Journal de l'École polytechnique
Mathématiques

Rémi Coulon & Markus Steenbock

Product set growth in Burnside groups

Tome 9 (2022), p. 463-504.

<http://jep.centre-mersenne.org/item/JEP_2022__9__463_0>

Journal de l'École polytechnique
Mathématiques

Tome 9, 2022, p. 463–504 DOI: 10.5802/jep.187

PRODUCT SET GROWTH IN BURNSIDE GROUPS

by Rémi Coulon & Markus Steenbock

Abstract. — Given a periodic quotient of a torsion-free hyperbolic group, we provide a fine lower estimate of the growth function of any sub-semi-group. This generalizes results of Razborov and Safin for free groups.

Résumé (Croissance des ensembles produit dans les groupes de Burnside)
Étant donné un quotient périodique d'un groupe hyperbolique sans torsion, nous donnons une estimation inférieure fine de la fonction de croissance pour chacun de tous ses sous-semi-groupes. Cet énoncé généralise des résultats de Razborov et Safin pour les groupes libres.

Contents

1. Introduction ... 463
2. Hyperbolic geometry ... 468
3. Periodic and aperiodic words ... 472
4. Power-free elements ... 480
5. Energy and quasi-center ... 481
6. Sets of diffuse energy ... 483
7. Sets of concentrated energy ... 488
8. Growth in groups acting on hyperbolic spaces ... 490
9. Small cancellation groups ... 492
10. Product set growth in Burnside groups of odd exponent ... 497
References ... 502

1. Introduction

If V is a subset in a group G, we denote by $V^r \subset G$ the set of all group elements that are represented by a product of exactly r elements of V. In this paper we are interested in the growth of V^r. Such a problem has a long history which goes back (at least) to the

---

Mathematical subject classification (2020). — 20F65, 20F67, 20F50, 20F06, 20F69.
Keywords. — Product sets, growth, hyperbolic groups, acylindrical actions, small cancellation, infinite periodic groups, Burnside problem.

The first author acknowledges support from the *Centre Henri Lebesgue* ANR-11-LABX-0020-01 and the *Agence Nationale de la Recherche* under Grant *Dagger* ANR-16-CE40-0006-01. The second author was supported in parts by the *Austrian Science Fund (FWF)* project J 4270-N35 and the *Austrian Science Fund (FWF)* project P 35079-N, and thanks the Université de Rennes 1 for hospitality during his stay in Rennes.

study of additive combinatorics. See for instance [Nat96, TV06]. In the context of non-abelian groups, it yields to the theory of approximate subgroups, see [Tao08, BGT12], and relates to spectral gaps in linear groups, see [Hel08, BG08, BG12], as well as exponential growth rates of negatively curved groups, [Kou98, AL06, BF21, FS20].

If G is a free group, Safin [Saf11], improving former results by Chang [Cha08] and Razborov [Raz14], proves that there exists $c > 0$ such that for every finite subset $V \subset G$, either V is contained in a cyclic subgroup, or for every $r \in \mathbb{N}$, we have

$$|V^r| \geqslant (c|V|)^{[(r+1)/2]}.$$

This estimate can be thought of as a quantified version of the Tits alternative in G. A similar statement holds for $\mathrm{SL}_2(\mathbb{Z})$ [Cha08], free products, limit groups [But13] and groups acting on δ-hyperbolic spaces [DS20]. All these groups display strong features of negative curvature, inherited from a non-elementary acylindrical action on a hyperbolic space. Some results are also available for solvable groups [Tao10, But13], as well as mapping class groups and right-angled Artin groups [Ker21].

By contrast, in this work, we focus on a class of groups which do not admit any non-elementary action on a hyperbolic space, namely the set of infinite groups with finite exponent, often referred to as of Burnside groups.

1.1. Burnside groups of odd exponent.

— Given a group G and an integer n, we denote by G^n the subgroup of G generated by all its n-th powers. We are interested in quotients of the form G/G^n which we call *Burnside groups* of exponent n. If $G = \mathbb{F}_k$ is the free group of rank k, then $B_k(n) = G/G^n$ is the *free Burnside group* of rank k and exponent n. The famous Burnside problem asks whether a finitely generated free Burnside group is necessarily finite.

Here, we focus on the case that the exponent n is odd. By Novikov's and Adian's solution of the Burnside problem, it is known that $B_k(n)$ is infinite provided $k \geqslant 2$ and n is a sufficiently large odd integer [Adi79]. See also [Ol'82, DG08]. More generally, if G is a non-cyclic, torsion-free, hyperbolic group, then the quotient G/G^n is infinite provided n is a sufficiently large odd exponent [Ol'91, DG08]. Our main theorem extends Safin's result to this class of Burnside groups of odd exponents.

Remark 1.1. — Note that free Burnside groups of sufficiently large even exponents are also infinite. This was independently proved by Ivanov [Iva94] and Lysenok [Lys96]. Moreover any non-elementary hyperbolic group admits infinite Burnside quotients, see [IO96, Cou18b]. Nevertheless in the remainder of this article we will focus on torsion-free hyperbolic groups and odd exponents. In Section 1.4 we discuss the difficulties to extend our results to the case of even exponents.

Theorem 1.2. — *Let G be a non-cyclic, torsion-free hyperbolic group. There are numbers $n_0 > 0$ and $c > 0$ such that for all odd integers $n \geqslant n_0$ the following holds. Given a finite subset $V \subset G/G^n$, either V is contained in a finite cyclic subgroup, or for all $r \in \mathbb{N}$, we have*

$$|V^r| \geqslant (c|V|)^{[(r+1)/2]}.$$

Observe that the constant c only depends on G and not on the exponent n. Recall that Burnside groups do not act, at least in any useful way, on a hyperbolic space. Indeed, any such action is either elliptic or parabolic. On the other hand, it is well-known that any linear representation of a finitely generated Burnside group has finite image. Thus our main theorem is not a direct application of previously known results.

Let us mention some consequences of Theorem 1.2. If V is a finite subset of a group G, one defines its *entropy* by

$$h(V) = \limsup_{r\to\infty} \frac{1}{r} \log |V^r|.$$

The group G has *uniform exponential growth* if there exists $\varepsilon > 0$ such that for every finite symmetric generating subset V of G, $h(V) > \varepsilon$. In addition, G has *uniform uniform exponential growth* if there exists $\varepsilon > 0$ such that for every finite symmetric subset $V \subset G$, either V generates a virtually nilpotent group, or $h(V) > \varepsilon$.

Corollary 1.3. — *Let G be a non-cyclic, torsion-free hyperbolic group. There are numbers $n_0 > 0$ and $\alpha > 0$ such that for all odd integers $n \geqslant n_0$, the following holds. Given a finite subset $V \subset G/G^n$ containing the identity, either V is contained in a finite cyclic subgroup, or*

$$h(V) \geqslant \alpha \ln |V| \geqslant \alpha \ln 3.$$

In particular, G/G^n has uniform uniform exponential growth.

It was already known that free Burnside groups of sufficiently large odd exponent have uniform exponential growth, see Osin [Osi07, Cor. 1.4] and Atabekyan [Ata09, Cor. 3]. Note that Theorem 2.7 in [Osi07] actually shows that free Burnside groups have uniform uniform exponential growth. Nevertheless, to the best of our knowledge, the result was not proved for Burnside quotients of hyperbolic groups. We shall also stress the fact that, unlike in Corollary 1.3, the growth estimates provided in [Osi07, Ata09] depend on the exponent n. The reason is that the parameter M given for instance by [Osi07, Th. 2.7] is a quadratic function of n.

Given a group G with uniform exponential growth, a natural question is whether or not there exists a finite generating set that realizes the minimal growth rate. The first inequality is a statement *à la* Arzhantseva-Lysenok for torsion groups, see [AL06, Th. 1]. The philosophy is the following: if the set V has a small entropy, then it cannot have a large cardinality. In particular, if we expect the minimal growth rate to be achieved, we can restrict our investigation to generating sets with fixed cardinality. Note that this is exactly the starting point of the work of Fujiwara and Sela in the context of hyperbolic groups, [FS20].

Let us discuss now the power arising in Theorem 1.2. We claim that, as our estimate is independent of the exponent n, the power $(r+1)/2$ is optimal. For this purpose we adapt an example of [Saf11].

Example 1.4. — Let g and h be two elements in $B_2(n)$ such that g generates a group of order n, that does not contain h. Consider the set

$$V_N = \{1, g, g^2, \dots, g^N, h\}.$$

Whenever the exponent n is sufficiently large compared to N, we have $|V_N^r| \sim N^{[(r+1)/2]}$ while $|V_N| = N+1$.

Button observed the following fact. Assume that there is $c > 0$ and $\varepsilon > 0$ with the following property: for all finite subsets V in a group G that are not contained in a virtually nilpotent subgroup, we have $|V^3| \geqslant c|V|^{2+\varepsilon}$. Then G is either virtually nilpotent, or of bounded exponent [But13, Prop. 4.1]. We do not know if such a non-virtually nilpotent group exists.

1.2. Groups acting on hyperbolic spaces.

— In the first part of our paper, we revisit product set growth for a group G acting on a hyperbolic space X, see [DS20, Th. 1.14]. For this purpose, we use the notion of an acylindrical action, see [Sel97, Bow08]. Given a subset U of G, we exploit its *ℓ^∞-energy* defined as

$$\lambda(U) = \inf_{x\in X} \sup_{u\in U} |ux - x|.$$

Remark 1.5. — Unlike in [DS20], we will not make use of the ℓ^1-energy. Our motivation is mostly technical. We explain this choice in Section 1.3.

Theorem 1.6 (see Theorem 8.1). — *Let G be a group acting acylindrically on a hyperbolic length space X. There exists a constant $C > 0$ such that for every finite subset $U \subset G$ with $\lambda(U) > C$,*

(1) *either $|U| \leqslant C$,*

(2) *or there is a subset $W \subset U^2$ freely generating a free sub-semigroup of cardinality*

$$|W| \geqslant \frac{1}{C}\,\frac{1}{\lambda(U)}\,|U|.$$

Remark 1.7. — For simplicity we stated here a weakened form of Theorem 8.1. Actually we prove that the constant C only depends on the hyperbolicity constant of the space X and the acylindricity parameters of the action of G. The set W is also what we called *strongly reduced*, see Definition 3.1. Roughly speaking this means that the orbit map from the free semi-group W^* to X is a quasi-isometric embedding.

There is quite some literature on finding free sub-semigroups in powers of symmetric subsets U in groups of negative curvature, see [Kou98, AL06, BF21]. We can for example compare Theorem 8.1 to Theorem 1.13 of [BF21]. In this theorem, under the additional assumption that U is symmetric, the authors construct a 2-element set in U^r that generates a free sub-semigroup, where the exponent r does only depend on the doubling constant of the space. Let us highlight two important differences. First we do not assume that the set U is symmetric. In particular, we cannot build the generators of a free sub-semigroup by conjugating a given hyperbolic element. Hence the proofs require different techniques. Moreover, for our purpose, it is important that the cardinality of W grows linearly with the one of U. For the optimality of our estimates discussed in the previous paragraph, we require that it is contained in U^2. The price that we pay for this is the correction term of the order of the ℓ^∞-energy of U.

As the set W constructed in Theorem 1.6 freely generates a free sub-semigroup, we obtain the following estimate on the growth of U^r.

Corollary 1.8 (see Corollary 8.2). — *Let G be a group acting acylindrically on a hyperbolic length space X. There exists a constant $C > 0$ such that for every finite $U \subset G$ with $\lambda(U) > C$, and for all integers $r \geqslant 0$, we have*

$$|U^r| \geqslant \left(\frac{1}{C\lambda(U)}|U|\right)^{[(r+1)/2]}.$$

As in the previous statement, the constant C actually only depends on the parameters of the action of G on X. Corollary 1.8 is a variant of [DS20, Th. 1.14], where the correction term of the order of $\log|U|$ in this theorem is replaced by a geometric quantity, the ℓ^∞-energy of U. Note that the conclusion is void whenever $|U| \leqslant C\lambda(U)$. This can be compared with Theorem 1.2 which is not relevant for small subsets V.

1.3. Strategy for Burnside groups.

— Let us explain the main idea behind the proof of Theorem 1.2. For simplicity we restrict ourselves to the case of free Burnside groups of rank 2. Let n be a sufficiently large odd exponent. Any known strategy to prove the infiniteness of $B_2(n)$ starts in the same way. One produces a sequence of groups

$$\mathbb{F}_2 = G_0 \longrightarrow G_1 \longrightarrow G_2 \longrightarrow \dots \longrightarrow G_i \longrightarrow G_{i+1} \longrightarrow \cdots \tag{1}$$

that converges to $B_2(n)$ where each G_i is a hyperbolic group obtained from G_{i-1} by means of small cancellation. The approach provided by Delzant and Gromov associates to each group G_i a hyperbolic space X_i on which it acts properly co-compactly. An important point is that the geometry of X_i is somewhat “finer” than the one of the Cayley graph of G_i. In particular, one controls uniformly along the sequence (G_i, X_i), the hyperbolicity constant of X_i as well as the acylindricity parameters of the action of G_i, see Proposition 10.1. As we stressed before the constant C involved in Theorem 1.6 only depends on those parameters. Thus it holds, with the same constant C, for each group G_i acting on X_i.

Consider now a subset $V \subset B_2(n)$ that is not contained in a finite subgroup. Our idea is to choose a suitable step j and a pre-image U_j in G_j such that the ℓ^∞-energy $\lambda(U_j)$ is greater than C and at the same time bounded from above by a constant C' that does not depend on j. The strategy for choosing j is the following. The metric spaces X_i defined above come with uniformly contracting maps $X_i \to X_{i+1}$. Hence if $\widetilde{V}$ stands for a finite pre-image of V in $\mathbb{F}_2$, then the energy of its image V_i in G_i is a decreasing sequence converging to zero. Hence there is a smallest index j such that V admits a pre-image U_{j+1} in G_{j+1} whose energy is at most C. Working with the ℓ^∞-energy plays now an important role. Indeed we have a control of the length of *every* element in U_j. This allow us to lift U_{j+1} to a finite subset $U_j \subset G_j$ whose energy is controlled (i.e. bounded above by some C'). It follows from the minimality of j that the energy of U_j is also bounded from below by C. The details of the construction are given in Section 10.3. By Theorem 1.6, we find a "large" subset $W \subset U_j^2$ that freely generates a free sub-semigroup. By large we mean that the cardinality of W is linearly bounded from below by the cardinality of U_j (hence of V).

At this point we get an estimate for the cardinality of W^r, hence for the one of $U_j^r \subset G_j$, see Corollary 1.8. However the map $G_j \to G/G^n$ is not one-to-one. Nevertheless there is a sufficient condition to check whether two elements g and g' in G_j have distinct images in G/G^n: roughly speaking, if none of them "contains a subword" of the form u^m, with $m \geqslant n/3$, then g and g' have distinct images in G/G^n. This formulation is purposely vague here. We refer the reader to Definition 4.1 for a rigorous definition of power-free elements in G_j. In particular, the projection $G_j \to G/G^n$ is injective when restricted to a suitable set of power-free elements. Hence it suffices to count the number of power-free elements in W^r. This is the purpose of Sections 3 and 4. The computation is done by induction on r following the strategy of the first author from [Cou13].

Again, we would like to draw the attention of the reader to the fact that in this procedure, we took great care to make sure that all the involved parameters do not depend on j.

1.4. Burnside groups of even exponent.

— Burnside groups of even exponent have a considerably different algebraic structure. For instance it turns out that the approximation groups G_j in the sequence (1) contain elementary subgroups of the form $D_\infty \times F$ where F is a finite subgroup with arbitrary large cardinality that embeds in a product of dihedral groups. In particular one cannot control acylindricity parameters along the sequence (G_i), which means that our strategy fails here. It is very plausible that Burnside groups of large even exponents have uniform uniform exponential growth. Nevertheless we wonder if Theorem 1.2 still holds for such groups.

Acknowledgements. — The second author thanks Thomas Delzant for related discussion during his stay in Strasbourg. We thank the coffeeshop *Bourbon d'Arsel* for welcoming us when the university was closed down during the pandemic, and for serving a wonderful orange cake. We thank the referees for their careful reading and helpful comments.

2. Hyperbolic geometry

We collect some facts on hyperbolic geometry in the sense of Gromov [Gro87], see also [CDP90, GdlH90].

2.1. Hyperbolic spaces.

— Let X be a metric length space. The distance of two points x and y in X is denoted by $|x-y|$, or $|x-y|_X$ if we want to indicate that we measure the distance in X. If $A \subset X$ is a set and x a point, we write $d(x,A) = \inf_{a\in A} |x-a|$ to denote the distance from x to A. Let $A^{+\alpha} = \{x \in X \mid d(x,A) \leqslant \alpha\}$ be the *α-neighborhood* of A. Given $x, y \in X$, we write $[x,y]$ for a geodesic from x

to y (provided that such a path exists). Recall that there may be multiple geodesics joining two points. We recall that the Gromov product of y and z at x is defined by

$$(y,z)_x = \frac{1}{2}\left(|y-x|+|z-x|-|y-z|\right).$$

We will often use the following facts each of which is equivalent to the triangle inequality: for every $x,y,z,t \in X$,

$$(x,y)_t \leqslant (x,z)_t + |y-z| \quad \text{and} \quad (x,y)_t \leqslant (x,y)_z + |t-z|.$$

A similar useful inequality is

$$(x,y)_t \leqslant (x,y)_z + (x,z)_t, \quad \forall x,y,z,t \in X. \tag{2}$$

Indeed, after unwrapping the definition of Gromov's products, it boils down to the triangle inequality.

Definition 2.1. — Let $\delta \geqslant 0$. The space X is δ-hyperbolic if for every x, y, z and $t \in X$, the *four point inequality* holds, that is

$$(x,z)_t \geqslant \min\left\{(x,y)_t,(y,z)_t\right\} - \delta. \tag{3}$$

If $\delta = 0$ and X is geodesic, then X is an $\mathbb{R}$-tree. From now on, we assume that $\delta > 0$ and that X is a δ-hyperbolic metric length space. We denote by ∂X the boundary at infinity of X. Hyperbolicity has the following consequences.

Lemma 2.2 ([Cou14, Lem. 2.2]). — *Let x, y, z, s and t be five points of X.*

(1) $(x,y)_t \leqslant \max\left\{|x-t|-(y,z)_x,(x,z)_t\right\}+\delta$,

(2) $|s-t| \leqslant \left||x-s|-|x-t|\right| + 2\max\left\{(x,y)_s,(x,y)_t\right\} + 2\delta$,

(3) *The distance $|s-t|$ is bounded above by*

$$\max\left\{\left||x-s|-|x-t|\right|+2\max\left\{(x,y)_s,(x,z)_t\right\}, |x-s|+|x-t|-2(y,z)_x\right\}+4\delta.$$

2.2. Quasi-geodesics.

— A rectifiable path $\gamma : [a,b] \to X$ is a (k,ℓ)-*quasi-geodesic* if for all $[a',b'] \subset [a,b]$

$$\operatorname{length}(\gamma[a',b']) \leqslant k|\gamma(a')-\gamma(b')| + \ell;$$

and γ is a L-*local* (k,ℓ)-*quasi-geodesic* if any subpath of γ whose length is at most L is a (k,ℓ)-quasi-geodesic. The next lemma is used to construct (bi-infinite) quasi-geodesics.

Lemma 2.3 (Discrete quasi-geodesics [AL06, Lem. 1]). — *Let $n \geqslant 3$. Let $x_1,\dots,x_n$ be n points of X. Assume that for every $i \in \{2,\dots,n-2\}$,*

$$(x_{i-1},x_{i+1})_{x_i} + (x_i,x_{i+2})_{x_{i+1}} < |x_i - x_{i+1}| - 3\delta.$$

Then the following holds

(1) $|x_1-x_n| \geqslant \displaystyle\sum_{i=1}^{n-1}|x_i-x_{i+1}| - 2\sum_{i=2}^{n-1}(x_{i-1},x_{i+1})_{x_i} - 2(n-3)\delta.$

(2) $(x_1,x_n)_{x_j} \leqslant (x_{j-1},x_{j+1})_{x_j} + 2\delta$, *for every $j \in \{2,\dots,n-1\}$.*

(3) *If, in addition, X is geodesic, then $[x_1, x_n]$ lies in the 5δ-neighborhood of the broken geodesic $\gamma = [x_1, x_2] \cup \dots \cup [x_{n-1}, x_n]$, while γ is contained in the r-neighborhood of $[x_1, x_n]$, where*

$$r = \sup_{2 \leqslant i \leqslant n-1} (x_{i-1}, x_{i+1})_{x_i} + 14\delta. \qquad \square$$

Remark 2.4. — Note that the result still holds if $n = 1$ or $n = 2$. Indeed the statement is mostly void, or follows from the definition of Gromov products. One just need to replace the error term $2(n-3)\delta$ in (1) by zero. Thus in the remainder of the article, we will invoke Lemma 2.3 regardless how points are involved.

We denote by L_0 the smallest positive number larger than 500 such that for every $\ell \in [0, 10^5\delta]$, the Hausdorff distance between any two $L_0\delta$-local $(1, \ell)$-quasi-geodesic with the same endpoints is at most $(2\ell + 5\delta)$. See [Cou14, Cor. 2.6 & 2.7].

2.3. Quasi-convex subsets.

— A subset $Y \subset X$ is *α-quasi-convex* if for all two points $x, y \in Y$, and for every point $z \in X$, we have $d(z, Y) \leqslant (x, y)_z + \alpha$. For instance, geodesics are 2δ-quasi-convex.

If $Y \subset X$, we denote by $|\cdot|_Y$ the length metric induced by the restriction of $|\cdot|_X$ to Y. A subset Y that is connected by rectifiable paths is *strongly-quasi-convex* if it is 2δ-quasi-convex and if for all $y, y' \in Y$,

$$|y - y'|_X \leqslant |y - y'|_Y \leqslant |y - y'|_X + 8\delta.$$

2.4. Isometries.

— Let G be a group that acts by isometries on X. Let $g \in G$. The *translation length* of g is

$$\|g\| = \inf_{x \in X} |gx - x|.$$

The *stable translation length* of g is

$$\|g\|^\infty = \lim_{n \to \infty} \frac{1}{n} |g^n x - x|.$$

Those two quantities are related by the following inequality: $\|g\|^\infty \leqslant \|g\| \leqslant \|g\|^\infty + 16\delta$. See [CDP90, Ch. 10, Prop. 6.4]. The isometry g is hyperbolic if, and only if, its stable translation length is positive, [CDP90, Ch. 10, Prop. 6.3].

Definition 2.5. — Let $d > 0$ and $U \subset G$. The set of *d-quasi-fixpoints* of U is defined by

$$\operatorname{Fix}(U, d) = \{x \in X \mid \text{for all } u \in U \ |ux - x| < d\}.$$

The *axis of $g \in G$* is the set $A_g = \operatorname{Fix}(g, \|g\| + 8\delta)$.

Lemma 2.6 ([Cou18b, Lem. 2.9]). — *Let $U \subset G$. If $d > 7\delta$, then the set of d-quasi-fixpoints of U is 10δ-quasi-convex. Moreover, assuming that $\operatorname{Fix}(U, d)$ is non-empty*

(1) *if $x \in X \setminus \operatorname{Fix}(U, d)$, then*

$$\sup_{u \in U} |ux - x| \geqslant 2d(x, \operatorname{Fix}(U, d)) + d - 14\delta.$$

(2) *given $x \in X$ and $L \geqslant 0$, if $\sup_{u \in U} |ux - x| \leqslant d + 2L$, then $x \in \operatorname{Fix}(U, d)^{+L+7\delta}$.*

$\square$

Corollary 2.7 ([DG08, Prop. 2.3.3]). — *Let g be an isometry of X. Then A_g is 10δ-quasi-convex and g-invariant. Moreover, for all $x \in X$,*

$$\|g\| + 2d(x, A_g) - 6\delta \leqslant |gx - x| \leqslant \|g\| + 2d(x, A_g) + 8\delta. \qquad \square$$

2.5. Acylindricity. — We recall the definition of an acylindrical action. The action of G on X is *acylindrical* if there exists two functions $N, \kappa \colon \mathbb{R}_+ \to \mathbb{R}_+$ such that for every $r \geqslant 0$, for all points x and y at distance $|x - y| \geqslant \kappa(r)$, there are at most $N(r)$ elements $g \in G$ such that $|x - gx| \leqslant r$ and $|y - gy| \leqslant r$.

Recall that we assumed X to be δ-hyperbolic, with $\delta > 0$. In this context, acylindricity satisfies a local-to-global phenomenon: if there exists $N_0, \kappa_0 \in \mathbb{R}_+$ such that for all points x and y at distance $|x - y| \geqslant \kappa_0$, there are at most N_0 elements $g \in G$ such that $|x - gx| \leqslant 100\delta$ and $|y - gy| \leqslant 100\delta$, then the action of G is acylindrical, with the following estimates for the functions N and κ:

$$\kappa(r) = \kappa_0 + 4r + 100\delta \quad \text{and} \quad N(r) = \Big(\frac{r}{5\delta} + 3\Big) N_0. \tag{4}$$

See [DGO17, Prop. 5.31]. This motivates the next definition.

Definition 2.8. — Let $N, \kappa \in \mathbb{R}_+$. The action of G on X is (N, κ)-*acylindrical* if for all points x and y at distance $|x - y| \geqslant \kappa$, there are at most N elements $g \in G$ such that $|x - gx| \leqslant 100\delta$ and $|y - gy| \leqslant 100\delta$.

We need the following geometric invariants of the action of G on X. The limit set of G acting on X consists of the accumulation points in the Gromov boundary ∂X of X of the orbit of one (and hence any) point in X. By definition, a subgroup E of G is elementary if the limit set of E consists of at most two points.

Definition 2.9. — The *injectivity radius* is defined as

$$\tau(G, X) = \inf\{\|g\|^\infty \mid g \in G \text{ is a hyperbolic isometry}\}.$$

Definition 2.10. — The *acylindricity parameter* is defined as

$$A(G, X) = \sup_U \operatorname{diam}(\operatorname{Fix}(U, 2L_0\delta)),$$

where U runs over the subsets of G that do not generate an elementary subgroup.

Definition 2.11. — The ν-*invariant* is the smallest natural number $\nu = \nu(G, X)$ such that for every $g \in G$ and every hyperbolic $h \in G$ the following holds: if g, hgh^{-1}, $\dots$, $h^\nu g h^{-\nu}$ generate an elementary subgroup, then so do g and h.

Remark 2.12. — In the above definitions we adopt the following conventions. The diameter of the empty set is zero. If G does not contain any hyperbolic isometry, then $\tau(G, X) = \infty$. If every subgroup of G is elementary, then $A(G, X) = 0$.

The parameters $A(G, X)$ and $\nu(G, X)$ allow us to state the following version of Margulis' lemma.

Proposition 2.13 ([Cou18b, Prop. 3.5]). — *Let U be a subset of G. If U does not generate an elementary subgroup, then, for every $d > 0$, we have*

$$\operatorname{diam}\left(\operatorname{Fix}(U, d)\right) \leqslant \left[\nu(G,X) + 3\right] d + A(G,X) + 209\delta. \qquad \square$$

If there is no ambiguity we simply write $\tau(G)$, $A(G)$, and $\nu(G)$ for $\tau(G,X)$, $A(G,X)$, and $\nu(G,X)$ respectively. Sometimes, if the context is clear, we even write τ, A, or ν.

If the action of G on X is (N,κ)-acylindrical, then $\tau \geqslant \delta/N$, while A and ν are finite. In fact, one could express upper bounds on A and ν in terms of N, κ, δ, and L_0. See for instance [Cou16, §6]. However, for our purpose we need a finer control on these invariants.

From now on we assume that $\kappa \geqslant \delta$ and that the action of G on X is (N,κ)-acylindrical.

2.6. Loxodromic subgroups.

— An elementary subgroup is loxodromic if it contains a hyperbolic element. Equivalently, an elementary subgroup is loxodromic if it has exactly two points in its limit set. If h is a hyperbolic isometry, we denote by $E(h)$ the maximal loxodromic subgroup containing h. Let $E^+(h)$ be the maximal subgroup of $E(h)$ fixing pointwise the limit set of $E(h)$. It is known that the set F of all elliptic elements of $E^+(h)$ forms a (finite) normal subgroup of $E^+(h)$ and the quotient $E^+(h)/F$ is isomorphic to $\mathbb{Z}$. We say that h is *primitive* if its image in $E^+(h)/F$ generates the quotient.

Definition 2.14 (Invariant cylinder). — Let E be a loxodromic subgroup with limit set $\{\xi,\eta\}$. The *E-invariant cylinder*, denoted by C_E, is the 20δ-neighborhood of all $L_0\delta$-local $(1,\delta)$-quasi-geodesics with endpoints ξ and η at infinity.

Lemma 2.15 (Invariant cylinder). — *Let E be a loxodromic subgroup. Then*

- *C_E is 2δ-quasi-convex and invariant under the action of E. If, in addition, X is proper and geodesic, then C_E is strongly quasi-convex* [Cou14, Lem. 2.31]*,*
- *if $g \in E$ and $\|g\| > L_0\delta$, then $A_g \subset C_E$,* [Cou14, Lem. 2.33],
- *if $g \in E$ is hyperbolic, then $C_E \subset A_g^{+52\delta}$. In particular, if $x \in C_E$, then $|gx - x| \leqslant \|g\| + 112\delta$,* [Cou14, Lem. 2.32].

3. Periodic and aperiodic words

Let U be a finite subset of G *containing at least two elements*. We denote by U^* the free monoid generated by U. We write $\pi\colon U^* \to G$ for the canonical projection. In case there is no ambiguity, we make an abuse of notations and still write w for an element in U^* and its image under π. We fix a base point $p \in X$. Recall that the action of G on X is (N,κ)-acylindrical.

Definition 3.1. — Let $\alpha > 0$. We say that the subset U is *α-reduced (at p)* if

- $(u_1^{-1}p, u_2p)_p \leqslant \alpha$ for every $u_1, u_2 \in U$,
- $|up - p| > 2\alpha + 300\delta$ for every $u \in U$.

The set U is *α-strongly reduced (at p)* if, in addition, for every distinct $u_1, u_2 \in U$, we have

$$(u_1p, u_2p)_p < \min\{|u_1p - p|, |u_2p - p|\} - \alpha - 150\delta.$$

We say that U is *reduced at p* (respectively *strongly reduced at p*) if there exists $\alpha > 0$ such that U is α-reduced at p (respectively α-strongly reduced at p).

In practice, the base point p is fixed once and for all. Thus we simply say that U is (α-)reduced or (α-)strongly reduced.

Lemma 3.2. — *If U is α-strongly reduced, then U freely generates a free sub-semi-group of G. Moreover U satisfies the* geodesic extension property*, that is if $w, w' \in U^*$ are such that $(p, w'p)_{wp} \leqslant \alpha + 145\delta$, then w is a prefix of w'.*

Remark 3.3. — Roughly speaking, the geodesic extension property has the following meaning: if the geodesic $[p, w'p]$ extends $[p, wp]$ as a path in X, then w' extends w as a word over U.

Proof. — We first prove the geodesic extension property. Let $w = u_1 \cdots u_m$ and $w' = u'_1 \cdots u'_{m'}$ be two words in U^* such that $(p, w'p)_{wp} \leqslant \alpha + 145\delta$. We denote by r the largest integer such that $u_i = u'_i$ for every $i \in \{1, \dots, r-1\}$. For simplicity we let

$$q = u_1 \cdots u_{r-1}p = u'_1 \cdots u'_{r-1}p.$$

Assume now that contrary to our claim w is not a prefix of w', that is $r - 1 < m$. We claim that $(wp, w'p)_q < |u_rp - p| - \alpha - 148\delta$. If $r - 1 = m'$, then $w'p = q$ and the claim holds. Hence we can suppose that $r - 1 < m'$. It follows from our choice of r that $u_r \neq u'_r$. We let

$$t = u_1 \cdots u_rp \quad \text{and} \quad t' = u'_1 \cdots u'_rp.$$

Since U is α-strongly reduced, we have

$$(t, t')_q = (u_rp, u'_rp)_p < \min\{|u_rp - p|, |u'_rp - p|\} - \alpha - 150\delta.$$

It follows then from the four point inequality that

$$\begin{aligned}\min\{(t, wp)_q, (wp, w'p)_q, (w'p, t')_q\} &\leqslant (t, t')_q + 2\delta \\ &< \min\{|u_rp - p|, |u'_rp - p|\} - \alpha - 148\delta.\end{aligned} \tag{5}$$

Applying Lemma 2.3(2) with the sequence of points

$$q = u_1 \cdots u_{r-1}p, \quad t = u_1 \cdots u_rp, \quad u_1 \cdots u_{r+1}p, \quad \dots \quad , wp = u_1 \cdots u_mp,$$

we get

$$(q, wp)_t \leqslant (q, u_1 \cdots u_{r+1}p)_t + 2\delta = (u_r^{-1}p, u_{r+1}p)_p + 2\delta \leqslant \alpha + 2\delta.$$

(note that the last inequality follows from the fact that U is α-reduced). Hence

$$(t, wp)_q = |q - t| - (q, wp)_t \geqslant |u_rp - p| - \alpha - 2\delta.$$

Thus the minimum in (5) cannot be achieved by $(t, wp)_q$. Similarly, it cannot be achieved by $(w'p, t')_q$ either. Thus

$$(wp, w'p)_q < \min\{|u_r p - p|, |u'_r p - p|\} - \alpha - 148\delta \leqslant |u_r p - p| - \alpha - 148\delta,$$

which completes the proof of our claim.

Using Lemma 2.3(1) with the sequence of points

$$q = u_1 \cdots u_{r-1} p, \quad t = u_1 \cdots u_r p, \quad u_1 \cdots u_{r+1} p, \quad \dots \quad , wp = u_1 \cdots u_m p,$$

we get

$$|wp - p| \geqslant \sum_{j=r}^{m} |u_j p - p| - 2 \sum_{j=r}^{m-1} (u_j^{-1} p, u_{j+1} p)_p - 2\max\{m - r - 1, 0\}\delta.$$

Since U is α-reduced, we have

$$\sum_{j=r}^{m} |u_j p - p| > |u_r p - p| + (m - r)(2\alpha + 300\delta),$$

while

$$2 \sum_{j=r}^{m-1} (u_j^{-1} p, u_{j+1} p)_p \leqslant 2(m - r)\alpha.$$

Consequently, $|wp - q| \geqslant |u_r p - p|$. Combined with the previous claim, it yields

$$(q, w'p)_{wp} = |wp - q| - (wp, w'p)_q \geqslant |u_r p - p| - (wp, w'p)_q > \alpha + 148\delta.$$

Applying again the four point inequality, we get

$$\min\{(p, q)_{wp}, (q, w'p)_{wp}\} \leqslant (p, w'p)_{wp} + \delta \leqslant \alpha + 146\delta. \tag{6}$$

It follows from our previous computation that the minimum cannot be achieved by $(q, w'p)_{wp}$. We proved previously that $|wp - q| \geqslant |u_r p - p|$. Reasoning as in our first claim, Lemma 2.3(2) yields $(p, wp)_q \leqslant \alpha + 2\delta$. Since U is α-reduced we get

$$(p, q)_{wp} = |wp - q| - (p, wp)_q \geqslant |u_r p - p| - \alpha - 2\delta > \alpha + 298\delta.$$

Hence the minimum in (6) cannot be achieved by $(p, q)_{wp}$ either, which is a contradiction. Consequently, w is a prefix of w'.

Let us prove now that U freely generates a free sub-semi-group of G. Let $w_1, w_2 \in U^*$ whose images in G coincide. In particular $(p, w_1 p)_{w_2 p} = 0 = (p, w_2 p)_{w_1 p}$. It follows from the geodesic extension property that w_1 is a prefix of w_2 and conversely. Thus $w_1 = w_2$ as words in U^*. ☐

3.1. Periodic words.

— From now on, we assume that U is α-strongly reduced (in the sense of Definition 3.1). We let $\lambda = \max_{u \in U} |up - p|$. We denote by $|w|_U$ the word metric of $w \in U^*$. Given an element $w = u_1 \cdots u_m$ in U^*, we let

$$[w] = \{p, u_1 p, u_1 u_2 p, \dots, wp\}.$$

Definition 3.4. — Let $m \geqslant 0$. Let E be a maximal loxodromic subgroup. We say that a word $v \in U^*$ is *m-periodic with period E* if $[v] \subset C_E^{+\alpha + 100\delta}$ and $|p - vp| > m\tau(E)$.

Remark 3.5. — Note that the definition does not require m to be an integer. Let E be a maximal loxodromic subgroup such that p belongs to the $(\alpha + 100\delta)$-neighborhood of C_E. Let $v \in U^*$ whose image in G is a hyperbolic element of E. Then for every integer $m \geqslant 0$, the element v^{m+1} is m-periodic with period E. The converse is not true; that is, an m-periodic word with period E is not necessarily contained in E.

If m is sufficiently large, then periods are unique in the following sense.

Proposition 3.6. — *There exists $m_0 \geqslant 0$ which only depends on δ, A, ν, τ and α such that for every $m \geqslant m_0$ the following holds. If $v \in U^*$ is m-periodic with periods E_1 and E_2, then $E_1 = E_2$.*

Proof. — Let $h_1 \in E_1$ realize $\tau(E_1)$, and $h_2 \in E_2$ realize $\tau(E_2)$. If v is m-periodic with period E_1 and E_2, then

$$\operatorname{diam}\left(C_{E_1}^{+\alpha+100\delta} \cap C_{E_2}^{+\alpha+100\delta}\right) > m \max\{\|h_1\|^\infty, \|h_2\|^\infty\}.$$

Recall that $C_{E_i} \subset A_{h_i}^{+52\delta}$, see Lemma 2.15. By [Cou14, Lem. 2.13] we have

$$\operatorname{diam}\left(C_{E_1}^{+\alpha+100\delta} \cap C_{E_2}^{+\alpha+100\delta}\right) \leqslant \operatorname{diam}\left(A_{h_1}^{+13\delta} \cap A_{h_2}^{+13\delta}\right) + 2\alpha + 308\delta.$$

Hence there exists $m_0 \geqslant 0$ which only depends on δ, A, ν, τ and α such that if $m \geqslant m_0$, we have

$$\operatorname{diam}\left(A_{h_1}^{+13\delta} \cap A_{h_2}^{+13\delta}\right) > (\nu + 2)\max\{\|h_1\|, \|h_2\|\} + A + 680\delta.$$

It follows from [Cou16, Prop. 3.44] that h_1 and h_2 generates an elementary subgroup, hence $E_1 = E_2$. ☐

Remark 3.7. — For all $w \in U^*$, we have $\lambda|w|_U \geqslant |wp - p|$. In particular, if w is an m-periodic word with period E, then

$$|w|_U > m\tau(E)/\lambda.$$

Consider now a general non-empty word $w = u_1 \cdots u_r$ in U^*. We claim that $|wp - p| > 2\alpha + 298\delta|w|_U$. Indeed applying Lemma 2.3(1) with the sequence of points

$$p, \quad u_1p, \quad u_1u_2p, \quad \dots \quad , wp = u_1 \cdots u_r p,$$

we get

$$|wp - p| \geqslant \sum_{j=1}^{r} |u_jp - p| - 2\sum_{j=1}^{r-1} (u_j^{-1}p, u_{j+1}p)_p - 2\max\{r-2, 0\}\delta.$$

Since U is α-reduced, we have

$$\sum_{j=1}^{r} |u_jp - p| > r(2\alpha + 300\delta),$$

while

$$2\sum_{j=1}^{r-1} (u_j^{-1}p, u_{j+1}p)_p \leqslant 2(r-1)\alpha.$$

Combining the previous inequalities we get the announced estimate. Consequently, if $[w] \subset C_E^{+\alpha+100\delta}$ but w is not m-periodic with period E, then

$$|w|_U < m\tau(E)/\delta.$$

Proposition 3.8. — *Let E be a maximal loxodromic subgroup. Let $m \geqslant 0$. There are at most two elements in U^* which are m-periodic with period E, but whose proper prefixes are not m-periodic.*

Proof. — Let E be a maximal loxodromic subgroup. Let P_E be the set of m-periodic words $w \in U^*$ with period E. Assume that P_E is non-empty, otherwise the statement is void. Let η^- and η^+ be the points of ∂X fixed by E and $\gamma\colon \mathbb{R} \to X$ be an $L_0\delta$-local $(1,\delta)$-quasi-geodesic from η^- to η^+. For any $w \in \mathsf{P}_E$, the points p and wp lie in the $(\alpha + 100\delta)$-neighborhood of C_E, hence in the $(\alpha+120\delta)$-neighborhood of γ. Without loss of generality, we can assume that $q = \gamma(0)$ is a projection of p on γ. We decompose P_E in two parts as follows: an element $w \in \mathsf{P}_E$ belongs to P_E^+ (respectively P_E^-) if there is a projection $\gamma(t)$ of wp on γ with $t \geqslant 0$ (respectively $t \leqslant 0$). Observe that a priori P_E^- and P_E^+ are not disjoint, but that will not be an issue for the rest of the proof.

We are going to prove that $\mathsf{P}_E^+ \cap U^*$ contains at most one word satisfying the proposition. Let w_1 and w_2 be two words in $\mathsf{P}_E^+ \cap U^*$ which are m-periodic with period E, and whose proper prefixes are not m-periodic. We write $q_1 = \gamma(t_1)$ and $q_2 = \gamma(t_2)$ for the respective projections of w_1p and w_2p on γ. Without loss of generality we can assume that $t_1 \leqslant t_2$. We are going to prove that $(p, w_2p)_{w_1p} \leqslant \alpha + 145\delta$. As a quasi-geodesic, γ is 9δ-quasi-convex [Cou14, Cor. 2.7(2)]. According to Remark 3.7, the word w_2 is not empty and $|w_2p - p| > 2\alpha + 298\delta$. Applying the triangle inequality we get $|q_2 - q| > 19\delta$. Recall that q and q_2 are respective projections of p and w_2p on the quasi-convex γ. Hence

$$|w_2p - p| \geqslant |w_2p - q_2| + |q_2 - q| + |q - p| - 38\delta,$$

see [Cou14, Cor. 2.12(2)]. Since q_1 lies on γ between q and q_2 we also have

$$|q_2 - q| = |q_2 - q_1| + |q_1 - q| - 2(q_2, q)_{q_1} \geqslant |q_2 - q_1| + |q_1 - q| - 12\delta,$$

see [Cou14, Cor. 2.7(1)]. Combining the previous two inequalities, we get

$$\begin{aligned} |w_2p - p| &\geqslant |w_2p - q_2| + |q_2 - q_1| + |q_1 - q| + |q - p| - 50\delta \\ &\geqslant |w_2p - q_1| + |q_1 - p| - 50\delta. \end{aligned}$$

Thus $(w_2p, p)_{q_1} \leqslant 25\delta$. According to the triangle inequality, we get

$$(p, w_2p)_{w_1p} \leqslant |w_1p - q_1| + (w_2p, p)_{q_1} \leqslant \alpha + 145\delta,$$

which completes the proof of our claim.

Applying the geodesic extension property (see Lemma 3.2) we get that w_1 is a prefix of w_2. As w_1 is m-periodic, it cannot be a proper prefix, hence $w_1 = w_2$. Similarly, $\mathsf{P}_E^- \cap U^*$ has at most one element satisfying the statement. ☐

3.2. The growth of aperiodic words

Definition 3.9. — Let $w \in U^*$ and let E be a maximal loxodromic subgroup. We say that the word w *contains an m-period of E* if w splits as $w = w_0w_1w_2$, where the word w_1 is m-periodic with period E. If the word w does not contain any m-period, we say that w is *m-aperiodic.*

Observe that containing a period is a property of the word $w \in U^*$ and not of its image $\pi(w)$ in G: one could find two words w_1 and w_2, where w_1 is m-aperiodic while w_2 is not, and that have the same image in G. However since U is strongly reduced, it freely generates a free sub-semigroup of G. Hence this pathology does not arise in our context.

We denote by U^*_m the set of m-aperiodic words in U^*. Recall that p is a base point of X and the parameter λ is defined by

$$\lambda = \max_{u \in U} |up - p|.$$

Example 3.10. — If $m \geqslant \lambda/\tau$, then $U \subseteq U^*_m$. Indeed, for all $u \in U$ and loxodromic subgroups E,

$$|u|_U \leqslant 1 \leqslant m\tau/\lambda \leqslant m\tau(E)/\lambda.$$

So, by Remark 3.7, u cannot be m-periodic.

We denote by $S(r)$ the *sphere* of radius r in U^*. Similarly $B(r) \subset U^*$ stands for the *ball* of radius r, that is the subset of elements $w \in U^*$ of word length $|w|_U \leqslant r$. We note that $|B(r)| \leqslant |U|^{r+1}$, since $|U| \geqslant 2$.

Proposition 3.11. — *Let U be a α-strongly reduced subset of G, with at least two elements. There exists m_1 which only depends on λ, α, A, ν, τ, and δ with the following property. For all $m \geqslant m_1$, and $r > 0$, we have*

$$|U^*_m \cap B(r+1)| \geqslant \frac{|U|}{2} |U^*_m \cap B(r)|.$$

Proof. — We adapt the counting arguments of [Cou13]. We firstly fix some notations. Let m_0 be the parameter given by Proposition 3.6. Recall that m_0 only depends on α, A, ν, τ, and δ. Let $U \subset G$ be an α-strongly reduced subset, with at least two elements. Let $m > m_0 + 5\lambda/\tau$. We let

$$Z = \{w \in U^* \mid w = w_0 u,\ w_0 \in U^*_m,\ u \in U\}.$$

We denote by $\mathcal{E}$ the set of all maximal loxodromic subgroups in G. For each $E \in \mathcal{E}$, let $Z_E \subset Z$ be the subset of all $w \in Z$ that split as a product $w = w_1 w_2$, where $w_1 \in U^*_m$ and $w_2 \in U^*$ is an m-periodic word with period E.

Lemma 3.12. — *The set $Z \setminus \bigcup_{E \in \mathcal{E}} Z_E$ is contained in U^*_m.*

Proof. — Let $w \in Z$ contain an m-period of a loxodromic subgroup $E \in \mathcal{E}$. By definition of Z, we have $w = w_0 u$, where $u \in U$ and the prefix $w_0 \in U^*$ does not contain any m-period. On the other hand w contains a subword w_2 which is an m-period with period E. Since w_2 cannot be a subword of w_0, it is a suffix of w. $\square$

Recall that if $W \subset U^*$, then $|W|$ stands for the cardinality of *the image* of W in G. However, since U freely generates a free sub-semi-group (Lemma 3.2), we can safely identify the elements of U^* with their images in G. It follows from Lemma 3.12, that for all natural numbers r,

$$|U_m^* \cap B(r)| \geqslant |Z \cap B(r)| - \sum_{E \in \mathsf{E}} |Z_E \cap B(r)|. \tag{7}$$

The next step is to estimate each term in the above inequality.

Lemma 3.13. — *For all real numbers r,*

$$|Z \cap B(r+1)| \geqslant |U||U_m^* \cap B(r)|.$$

Proof. — It is a direct consequence of the fact that U freely generates a free sub-semi-group. ☐

Lemma 3.14. — *Let $E \in \mathsf{E}$. For all real numbers r,*

$$|Z_E \cap B(r)| \leqslant 2|\,U_m^* \cap B\left(r - m\tau(E)/\lambda\right)|.$$

Proof. — Let $w \in Z_E \cap B(r)$. By definition, w splits as a product $w = w_1 w_2$, where $w_1 \in U_m^*$ and $w_2 \in U^*$ is m-periodic with period E. By Remark 3.7, $|w_2|_U > m\tau(E)/\lambda$, so that $w_1 \in U_m^* \cap B(r - m\tau(E)/\lambda)$.

Since w also belongs to Z, the prefix consisting of all but the last letter does not contain m-periods. Thus every proper prefix of w_2 cannot be m-periodic. It follows from Lemma 3.8 that there are at most two possible choices for w_2. Hence the result. ☐

Lemma 3.15. — *For all real numbers r, the following inequality holds:*

$$\sum_{E \in \mathsf{E}} |Z_E \cap B(r)| \leqslant 2|U|^{m_0\tau/\delta + 2} \sum_{j \geqslant 1} |\,U_m^* \cap B\left(r - jm\tau/\lambda\right)|\;|U|^{jm_0\tau/\delta}.$$

Remark 3.16. — Note that the terms in the series on the right hand side are all non-negative. Hence if the series diverges, the statement is void. Later we will apply this lemma in a setting where the series actually converges.

Proof. — Given $j \geqslant 1$, we define E_j as the set of all maximal loxodromic subgroups $E \in \mathsf{E}$, such that $j\tau \leqslant \tau(E) < (j+1)\tau$ and U^* contains a word that is m-periodic with period E. We split the left-hand sum as follows

$$\sum_{E \in \mathsf{E}} |Z_E \cap B(r)| = \sum_{j \geqslant 1} \sum_{E \in \mathsf{E}_j} |Z_E \cap B(r)|.$$

Indeed if U^* does not contain a word that is m-periodic with period E, then the set Z_E is empty. Observe that for every $E \in \mathsf{E}_j$ we have by Lemma 3.14

$$|Z_E \cap B(r)| \leqslant 2|\,U_m^* \cap B\left(r - jm\tau/\lambda\right)|.$$

Thus it suffices to bound the cardinality of E_j for every $j \geqslant 1$.

Let $j \geqslant 1$. For simplicity we let $d_j = (j+1)m_0\tau/\delta + 1$. We claim that $|\mathsf{E}_j| \leqslant |U|^{d_j+1}$. To that end we are going to build a one-to-one map from $\chi\colon \mathsf{E}_j \to B(d_j)$. Indeed the cardinality of the ball $B(d_j)$ is at most $|U|^{d_j+1}$. Let $E \in \mathsf{E}_j$. By definition there exists $w \in U^*$ which is m-periodic with period E. Let w' be the shortest prefix of w that is m_0-periodic with period E. Note that such prefix always exists since $m \geqslant m_0$. By Remark 3.7, w' belongs to $B(m_0\tau(E)/\delta + 1)$ hence to $B(d_j)$. We define $\chi(E)$ to be w'. Observe that there is at most one E such that w' is m_0-periodic with period E (Proposition 3.6). Hence χ is one-to-one. This completes the proof of our claim and the lemma. □

We now complete the proof of Proposition 3.11. Let us define first some auxiliary parameters. We fix once for all an arbitrary number $\varepsilon \in (0, 1/2)$. In addition we let

$$\mu = (1-\varepsilon)|U|,\ \gamma = |U|^{m_0\tau/\delta},\ \xi = 2|U|^{m_0\tau/\delta+2},\ \sigma = \frac{\varepsilon}{2(1-\varepsilon)\xi},\ \text{and } M = \left\lfloor \frac{m\tau}{\lambda} \right\rfloor.$$

Since $|U| \geqslant 2$, we observe that $\sigma \leqslant 1/2$. We claim that there exists $m_1 \geqslant m_0$ which only depends on λ, α, A, ν, τ, and δ such that

$$\frac{\gamma}{\mu^M} \leqslant \sigma,$$

provided that $m \geqslant m_1$. The computation shows that

$$\ln\Bigl(\frac{\gamma}{\sigma\mu^M}\Bigr) \leqslant \Bigl(\frac{2m_0\tau}{\delta} + 3 - \frac{m\tau}{\lambda}\Bigr)\ln|U| - \ln\Bigl(\frac{\varepsilon}{4(1-\varepsilon)}\Bigr) - \frac{m\tau}{\lambda}\ln(1-\varepsilon).$$

Recall that $|U| \geqslant 2$. Hence, if

$$m \geqslant \frac{2m_0\lambda}{\delta} + \frac{3\lambda}{\tau},$$

then the previous inequality yields

$$\ln\Bigl(\frac{\gamma}{\sigma\mu^M}\Bigr) \leqslant -\frac{m\tau}{\lambda}\left[\ln 2 + \ln(1-\varepsilon)\right] + \Bigl(\frac{2m_0\tau}{\delta} + 3\Bigr)\ln 2 - \ln\Bigl(\frac{\varepsilon}{4(1-\varepsilon)}\Bigr). \tag{8}$$

We can see from there, that there exists $m_1 \geqslant m_0$ which only depends on λ, m_0, τ, and δ, such that as soon as $m \geqslant m_1$ the right hand side of Inequality (8) is non-positive, which completes the proof of our claim. Up to increasing the value of m_1, we can assume that $M \geqslant 1$, provided $m \geqslant m_1$.

Let us now estimate the number of aperiodic words in U^*. From now on we assume that $m \geqslant m_1$. For every integer r, we let

$$c(r) = |U^*_m \cap B(r)|.$$

We claim that for every integer r, we have $c(r) \geqslant \mu c(r-1)$. The proof goes by induction on r. In view of Example 3.10, the inequality holds true for $r = 1$. Assume that our claim holds for every $s \leqslant r$. In particular for every integer $t \geqslant 0$, we get $c(r-t) \leqslant \mu^{-t} c(r)$. It follows from (7) that

$$c(r+1) \geqslant |Z \cap B(r+1)| - \sum_{E \in \mathsf{E}} |Z_E \cap B(r+1)|.$$

Applying Lemmas 3.13 and 3.15, we get

$$c(r+1) \geqslant |U| c(r) - \xi \sum_{j \geqslant 1} c(r+1-jM)\gamma^j.$$

Note that $jM - 1 \geqslant 0$, for every $j \geqslant 1$. Thus applying the induction hypothesis we get

$$c(r+1) \geqslant \left(1 - \frac{\xi\mu}{|U|}\sum_{j\geqslant 1}\Big(\frac{\gamma}{\mu^M}\Big)^j\right)|U|\, c(r).$$

We defined μ as $\mu = (1-\varepsilon)|U|$, hence it suffices to prove that

$$\frac{\xi\mu}{|U|}\sum_{j\geqslant 1}\Big(\frac{\gamma}{\mu^M}\Big)^j \leqslant \varepsilon.$$

Recall that $\gamma/\mu^M \leqslant \sigma \leqslant 1/2$. Hence the series converges. Moreover

$$\frac{\xi\mu}{|U|}\sum_{j\geqslant 1}\Big(\frac{\gamma}{\mu^M}\Big)^j \leqslant \frac{\xi\mu}{|U|}\,\frac{\sigma}{1-\sigma} \leqslant \frac{2\xi\mu\sigma}{|U|} \leqslant \varepsilon.$$

This completes the proof of our claim for $r+1$. ☐

4. Power-free elements

Let G be a group that acts (N,κ)-acylindrically on a δ-hyperbolic geodesic space X. We fix a basepoint $p \in X$. Recall our convention: the diameter of the empty set is zero, see Remark 2.12.

Definition 4.1. — Let $m \geqslant 0$. An element $g \in G$ *contains an m-power* if there is a maximal loxodromic subgroup E and a geodesic $[p, gp]$ such that

$$\operatorname{diam}\left([p,gp]^{+5\delta} \cap C_E^{+5\delta}\right) > m\tau(E).$$

If $g \in G$ does not contain any m-power, we say that g is *m-power-free.*

Let $U \subset G$ be a finite subset. We recall that $\lambda = \max_{u\in U}|up - p|$ and that U^* is the set of all words over the alphabet U. The idea of the next statement is the following. Take a word $w \in U^*$. If w, seen as an element of G, contains a sufficiently large power, then the *word* w already contains a large period.

Proposition 4.2. — *Let $m \geqslant (2\lambda + 20\delta)/\tau$. Let $U \subset G$ be a finite α-reduced subset. Let $w \in U^*$. If w contains an m-power (as an element of G), then w contains an m'-period (as a word over U), where $m' = m - (2\lambda+20\delta)/\tau$.*

Proof. — Let $w = u_1\cdots u_l$. As w contains a m-power, there is a loxodromic subgroup E and a geodesic $[p, wp]$ such that

$$\operatorname{diam}([p,wp]^{+5\delta} \cap C_E^{+5\delta}) > m\tau(E).$$

Let x_1, x_2 in $[p,wp]^{+5\delta} \cap C_E^{+5\delta}$ such that $|x_1 - x_2| > m\tau(E)$. Let $\gamma_w = [p, u_1p] \cup u_1[p, u_2p] \cup \cdots \cup (u_1\cdots u_{\ell-1})[p, u_\ell p]$ be a broken geodesic joining p to wp. Let p_1 and p_2 be the respective projections of x_1 and x_2 on γ_w. By Lemma 2.3, the geodesic $[p, wp]$ is contained in the 5δ-neighborhood of γ_w. Hence p_1 and p_2 are 15δ-close to C_E. Moreover,

$$|p_1 - p_2| \geqslant |x_1 - x_2| - 20\delta > m\tau(E) - 20\delta.$$

Up to permuting x_1 and x_2 we can assume that p, p_1, p_2 and wp are ordered in this way along γ_w. In particular, there is $i \leqslant \ell - 1$ such that $p_1 \in (u_1 \cdots u_i) \cdot [p, u_{i+1}p]$, and $j \leqslant \ell - 1$ such that $p_2 \in (u_1 \cdots u_j) \cdot [p, u_{j+1}p]$. Since p_1 comes before p_2 on γ_w, we have $i \leqslant j$. Note that actually $i < j$. Indeed if $i = j$, we would have

$$\lambda \geqslant |u_{i+1}p - p| \geqslant |p_1 - p_2| > m\tau(E) - 20\delta \geqslant m\tau - 20\delta,$$

which contradicts our assumption. Let us set $w_0 = u_1 \cdots u_{i+1}$ and take the word w_1 such that $u_1 \cdots u_j = w_0 w_1$. At this stage w_1 could be the empty word. But we will see that this is not the case. Indeed

$$|p_1 - p_2| \leqslant |p_1 - w_0 p| + |w_0 p - w_0 w_1 p| + |w_0 w_1 p - p_2| \leqslant |p - w_1 p| + 2\lambda.$$

Thus,

$$|p - w_1 p| > m\tau(E) - 2\lambda - 20\delta \geqslant m'\tau(E).$$

Applying Lemma 2.3 to the subpath γ' of γ_w bounded by p_1 and p_2, we get that γ' lies in the $(\alpha + 14\delta)$-neighborhood of the geodesic $[p_1, p_2]$. However p_1 and p_2 are in the 15δ-neighborhood of C_E which is 2δ-quasi-convex. Thus γ' is contained in the $(\alpha + 31\delta)$-neighborhood of C_E. We conclude that w_1 is m'-periodic with period $w_0^{-1} E w_0$. ☐

5. Energy and quasi-center

Let G be a group acting by isometries on a δ-hyperbolic length space X. Recall that we assume for simplicity that $\delta > 0$. In next sections, we denote by $S(x, r)$ the sphere in X of radius r centered at x. (This should not be confused with the spheres in U^* used in the previous section.) Let $U \subset G$ be a finite subset. In order to apply the counting results from Section 3, we explain in this section and the followings how to build a strongly reduced subset of U^2. To that end we define the notion of energy of U.

Definition 5.1. — The *ℓ^∞-energy $\lambda(U,x)$ of U at x* is defined by $\lambda(U,x) = \max_{u \in U} |ux - x|$. The *$\ell^\infty$-energy* of U is given by

$$\lambda(U) = \inf_{x \in X} \lambda(U,x).$$

A point $q \in X$ is *almost-minimizing the ℓ^∞-energy* if $\lambda(U,q) \leqslant \lambda(U) + \delta$.

Let $x \in X$ and $A, B \subseteq X$. Define $U_x(A,B)$ to be the set of elements $u \in U$ satisfying the following conditions

- $|x - ux| \geqslant 4 \cdot 10^3 \delta$,
- there exists $a \in A \cap S(x, 10^3\delta)$, such that $(x, ux)_a \leqslant \delta$,
- there exists $b \in B \cap S(x, 10^3\delta)$, such that $(u^{-1}x, x)_b \leqslant \delta$.

We write $U_x(A) = U_x(A,A)$, and, if there is no ambiguity, $U(A,B) = U_x(A,B)$ for short.

Definition 5.2 (Quasi-centre). — A point $x \in X$ is a *quasi-center for* U if, for all $y \in S(x, 10^3\delta)$, we have

$$\left|U_x\left(y^{+100\delta}\right)\right| \leqslant \frac{3}{4}|U|.$$

Proposition 5.3. — *Let* q *be a point that almost-minimizes the* ℓ^∞*-energy of* U*. There exists a quasi-center* p *for* U *such that* $|p-q| \leqslant \lambda(U)$.

Remark 5.4. — The existence of a quasi-center is already known by [DS20]. The authors prove there that any point almost-minimizing the ℓ^1-energy is a quasi-center. However such a point could be very far from any point almost-minimizing the ℓ^∞-energy.

Proof. — We describe a recursive procedure to find a quasi-center p. The idea is to construct a quasi-geodesic from q to a quasi-center p. Let $x_0 = q$ and suppose that $x_0, \dots, x_{i-1}, x_i \in X$ are already defined. If x_i is a quasi-center for U, we let $p = x_i$ and stop the induction. Otherwise, there is a point $x_{i+1} \in S(x_i, 10^3\delta)$ such that $|U_{x_i}(x_{i+1}^{+100\delta})| > \frac{3}{4}|U|$.

Our idea is to apply Lemma 2.3 to the sequence of points $x_0, x_1, \dots, x_i, x_{i+1}, ux_{i+1}, ux_i, \dots, ux_1, ux_0$ for some $u \in U_{x_i}(x_{i+1}^{+100\delta})$. Like this we can write the distance from x_0 to ux_0 as a function of the index i. We will observe that this function diverges to infinity, which forces the procedure to stop. To do this, we collect the following observations. By construction, we have:

Lemma 5.5. — *For all* $u \in U_{x_{i-1}}(x_i^{+100\delta})$*, the following holds*

1. $(x_{i-1}, ux_{i-1})_{x_i} \leqslant 101\delta$ *and* $(x_{i-1}, ux_{i-1})_{ux_i} \leqslant 101\delta$,
2. $(x_{i-1}, ux_i)_{x_i} \leqslant 102\delta$ *and* $(x_i, ux_{i-1})_{ux_i} \leqslant 102\delta$.

Remark 5.6. — Roughly speaking, this lemma tells us that x_{i-1}, x_i, ux_i and ux_{i-1} are aligned in the order of their listing along the neighborhood of the geodesic $[x_{i-1}, ux_{i-1}]$.

Proof

The first point is just a reformulation of the definition of the set $U_{x_{i-1}}(x_i^{+100\delta})$. Let $u \in U_{x_{i-1}}(x_i^{+100\delta})$. By Lemma 2.2 (1) we have

$$(x_{i-1}, ux_i)_{x_i} \leqslant \max\left\{|x_{i-1} - x_i| - (ux_i, ux_{i-1})_{x_{i-1}}, (x_{i-1}, ux_{i-1})_{x_i}\right\} + \delta. \tag{9}$$

According to the triangle inequality we have

$$(ux_i, ux_{i-1})_{x_{i-1}} \geqslant |ux_{i-1} - x_{i-1}| - |x_{i-1} - x_i|.$$

However, by construction $|ux_{i-1} - x_{i-1}| > 2|x_{i-1} - x_i| + 2\delta$. Hence the maximum in (9) has to be achieved by $(x_{i-1}, ux_{i-1})_{x_i}$. The same argument works for $(x_i, ux_{i-1})_{ux_i}$. □

Lemma 5.7. — *If* x_i *is not a quasi-center for* U*, then* $(x_{i-1}, x_{i+1})_{x_i} \leqslant 103\delta$.

Proof. — We note that $|U_{x_{i-1}}(x_i^{+100\delta})\cap U_{x_i}(x_{i+1}^{+100\delta})| > |U|/2$. Let us fix an element u in this intersection. By Lemma 5.5, $(x_{i-1}, ux_i)_{x_i} \leqslant 102\delta$ and $(x_i, ux_i)_{x_{i+1}} \leqslant 101\delta$. According to the four point inequality we have

$$102\delta \geqslant (x_{i-1}, ux_i)_{x_i} \geqslant \min\left\{(x_{i-1}, x_{i+1})_{x_i}, (x_{i+1}, ux_i)_{x_i}\right\} - \delta.$$

Observe that

$$|x_i - x_{i+1}| = (x_{i+1}, ux_i)_{x_i} + (x_i, ux_i)_{x_{i+1}} \leqslant (x_{i+1}, ux_i)_{x_i} + 101\delta.$$

Since $|x_i - x_{i+1}| = 10^3\delta$, the minimum cannot be achieved by $(x_{i+1}, ux_i)_{x_i}$, whence the result. ☐

Lemma 5.8. — *If x_i is not a quasi-center, then, for all $u \in U_{x_i}(x_{i+1}^{+100\delta})$,*

$$|ux_0 - x_0| \geqslant |x_{i+1} - ux_{i+1}| + 10^3(i+1)\delta.$$

Proof. — Let u be in $U_{x_i}(x_{i+1}^{+100\delta})$. By Lemma 5.5,

$$x_i, ux_{i+1})_{x_{i+1}} \leqslant 102\delta \text{ and } (x_{i+1}, ux_i)_{ux_{i+1}} \leqslant 102\delta.$$

On the other hand, by Lemma 5.7, we have

$$(x_{j-1}, x_{j+1})_{x_j} \leqslant 103\delta \text{ and } (ux_{j-1}, ux_{j+1})_{ux_j} \leqslant 103\delta, \text{ for all } 0 < j \leqslant i.$$

The claim follows from Lemma 2.3 applied to the sequence of points x_0, x_1, …, x_i, x_{i+1}, ux_{i+1}, ux_i, …, ux_1, ux_0. ☐

Suppose that x_i is not a quasi-center. Fix $u \in U_{x_i}(x_{i+1}^{+100\delta})$. By construction we have $|x_{i+1} - ux_{i+1}| \geqslant 4\cdot 10^3\delta$. Recall that $x_0 = q$ almost-minimizes the energy. By Lemma 5.8, we get

$$\lambda(U) \geqslant 10^3(i+5)\delta.$$

This means that the induction used to build the sequence (x_i) stops after finitely many steps. Moreover, when the process stops we have $x_i = p$ and $\lambda(U) \geqslant 10^3(i+5)\delta$. For every $j \leqslant i-1$ we have $|x_j - x_{j+1}| \leqslant 10^3\delta$, thus $|p-q| \leqslant \lambda(U)$. ☐

6. Sets of diffuse energy

In this section we assume that the action of G on X is (N,κ)-acylindrical, with $\kappa > 50\cdot 10^3\delta$. Let $U \subset G$ be a finite subset. Let p be a quasi-center of U. In this section we assume that U is of *diffuse energy* (at p) that is for at least 99/100 of the elements of $U \subset G$, we have $|up - p| > 2\kappa$.

6.1. Reduction lemma. — We first prove the following variant of the reduction lemmas in [DS20].

Proposition 6.1 (Reduction). — *There is $v \in U$, and $U_1 \subset U$ of cardinality $|U_1| \geqslant \frac{1}{100}|U|$ such that for all $u_1 \in U_1$,*

- $(u_1^{-1}p, vp)_p \leqslant 10^3\delta$ *and* $(v^{-1}p, u_1p)_p \leqslant 10^3\delta$, *and*
- $2\kappa \leqslant |u_1p - p| \leqslant |vp - p|$.

Remark 6.2. — In the case of trees, Proposition 6.1 follows directly from [DS20, Lem. 6.4], and the proof of this lemma is due to Button [But13]. The situation is different in the case of hyperbolic spaces. Indeed, in contrast to the reduction lemmas in [DS20, §6.1], the cardinality of U_1 in Proposition 6.1 does not depend the cardinality of balls in X, as in [DS20, Lem. 6.3], and the estimates on the Gromov products do not depend on the logarithm of the cardinality of U, as in [DS20, Lem. 6.8].

Proof. — For simplicity we let $\eta = 1/100$. Let $U' = \{u \in U \mid |up - p| \geqslant 2\kappa\}$. As the energy of U is diffuse at p, we have $|U'| \geqslant (1-\eta)|U|$. Let us fix $u_0 \in U'$ such that $|u_0 p - p|$ is maximal. We claim the following result.

Lemma 6.3. — *At least one of the following holds:*

(1) *there is $U_1 \subset U'$ of cardinality $|U_1| > \eta|U|$ such that for all $u_1 \in U_1$,*

$$(u_1^{-1}p, u_0 p)_p \leqslant 10^3\delta \text{ and } (u_0^{-1}p, u_1 p)_p \leqslant 10^3\delta;$$

(2) *there are U_1, $U_2 \subset U'$ of cardinalities $|U_1| > \eta|U|$ and $|U_2| > \eta|U|$ such that for all $u_1 \in U_1$, $u_2 \in U_2$,*

$$(u_1^{-1}p, u_2 p)_p \leqslant 10^3\delta \text{ and } (u_2^{-1}p, u_1 p)_p \leqslant 10^3\delta.$$

We postpone for the moment the proof of this lemma and complete first the demonstration of Proposition 6.1. In case (1) of Lemma 6.3, we set $v = u_0$. In case (2) of Lemma 6.3, we may assume, up to exchanging the roles of U_1 and U_2, that there is $v \in U_2$ such that for all $u_1 \in U_1$, $|u_1 p - p| \leqslant |vp - p|$. This yields Proposition 6.1. ☐

Proof of Lemma 6.3. — We write $S = S(p, 10^3\delta)$ for short. For simplicity, in this proof we write

$$U'(A,B) = U' \cap U_p(A,B).$$

See Section 5 for the definition or $U_p(A,B)$.

The definition of hyperbolicity implies the following useful lemma.

Lemma 6.4 ([DS20, Lem. 6.1]). — *Let $y_1, z_1, y_2, z_2 \in S$. If $|z_1 - y_2| > 6\delta$, then for every $u_1 \in U(y_1, z_1)$ and $u_2 \in U(y_2, z_2)$, we have $(u_1^{-1}p, u_2 p)_p \leqslant 10^3\delta$.* ☐

By construction $|u_0 p - p| \geqslant 4 \cdot 10^3\delta$. Thus there exists y_0 and $z_0 \in S$ such that $(u_0 p, p)_{y_0} \leqslant \delta$ and $(p, u_0^{-1}p)_{z_0} \leqslant \delta$, so that $u_0 \in U'(y_0, z_0)$. Assume first that $|U'(S \smallsetminus z_0^{+6\delta}, S \smallsetminus y_0^{+6\delta})| > \eta|U|$. Then we let $U_1 = U'(S \smallsetminus z_0^{+6\delta}, S \smallsetminus y_0^{+6\delta})$. Using Lemma 6.4 we conclude that (1) holds.

Observe that the complement in U' of the previous set is the union of $U'(z_0^{+6\delta}, S)$ and $U'(S, y_0^{+6\delta})$. Recall that $|U'| > (1-\eta)|U|$. Thus we can now assume that

$$\left|U'(z_0^{+6\delta}, S) \cup U'(S, y_0^{+6\delta})\right| > (1-2\eta)|U|. \tag{10}$$

Let us now assume that $|U'(z_0^{+6\delta}, S \smallsetminus z_0^{+12\delta}) > \eta|U|$. In this case we let $U_1 = U_2 = U'(z_0^{+6\delta}, S \smallsetminus z_0^{+12\delta})$. Using Lemma 6.4 we conclude that (2) holds. The same argument works if $|U'(S \smallsetminus y_0^{+12\delta}, y_0^{+6\delta})| > \eta|U|$. Suppose now that the cardinality of

$U'(z_0^{+6\delta}, S \setminus z_0^{+12\delta})$ and $U'(S \setminus y_0^{+12\delta}, y_0^{+6\delta})$ are both bounded above by $\eta|U|$. It follows from (10) that

$$(11)\quad |U_1 \cup U_2| > (1-4\eta)|U|, \quad \text{where } U_1 = U'(z_0^{+6\delta}, z_0^{+12\delta}) \text{ and } U_2 = U'(y_0^{+12\delta}, y_0^{+6\delta}).$$

Since p is a quasi-center, the cardinality of both U_1 and U_2 is bounded above by $3|U|/4$. It follows from (11) that each of them contains at least $(1/4-4\eta)|U|$ elements. Observe also that $|y_0 - z_0| > 30\delta$. Indeed otherwise both U_1 and U_2 are contained in $U(y_0^{+100\delta})$. Hence (11) contradicts the fact that p is a quasi-center. Applying Lemma 6.4 we conclude that U_1 and U_2 satisfy (2). ☐

6.2. Construction of free sub-semi-groups. — We recall that $\lambda(U)$ denotes the ℓ^∞-energy of the finite subset $U \subset G$. By Proposition 5.3, we can assume that the quasi-center p, which we fixed at the beginning of this section, is at distance at most $\lambda(U)$ from a point almost-minimizing the ℓ^∞-energy of U. We still assume that the energy of U is diffuse (at p). We treat p as the base point of X.

Remark 6.5. — According to the triangle inequality, we have $|up - p| \leqslant 3\lambda(U) + \delta$, for every $u \in U$. Since the energy of U is diffuse at p, there is an element $u \in U$ that moves p by a large distance. As a consequence $\lambda(U) \geqslant \delta$, and thus $|up - p| \leqslant 4\lambda(U)$, for every $u \in U$. This estimates are far from being optimal, but sharp enough for our purpose.

Proposition 6.6. — *There exists $v \in U$ and a subset $W \subset Uv$ such that W is 1002δ-strongly reduced and*

$$|W| \geqslant \frac{1}{10^6 N} \frac{\delta}{\lambda(U)} |U|.$$

Proof. — For simplicity we let $\alpha = 1002\delta$. We fix U_1 and v given by Proposition 6.1. We set $T = U_1 v$.

Lemma 6.7. — *For every $t, t' \in T$, we have $(t^{-1}p, t'p)_p \leqslant \alpha$ and $|tp - p| > 2\alpha + 300\delta$.*

Proof. — We write $t = uv$ and $t' = u'v$ with $u, u' \in U_1$. Applying twice the four point inequality (3) we have

$$(12)\qquad \min\left\{(v^{-1}p, t^{-1}p)_p, (t^{-1}p, t'p)_p, (t'p, u'p)_p\right\} \leqslant (v^{-1}p, u'p)_p + 2\delta \leqslant \alpha.$$

Observe that

$$(v^{-1}p, t^{-1}p)_p = |p - v^{-1}p| - (p, t^{-1}p)_{v^{-1}p} = |p - vp| - (vp, u^{-1}p)_p \geqslant 2\kappa - 1000\delta > \alpha.$$

Similarly we prove that $(t'p, u'p)_p > \alpha$. Hence the minimum in (12) is achieved by $(t^{-1}p, t'p)_p$ which proves the first point. By definition of Gromov products we have

$$|tp - p| = |up - p| + |vp - p| - 2(u^{-1}p, vp)_p \geqslant 4\kappa - 2000\delta. \qquad ☐$$

For every $w \in T$, we set

$$A_w = \{t \in T \mid |p - tp| \leqslant |p - wp| \text{ and } (wp, p)_{tp} \leqslant \alpha + 150\delta\}.$$

Note that $w \in A_w$.

In order to define W, we construct by induction an increasing sequence (W_i) of subsets of T. We first let $W_0 = \varnothing$. Assume that now that W_i has been defined for some integer $i \geqslant 0$. If the set

$$T \setminus \bigcup_{w \in W_i} A_w$$

is empty, then the process stops and we let $W = W_i$ (note that this will ineluctably happen as T is finite). Otherwise, we choose an element w_{i+1} in this set for which $|p - w_{i+1}p|$ is maximal and let $W_{i+1} = W_i \cup \{w_{i+1}\}$.

Lemma 6.8. — *The set W is α-strongly reduced.*

Proof. — By Lemma 6.7, the set T (hence W) is α-reduced. It suffices to prove that for every distinct $w, w' \in W$ we have

$$(wp, w'p)_p \leqslant \min\{|wp - p|, |w'p - p|\} - \alpha - 150\delta.$$

Using the notation above, we write, $w_1, w_2, \dots, w_n$ for the elements W in the order they have been constructed. Let $i, j \in \{1, \dots, n\}$ such that $|p - w_jp| \leqslant |p - w_ip|$. If $i < j$, then w_j does not belong to A_{w_i}, thus

$$\begin{aligned}(w_ip, w_jp)_p &= |w_jp - p| - (w_ip, p)_{w_jp} < |w_jp - p| - \alpha - 150\delta \\ &\leqslant \min\{|w_ip - p|, |w_jp - p|\} - \alpha - 150\delta.\end{aligned}$$

Assume now that $j < i$. Note that the sequence $\{|p - w_kp|\}$ is non-increasing, hence $|p - w_jp| = |p - w_ip|$. Since w_i does not belong to A_{w_j}, thus

$$\begin{aligned}(w_ip, w_jp)_p &= |w_ip - p| - (w_jp, p)_{w_ip} < |w_ip - p| - \alpha - 150\delta \\ &\leqslant \min\{|w_ip - p|, |w_jp - p|\} - \alpha - 150\delta.\end{aligned}$$ ☐

Lemma 6.9. — *For every $w \in T$, we have*

$$|A_w| \leqslant \frac{2065N}{\delta} \lambda(U).$$

Proof. — Let $w \in T$. The proof goes in two steps. First we give an upper bound for subsets of sparse elements in A_w. Let $m \geqslant 0$ be an integer. Let $t_0 = u_0v$, $t_1 = u_1v$, …, $t_m = u_mv$ be m pairwise distinct elements in A_w. We assume in addition that $|u_ip - u_jp| > 6 \cdot 10^3\delta$, for every distinct $i, j \in \{0, \dots, m\}$. Let $\gamma\colon [a,b] \to X$ be a $(1,\delta)$-quasi-geodesic from p to wp. We are going to give an upper bound for m. To that end we claim that the points $u_0p, \dots, u_mp$ lie close to γ. Since the points u_ip are sparse, this will roughly say that $m \lesssim |wp - p| / \max\{|u_ip - u_jp|\}$. More precisely, the argument goes as follows. For every $i \in \{0, \dots, m\}$, we write p_i for a projection of u_ip onto γ. Up to reindexing the elements we can suppose that the points $p, p_0, p_1, \dots, p_m, wp$ are aligned in this order along γ.

Since t_i belongs to A_w, we have

$$(wp, p)_{t_ip} \leqslant \alpha + 150\delta \leqslant 1152\delta.$$

On the other hand, we know by construction of U_1 and v that $(p, t_ip)_{u_ip} = (u_i^{-1}p, vp)_p$ is at most $10^3\delta$, see Proposition 6.1. Hence the triangle inequality yields, see (2),

$$(p, wp)_{u_ip} \leqslant (wp, p)_{t_ip} + (p, t_ip)_{u_ip} \leqslant 2152\delta.$$

Since γ is $(1, \delta)$-quasi-geodesic, it is 9δ-quasi-convex, see [Cou14, Cor. 2.7(2)]. It follows that $|u_ip - p_i| = d(u_ip, \gamma)$ is at most 2161δ. According to the triangle inequality we get

$$|p_i - p_j| > 1678\delta, \quad \forall i \neq j.$$

Observe now that

$$1678m\delta \leqslant \sum_{i=0}^{m-1} |p_i - p_{i+1}| \leqslant \operatorname{length}(\gamma) \leqslant |p - wp| + \delta.$$

Recall that w is a two letter word in U, while $\lambda(U)$ is very large compare to δ. Hence $1678m\delta \leqslant 9\lambda(U)$. To simply the rest of the computations, we will use the following generous estimate

$$m \leqslant \frac{\lambda(U)}{\delta}.$$

We now start the second step of the proof. Using acylindricity we reduce the counting of elements in A_w to the case of a sparse subset. Any element $t \in A_w$ can be written $t = u_tv$ with $u_t \in U_1$. Consider now $t, t' \in A_w$.

We claim that $|u_tvp - u_{t'}vp| \leqslant |u_tp - u_{t'}p| + 4306\delta$. Indeed, by definition of A_w, we have $(wp, p)_{u_tvp} \leqslant \alpha + 150\delta$ and $(wp, p)_{u_{t'}vp} \leqslant \alpha + 150\delta$. By Lemma 2.2(2) we have

$$\begin{aligned} |u_tvp - u_{t'}vp| &\leqslant \big||p - u_tvp| - |p - u_{t'}vp|\big| + 2\max\{(wp, p)_{u_tvp}, (wp, p)_{u_{t'}vp}\} + 2\delta \\ &\leqslant \big||p - u_tvp| - |p - u_{t'}vp|\big| + 2306\delta. \end{aligned} \tag{13}$$

Note that

$$|p - u_tp| + |vp - p| - 2 \cdot 10^3\delta \leqslant |p - u_tvp| \leqslant |p - u_tp| + |vp - p|.$$

Indeed the second inequality is just the triangle inequality, while the first one is equivalent to the following known fact $(u_t^{-1}p, vp)_p \leqslant 10^3\delta$. Similarly we have

$$|p - u_{t'}p| + |vp - p| - 2 \cdot 10^3\delta \leqslant |p - u_{t'}vp| \leqslant |p - u_{t'}p| + |vp - p|.$$

The difference of the previous two inequalities yields

$$\big||p - u_tvp| - |p - u_{t'}vp|\big| \leqslant \big||p - u_tp| - |p - u_{t'}p|\big| + 2 \cdot 10^3\delta.$$

Plugging this inequality in (13) we obtain

$$|u_tvp - u_{t'}vp| \leqslant \big||p - u_tp| - |p - u_{t'}p|\big| + 4306\delta.$$

Finally, by the triangle inequality $\big||p - u_tp| - |p - u_{t'}p|\big| \leqslant |u_tp - u_{t'}p|$. This implies the claim.

We can now take advantage of acylindricity. Recall that $|vp - p| \geqslant 2\kappa$, with $\kappa > 50 \cdot 10^3\delta$. In particular,

$$|vp - p| \geqslant \kappa + 41324\delta.$$

We let $M = 2065N$. According to acylindricity – see (4) applied with $r = 10306\delta$ – the set

$$F = \{g \in G \mid |gp - p| \leqslant 6000\delta \text{ and } |gvp - vp| \leqslant 10306\delta\}$$

contains at most M elements. It follows that for every $t \in A_w$, there are at most M elements $t' \in A_w$ such that $|u_t p - u_{t'} p| \leqslant 6 \cdot 10^3\delta$. Indeed, if $|u_t p - u_{t'} p| \leqslant 6 \cdot 10^3\delta$, our previous claim implies that $u_t^{-1} u_{t'}$ belongs to F.

So we can extract a subset $B \subset A_w$ containing $m \geqslant |A_w|/M$ elements such that for every distinct $t, t' \in B$ we have $|u_t p - u_{t'} p| > 6 \cdot 10^3\delta$. It follows from the previous discussion that $m \leqslant \lambda(U)/\delta$. Consequently,

$$|A_w| \leqslant \frac{2065N}{\delta}\,\lambda(U). \qquad \square$$

Lemma 6.10. — *The cardinality of W is bounded from below as follows:*

$$|W| \geqslant \frac{1}{2065N}\,\frac{\delta}{\lambda(U)}\,|T|.$$

Proof. — Recall that $w \in A_w$ for every $w \in T$. Thus, by construction, the collection of sets $\{A_w\}_{w \in W}$ covers T. We have seen in Lemma 6.9 that the cardinality of each of them is at most $2065N\lambda(U)/\delta$. Hence the result. ☐

The previous lemma completes the proof of Proposition 6.6. ☐

7. Sets of concentrated energy

We still assume here that the action of G on X is (N, κ)-acylindrical, with $\kappa > 50 \cdot 10^3\delta$. Let $U \subset G$ be a finite subset and $p \in X$ a base point. In this section we also assume that U has *concentrated energy* (at p) that is, there exists $U_1 \subset U$ with $|U_1| \geqslant |U|/100$ such that $|up - p| \leqslant 2\kappa$, for all $u \in U_1$. The goal of the section is to prove the following statement.

Proposition 7.1. — *Let $M = 2\kappa N/\delta$. If $\lambda(U, p) > 100\kappa$, then one of the following holds:*

(1) *either $|U| \leqslant 100M$;*

(2) *or there exists $v \in U$ and a subset $W \subset Uv$ such that W is 25κ-strongly reduced and*

$$|W| \geqslant |U|/100M - 1.$$

Proof. — We assume that $|U| > 100M$, so that $|U_1| > M$. The proof follows the exact same ideas as Lemmas 5.2 and 5.3 of [DS20]. Since the energy $\lambda(U, p)$ at p is larger than 100κ, there exists $v \in U$ satisfying $|vp - p| > 100\kappa$. For every $u \in U_1$, we let

$$B_u = \{u' \in U_1 \mid (uvp, u'vp)_p \geqslant 23\kappa - \delta\}.$$

Note that by the triangle inequality, $|uvp - p| > |vp - p| - |up - p| \geqslant 98\kappa$, for every $u \in U_1$. Hence $u \in B_u$.

Let us fix first an element $u \in U_1$. We claim that the cardinality of B_u is at most M. Recall that X is a a length space, hence there is a point m in X such that $|p-m| = 21\kappa - \delta$ and $(p, vp)_m \leqslant \delta$. Let $u' \in B_u$. The element $u'u^{-1}$ moves the point up by at most 4κ. We now show that $u'u^{-1}$ moves um by at most $4\kappa + 8\delta$. By Lemma 2.2(1) we have

$$\begin{aligned}(p, uvp)_{um} &\leqslant \max\left\{|uvp - um| - (p, up)_{uvp}, (up, uvp)_{um}\right\} + \delta \\ &\leqslant \max\left\{|vp - m| - (p, up)_{uvp}, (p, vp)_m\right\} + \delta.\end{aligned} \tag{14}$$

On the one hand, we have

$$|vp - m| = |p - vp| - |p - m| + 2(p, vp)_m \leqslant |p - vp| - 21\kappa + 3\delta.$$

On the other hand, the triangle inequality yields

$$(p, up)_{uvp} \geqslant |up - uvp| - |p - up| \geqslant |p - vp| - 2\kappa.$$

If we plug in the last two inequalities in (14) we get $(p, uvp)_{um} \leqslant 2\delta$. Now observe that

$$\big||p - um| - |p - m|\big| = \big||p - um| - |up - um|\big| \leqslant |p - up| \leqslant 2\kappa.$$

Similarly $(p, u'vp)_{u'm} \leqslant 2\delta$ and $\big||p - u'm| - |p - m|\big| \leqslant 2\kappa$. In particular both $|p - um|$ and $|p - u'm|$ are at most $(uvp, u'vp)_p$. By Lemma 2.2(3) we have

$$|um - u'm| \leqslant \max\Big\{\big||p - um| - |p - u'm|\big| + 4\delta, 0\Big\} + 4\delta \leqslant 4\kappa + 8\delta,$$

which corresponds to our announcement.

Note that the point up and um, which are "hardly" moved by $u'u^{-1}$, are far away. More precisely

$$|up - um| = |p - m| = 21\kappa - \delta.$$

Recall that $M = 2\kappa N/\delta$. Using acylindricity – see (4) with $r = 4\kappa + 8\delta$ – we get that B_u contains at most M elements, which completes the proof of our claim.

Recall that $u \in B_u$, for every $u \in U_1$. We now fix a maximal subset $U_2 \subset U_1$ such that for every $u \in U_1$, any two distinct $u_1, u_2 \in U_2$ never belong to the same subset B_u. The cardinality of U_2 is at least $|U_2| \geqslant |U_1|/M$. Indeed by maximality of U_2, the U_1 is covered by the collection $(B_u)_{u \in U_2}$.

We claim that there is at most one element $u \in U_2$ such that $(v^{-1}p, uvp)_p > 23\kappa$. Assume on the contrary that it is not the case. We can find two distinct element $u, u' \in U_2$ such that

$$(uvp, u'vp)_p \geqslant \min\{(v^{-1}p, uvp)_p, (v^{-1}p, u'vp)_p\} - \delta > 23\kappa - \delta.$$

Thus u' belongs to B_u which contradicts the definition of U_2. Recall that $|U_1| > M$, hence U_2 contains at least 2 elements. We define then U_3 from U_2 by removing if necessary the element $u \in U_2$ such that $(v^{-1}p, uvp)_p > 23\kappa$. Note that

$$|U_3| \geqslant \frac{|U_1|}{M} - 1 \geqslant \frac{|U|}{100M} - 1.$$

We now let $W = U_3 v$. We are going to prove that W is 25κ-strongly reduced. Note first that

$$|wp - p| \geqslant |vp - p| - 2\kappa > 98\kappa > 50\kappa + 300\delta$$

for every $w \in W$. Let $w = uv$ and $w' = u'v$ be two elements in W. It follows from the triangle inequality that

$$(w^{-1}p, w'p)_p \leqslant (v^{-1}p, w'p)_p + |up - p| \leqslant (v^{-1}p, w'p)_p + 2\kappa.$$

By construction of U_3, no element $w' \in W$ has a large Gromov product with v^{-1}. Hence $(w^{-1}p, w'p)_p \leqslant 25\kappa$. Thus the set W is 25κ-reduced. By choice of U_2 we also have $(wp, w'p)_p < 23\kappa - \delta$ for every distinct $w, w' \in W$. Recall that

$$\min \{|wp - p|, |w'p - p|\} \geqslant |vp - p| - 2\kappa > 98\kappa.$$

Consequently, W is 25κ-strongly reduced. ☐

8. Growth in groups acting on hyperbolic spaces

As a warm-up for the study of Burnside groups we first prove the following statement.

Theorem 8.1. — *Let $\delta > 0$, $\kappa \geqslant 50 \cdot 10^3\delta$, and $N > 0$. Assume that the group G acts (N, κ)-acylindrically on a δ-hyperbolic length space. For every finite $U \subset G$ such that $\lambda(U) > 100\kappa$, one of the following holds.*

(1) $|U| \leqslant 400\kappa N/\delta$.

(2) *There exists $v \in U$ and a subset $W \subset Uv$ such that W is α-strongly reduced, with $\alpha \leqslant 25\kappa$, and*

$$|W| \geqslant \frac{1}{10^6 N} \frac{\delta}{\lambda(U)} |U|.$$

Proof of Theorem 8.1. — Let $U \subset G$ be a finite subset such that $\lambda(U) > 100\kappa$.

Choice of the base-point. — Let q be a point almost-minimizing the ℓ^∞-energy of U. We now fix the base-point p to be a quasi-center for U. By Proposition 5.3, we can assume that $|p - q| \leqslant \lambda(U)$.

Case 1: diffuse energy. — Let us first assume that U is of diffuse energy at p. That is, there is a subset $U' \subset U$ such that $|U'| \geqslant 99|U|/100$ and such that for all $u' \in U'$ we have $|u'p - p| > 2\kappa$. Then, by Proposition 6.6, there exists $v \in U$ and a subset $W \subset Uv$ such that W is α-strongly reduced (with $\alpha = 1002\delta$) and whose cardinality satisfies

$$|W| \geqslant \frac{1}{10^6 N} \frac{\delta}{\lambda(U)} |U|.$$

Case 2: concentrated energy. — Otherwise U is of concentrated energy at p. Indeed, there is a subset $U' \subset U$ of cardinality $|U'| \geqslant |U|/100$ such that $|u'p - p| \leqslant 2\kappa$, for all $u' \in U'$. Recall that $\lambda(U) > 100\kappa$. Assume that $|U| > 400\kappa N/\delta$. By Proposition 7.1, there exists $v \in U$ and a subset $W \subset Uv$ such that W is α-strongly reduced (with $\alpha = 25\kappa$) and whose cardinality satisfies

$$|W| \geqslant \frac{1}{200N} \frac{\delta}{\kappa} |U| - 1 \geqslant \frac{1}{10^6 N} \frac{\delta}{\lambda(U)} |U|.$$

This completes the proof of Theorem 8.1. ☐

Corollary 8.2. — *Let $\delta > 0$, $\kappa \geqslant 50\cdot 10^3\delta$, and $N > 0$. Assume that the group G acts (N,κ)-acylindrically on a δ-hyperbolic length space. For every finite $U \subset G$ such that $\lambda(U) > 100\kappa$ and for all integers $r \geqslant 0$, we have*

$$|U^r| \geqslant \Bigl(\frac{1}{10^6 N}\,\frac{\delta}{\lambda(U)}\,|U|\Bigr)^{[(r+1)/2]}.$$

Proof. — Without loss of generality we can assume that $|U| > 400\kappa N/\delta$. Indeed otherwise the base of the exponential function on the right hand side of the stated inequality is less than one, hence the statement is void. According to Theorem 8.1, there exists $v \in U$ and a subset $W \subset Uv$ such that W is α-strongly reduced and

$$|W| \geqslant \frac{1}{10^6 N}\,\frac{\delta}{\lambda(U)}\,|U|.$$

Let $s \geqslant 0$ be an integer. On the one hand, $(Uv)^s$ is contained in U^{2s}, hence $|U^{2s}| \geqslant |(Uv)^s|$. On the other hand $(Uv)^sU$ is contained in U^{2s+1}. Right multiplication by v induces a bijection from G to itself. Hence

$$|U^{2s+1}| \geqslant |(Uv)^sU| = |(Uv)^{s+1}|.$$

Recall that W is contained in Uv and freely generates a free-sub-semigroup of G by Lemma 3.2. It follows that for every integer $r \geqslant 0$,

$$|U^r| \geqslant |(Uv)^{[(r+1)/2]}| \geqslant |W|^{[(r+1)/2]} \geqslant \Bigl(\frac{1}{10^6 N}\,\frac{\delta}{\lambda(U)}\,|U|\Bigr)^{[(r+1)/2]}. \qquad \square$$

We now combine Theorem 8.1 with our estimates on the growth of aperiodic words, see Proposition 3.11. If we use Proposition 4.2 to compare the notion of aperiodic words and power-free elements we obtain the following useful growth estimate.

Corollary 8.3. — *Let $\delta > 0$, $\kappa \geqslant 50\cdot 10^3\delta$, $N > 0$ and $\lambda_0 \geqslant 0$. There exists a parameter $m_2 > 0$ with the following properties. Assume that the group G acts (N,κ)-acylindrically on a δ-hyperbolic geodesic space. Let $U \subset G$ such that $100\kappa < \lambda(U) \leqslant \lambda_0$. One of the following holds.*

(1) $|U| \leqslant \max\{4\kappa N/\delta, 4\cdot 10^6 N\lambda(U)/\delta\}$.

(2) *There is $v \in U$ with the following property. For every $r > 0$ and $m \geqslant m_2$, denote by $K(m,r)$ the set of all m-power-free elements in $(Uv)^r$. Then,*

$$|K(m,r)| \geqslant \Bigl(\frac{1}{4\cdot 10^6 N}\,\frac{\delta}{\lambda(U)}\,|U|\Bigr)^r.$$

Proof. — Let $U \subset G$ be a finite subset such that $\lambda(U) > 100\kappa$. Without loss of generality we can assume that $|U| > \max\{4\kappa N/\delta, 4\cdot 10^6 N\lambda(U)/\delta\}$. By Theorem 8.1 there exists $v \in U$ and a subset $W \subset Uv$ such that W is α-strongly reduced with $\alpha \leqslant 25\kappa$ and

$$|W| \geqslant \frac{1}{10^6 N}\,\frac{\delta}{\lambda(U)}\,|U|.$$

It follows from our choice that $|W| \geqslant 4$ and $\lambda(W) \leqslant 2\lambda(U)$.

Before moving on, let us recall some notations from Section 3. For every integer m, the set W_m^* stands for the collection of m-aperiodic words in W^*. In addition $S(r)$

and $B(r)$ are respectively the sphere and the ball of radius r in W^* (for the word metric with respect to W).

In view of Proposition 3.11, there exists $m_1 > 0$, which only depends on δ, N, κ and λ_0 such that for every $m \geqslant m_1$, for every $r \geqslant 0$, we have,

$$|W^*_m \cap B(r+1)| \geqslant \frac{|W|}{2}\,|W^*_m \cap B(r)|. \tag{15}$$

Let us now focus on the cardinality of *spheres*. As W is α-strongly reduced, it generates a free sub-semi-group (Lemma 3.2). Thus

$$|W^*_m \cap S(r+1)| = |W^*_m \cap B(r+1)| - |W^*_m \cap B(r)|.$$

If we combine this inequality with (15) and the fact that $|W|/4 \geqslant 1$, we obtain that

$$\begin{aligned}|W^*_m \cap S(r+1)| \geqslant \frac{|W|}{2}\,|W^*_m \cap B(r)| - |W^*_m \cap B(r)| &\geqslant \Big(\frac{|W|}{2} - 1\Big)|W^*_m \cap B(r)| \\ &\geqslant \frac{|W|}{4}\,|W^*_m \cap B(r)| \\ &\geqslant \frac{|W|}{4}\,|W^*_m \cap S(r)|.\end{aligned}$$

By an inductive argument, we obtain that, for all $r \geqslant 0$,

$$|W^*_m \cap S(r)| \geqslant (|W|/4)^r.$$

Now let $m_2 = m_1 + (2\lambda(S) + 20\delta)/\tau$ and let $m \geqslant m_2$. Then, by Proposition 4.2, every element in $W^*_{m'}$ (seen as an element of G) is m-power-free, where $m' = m - (2\lambda(S) + 20\delta)/\tau$ is larger than m_1. Thus,

$$|K(m,r)| \geqslant |W^*_{m'} \cap S(r)| \geqslant (|W|/4)^r.$$

This completes the proof. ☐

9. Small cancellation groups

In this section we recall the necessary background on small cancellation theory with a special attention on acylindricity, see Proposition 9.9. The presentation follows [Cou14] in content and notations.

9.1. Cones. — Let Y be a metric length space and let $\rho > 0$. The *cone of radius ρ over Y* is the set

$$Z(Y) = Y \times [0,\rho]/\sim,$$

where $\sim$ is the equivalence relation which identifies all the points of the form $(y,0)$ for $y \in Y$. If $x \in Z(Y)$, we write $x = (y,r)$ to say that (y,r) represents x. We let $v = (y,0)$ be the *apex* of the cone.

If y, y' are in Y, we let $\theta(y,y') = \min\{\pi,\ |y-y'|/\sinh\rho\}$ be their angle at v. There is a metric on $Z(Y)$ that is characterized as follows, see [BH99, Ch. I.5]. Let $x = (y,r)$ and $x' = (y',r')$ in $Z(Y)$. Then

$$\cosh|x-x'| = \cosh r\cosh r' - \sinh r\sinh r'\cos\theta(y,y').$$

It turns out that $Z(Y)$ is a hyperbolic space [Cou14, Prop. 4.6].

We let $\iota : Y \to Z(Y)$ be the embedding defined as $\iota(y) = (y, \rho)$. The metric distortion of ι is controlled by a function $\mu : \mathbb{R}_+ \to [0, 2\rho]$ that is characterized as follows: for every $t \in \mathbb{R}_+$,

$$\cosh \mu(t) = \cosh^2 \rho - \sinh^2 \rho \cos(\min\{\pi, t/\sinh \rho\}).$$

For all y, $y' \in Y$, we have

$$|\iota(y) - \iota(y')|_{Z(Y)} = \mu(|y - y'|_Y). \tag{16}$$

Let us mention some properties of μ for later use.

Proposition 9.1 ([Cou14, Prop. 4.4]). — *The map μ is continuous, concave, non-decreasing. Moreover, if $\mu(t) < 2\rho$, then $t \leqslant \pi \sinh(\mu(t)/2)$.* ☐

Let H be a group that acts by isometries on Y. Then H acts by isometries on $Z(Y)$ by $hx = (hy, r)$. We note that H fixes the apex of the cone.

9.2. The cone off space. — From now, we assume that X is a proper, geodesic, δ-hyperbolic space, where $\delta > 0$. We fix a parameter $\rho > 0$, whose value will be made precise later. In addition, we consider a group G that acts properly co-compactly by isometries on X. We assume that this action is (N, κ)-acylindrical.

We let $\mathcal{Q}$ be a collection of pairs (H, Y) such that Y is closed strongly-quasi-convex in X and H is a subgroup of $\operatorname{Stab}(Y)$ acting co-compactly on Y. Suppose that $\mathcal{Q}$ is closed under the action of G given by the rule $g(H, Y) = (gHg^{-1}, gY)$. In addition we assume that $\mathcal{Q}/G$ is finite. Furthermore, we let

$$\Delta(\mathcal{Q}) = \sup\left\{\operatorname{diam}\left(Y_1^{+5\delta} \cap Y_2^{+5\delta}\right) \mid (H_1, Y_1) \neq (H_2, Y_2) \in \mathcal{Q}\right\}$$

and

$$\mathcal{T}\,(\mathcal{Q}) = \inf\{\|h\| \mid h \in H \setminus \{1\},\, (H, Y) \in Q\}.$$

Observe that if $\Delta(\mathcal{Q})$ is finite, then H is normal in $\operatorname{Stab}(Y)$, for every $(H, Y) \in \mathcal{Q}$.

Let $(H, Y) \in \mathcal{Q}$. We denote by $|\cdot|_Y$ the length metric on Y induced by the restriction of $|\cdot|$ to Y. As Y is strongly quasi-convex, for all $y, y' \in Y$,

$$|y - y'|_X \leqslant |y - y'|_Y \leqslant |y - y'|_X + 8\delta.$$

We write $Z(Y)$ for the cone of radius ρ over the metric space $(Y, |\cdot|_Y)$.

We let the *cone-off space* $\dot{X} = \dot{X}(Y, \rho)$ be the space obtained by gluing, for each pair $(H, Y) \in \mathcal{Q}$, the cone $Z(Y)$ on Y along the natural embedding $\iota : Y \to Z(Y)$. We let $\mathcal{V}$ denote the set of apices of $\dot{X}$. We endow $\dot{X}$ with the largest metric $|\cdot|_{\dot{X}}$ such that the map $X \to \dot{X}$ and the maps $Z(Y) \to \dot{X}$ are 1-Lipschitz, see [Cou14, §5.1]. It has the following properties.

Lemma 9.2 ([Cou14, Lem. 5.7]). — *Let $(H, Y) \in \mathcal{Q}$. Let $x \in Z(Y)$ and $x' \in \dot{X}$. Let $d(x, Y)$ be the distance from x to $\iota(Y)$ computed in $Z(Y)$. If $|x - x'|_{\dot{X}} < d(x, Y)$, then $x' \in Z(Y)$ and $|x - x'|_{\dot{X}} = |x - x'|_{Z(Y)}$.*

We recall that μ is the map that controls the distortion of the embedding ι of Y in its cone, see (16). It also controls the distortion of the map $X \to \dot{X}$.

Lemma 9.3 ([Cou14, Lem. 5.8]). — *For all x, $x' \in X$, we have*

$$\mu(|x-x'|_X) \leqslant |x-x'|_{\dot{X}} \leqslant |x-x'|_X. \qquad \square$$

The action of G on X then extends to an action by isometries on $\dot{X}$: given any $g \in G$, a point $x=(y,r)$ in $Z(Y)$ is sent to the point $gx=(gy,r)$ in $Z(gY)$. We denote by K the normal subgroup generated by the subgroups H such that $(H,Y) \in \mathsf{Q}$.

9.3. The quotient space. — We let $\overline{X} = \dot{X}/K$ and $\overline{G} = G/K$. We denote by ζ the projection of $\dot{X}$ onto $\overline{X}$ and write $\overline{x}$ for $\zeta(x)$ for short. Furthermore, we denote by $\overline{\mathsf{V}}$ the image in $\overline{X}$ of the apices V. We consider $\overline{X}$ as a metric space equipped with the quotient metric, that is for every $x, x' \in \dot{X}$

$$|\overline{x}-\overline{x}'|_{\overline{X}} = \inf_{h\in K} |hx-x'|_{\dot{X}}.$$

We note that the action of G on $\dot{X}$ induces an action by isometries of $\overline{G}$ on $\overline{X}$. The following theorem summarizes Proposition 3.15 and Theorem 6.11 of [Cou14].

Theorem 9.4 (Small Cancellation Theorem [Cou14]). — *There are distances δ_0, δ_1, Δ_0 and ρ_0 (that do not depend on X or Q) such that, if $\delta \leqslant \delta_0$, $\rho > \rho_0$, $\Delta(\mathsf{Q}) \leqslant \Delta_0$, and $\mathsf{T}(\mathsf{Q}) > 4\pi \sinh \rho$, then the following holds:*

(1) *$\overline{X}$ is a proper geodesic δ_1-hyperbolic space on which $\overline{G}$ acts properly co-compactly.*

(2) *Let $r \in (0, \rho/20]$. If for all $v \in \mathsf{V}$, the distance $|x-v| \geqslant 2r$ then the projection $\zeta\colon \dot{X} \to \overline{X}$ induces an isometry from $B(x,r)$ onto $B(\overline{x},r)$.*

(3) *Let $(H,Y) \in \mathsf{Q}$. If $v \in \mathsf{V}$ stands for the apex of the cone $Z(Y)$, then the projection from G onto $\overline{G}$ induces an isomorphism from* $\mathrm{Stab}(Y)/H$ *onto* $\mathrm{Stab}(\overline{v})$. $\square$

Let us now fix δ_0, δ_1, Δ_0 and ρ_0 as in Theorem 9.4. We assume that $\delta \leqslant \delta_0$, $\Delta(\mathsf{Q}) \leqslant \Delta_0$, $\mathsf{T}(\mathsf{Q}) > 4\pi \sinh \rho$, and $\rho > \rho_0$, so that $\overline{X}$ is $\overline{\delta}$-hyperbolic, with $\overline{\delta} \leqslant \delta_1$.

We use point (2) of Theorem 9.4 to compare the local geometry of $\dot{X}$ and $\overline{X}$. To compare the global geometry, we use the following proposition.

Proposition 9.5 ([Cou14, Prop. 3.21]). — *Let $\overline{Z} \subset \overline{X}$ be $10\overline{\delta}$-quasi-convex and $d \geqslant 10\overline{\delta}$. If, for all $\overline{v} \in \overline{\mathsf{V}}$, we have $\overline{Z} \cap B(\overline{v}, \rho/5 + d + 1210\overline{\delta}) = \varnothing$, then there is a pre-image $Z \subset \dot{X}$ such that the projection ζ induces an isometry from Z onto $\overline{Z}$.*

In addition, if $\overline{S} \subset \overline{G}$ such that $\overline{S}\,\overline{Z} \subseteq \overline{Z}^{+d}$, then there is a pre-image $S \subset G$ such that for every $g \in S$, $z, z' \in Z$, we have $|\overline{g}\,\overline{z} - \overline{z}'| = |gz-z'|_{\dot{X}}$. $\square$

9.4. Group action on $\overline{X}$. — We collect some properties of the action of $\overline{G}$.

Lemma 9.6 (Lemma 6.8 of [Cou14]). — *If $v \in \mathsf{V}$ and $g \in \overline{G} \setminus \mathrm{Stab}(\overline{v})$, then for every $\overline{x} \in \overline{X}$ we have*

$$|\overline{g}\,\overline{x} - \overline{x}| \geqslant 2(\rho - |\overline{x} - \overline{v}|).$$

$\square$

In combination with assertion (2) of Theorem 9.4, the previous lemma implies that local properties of the action are often inherited from the action of G on the cone-off space. For example, if $\overline{F}$ is an elliptic subgroup of $\overline{G}$, then either $\overline{F} \subseteq \operatorname{Stab}(\overline{v})$ for some $v \in \mathsf{V}$ or it is the image of an elliptic subgroup of G, see [Cou14, Prop. 6.12].

There is a lower bound on the injectivity radius of the action on $\overline{X}$, and an upper bound on the acylindricity parameter.

Proposition 9.7 ([Cou14, Prop. 6.13]). — *Let* $\ell = \inf\{\|g\|^\infty \mid g \notin \operatorname{Stab}(Y), (H,Y) \in \mathsf{Q}\}$. *Then*

$$\tau(\overline{G},\overline{X}) \geqslant \min\left\{\frac{\rho\ell}{4\pi\sinh\rho}, \overline{\delta}\right\}. \qquad \square$$

We recall that L_0 is the number fixed in Section 2.2 using stability of quasi-geodesics.

Proposition 9.8 ([Cou14, Cor. 6.15]). — *Assume that all elementary subgroups of G are cyclic infinite or finite with odd order. If* $\operatorname{Stab}(Y)$ *is elementary for every* $(H,Y) \in \mathsf{Q}$, *then* $A(\overline{G},\overline{X}) \leqslant A(G,X) + 5\pi\sinh(2L_0\overline{\delta})$. $\square$

Note that the proposition actually does not require that finite subgroups of G have odd order. This assumption in [Cou14, Prop. 6.15] was mainly made to simplify the overall exposition in this paper. The error of the order of $\pi\sinh(2L_0\overline{\delta})$ in the above estimates is reminiscent of the distortion of the embedding of X into $\dot{X}$, measured by the map μ, see Proposition 9.1.

9.5. Acylindricity. — Let us assume that all elementary subgroups of G are cyclic (finite or infinite). In particular, it follows that $\nu(G,X) = 1$, see for instance [Cou14, Lem. 2.40]. Moreover, we assume that for every pair $(H,Y) \in \mathsf{Q}$, there is a primitive hyperbolic element $h \in G$ and a number n such that $H = \langle h^n \rangle$ and Y is the cylinder C_H of H.

Proposition 9.9. — *The action of* $\overline{G}$ *on* $\overline{X}$ *is* $(\overline{N},\overline{\kappa})$*-acylindrical, where*

$$\overline{N} \leqslant \max\left\{N, \frac{3\pi\sinh\rho}{\tau(G,X)} + 1\right\} \quad \text{and} \quad \overline{\kappa} = \max\{A(G,X), \kappa\} + 5\pi\sinh(150\overline{\delta}).$$

Remark 9.10. — It is already known that if G acts acylindrically on X, then so does $\overline{G}$ on $\overline{X}$, see Dahmani-Guirardel-Osin [DGO17, Prop. 2.17, 5.33]. However in their proof $\overline{\kappa}$ is much larger than ρ. For our purpose we need a sharper control on the acylindricity parameters. With our statement, we will be able to ensure that $\overline{\kappa} \ll \rho$.

Later we will use this statement during an induction process for which we also need to control *uniformly* the value of N. Unlike in [DGO17], if N is very large, our estimates tells us that $\overline{N} \leqslant N$.

Proof. — Let $\overline{S} \subset \overline{G}$, let

$$\overline{Z} = \operatorname{Fix}(\overline{S}, 100\overline{\delta})$$

and let us assume that $\operatorname{diam}\overline{Z} \geqslant \overline{\kappa}$. We are going to prove that $\overline{S}$ contains at most $\overline{N}$ elements. We distinguish two cases: either $\overline{S}$ fixes an apex $\overline{v} \in \overline{\mathsf{V}}$ or not.

Lemma 9.11. — *If there is $\overline{v} \in \overline{V}$, such that $\overline{S} \subset \mathrm{Stab}(\overline{v})$, then*

$$|\overline{S}| \leqslant 3\pi \sinh \rho / \tau(G,X) + 1.$$

Proof. — If $\overline{S} \subset \mathrm{Stab}(\overline{v})$, then $\overline{v} \in \overline{Z}$. As $\operatorname{diam}(\overline{Z}) \geqslant \overline{\kappa}$, there is a point $\overline{x} \in \overline{Z}$ such that $|\overline{v} - \overline{x}| \geqslant \overline{\kappa} - \overline{\delta}$. Recall that $\overline{\kappa} > 100\overline{\delta}$. Denote by $\overline{z}$ the point on the geodesic $[\overline{v}, \overline{x}]$ at distance $100\overline{\delta}$ from $\overline{v}$, so that $\overline{z} \in B(\overline{v}, \rho/2)$. Since $\overline{Z}$ is $10\overline{\delta}$-quasi-convex, $\overline{z}$ lies in the the $10\overline{\delta}$-neighborhood of $\overline{Z}$. In particular, for all $\overline{s} \in \overline{S}$, we have $|\overline{s}\,\overline{z} - \overline{z}| \leqslant 120\overline{\delta}$. Let v be a pre-image of $\overline{v}$ and z a pre-image of $\overline{z}$ in the ball $B(v, \rho/2)$. For every $\overline{s} \in \overline{S}$, we choose a pre-image $s \in G$ such that $|sz - z|_{\dot{X}} \leqslant 120\overline{\delta}$ and write S for the set of all pre-images obtained in this way. Observe that by the triangle inequality, $|sv - v|_{\dot{X}} \leqslant \rho + 120\overline{\delta}$, for every $s \in S$. However any two distinct apices in $\dot{X}$ are at a distance at least 2ρ. Thus S is contained in $\mathrm{Stab}(v)$. If $(H, Y) \in \mathcal{Q}$ is such that v is the apex of the cone $Z(Y)$, then, by Lemma 9.2, $|sz - z|_{Z(Y)} \leqslant 120\overline{\delta} < |z - v|_{Z(Y)} + |sz - v|_{Z(Y)}$. Let y be a radial projection of z on Y. By the very definition of the metric on $Z(Y)$, we get that $|sy - y| < \pi \sinh \rho$. Recall that every elementary subgroup is cyclic, in particular so is $\mathrm{Stab}(Y)$. Consequently, the number of elements $g \in \mathrm{Stab}(Y)$ such that $|gy - y| \leqslant r$ is linear in r. More precisely, using Lemma 2.15, we have

$$|S| \leqslant \frac{2(\pi \sinh \rho + 112\delta)}{\tau(G,X)} + 1 \leqslant \frac{3\pi \sinh \rho}{\tau(G,X)} + 1,$$

which yields the claim. □

Lemma 9.12. — *If $\overline{S}$ does not stabilize any $\overline{v} \in \overline{V}$, then $|\overline{S}| \leqslant \overline{N}$.*

Proof. — By Lemma 9.6, $\overline{Z} \cap B(\overline{v}, \rho - 100\overline{\delta}) = \varnothing$, for every $\overline{v} \in \overline{V}$. By Lemma 2.6, $\overline{Z}$ is $10\overline{\delta}$-quasi-convex. By Lemma 9.5, there exists pre-images $Z \subset \dot{X}$ and $S \subset G$ such that $\operatorname{diam}(Z) > \overline{\kappa}$ and for all $s \in S$ and all $z \in Z$, we have $|sz - z|_{\dot{X}} \leqslant 100\overline{\delta}$.

Let us write $d = \pi \sinh(150\overline{\delta})$. We now focus on the subset $\mathrm{Fix}(S, d) \subset X$. Let $x, y \in Z$ such that $|x - y| > \overline{\kappa}$. Let p, q be projections of x, y in X. Then, as $|p - x|_{\dot{X}} \leqslant 100\overline{\delta}$ and $|q - y|_{\dot{X}} \leqslant 100\overline{\delta}$, $|p - q|_{\dot{X}} \geqslant \overline{\kappa} - 200\overline{\delta}$. As $|p - q|_X \geqslant |p - q|_{\dot{X}}$, the distance $|p - q|_X \geqslant \overline{\kappa} - 200\overline{\delta}$. On the other hand, $\mu(|sp - p|_X) \leqslant |sp - p|_{\dot{X}} < 300\overline{\delta} < 2\rho$. Thus, by Proposition 9.1, $|sp - p|_X < d$. Similarly, $|sq - q|_X < d$. This means that the diameter of $\mathrm{Fix}(S, d) \subset X$ is at least $\overline{\kappa} - 200\overline{\delta}$, hence, larger than $A(G, X) + 4d + 209\delta$. It follows by Proposition 2.13 that S generates an elementary subgroup E.

Suppose first that this subgroup E is loxodromic. It is infinite cyclic by assumption. Recall that the translation length of any element in S is at most d. Hence, as previously we get

$$|S| \leqslant \frac{2(d + 112\delta)}{\tau(G,X)} + 1 \leqslant \frac{3\pi \sinh \rho}{\tau(G,X)} + 1.$$

Suppose now that E is an elliptic subgroup. In particular, the set $\mathrm{Fix}(S, 14\delta) \subset X$ is non-empty, and by Lemma 2.6, $\mathrm{Fix}(S, d)$ is contained in the $d/2$-neighborhood of $\mathrm{Fix}(S, 14\delta)$. In particular the diameter of $\mathrm{Fix}(S, 14\delta)$ is larger that $\overline{\kappa} - 200\overline{\delta} - d$, hence, larger than κ. Consequently, by acylindricity, $|S| \leqslant N$. □

This completes the proof of Proposition 9.9. □

9.6. ℓ^∞-energy. — In this section we compare the ℓ^∞-energy of finite subset $U \subset G$ and its image $\overline{U} \subset \overline{G}$ respectively.

Proposition 9.13. — *Let $\overline{U} \subset \overline{G}$ be a finite set such that $\lambda(\overline{U}) \leqslant \rho/5$. If, for all $\overline{v} \in \overline{V}$, the set $\overline{U}$ is not contained in* $\operatorname{Stab}(\overline{v})$*, then there is a pre-image $U \subset G$ of $\overline{U}$ of energy $\lambda(U) \leqslant \pi \sinh \lambda(\overline{U})$.*

Proof. — Let $\varepsilon > 0$. Let $\overline{q} \in \overline{X}$ such that $\lambda(\overline{U}, \overline{q}) \leqslant \lambda(\overline{U}) + \varepsilon$. By Lemma 9.6, $|\overline{q} - \overline{v}| > \rho - (\lambda(\overline{U}) + \varepsilon)/2 > 4\rho/5$, for all $v \in V$. Let q be a pre-image of $\overline{q}$ in $\dot{X}$. We choose a pre-image $U \subset G$ of $\overline{U}$ such that for every $u \in U$, we have $|uq - q|_{\dot{X}} = |\overline{uq} - \overline{q}|$. Let $x \in X$ be a projection of q onto X. We note that $\mu(|ux - x|_X) \leqslant |ux - x|_{\dot{X}} \leqslant 2(\lambda(\overline{U}) + \varepsilon) < 2\rho$. Thus $|ux - x|_X \leqslant \pi \sinh(\lambda(\overline{U}) + \varepsilon)$, see Proposition 9.1. We just proved that $\lambda(U) \leqslant \pi \sinh(\lambda(\overline{U}) + \varepsilon)$ for every $\varepsilon > 0$, whence the result. ☐

10. Product set growth in Burnside groups of odd exponent

We finally prove Theorem 1.2.

10.1. The induction step. — We will use the following.

Proposition 10.1 (cf. [Cou14, Prop. 6.18]). — *There are distances $\rho_0, \delta_1 > 0$, and $A_0 \in [50 \cdot 10^3 \delta_1, \rho_0/500]$, as well as natural numbers L_0 and n_0 such that the following holds.*

Let $n_1 \geqslant n_0$ and $n \geqslant n_1$ be an odd integer. Let G act properly co-compactly by isometries on a proper geodesic δ_1-hyperbolic space X such that

(1) *the elementary subgroups of G are cyclic or finite of odd order n,*
(2) *$A(G, X) \leqslant A_0$ and $\tau(G, X) \geqslant \sqrt{\rho_0 L_0 \delta_1 / 4 n_1}$, and*
(3) *the action of G is (N, A_0)-acylindrical, for some integer N.*

Let P be the set of primitive hyperbolic elements h of translation length $\|h\| \leqslant L_0 \delta_1$. Let K be the normal closure of the set $\{h^n \mid h \in P\}$ in G.

Then there is proper geodesic δ_1-hyperbolic space $\overline{X}$ on which $\overline{G} = G/K$ acts properly co-compactly by isometries. Moreover,

- (1) *and* (2) *hold for the action of $\overline{G}$ on $\overline{X}$;*
- *the action of $\overline{G}$ on $\overline{X}$ is $(\overline{N}, A_0)$-acylindrical where $\overline{N} = \max\{N, n_1\}$;*
- *if $\overline{U}$ is a subset of $\overline{G}$ with $\lambda(\overline{U}) \leqslant \rho_0/5$ that does not generated a finite subgroup, then there exists a pre-image $U \subset G$ of $\overline{U}$ such that $\lambda(U) \leqslant \sqrt{n_1} \sinh \lambda(\overline{U})$.*

Remark 10.2. — Note that Assumptions (2) and (3) are somewhat redundant. Indeed, if the action of G on X is (N, κ)-acylindrical, then the parameters $A(G, X)$ and $\tau(G, X)$ can be estimated in terms of δ, N and κ only. However, we chose to keep them both, to make it easier to apply existing results in the literature.

Proof. — This is essentially [Cou14, Prop. 7.1]. The only additional observation is point (3). For details of the proof, we refer the reader to [Cou14]. Here, we only give a rough idea of the proof and fix some notation for later use.

We choose for δ_0, Δ_0, δ_1, and ρ_0 the constants given by the Small Cancellation Theorem, see Theorem 9.4. We fix

$$A_0 = \max\{6\pi\sinh(2L_0\delta_1), 50\cdot 10^3\delta_1\}.$$

Without loss of generality we can assume that $\delta_0, \Delta_0 \ll \delta_1$ while $\rho_0 \gg L_0\delta_1$. In particular $A_0 \leqslant \rho_0/500$. Following [Cou14, pp. 319], we define a rescaling constant as follows. Let

$$\varepsilon_n = \frac{8\pi\sinh\rho_0}{\sqrt{\rho_0 L_0\delta_1}}\frac{1}{\sqrt{n}}.$$

We note for later use that if ρ_0 is sufficiently large (which we assume here) we have $\varepsilon_n \geqslant 1/\sqrt{n}$, for every $n > 0$. We then choose n_0 such that for all $n \geqslant n_0$, the following holds

$$\varepsilon_n\delta_1 \leqslant \delta_0, \tag{17}$$

$$\varepsilon_n(A_0 + 118\delta_1) \leqslant \min\{\Delta_0, \pi\sinh(2L_0\delta_1)\}, \tag{18}$$

$$\frac{\varepsilon_n\rho_0L_0\delta_1}{16\pi\sinh\rho_0} \leqslant \delta_1, \tag{19}$$

$$\varepsilon_n < 1. \tag{20}$$

These are the same conditions as in [Cou14, p. 319] (in this reference, ε is denoted by λ). We now fix $n_1 \geqslant n_0$ and an odd integer $n \geqslant n_1$. For simplicity we let $\varepsilon = \varepsilon_{n_1}$. Moreover, let

$$\mathcal{Q} = \left\{\left(\langle h^n\rangle, C_{E(h)}\right) \mid h \in P\right\}.$$

As explained in [Cou14, Lem. 7.2], the small cancellation hypothesis needed to apply Theorem 9.4 are satisfied by $\mathcal{Q}$ for the action of G on εX. We let $\overline{G}$ and $\overline{X}$ as in Section 9.3 (applied to G acting on εX). Observe, for later use, that the map

$$X \longrightarrow \varepsilon X \xrightarrow{\zeta} \overline{X}$$

is ε-Lipschitz. Assertions (1) and (2) follows from Lemmas 7.3 and 7.4 in [Cou14]. By Proposition 9.9 the action of $\overline{G}$ on $\overline{X}$ is $(\overline{N}, \overline{\kappa})$-acylindrical where

$$\overline{N} \leqslant \max\left\{N, \frac{3\pi\sinh\rho}{\tau(G,\varepsilon X)} + 1\right\} \quad \text{and} \quad \overline{\kappa} = \max\{A(G,\varepsilon X), \varepsilon A_0\} + 5\pi\sinh(150\delta_1).$$

It follows from the definition of ε and our hypothesis on $\tau(G,X)$ that $\overline{N} \leqslant \max\{N, n_1\}$. On the other hand by (18) we have

$$\overline{\kappa} \leqslant \varepsilon A_0 + 5\pi\sinh(150\delta_1) \leqslant A_0.$$

Hence the action of the $\overline{G}$ on $\overline{X}$ is $(\overline{N}, A_0)$-acylindrical as we announced.

Consider now a subset $\overline{U}$ of $\overline{G}$ such that $\lambda(\overline{U}) \leqslant \rho_0/5$ and $\overline{U}$ does not generate a finite subgroup. Hence, applying Proposition 9.13, we see that there exists a pre-image $U \subset G$ of $\overline{U}$ such that the ℓ^∞-energy of U for the action of G on εX is bounded above by $\pi\sinh\lambda(\overline{U})$. Thus, for the action of G on X, we obtain that

$$\lambda(U) \leqslant \varepsilon^{-1}\pi\sinh\lambda(\overline{U}) < \sqrt{n_1}\sinh\lambda(\overline{U}).$$

This is the lifting property stated at the end of Proposition 10.1. $\square$

Assume now that G is a non-elementary, torsion-free hyperbolic group. Proposition 10.1 can be used as the induction step to build from G a sequence of hyperbolic groups (G_i) that converges to the infinite periodic quotient G/G^n, provided n is a sufficiently large odd exponent. For our purpose, we need a sufficient condition to detect whenever an element $g \in G$ has a trivial image in G/G^n. This is the goal of the next statement, see [Cou18a, Th. 4.13]. The result is reminiscence of the key argument used by Ol'shanskiĭ in [Ol'91, §10]. Recall that the definition of containing a (large) power (Definition 4.1) involves the choice of a basepoint $p \in X$.

Theorem 10.3. — *Let G be a non-elementary torsion-free group acting properly co-compactly by isometries on a hyperbolic geodesic space X. We fix a basepoint $p \in X$. There are n_0 and ξ such that for all odd integers $n \geqslant n_0$ the following holds. If g_1 and g_2 are two elements of G whose images in G/G^n coincide, then one of them contains a $(n/2 - \xi)$-power.* ☐

Here, we need a stronger result. Indeed we will have to apply this criterion for any group (G_i) approximating G/G^n. In particular we need to make sure that the critical exponent n_0 appearing in Theorem 10.3 does not depend on i. For this reason, we use instead the following statement.

Theorem 10.4. — *There are distances $\rho_0, \delta_1 > 0$, and $A_0 \in [50 \cdot 10^3\delta_1, \rho_0/500]$, as well as natural numbers L_0, n_0 such that the following holds.*

Let $n_1 \geqslant n_0$ and set $\xi = n_1 + 1$. Fix an odd integer $n \geqslant \max\{100, 50n_1\}$. Let G be a group acting properly, co-compactly by isometries on a proper, geodesic, δ_1-hyperbolic space X with a basepoint $p \in X$, such that

(1) *the elementary subgroups of G are cyclic or finite of odd order n,*

(2) *$A(G,X) \leqslant A_0$ and $\tau(G,X) \geqslant \sqrt{\rho_0 L_0 \delta_1/4n_1}$.*

If g_1 and g_2 are two elements of G whose images in G/G^n coincide, then one of them contains a $(n/2 - \xi)$-power.

Remark 10.5. — The "novelty" of Theorem 10.4 compared to Theorem 10.3 is that the critical exponent n_0 does not depend on G but only on the parameters of the action of G on X (acylindricity, injectivity radius, etc). Note that the critical exponent given by Ol'shanskiĭ in [Ol'91] only depends on the hyperbolicity constant of the *Cayley graph* of G. However this parameter will explode along the sequence (G_i). Thus we cannot formally apply this result. Although it is certainly possible to adapt Ol'shanskiĭ's method, we rely here on the material of [Cou18a].

Sketch of proof. — The arguments follow verbatim the ones of [Cou18a, §4]. Observe first that the parameters δ_1, L_0, ρ_0, A_0 and n_0 in [Cou18a, p. 797] are chosen in a similar way as we did in the proof of Proposition 10.1 (note that the rescaling parameter that denote ε_n is called λ_n there). Once $n_1 \geqslant n_0$ has been fixed, we set, exactly as in [Cou18a, p. 797], $\xi = n_1+1$ and $n_2 = \max\{100, 50n_1\}$. We now fix an odd integer $n \geqslant n_2$. At this point in the proof of [Cou18a] one chooses a non-elementary torsion-free group G acting properly co-compactly on a hyperbolic space X with a basepoint $p \in X$. Note in particular that the base point p is chosen after fixing all the other parameters. Next one uses an analogue of Proposition 10.1 to build a sequence of hyperbolic groups (G_i) converging to G/G^n. The final statement, that is Theorem 10.3, is then proved using an induction on i, see [Cou18a, Prop. 4.6].

Observe that the fact that G is torsion-free is not necessary here. We only need that the initial group G satisfies the induction hypothesis, that is:

(1) X is a geodesic δ_1-hyperbolic space on which G acts properly co-compactly by isometries.

(2) the elementary subgroups of G are cyclic or finite of odd order n,

(3) $A(G,X) \leqslant A_0$ and $\tau(G,X) \geqslant \sqrt{\rho_0 L_0 \delta_1 / 4n_1}$.

These are exactly the assumptions stated in Theorem 10.4. In particular, we can build as in [Cou18a] a sequence of hyperbolic (G_i) converging to G/G^n. The theorem is proved using an induction on i just as in [Cou18a]. Actually the proof is even easier, since we only need a sufficient condition to detect elements of G which are not trivial in G/G^n, while [Cou18a] provides a sufficient and necessary condition for this property. ☐

10.2. The approximating sequence.

— Let G be a non-elementary torsion-free hyperbolic group. The periodic quotient G/G^n is the direct limit of a sequence of infinite hyperbolic groups G_i that can be recursively constructed as follows. We let δ_1, ρ_0, L_0, n_0, and $A_0 \geqslant 50 \cdot 10^3 \delta_1$ be the parameters given by Proposition 10.1.

Let $G_0 = G$ and let X_0 be its Cayley graph. Up to rescaling X_0 we can assume that X_0 is a δ_1-hyperbolic metric geodesic space and $A(G_0, X_0) \leqslant A_0$. We choose $n_1 \geqslant n_0$ such that

$$\tau(G_0, X_0) \geqslant \sqrt{\frac{\rho_0 L_0 \delta_1}{4n_1}}.$$

Recall that the action of G_0 on X_0 is proper and co-compact. Thus there exists $N \geqslant n_1$, such that every subset $S \subset G_0$ for which $\operatorname{Fix}(S, 100\delta_1)$ is non-empty contains at most N elements. Consequently, the action is (N, A_0)-acylindrical. For simplicity we let $\lambda_0 = \sqrt{n_1}\pi \sinh(100A_0)$ and denote by $m_2 = m_2(\delta_1, N, A_0, \lambda_0)$ the parameter given by Corollary 8.3. In addition, we set $\xi = n_1 + 1$ and

$$n_2 = \max\{100, 50n_1, 2(m_2 + \xi)\}.$$

Let $n \geqslant n_2$ be an odd integer. It follows from our choices that the assumptions of Proposition 10.1 are then satisfied for the action of G_0 on X_0.

Let us suppose that G_i is already given, and acts on a δ_1-hyperbolic space X_i such that the assumptions of Proposition 10.1 are satisfied. Then $G_{i+1} = \overline{G_i}$ and $X_{i+1} = \overline{X_i}$ are given by Proposition 10.1. In particular, the action of G_{i+1} on X_{i+1} is $(\overline{N}, A_0)$-acylindrical, with $\overline{N} = \max\{N, n_1\}$. However we chose $N \geqslant n_1$. Hence the action of G_{i+1} on X_{i+1} is (N, A_0)-acylindrical. It follows from the construction that G/G^n is the direct limit of the sequence (G_i). Compare with [Cou14, Th. 7.7].

Remark 10.6. — As the quotient G/G^n is a direct limit of non-elementary hyperbolic groups, it is an infinite group itself. In fact, for the same reason, it is not finitely presented either, see [Cou14, Th. 7.7].

10.3. Growth estimates. — As before, we write $\varepsilon = \varepsilon_{n_1}$ for the renormalisation parameter that we used in the proof of Proposition 10.1. The action of G_0 on X_0 is proper and co-compact, hence there exists an integer M_0 such that for every $x \in X_0$,

$$|\{g \in G_0 \mid |gx - x| \leqslant 100A_0\}| \leqslant M_0.$$

We now let

$$M = \max\left\{M_0, \frac{4A_0N}{\delta_1}, \frac{4 \cdot 10^6 N\lambda_0}{\delta_1}\right\}, \quad \text{and} \quad a = \frac{1}{M}.$$

Let $V \subset G/G^n$ be finite and not contained in a finite subgroup. Recall that if $U_i \subset G_i$ is a pre-image of V, its energy measured in X_i is defined by

$$\lambda(U_i) = \inf_{x \in X_i} \max_{u \in U_i} |ux - x|_{X_i}.$$

We now let

$$j = \inf\{i \in \mathbb{N} \mid \text{there is a pre-image } U \subset G_i \text{ of } V \text{ such that } \lambda(U) \leqslant 100A_0\}. \tag{21}$$

Recall that the map $X_i \to X_{i+1}$ is ε-Lipschitz. Hence, if $U_{i+1} \subset G_{i+1}$ is the image of a subset $U_i \subset G_i$, we have $\lambda(U_{i+1}) \leqslant \varepsilon\lambda(U_i)$. Since $\varepsilon < 1$, the index j is well-defined. Let us fix a pre-image U_j of V in G_j such that $\lambda(U_j) \leqslant 100A_0$. We now distinguish two cases.

Case 1. — Assume that $j = 0$. It follows from our choice of M_0, that $|V| \leqslant |U_0| \leqslant M_0$. Thus for every $r \geqslant 0$ we have

$$|V^r| \geqslant 1 \geqslant \left(\frac{1}{M_0}|V|\right)^{[(r+1)/2]} \geqslant (a|V|)^{[(r+1)/2]}.$$

Case 2. — Assume that $j > 0$. Note that U_j cannot generate a finite subgroup G_j, otherwise so would V in G/G^n. Recall that $100A_0 \leqslant \rho_0/5$. By Proposition 10.1, there exists a pre-image $U_{j-1} \subset G_{j-1}$ of U_j such that the energy of U_{j-1} satisfies $\lambda(U_{j-1}) \leqslant \lambda_0$. By definition of j, we also have $\lambda(U_{j-1}) > 100A_0$. For simplicity we let $m = n/2 - \xi$. It follows from our choice of n that $m \geqslant m_2$. Hence we can apply Corollary 8.3 so that one of the following holds.

- The cardinality of U_{j-1} is at most $\max\{4A_0N/\delta_1, 4 \cdot 10^6 N\lambda_0/\delta_1\}$, which is by definition bounded above by M. In particular, the same holds for V and we prove as in Case 1 that for every $r \geqslant 0$,

$$|V^r| \geqslant (a|V|)^{[(r+1)/2]}.$$

- There is $v \in U_{j-1}$ such that if $K(m,r)$ stands for the set of all m-power elements in $(U_{j-1}v)^r$, then

$$|K(m,r)| \geqslant \left(\frac{1}{4 \cdot 10^6 N}\frac{\delta_1}{\lambda(U_{j-1})}|U_{j-1}|\right)^r.$$

On the one hand $\lambda(U_{j-1}) \leqslant \lambda_0$. On the other hand, $M \geqslant 4\cdot 10^6 N\lambda_0/\delta_1$, while $a = 1/M$. Consequently,

$$|K(m,r)| \geqslant \Big(\frac{1}{4\cdot 10^6 N}\,\frac{\delta_1}{\lambda_0}\,|U_{j-1}|\Big)^r \geqslant (a|V|)^r.$$

According to our choice of n, we have $n \geqslant \max\{100, 50n_1\}$. Moreover, by construction G_{j-1} satisfies the assumptions of Theorem 10.4. Hence, the projection $\pi\colon G_{j-1}\to G/G^n$ induces an embedding from $K(m,r)$ into $(V\pi(v))^r$. Consequently, $|(V\pi(v))^r| \geqslant (a|V|)^r$.

The proof now goes as in Corollary 8.2. Let $s \geqslant 0$ be an integer. On the one hand, $(V\pi(v))^s$ is contained in V^{2s}, hence $|V^{2s}| \geqslant |(V\pi(v))^s|$. On the other hand $(V\pi(v))^s V$ is contained in V^{2s+1}. Right multiplication by $\pi(v)$ induces a bijection from G/G^n to itself. Hence

$$|V^{2s+1}| \geqslant |(V\pi(v))^s V| = |(V\pi(v))^{s+1}| \geqslant (a|V|)^{s+1}.$$

It follows that $|V^r| \geqslant (a|V|)^{[(r+1)/2]}$, for every integer $r \geqslant 0$.

This completes the proof of Theorem 1.2. ☐

Proof of Corollary 1.3. — Let $n_0 > 0$ and $a > 0$ be the constants given by Theorem 1.2. We fix N such that $a^3N > 1$. Let $n \geqslant n_0$. Let us take a subset $V \subset G/G^n$ that is not contained in a finite subgroup and that contains the identity. Then, for all $k \geqslant 1$, we have $V^{k-1} \subseteq V^k$. As V is not contained in a finite subgroup, this implies that $|V^k| > |V^{k-1}|$. Thus $a^3|V^N| > 1$. We now apply twice Theorem 1.2, first with the set V^{3N}, and second with V^N. For every integer $r \geqslant 0$, we have

$$|V^{3rN}| \geqslant \big(a|V^{3N}|\big)^{[(r+1)/2]} \geqslant \big(a\big(a|V^N|\big)^2\big)^{[(r+1)/2]} \geqslant \big(a^3|V^N|\cdot|V^N|\big)^{[(r+1)/2]}.$$

Recall that $a^3|V^N| > 1$. Hence, for every integer $r \geqslant 0$,

$$|V^{3rN}| \geqslant |V^N|^{[(r+1)/2]} \geqslant |V|^{[(r+1)/2]}.$$

Taking the logarithm and passing to the limit we get

$$h(V) \geqslant \frac{1}{6N}\ln(|V|).$$

Since V does not lie in a cyclic subgroup and contains the identity, it has at least three elements, whence the second inequality in our statement. ☐

References

[Adi79] S. I. Adian – *The Burnside problem and identities in groups*, Ergeb. Math. Grenzgeb. (3), vol. 95, Springer-Verlag, Berlin-New York, 1979.

[AL06] G. N. Arzhantseva & I. G. Lysenok – “A lower bound on the growth of word hyperbolic groups”, *J. London Math. Soc. (2)* **73** (2006), no. 1, p. 109–125.

[Ata09] V. S. Atabekyan – “Uniform nonamenability of subgroups of free Burnside groups of odd period”, *Mat. Zametki* **85** (2009), no. 4, p. 516–523.

[BG08] J. Bourgain & A. Gamburd – “On the spectral gap for finitely-generated subgroups of SU(2)”, *Invent. Math.* **171** (2008), no. 1, p. 83–121.

[BG12] ———, “A spectral gap theorem in SU(d)”, *J. Eur. Math. Soc. (JEMS)* **14** (2012), no. 5, p. 1455–1511.

[Bow08] B. H. Bowditch – "Tight geodesics in the curve complex", *Invent. Math.* **171** (2008), no. 2, p. 281–300.

[BF21] E. Breuillard & K. Fujiwara – "On the joint spectral radius for isometries of non-positively curved spaces and uniform growth", *Ann. Inst. Fourier (Grenoble)* **71** (2021), no. 1, p. 317–391.

[BGT12] E. Breuillard, B. Green & T. Tao – "The structure of approximate groups", *Publ. Math. Inst. Hautes Études Sci.* **116** (2012), p. 115–221.

[BH99] M. R. Bridson & A. Haefliger – *Metric spaces of non-positive curvature*, Grundlehren Math. Wiss., vol. 319, Springer-Verlag, Berlin, 1999.

[But13] J. O. Button – "Explicit Helfgott type growth in free products and in limit groups", *J. Algebra* **389** (2013), p. 61–77.

[Cha08] M.-C. Chang – "Product theorems in SL_2 and SL_3", *J. Inst. Math. Jussieu* **7** (2008), no. 1, p. 1–25.

[CDP90] M. Coornaert, T. Delzant & A. Papadopoulos – *Géométrie et théorie des groupes. Les groupes hyperboliques de Gromov*, Lect. Notes in Math., vol. 1441, Springer-Verlag, Berlin, 1990.

[Cou13] R. Coulon – "Growth of periodic quotients of hyperbolic groups", *Algebraic Geom. Topol.* **13** (2013), no. 6, p. 3111–3133.

[Cou14] ———, "On the geometry of Burnside quotients of torsion free hyperbolic groups", *Internat. J. Algebra Comput.* **24** (2014), no. 3, p. 251–345.

[Cou16] ———, "Partial periodic quotients of groups acting on a hyperbolic space", *Ann. Inst. Fourier (Grenoble)* **66** (2016), no. 5, p. 1773–1857.

[Cou18a] ———, "Detecting trivial elements of periodic quotient of hyperbolic groups", *Bull. Soc. math. France* **146** (2018), no. 4, p. 745–806.

[Cou18b] ———, "Infinite periodic groups of even exponents", 2018, `arXiv:1810.08372`.

[DGO17] F. Dahmani, V. Guirardel & D. Osin – *Hyperbolically embedded subgroups and rotating families in groups acting on hyperbolic spaces*, Mem. Amer. Math. Soc., vol. 245, no. 1156, American Mathematical Society, Providence, RI, 2017.

[DG08] T. Delzant & M. Gromov – "Courbure mésoscopique et théorie de la toute petite simplification", *J. Topology* **1** (2008), no. 4, p. 804–836.

[DS20] T. Delzant & M. Steenbock – "Product set growth in groups and hyperbolic geometry", *J. Topology* **13** (2020), no. 3, p. 1183–1215.

[FS20] K. Fujiwara & Z. Sela – "The rates of growth in a hyperbolic group", 2020, `arXiv:2002.10278`.

[GdlH90] É. Ghys & P. de la Harpe (eds.) – *Sur les groupes hyperboliques d'après Mikhael Gromov*, Progress in Math., vol. 83, Boston, MA, Birkhäuser Boston, Inc., 1990.

[Gro87] M. Gromov – "Hyperbolic groups", in *Essays in group theory*, Math. Sci. Res. Inst. Publ., vol. 8, Springer, New York, 1987, p. 75–263.

[Hel08] H. A. Helfgott – "Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$", *Ann. of Math. (2)* **167** (2008), no. 2, p. 601–623.

[Iva94] S. V. Ivanov – "The free Burnside groups of sufficiently large exponents", *Internat. J. Algebra Comput.* **4** (1994), no. 1-2, p. 1–308.

[IO96] S. V. Ivanov & A. Y. Ol'shanskiĭ – "Hyperbolic groups and their quotients of bounded exponents", *Trans. Amer. Math. Soc.* **348** (1996), no. 6, p. 2091–2138.

[Ker21] A. Kerr – "Product set growth in mapping class groups", 2021, `arXiv:2103.12643`.

[Kou98] M. Koubi – "Croissance uniforme dans les groupes hyperboliques", *Ann. Inst. Fourier (Grenoble)* **48** (1998), no. 5, p. 1441–1453.

[Lys96] I. G. Lysenok – "Infinite Burnside groups of even period", *Izv. Akad. Nauk SSSR Ser. Mat.* **60** (1996), no. 3, p. 3–224.

[Nat96] M. B. Nathanson – *Additive number theory. Inverse problems and the geometry of sumsets*, Graduate Texts in Math., vol. 165, Springer-Verlag, New York, 1996.

[Ol'82] A. Y. Ol'shanskiĭ – "The Novikov-Adyan theorem", *Mat. Sb. (N.S.)* **118** (1982), no. 2, p. 203–235, 287.

[Ol'91] ———, "Periodic quotient groups of hyperbolic groups", *Mat. Sb. (N.S.)* **182** (1991), no. 4, p. 543–567.

[Osi07] D. V. Osin – "Uniform non-amenability of free Burnside groups", *Arch. Math. (Basel)* **88** (2007), no. 5, p. 403–412.

[Raz14] A. A. Razborov – "A product theorem in free groups", *Ann. of Math. (2)* **179** (2014), no. 2, p. 405–429.

[Saf11] S. R. Safin – "Powers of subsets of free groups", *Mat. Sb. (N.S.)* **202** (2011), no. 11, p. 97–102.

[Sel97] Z. Sela – "Acylindrical accessibility for groups", *Invent. Math.* **129** (1997), no. 3, p. 527–565.

[Tao08] T. Tao – "Product set estimates for non-commutative groups", *Combinatorica* **28** (2008), no. 5, p. 547–594.

[Tao10] ———, "Freiman's theorem for solvable groups", *Contrib. Discrete Math.* **5** (2010), no. 2, p. 137–184.

[TV06] T. Tao & V. Vu – *Additive combinatorics*, Cambridge Studies in Advanced Math., vol. 105, Cambridge University Press, Cambridge, 2006.

Manuscript received 26th February 2021
accepted 15th February 2022

Rémi Coulon, IRMAR, Univ Rennes et CNRS
35000 Rennes, France
E-mail : `remi.coulon@univ-rennes1.fr`
Url : `http://rcoulon.perso.math.cnrs.fr/`

Markus Steenbock, Fakultät für Mathematik, Universität Wien
1090 Wien, Austria
E-mail : `markus.steenbock@univie.ac.at`
Url : `https://sites.google.com/view/msteenbock`