := (\prod_{j=0}^i g_j c^{2p})c^{-2p}.o$ , and  $Z_{2i} := (\prod_{j=0}^i g_j c^{2p})E$ . For  $0 < i \leq k$ , set  $z_{2i-1} := (\prod_{j=0}^{i-1} g_j c^{2p})e_{2i-1}^{-1}.o$ ,  $z'_{2i-1} := (\prod_{j=0}^{i-1} g_j c^{2p})e_{2i-1}^{-1}c^p.o$ , and  $Z_{2i-1} := (\prod_{j=0}^{i-1} g_j c^{2p})e_{2i-1}^{-1}E$ . See Figure 3.

Let us complete the proof assuming the following claim, to which we shall return:

$$\text{for all } 0 \leq i < j \leq 2k, \quad d_{Z'_i}^\pi(z'_i, Z_j) < 5K, \quad \text{and} \quad d_{Z'_j}^\pi(z_j, Z_i) < 5K. \tag{11}$$

When  $p$  is sufficiently large,  $d(z_i, z'_i) \gg 10K$  for all  $i$ , so (11) implies that  $Z_i \sqsubset Z_j$  for all  $0 \leq i < j \leq 2k$ , where  $\sqsubset$  is the order of Proposition 2.3 on  $\mathbf{Y}[Z_0, Z_{2k}]$ .

Suppose that the map  $\hat{G}_{4,p,D} \rightarrow X$  is not an injection; there exist distinct elements  $(g_1, \dots, g_m)$  and  $(h_1, \dots, h_n)$  of  $\hat{G}_{4,p,D}$  with the same image  $z \in X$ . Suppose that  $m + n$  is minimal among such tuples. If  $h_1E = g_1E$ , then  $h_1 = g_1$ , by Lemma 3.5. This contradicts minimality of  $m + n$ , so we must have  $h_1E \neq g_1E$ . Let  $Z_0, \dots, Z_{2m}$  be as in Figure 3 for  $(g_1, \dots, g_m)$ . By definition,  $o \in Z_0$  and  $z \in Z_{2m}$ . By (11),  $\pi_{Z_{2m}}(o)$  is close to  $z_{2m}$ . By Corollary 2.1, any geodesic from  $o$  to  $z$  ends with a segment that stays close to the subsegment of  $Z_{2m}$  between  $z_{2m}$  and  $z = z'_{2m}$ . However, if  $Z'_0, \dots, Z'_{2n}$  are as in Figure 3 for  $(h_1, \dots, h_n)$ , then the same is true for  $Z'_{2n}$ , which implies that

† Replace each ‘5’ in (8) with a ‘7’. This accounts for the restriction that  $r - 7|c^p| > 0$ .

$d_{\mathcal{Z}_{2m}}^\pi(\mathcal{Z}'_{2n}, \mathcal{Z}'_{2n}) \stackrel{+}{\succ} d(z_{2m}, z'_{2m}) = |c^{2p}|$ . Once  $p$  is sufficiently large, (P 0) requires that  $\mathcal{Z}_{2m} = \mathcal{Z}'_{2n}$ . Thus,  $\mathbf{Y}[\mathcal{Z}_0, \mathcal{Z}_{2m}] = \mathbf{Y}[\mathcal{Z}'_0, \mathcal{Z}'_{2n}]$  and all of the  $\mathcal{Z}_i$  and  $\mathcal{Z}'_j$  are comparable in the order  $\sqsubset$  on  $\mathbf{Y}[\mathcal{Z}_0, \mathcal{Z}_{2m}]$ . In particular,  $\mathcal{Z}'_2 = h_1\mathcal{E} \neq g_1\mathcal{E} = \mathcal{Z}_2$ , so one of them comes before the other. Suppose, without loss of generality, that  $h_1\mathcal{E} \sqsubset g_1\mathcal{E}$ . Then  $d_{h_1\mathcal{E}}^\pi(g_1\mathcal{E}, \mathcal{Z}_{2m}) \leq \theta'$ , by Proposition 2.3, and  $d_{h_1\mathcal{E}}^\pi(\mathcal{Z}_{2m}, h_1c^{2p}.o) < 5K$ , by (11), so

$$\begin{aligned} d_{h_1\mathcal{E}}^\pi(g_1.o, h_1c^{2p}.o) &\leq d_{h_1\mathcal{E}}^\pi(g_1\mathcal{E}, \mathcal{Z}_{2m}) + d_{h_1\mathcal{E}}^\pi(\mathcal{Z}_{2m}, h_1c^{2p}.o) \\ &< \theta' + 2\theta + 5K < 7K. \end{aligned}$$

On the other hand,  $d_{h_1\mathcal{E}}^\pi(o, h_1.o) < 3K$ , by (3), so  $d_{h_1\mathcal{E}}^\pi(o, g_1.o) \geq |c^{2p}| - 10K \gg C'$ . By Corollary 2.1, any geodesic from  $o$  to  $g_1.o$  passes within distance  $2C'$  of  $\pi_{h_1\mathcal{E}}(g_1.o)$ , which is less than  $7K$  from  $h_1c^{2p}.o$ . This means that  $g_1 \in G'_{4,p.(7K+2C')}$ , which is a contradiction if  $D \geq 7K + 2C'$ . Thus, if  $D \geq 7K + 2C'$ , then for sufficiently large  $p$  the map is injective.

We prove (11) by induction on  $m = j - i$ . For each  $0 \leq i < 2k$ , we have that  $z'_i$  and  $z_{i+1}$  differ by an element of  $G_1$ , so  $\mathcal{Z}_i \neq \mathcal{Z}_{i+1}$  and  $d_{\mathcal{Z}_{i+1}}^\pi(z_{i+1}, z'_i) \leq 2K$ . Furthermore, by (P 0),  $d_{\mathcal{Z}_{i+1}}^\pi(\mathcal{Z}_i, \mathcal{Z}_i) \leq \theta$ . Thus,

$$\begin{aligned} d_{\mathcal{Z}_{i+1}}^\pi(z_{i+1}, \mathcal{Z}_i) &\leq d_{\mathcal{Z}_{i+1}}^\pi(z_{i+1}, z'_i) + d_{\mathcal{Z}_{i+1}}^\pi(z'_i, \mathcal{Z}_i) \\ &\leq d_{\mathcal{Z}_{i+1}}^\pi(z_{i+1}, z'_i) + d_{\mathcal{Z}_{i+1}}^\pi(\mathcal{Z}_i, \mathcal{Z}_i) \\ &\leq 2K + \theta < 3K. \end{aligned}$$

Similarly,  $d_{\mathcal{Z}_i}^\pi(z'_i, \mathcal{Z}_{i+1}) < 3K$ .

Now, extend  $m$  to  $m + 1$ : suppose that for some  $m \geq 1$  and all  $0 < j - i \leq m$ , we have  $d_{\mathcal{Z}_j}^\pi(z_j, \mathcal{Z}_i) < 5K$  and  $d_{\mathcal{Z}_i}^\pi(z'_i, \mathcal{Z}_j) < 5K$ . (Note that this implies that  $\mathcal{Z}_i \neq \mathcal{Z}_j$ .) Then, for all  $0 \leq i \leq 2k - m - 1$ ,

$$d_{\mathcal{Z}_{i+1}}^\pi(\mathcal{Z}_{i+m+1}, \mathcal{Z}_i) \geq d_{\mathcal{Z}_{i+1}}^\pi(\mathcal{Z}_{i+m+1}, \mathcal{Z}_i) - 2\theta > d(z_{i+1}, z'_{i+1}) - 10K - 2\theta \gg \theta'.$$

The final inequality is true for sufficiently large  $p$ , because the distance between  $z_{i+1}$  and  $z'_{i+1}$  is either  $|c^p|$  or  $|c^{2p}| \stackrel{+}{\succ} 2|c^p|$ , according to whether  $i$  is even or odd. Thus, by (SP 3) and (SP 4),

$$\begin{aligned} d_{\mathcal{Z}_i}^\pi(\mathcal{Z}_{i+m+1}, \mathcal{Z}_{i+1}) &\leq d_{\mathcal{Z}_i}^\pi(\mathcal{Z}_{i+m+1}, \mathcal{Z}_{i+1}) + 2\theta \\ &= d_{\mathcal{Z}_i}^\pi(\mathcal{Z}_{i+1}, \mathcal{Z}_{i+1}) + 2\theta \leq \theta' + 2\theta < 2K, \end{aligned}$$

which implies that

$$d_{\mathcal{Z}_i}^\pi(z'_i, \mathcal{Z}_{i+m+1}) \leq d_{\mathcal{Z}_i}^\pi(z'_i, \mathcal{Z}_{i+1}) + d_{\mathcal{Z}_i}^\pi(\mathcal{Z}_{i+1}, \mathcal{Z}_{i+m+1}) < 3K + 2K = 5K.$$

A similar argument gives  $d_{\mathcal{Z}_{i+m+1}}^\pi(z_{i+m+1}, \mathcal{Z}_i) < 5K$ . This completes the induction. □

*Proof of Proposition 3.1.* Take  $D$  and  $p$  as in Lemma 3.7. For this  $D$ , enlarge  $p$  if necessary to satisfy the hypotheses of Lemma 3.6. Set  $G_4 := G_{4,p,D}$ . □



## 4. Questions

QUESTION 4.1. *Can we replace purely exponential growth of  $G$  by divergence of  $G$  in Theorem 1.1?*

By [19], the answer is ‘yes’ when  $X$  is hyperbolic.

Recall in (5) that we showed that  $\Theta_G(s)$  is comparable to  $\Theta_{G_{3,p}}(s/2)$ , while it is clear that  $\Theta_{G_{3,p}}(s/2) \leq \Theta_N(s/2)$ . If  $G$  is divergent, then  $\Theta_G(s)$  diverges at  $s = \delta_G$ , which means that  $\Theta_N(t)$  diverges at  $t = \delta_G/2$ . There are two possible circumstances in which  $\Theta_N(t)$  diverges at  $t = \delta_G/2$ :

$$\text{Either } \delta_N > \delta_G/2, \text{ or } \delta_N = \delta_G/2 \text{ and } N \text{ is divergent.} \quad (12)$$

We proved the first case of (12) directly, with the additional assumption of purely exponential growth of  $G$ . The approach of [19] is to prove, if  $X$  is hyperbolic, that  $\delta_N = \delta_G$  when  $N$  is divergent, so, since  $\delta_G > \delta_G/2$ , the second case of (12) is impossible. Thus, a positive answer to Question 4.1 would be implied by a positive answer to the following question, which is also interesting in its own right.

QUESTION 4.2. *If  $G$  is a group acting properly by isometries with a strongly contracting element on a geodesic metric space  $X$  and  $G \curvearrowright X$  is divergent, is it true that for every divergent normal subgroup  $N$  of  $G$  we have  $\delta_N = \delta_G$ ?*

Jaerisch and Matsuzaki [17] showed that if  $F$  is a finite-rank free group and  $N$  is a non-trivial normal subgroup of  $F$ , then, with respect to a word metric defined by a free generating set of  $F$ , there is an inequality  $\delta_N + \frac{1}{2}\delta_{F/N} \geq \delta_F$ . Notice that  $\delta_N > \delta_F/2$ , by the lower cogrowth bound, and  $\delta_{F/N} < \delta_F$ , by growth tightness of  $F$ .

QUESTION 4.3. *Is there an analogue of Jaerisch and Matsuzaki’s inequality for  $G$  acting with a strongly contracting element and complementary growth gap? Note that we know both growth tightness, by [2], and the lower cogrowth bound, by Theorem 1.1, for such actions.*

For  $G = X = F_n$  [7, 13, 20],  $X = \mathbb{H}^2$ , and  $G$  a closed surface group [5], there exists a sequence  $(N_i)_{i \in \mathbb{N}}$  of normal subgroups of  $G$  such that  $\delta_{N_i}/\delta_G$  limits to  $1/2$ , so the lower cogrowth bound is optimal.

QUESTION 4.4. *Is the lower cogrowth bound optimal in Theorem 1.1?*

We must mention that the *upper* cogrowth bound is also very interesting. Grigorchuk [13] and Cohen [8] showed that when  $F$  is a finite-rank free group, with respect to a word metric defined by a free generating set, the upper cogrowth bound  $\delta_N/\delta_F = 1$  is achieved for  $N \triangleleft F$  if and only if  $F/N$  is amenable. There have been several generalizations [6, 10, 12, 21, 22] to growth rates defined with respect to an action  $G \curvearrowright X$ , but the most general to date [10] still requires  $G$  to be hyperbolic, the action to be cocompact, and  $X$  to be either a Cayley graph of  $G$  or a CAT(−1) space. In the vein of our theorem, it would be very interesting to generalize such a result to a non-hyperbolic setting.

*Acknowledgements.* We thank the referee for the careful reading and helpful comments. This research was partially supported by the European Research Council (ERC) grant of the first author, ‘ANALYTIC’ grant agreement no. 259527. The second author is supported by the Austrian Science Fund (FWF): P 30487-N35.

## REFERENCES

- [1] G. N. Arzhantseva, C. H. Cashen, D. Gruber and D. Hume. Negative curvature in graphical small cancellation groups. *Groups Geom. Dyn.* to appear.
- [2] G. N. Arzhantseva, C. H. Cashen and J. Tao. Growth tight actions. *Pacific J. Math.* **278**(1) (2015), 1–49.
- [3] M. Bestvina, K. Bromberg and K. Fujiwara. Constructing group actions on quasi-trees and applications to mapping class groups. *Publ. Math. Inst. Hautes Études Sci.* **122**(1) (2015), 1–64.
- [4] M. Bestvina, K. Bromberg, K. Fujiwara and A. Sisto. Acylindrical actions on projection complexes. *Preprint*, 2017, arXiv:1711.08722v1.
- [5] P. Bonfert-Taylor, K. Matsuzaki and E. C. Taylor. Large and small covers of a hyperbolic manifold. *J. Geom. Anal.* **22**(2) (2012), 455–470.
- [6] R. Brooks. The bottom of the spectrum of a Riemannian covering. *J. Reine Angew. Math.* **357** (1985), 101–114.
- [7] C. Champetier. Cocroissance es groupes à petite simplification. *Bull. Lond. Math. Soc.* **25**(5) (1993), 438–444.
- [8] J. M. Cohen. Cogrowth and amenability of discrete groups. *J. Funct. Anal.* **48**(3) (1982), 301–309.
- [9] M. Coornaert. Mesures de Patterson–Sullivan sur le bord d’un espace hyperbolique au sens de Gromov. *Pacific J. Math.* **159**(2) (1993), 241–270.
- [10] R. Coulon, F. Dal’Bo and A. Sambusetti. Growth gap in hyperbolic groups and amenability. *Preprint*, 2017, arXiv:1709.07287v1.
- [11] F. Dal’Bo, M. Peigné, J.-C. Picaud and A. Sambusetti. On the growth of quotients of Kleinian groups. *Ergod. Th. & Dynam. Sys.* **31**(3) (2011), 835–851.
- [12] R. Dougall and R. Sharp. Amenability, critical exponents of subgroups and growth of closed geodesics. *Math. Ann.* **365**(3–4) (2016), 1359–1377.
- [13] R. I. Grigorchuk. Symmetrical random walks on discrete groups. *Multicomponent Random Systems (Advances in Probability and Related Topics, 6)*. Dekker, New York, 1980, pp. 285–325.
- [14] R. Grigorchuk and P. de la Harpe. On problems related to growth, entropy, and spectrum in group theory. *J. Dyn. Control Syst.* **3**(1) (1997), 51–89.
- [15] P. de la Harpe. *Topics in Geometric Group Theory (Chicago Lectures in Mathematics)*. University of Chicago Press, Chicago, IL, 2000.
- [16] J. Jaerisch. A lower bound for the exponent of convergence of normal subgroups of Kleinian groups. *J. Geom. Anal.* **25**(1) (2015), 298–305.
- [17] J. Jaerisch and K. Matsuzaki. Growth and cogrowth of normal subgroups of a free group. *Proc. Amer. Math. Soc.* **145**(10) (2017), 4141–4149.
- [18] K. Matsuzaki. Growth and cogrowth tightness of Kleinian and hyperbolic groups. *Geometry and Analysis of Discrete Groups and Hyperbolic Spaces (RIMS Kôkyûroku Bessatsu, B66)*. Eds. M. Fujii, N. Kawazumi and K. Ohshika. Research Institute for Mathematical Sciences, Kyoto, 2017, pp. 21–36.
- [19] K. Matsuzaki, Y. Yabuki and J. Jaerisch. Normalizer, divergence type and Patterson measure for discrete groups of the Gromov hyperbolic space. *Preprint*, 2015, arXiv:1511.02664v1.
- [20] Y. Ollivier. Cogrowth and spectral gap of generic groups. *Ann. Inst. Fourier (Grenoble)* **55**(1) (2005), 289–317.
- [21] T. Roblin. Un théorème de Fatou pour les densités conformes avec applications aux revêtements Galoisien en courbure négative. *Israel J. Math.* **147** (2005), 333–357.
- [22] T. Roblin and S. Tapie. Exposants critiques et moyennabilité. *Géométrie Ergodique (Monographs of L’Enseignement Mathématique, 43)*. L’Enseignement Mathématique, Geneva, 2013, pp. 61–92.
- [23] Y. Shalom. Rigidity, unitary representations of semisimple groups, and fundamental groups of manifolds with rank one transformation group. *Ann. of Math. (2)* **152**(1) (2000), 113–182.
- [24] D. Sullivan. Related aspects of positivity in Riemannian geometry. *J. Differential Geom.* **25**(3) (1987), 327–351.
- [25] W. Yang. Statistically convex-cocompact actions of groups with contracting elements. *Int. Math. Res. Not.* to appear.