ANNALS OF MATHEMATICS

Small cancellations over relatively hyperbolic groups and embedding theorems

By Denis Osin

Small cancellations over relatively hyperbolic groups and embedding theorems

By Denis Osin

Abstract

We generalize the small cancellation theory over ordinary hyperbolic groups to relatively hyperbolic settings. This generalization is then used to prove various embedding theorems for countable groups. For instance, we show that any countable torsion free group can be embedded into a finitely generated group with exactly two conjugacy classes. In particular, this gives the affirmative answer to the well-known question of the existence of a finitely generated group G other than $\mathbb{Z}/2\mathbb{Z}$ such that all nontrivial elements of G are conjugate.

1. Introduction

Originally the notion of relative hyperbolicity was proposed by Gromov [11] in order to generalize various examples of algebraic and geometric nature such as Kleinian groups, fundamental groups of hyperbolic manifolds of pinched negative curvature, small cancellation quotients of free products, etc. It has been extensively studied in the last several years from different points of view. The main aim of this paper is to generalize the small cancellation theory over hyperbolic groups developed by Olshanskii [23] to relatively hyperbolic settings. Our approach is based on author's papers [28], [26], [27], where the necessary background is provided. In the present paper we apply small cancellations over relatively hyperbolic groups to prove embedding theorems for countable groups. Further applications of our methods can be found in [2], [1], [4], [19], [20], [25].

In the paper [13], Higman, B. H. Neumann, and H. Neumann proved that any countable group G can be embedded into a countable group B such that every two elements of the same order are conjugate in B. We notice that the group B in [13] is constructed as a union of infinite number of subsequent HNN-extensions and thus B is never finitely generated. On the other hand, any countable group

---

can be embedded into a 2-generated group [13]. Our first theorem is a natural generalization of both of these results. For a group G, we denote by $\pi(G)$ the set of all finite orders of elements of G.

THEOREM 1.1. *Any countable group G can be embedded into a 2-generated group C such that any two elements of the same order are conjugate in C and $\pi(G) = \pi(C)$.*

COROLLARY 1.2. *Any countable torsion-free group can be embedded into a (torsion-free) 2-generated group with exactly 2 conjugacy classes.*

Since the number of finitely generated subgroups in any 2-generated groups is at most countable and the number of all torsion-free finitely generated groups is uncountable, we have

COROLLARY 1.3. *There exists an uncountable set of pairwise nonisomorphic torsion-free 2-generated groups with exactly 2 conjugacy classes.*

We note that the question of the existence of *any* finitely generated group with exactly two conjugacy classes other than $\mathbb{Z}/2\mathbb{Z}$ was open until now. It can be found, for example, in [17, Prob. 9.10] or in [3, Prob. FP20]. (A positive solution has been announced by Ivanov in 1989 [14], [15] but the complete proof has never been published.) Corollary 1.3 provides the first examples of such groups. Starting with the group $G = \mathbb{Z}/p^{n-2}\mathbb{Z} \times H$ for $n \geq 3$, where p is a prime number and H is a torsion-free group, we can generalize the previous result.

COROLLARY 1.4. *For any $n \in \mathbb{N}$, $n \geq 2$, there is an uncountable set of pairwise nonisomorphic finitely generated groups with exactly n conjugacy classes.*

For large enough prime numbers n, the first examples of finitely generated infinite periodic groups with exactly n conjugacy classes were constructed by Ivanov (see [21, Th. 41.2]) as limits of hyperbolic groups (although hyperbolicity was not used explicitly). Here we say that G is a *limit of hyperbolic groups* if there exists a finitely generated free group F and a series of normal subgroups $N_1 \lhd N_2 \lhd \ldots$ of F such that $G \cong F/N$ for $N = \bigcup\limits_{i=1}^{\infty} N_i$ and each of the groups F/N_i, $i = 1, 2, \ldots$ is hyperbolic. In contrast it is impossible to construct a finitely generated group other than $\mathbb{Z}/2\mathbb{Z}$ with exactly two conjugacy classes in this way.

Indeed suppose that a finitely generated group G has exactly two conjugacy classes. If G is not torsion-free, then G is a group of exponent p for some prime p as the orders of all nontrivial elements of G are equal. If $p = 2$, G is abelian and hence is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. In case $p > 2$, there exist nontrivial elements $g, t \in G$ such that

$$(1) \qquad\qquad\qquad t^{-1}gt = g^2.$$

The equality $g^{2^p - 1} = t^{-p}gt^p g^{-1} = gg^{-1} = 1$ implies $2^p - 1 \equiv 0 \pmod{p}$. However, by the Fermat Little Theorem, we have $2^p - 2 \equiv 0 \pmod{p}$, which

contradicts the previous equality. Assume now that G is torsion-free. If G is a limit of hyperbolic groups F/N_i, $i = 1, 2, \ldots$, then for some i large enough, there are elements $g, t \in F/N_i$ of infinite order that satisfy (1). This leads to a contradiction again since the equality of type (1) is impossible in a hyperbolic group if the order of g is infinite [10], [11].

Another theorem from [13] states that any countable group G can be embedded into a countable divisible group D. We recall that a group D is said to be *divisible* if for every element $d \in D$ and every positive integer n, the equation $x^n = d$ has a solution in D. A natural example of a divisible group is $\mathbb{Q}$. The question of the existence of a finitely generated divisible group was open during a long time. The first examples of such a type were constructed by Guba [12] (see also [21]).

Later Mikhajlovskii and Olshanskii [18] constructed a more general example of a finitely generated *verbally complete group*, that is a group W such that for every nontrivial freely reduced word $w(x_i)$ in the alphabet $x_1^{\pm 1}, x_2^{\pm 1}, \ldots$ and every $v \in W$, the equation $w(x_i) = v$ has a solution in W. That is, there are elements $v_1, v_2, \cdots \in W$ such that $w(v_i) = v$ in W, where $w(v_i)$ is the word obtained from $w(x_i)$ by substituting v_i for x_i, $i = 1, 2 \ldots$. Comparing these results one may ask whether any countable group can be embedded into a finitely generated divisible (or verbally complete) group. The next theorem provides the affirmative answer.

THEOREM 1.5. *Any countable group H can be embedded into a 2-generated verbally complete group W. Moreover, if H is torsion-free, then W can be chosen to be torsion-free.*

Note that the condition $\pi(G) = \pi(W)$ can not be ensured in Theorem 1.5. Indeed, it is easy to show that if a divisible group W contains a nontrivial element of finite order, then $\pi(W) = \mathbb{N}$. As above, we obtain

COROLLARY 1.6. *There exists an uncountable set of pairwise nonisomorphic 2-generated verbally complete groups.*

2. **Outline of the method**

In this section we give the proofs of Theorem 1.5 and Theorem 1.1 modulo technical results which are obtained in Sections 4–8. We assume the reader to be familiar with the notion of a relatively hyperbolic group and refer to the next section for precise definitions.

Let G be a group that is hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda}$. We divide the set of all elements of G into two subsets as follows. An element $g \in G$ is said to be *parabolic* if g is conjugate to an element of H_λ for some $\lambda \in \Lambda$. Otherwise g is said to be *hyperbolic*. Recall also that a group is *elementary* if it contains a cyclic subgroup of finite index. The following result concerning maximal elementary subgroups is proved in [26, Th. 4.3, Cor. 1.7].

THEOREM 2.5. *Suppose that a group G is hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda} \cup \{K\}$ and for some $\nu \in \Lambda$, there exists a monomorphism $\iota \colon K \to H_\nu$. Then the* HNN-*extension*

$$(2) \qquad \widetilde{G} = \langle G, t \mid t^{-1}kt = \iota(k), \; k \in K \rangle$$

is hyperbolic relative to $\{H_\lambda\}_{\lambda \in \Lambda}$.

Theorems 1.5 and 1.1 can be obtained in a uniform way from the following result.

THEOREM 2.6. *Suppose that R is a countable group such that for any elementary group E satisfying the condition $\pi(E) \subseteq \pi(R)$, there exists a subgroup of R isomorphic to E. Then there is an embedding of R into a 2-generated group $S = S(R)$ such that any element of S is conjugate to an element of R in S. In particular, $\pi(S) = \pi(R)$.*

Proof. The desired group S is constructed as an inductive limit of relatively hyperbolic groups as follows. Let us set

$$G(0) = R * F(x, y),$$

where $F(x, y)$ is the free group of rank 2 generated by x and y. We enumerate all elements of

$$R = \{1 = r_0, r_1, r_2, \dots\}$$

and

$$G(0) = \{1 = g_0, g_1, g_2, \dots\}.$$

Suppose that for some $i \geq 0$, the group $G(i)$ has already been constructed together with an epimorphism $\xi_i \colon G(0) \to G(i)$. We use the same notation for elements $x, y, r_0, r_1, \dots, g_0, g_1, \dots$ and their images under ξ_i in $G(i)$. Assume that $G(i)$ satisfies the following conditions. (It is straightforward to check these conditions for $G(0)$ and the identity map $\xi_0 \colon G(0) \to G(0)$.)

  (i) The restriction of ξ_i to the subgroup R is injective. In what follows we identify R with its image in $G(i)$.
 (ii) $G(i)$ is hyperbolic relative to R.
(iii) The elements x and y generate a suitable subgroup of $G(i)$.
(iv) All hyperbolic elements of $G(i)$ have infinite order. In particular, $\pi(G(i)) = \pi(R)$.
 (v) The elements $g_0, \dots, g_i$ are parabolic in $G(i)$.
(vi) In the group $G(i)$, the elements $r_0, \dots, r_i$ are contained in the subgroup generated by x and y.

The group $G(i + 1)$ is obtained from $G(i)$ in two steps.

Step 1. Let us take the element g_{i+1} and construct a group $G(i + 1/2)$ as follows. If g_{i+1} is a parabolic element of $G(i)$, we set $G(i + 1/2) = G(i)$. If g_{i+1} is hyperbolic, the order of g_{i+1} is infinite by (iv). Furthermore, since

$\pi(E_{G(i)}(g_{i+1})) \subseteq \pi(G(i)) = \pi(R)$, there is a monomorphism $\iota : E_{G(i)}(g_{i+1}) \to R$. Then we take the HNN-extension

$$G(i + 1/2) = \langle G(0), t \mid t^{-1}et = \iota(e), e \in E_{G(i)}(g_{i+1}) \rangle.$$

In both cases $G(i + 1/2)$ is hyperbolic relative to R. Indeed this is obvious in the first case and follows from the second assertion of Theorem 2.1 and Theorem 2.5 in the second one. Note also that all hyperbolic elements of $G(i + 1/2)$ have infinite order. (In the second case this immediately follows from the description of periodic elements in HNN-extensions [16, Ch. IV, Th. 2.4].)

Step 2. First we wish to show that the subgroup generated by x and y is suitable in $G(i + 1/2)$. This is obvious in case g_{i+1} is parabolic in $G(i)$, so we consider the second case only. Since $\langle x, y \rangle$ is suitable in $G(i)$ by (iii), Lemma 2.3 yields the existence of infinitely many pairwise noncommensurable (in $G(i)$) hyperbolic elements $h_j \in \langle x, y \rangle$ of infinite order, $j = 1, 2 \ldots$, such that $E_{G(i)}(h_j) = \langle h_j \rangle$. At most one of these elements is commensurable with g_{i+1} in $G(i)$. Therefore, there exist two noncommensurable in $G(i)$ hyperbolic elements of infinite order, say $h_1, h_2 \in \langle x, y \rangle$, such that h_j is not commensurable with g_{i+1} in $G(i)$ for $j = 1, 2$. In particular, h_j, $j = 1, 2$, is not conjugate to an element of $E_{G(i)}(g_{i+1})$ as $\langle g_{i+1} \rangle$ has finite index in $E_{G(i)}(g_{i+1})$. According to Britton's Lemma on HNN-extensions [16, Ch. 5, §2], this implies that h_1 and h_2 are hyperbolic and noncommensurable in $G(i + 1/2)$. Furthermore, if for some $j = 1, 2, n \in \mathbb{N}$, and $u \in G(i + 1/2)$, we have $u^{-1}h_j^n u = h_j^{\pm n}$, then $u \in G(i)$ by Britton's Lemma. Thus the explicit description of maximal elementary subgroups from the first assertion of Theorem 2.1 yields the equality $E_{G(i+1/2)}(h_j) = E_{G(i)}(h_j)$ for $j = 1, 2$. Finally since $E_{G(i)}(h_j) = \langle h_j \rangle$ and h_1, h_2 are noncommensurable, we have

$$E_{G(i+1/2)}(h_1) \cap E_{G(i+1/2)}(h_2) = E_{G(i)}(h_1) \cap E_{G(i)}(h_2) = \langle h_1 \rangle \cap \langle h_2 \rangle = \{1\}.$$

By Definition 2.2 this means that the subgroup generated by x and y is suitable in $G(i + 1/2)$.

We now apply Theorem 2.4 to the group $G = G(i + 1/2)$, the subgroup $H = \langle x, y \rangle \le G(i + 1/2)$, and the set of elements $\{t, r_{i+1}\}$. Let $G(i + 1) = \bar{G}$, where $\bar{G}$ is the quotient group provided by Theorem 2.4. Since t becomes an element of $\langle x, y \rangle$ in $G(i + 1)$, there is a naturally defined epimorphism $\xi_{i+1} : G(0) \to G(i + 1)$. Using Theorem 2.4 it is straightforward to check properties (i)–(vi) for $G(i + 1)$. This completes the inductive step.

Let N_i denote the kernel of ξ_i. Observe that $N_1, N_2, \ldots$ form an increasing normal series and set $S = G(0)/N$, where $N = \bigcup_{i=1}^{\infty} N_i$. By (i) the subgroup R is embedded into S. Further it is easy to see that S is 2-generated. Indeed, $G(0)$ is generated by $x, y, r_1, r_2, \ldots$. Condition (vi) yields $r_i \in \langle x, y \rangle$ in S for any $i \in \mathbb{N}$. Thus S is generated by x and y. Finally let s be an element of S. We take an arbitrary preimage $g \in G(0)$ of s. Then the image of the element g becomes

parabolic at a certain step according to (v). Thus s is conjugate to an element of R in S. The theorem is proved. □

It remains to derive Theorems 1.1 and 1.5.

Proof of Theorem 1.1. Let $\mathscr{E}$ denote the free product of all elementary groups E (taken up to isomorphism) such that $\pi(E) \subseteq \pi(G)$. We set $G^* = G * \mathscr{E}$. By a theorem from [13], we can embed G^* into an (infinitely generated) group R such that all elements of the same order are conjugate in R and

$$(3) \qquad \pi(R) = \pi(G^*) = \pi(G).$$

We now apply Theorem 2.6 and embed the group R into a 2-generated group $C = S(R)$ such that any element of C is conjugate to an element of R. As all elements of the same order are conjugate in R, this is so in C. The equality $\pi(C) = \pi(G)$ follows from (3) as $\pi(C) = \pi(R)$ by Theorem 2.6. □

Proof of Theorem 1.5. First note that any countable group G can be embedded into an infinitely generated countable verbally complete group R in the following way. (The idea comes from the proof of the Higman-Neumann-Neumann theorem on embeddings into divisible groups.) We denote by $F = F(a_1, a_2, \dots)$ the free group with basis $a_1, a_2, \dots$. Let us enumerate the set of all pairs

$$\{p_1, p_2, \dots\} = \{(v, g) \; : \; v \in F \setminus \{1\}, \, g \in G \setminus \{1\}\}.$$

Starting with the group G we first set $G^* = G$ if G is torsion-free, and $G^* = G * E_1 * E_2 * \dots$, where $\{E_1, E_2, \dots\}$ is the set of all elementary groups (up to isomorphism), otherwise. Further we construct a sequence of groups $G^* = U_0 \leq U_1 \leq \dots$ as follows. Suppose that for some $i \geq 0$, the group U_i has already been constructed and take $p_{i+1} = (v, g)$. There are two possibilities to consider.

1) The element g has infinite order. Then we define U_{i+1} to be the free product of U_i and F with the amalgamated subgroups $\langle g \rangle$ and $\langle v \rangle$.

2) The order of g is $n < \infty$. It is well-known [16, Ch. 4, Th. 5.2] that the order of the element v in the group $H = \langle a_1, a_2, \dots \mid v^n = 1 \rangle$ equals n. Thus the free product of U_i and H with amalgamated subgroups $\langle g \rangle$ and $\langle v \rangle$ is well-defined. We set $U_{i+1} = U_i *_{\langle g \rangle = \langle v \rangle} H$.

Now let $U(G^*) = \bigcup_{i=0}^{\infty} U_i$. Obviously G^* embeds in $U(G^*)$, $U(G^*)$ is countable and torsion-free whenever G^* is torsion-free, and any equation of type $w(x_i) = g$, where $w(x_i)$ is a word in the alphabet $x_1^{\pm 1}, x_2^{\pm 1}, \dots$ and $g \in G$, has a solution in $U(G^*)$. Finally we consider the sequence of groups $R_1 \leq R_2 \leq \dots$, where $R_1 = U(G^*)$ and $R_{i+1} = U(R_i)$, $i = 1, 2 \dots$. Clearly the group $R = \bigcup_{i=0}^{\infty} R_i$ is countable, verbally complete, torsion-free whenever G is torsion-free, and contains

a copy of every elementary group E such that $\pi(E) \subseteq \pi(G)$. Let $W = S(R)$ be the group provided by Theorem 2.6.

Consider an equations $w(x_i) = v$ for some $v \in W$. By Theorem 2.6, there is an element $t \in W$ such that $t^{-1}vt \in R$. Since R is verbally complete, there is a solution $x_1 = r_1$, $x_2 = r_2$, $\ldots$ to the equation $w(x_i) = t^{-1}vt$ in R. Clearly $x_1 = tr_1t^{-1}$, $x_2 = tr_2t^{-1}$, $\ldots$ is a solution to the equation $w(x_i) = v$. □

3. **Preliminaries**

Some conventions and notation. We write $W \equiv V$ to express the letter-for-letter equality of words W and V in some alphabet. If a word W decomposes as $W \equiv V_1 U V_2$, we call V_1 (respectively, V_2) a *prefix* (respectively, *suffix*) of W. For elements g, t of a group G, g^t denotes the element $t^{-1}gt$. Recall that a subset X of a group G is said to be *symmetric* if for any $x \in X$, we have $x^{-1} \in X$. In this paper all generating sets of groups under consideration are supposed to be symmetric.

All paths considered in this paper are combinatorial paths. Recall that a *combinatorial path* p in a CW-complex is a sequence of edges (i.e., 1-dimensional cells) $e_1 e_2 \ldots e_k$, where $(e_i)_+ = (e_{i+1})_-$. If edges of the complex are labeled, we define *the label of* p by $\mathbf{Lab}\,(p) \equiv \mathbf{Lab}\,(e_1)\mathbf{Lab}\,(e_2)\ldots\mathbf{Lab}\,(e_k)$, where $\mathbf{Lab}\,(e_i)$ is the label of e_i. We also denote by $p_- = (e_1)_-$ and $p_+ = (e_k)_+$ the origin and the terminus of p respectively. The length $l(p)$ of p is the number of edges of p.

Word metrics and Cayley graphs. Let G be a group generated by a (symmetric) set $\mathscr{A}$. Recall that the *Cayley graph* $\Gamma(G, \mathscr{A})$ of a group G with respect to the set of generators $\mathscr{A}$ is an oriented labeled 1-complex with the vertex set $V(\Gamma(G, \mathscr{A})) = G$ and the edge set $E(\Gamma(G, \mathscr{A})) = G \times \mathscr{A}$. An edge $e = (g, a)$ goes from the vertex g to the vertex ga and has label $\mathbf{Lab}\,(e) \equiv a$. As usual, we denote the origin and the terminus of the edge e, i.e., the vertices g and ga, by e_- and e_+ respectively.

Associated to $\mathscr{A}$ is the so-called *word metric* on G. More precisely, the length $|g|_{\mathscr{A}}$ of an element $g \in G$ is defined to be the length of a shortest word in $\mathscr{A}$ representing g in G. By abuse of notation, we also write $|W|_{\mathscr{A}}$ to denote the lengths of the element of G represented by a word W in the alphabet $\mathscr{A}$. This is to be distinguished from the length of the word W itself, which is denoted by $\|W\|$.

The word metric on G is defined by $\mathrm{dist}_{\mathscr{A}}(f, g) = |f^{-1}g|_{\mathscr{A}}$. We also denote by $\mathrm{dist}_{\mathscr{A}}$ the natural extension of the word metric to the Cayley graph $\Gamma(G, \mathscr{A})$.

van Kampen diagrams. Recall that a *van Kampen diagram* Δ over a presentation

$$(4) \qquad\qquad\qquad G = \langle \mathscr{A} \mid \mathbb{O} \rangle$$

is a finite oriented connected 2-complex endowed with a labeling function $\mathbf{Lab} :$ $E(\Delta) \to \mathscr{A}$, where $E(\Delta)$ denotes the set of oriented edges of Δ, such that $\mathbf{Lab}\,(e^{-1})$ $\equiv (\mathbf{Lab}\,(e))^{-1}$. Labels and lengths of paths are defined as in the case of Cayley
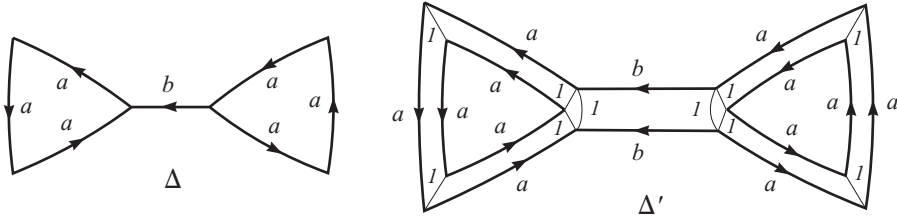
Figure 1. Δ' is a 0-refinement of the diagram Δ over the presentation $\langle a, b \mid a^3 = 1 \rangle$.

graphs. Given a cell Π of Δ, we denote by $\partial \Pi$ the boundary of Π; similarly, $\partial \Delta$ denotes the boundary of Δ. The labels of $\partial \Pi$ and $\partial \Delta$ are defined up to a cyclic permutation. An additional requirement is that for any cell Π of Δ, the boundary label $\mathbf{Lab}\,(\partial \Pi)$ is equal to a cyclic permutation of a word $P^{\pm 1}$, where $P \in \mathbb{O}$.

Sometimes it is convenient to use the notion of the so-called 0-*refinement*, which enables us to assume that all diagrams are homeomorphic to a disc. Roughly speaking, making a 0-refinement of a diagram Δ just means replacing every edge $e \in E(\Delta)$ with a rectangle labeled $\mathbf{Lab}\,(e) 1 \mathbf{Lab}\,(e)^{-1} 1$, and replacing every vertex of Δ with a polygon (2-cell) labeled by $11 \ldots 1$. Here 1 means the empty word. The rectangles are then attached to the polygons along edges labeled by 1 (see Figure 1). The lengths of edges labeled by 1 is supposed to be 0. This notion is quite standard and we refer the reader to [21, Ch. 4] for details.

The van Kampen Lemma states that a word W over the alphabet $\mathcal{A}$ represents the identity in the group given by (4) if and only if there exists a simply-connected planar diagram Δ over (4) such that $\mathbf{Lab}\,(\partial \Delta) \equiv W$ [16, Ch. 5, Th. 1.1].

For every van Kampen diagram Δ over (4) and any fixed vertex o of Δ, there is a (unique) combinatorial map $\gamma \colon Sk^{(1)}(\Delta) \to \Gamma(G, \mathcal{A})$ that preserves labels and orientation of edges and maps o to the vertex 1 of $\Gamma(G, \mathcal{A})$.

Hyperbolic spaces. Here we briefly discuss some properties of hyperbolic spaces used in this paper. For more details we refer to [6], [10], [11].

One says that a metric space M is δ-*hyperbolic* for some $\delta \geq 0$ (or simply *hyperbolic*) if for any geodesic triangle T in M, any side of T belongs to the union of the closed δ-neighborhoods of the other two sides.

Recall that a path p in a metric space is called (λ, c)-*quasi-geodesic* for some $\lambda > 0$, $c \geq 0$, if $\mathrm{dist}(q_-, q_+) \geq \lambda l(q) - c$ for any subpath q of p. Let p be a path in a van Kampen diagram Δ over (4). We need the following result about quasi-geodesics in hyperbolic spaces (see, for example, [6, Ch. III. H, Th. 1.7]).

LEMMA 3.1. *For any $\delta \geq 0$, $\lambda > 0$, $c \geq 0$, there exists a constant $\kappa = \kappa(\delta, \lambda, c)$ with the following property. If M is a δ-hyperbolic space and p, q are (λ, c)-quasi-geodesic paths in M with same endpoints, then p and q belong to the closed κ-neighborhoods of each other.*

The next result can be easily derived from the definition of a hyperbolic space by drawing the diagonal.

LEMMA 3.2. *Let M be a δ-hyperbolic metric space, Q a geodesic quadrangle in M. Then each side of Q belongs to the closed 2δ-neighborhood of the other three sides.*

From Lemma 3.1 and Lemma 3.2, we immediately obtain

COROLLARY 3.3. *For any $\delta \geq 0$, $\lambda > 0$, $c \geq 0$, there exists a constant $K = K(\delta, \lambda, c)$ with the following property: Let Q be a quadrangle in a δ-hyperbolic space whose sides are (λ, c)-quasi-geodesic. Then each side of Q belongs to the closed K-neighborhood of the union of the other three sides.*

Indeed it suffices to set $K = 2(\kappa + \delta)$, where κ is provided by Lemma 3.1.

The proof of the next lemma is also straightforward (see [23, Lemma 1.7]).

LEMMA 3.4. *Let Q be a geodesic quadrangle with sides a, b, c, d in a δ-hyperbolic space. Suppose that $l(a) \geq 4 \max\{l(b), l(d)\}$. Then there exist subsegments p, q of the sides a and b, respectively, such that $\min\{l(p), l(q)\} \geq \frac{7}{20}l(a) - 8\delta$ and $\operatorname{dist}(p_\pm, q_\pm) \leq 8\delta$.*

Passing from geodesic to quasi-geodesic quadrangles one can easily obtain the following. The proof is straightforward and consists of replacing the quasi-geometric quadrangle with a geodesic one having the same vertices, application of Lemma 3.4, and then Lemma 3.1. We leave details to the reader.

COROLLARY 3.5. *Let Q be a (λ, c)-quasi-geodesic quadrangle with sides a, b, c, d in a δ-hyperbolic space, $\kappa = \kappa(\lambda, c)$ the constant provided by Lemma 3.1. Suppose that $l(a) \geq (4 \max\{l(b), l(d)\} + c)/\lambda$. Then there exist subsegments p, q of the sides a and b, respectively, such that $\min\{l(p), l(q)\} \geq \frac{7}{20}(\lambda l(a) - c) - 8\delta - 2\kappa$ and $\operatorname{dist}(p_\pm, q_\pm) \leq 8\delta + 2\kappa$.*

The following lemma is also well-known (see, for example, [6, Ch. III. H, Th. 1.13]). Recall that a path in a metric space is said to be k-local geodesic if any of its subpaths of length at most k is geodesic.

LEMMA 3.6. *Let r be a k-local geodesic in a δ-hyperbolic metric space for some $k > 8\delta$. Then r is $(1/3, 2\delta)$-quasi-geodesic.*

The next lemma can be found in [22, Lemma 25].

LEMMA 3.7. *There are positive constants $c_1 = c_1(\delta)$ and $c_2 = c_2(\delta)$ such that for any geodesic r-gon P in a δ-hyperbolic space, the following holds: Suppose that the set of sides of P is divided into three subsets N_1, N_2, N_3 and σ_i is the sum of lengths of sides from N_i, $i = 1, 2, 3$. Assume that $\sigma_1 > ar$ and $\sigma_3 < 10^{-3}ar$ for some $a \geq c_2$. Then there exist different sides $p_1 \in N_1$ and $p_2 \in N_1 \cup N_2$ and subsegments q_j of p_j, $j = 1, 2$, of lengths at least $10^{-3}a$ such that*

$$\max\{\operatorname{dist}((q_1)_-, (q_2)_-), \operatorname{dist}((q_1)_+, (q_2)_+)\} \leq c_1.$$

Relatively hyperbolic groups. There are many equivalent definitions of relatively hyperbolic groups (see [5], [8], [28] and references therein). In this paper we use the isoperimetric characterization given in [28].

More precisely, let G be a group, $\{H_\lambda\}_{\lambda \in \Lambda}$ a collection of subgroups of G, X a subset of G. We say that X is a *relative generating set of G with respect to* $\{H_\lambda\}_{\lambda \in \Lambda}$ if G is generated by X together with the union of all H_λ. (In what follows, we assume X to be symmetric.) In this situation the group G can be regarded as a quotient group of the free product

$$(5) \qquad F = (*_{\lambda \in \Lambda} H_\lambda) * F(X),$$

where $F(X)$ is the free group with the basis X. Let N denote the kernel of the natural homomorphism $F \to G$. If N is the normal closure of a subset $\mathscr{Q} \subseteq N$ in the group F, we say that G has *relative presentation*

$$(6) \qquad \langle X, H_\lambda, \lambda \in \Lambda \mid \mathscr{Q} \rangle.$$

If $\sharp X < \infty$ and $\sharp \mathscr{Q} < \infty$, the relative presentation (6) is said to be *finite* and the group G is said to be *finitely presented relative to the collection of subgroups* $\{H_\lambda\}_{\lambda \in \Lambda}$.

Set

$$(7) \qquad \mathscr{H} = \bigsqcup_{\lambda \in \Lambda} (H_\lambda \setminus \{1\}).$$

Given a word W in the alphabet $X \cup \mathscr{H}$ such that W represents 1 in G, there exists an expression

$$(8) \qquad W =_F \prod_{i=1}^{k} f_i^{-1} Q_i^{\pm 1} f_i$$

with the equality in the group F, where $Q_i \in \mathscr{Q}$ and $f_i \in F$ for $i = 1, \ldots, k$. The smallest possible number k in a representation of the form (8) is called the *relative area* of W and is denoted by $\mathrm{Area}^{\mathrm{rel}}(W)$.

Definition 3.8. A group G is *hyperbolic relative to a collection of subgroups* $\{H_\lambda\}_{\lambda \in \Lambda}$ if G is finitely presented relative to $\{H_\lambda\}_{\lambda \in \Lambda}$ and there is a constant $L > 0$ such that for any word W in $X \cup \mathscr{H}$ representing the identity in G, we have $\mathrm{Area}^{\mathrm{rel}}(W) \leq L\|W\|$.

In particular, G is an ordinary *hyperbolic group* if G is hyperbolic relative to the trivial subgroup. An equivalent definition says that G is hyperbolic if it is generated by a finite set X and the Cayley graph $\Gamma(G, X)$ is hyperbolic. In the relative case these approaches are not equivalent, but we still have the following [28, Th. 1.7].

LEMMA 3.9. *Suppose that G is a group hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda}$. Let X be a finite relative generating set of G with respect to*

$\{H_\lambda\}_{\lambda \in \Lambda}$. *Then the Cayley graph* $\Gamma(G, X \cup \mathcal{H})$ *of* G *with respect to the generating set* $X \cup \mathcal{H}$ *is a hyperbolic metric space.*

Observe also that the relative area of a word W representing 1 in G can be defined geometrically via van Kampen diagrams. Let G be a group given by the relative presentation (6) with respect to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda}$. We denote by $\mathcal{S}$ the set of all words in the alphabet $\mathcal{H}$ representing the identity in the groups F defined by (5). Then G has the ordinary (nonrelative) presentation

$$(9) \qquad\qquad G = \langle X \cup \mathcal{H} \mid \mathcal{S} \cup \mathcal{Q} \rangle.$$

A cell in van Kampen diagram Δ over (9) is called a $\mathcal{Q}$-*cell* if its boundary is labeled by a word from $\mathcal{Q}$. We denote by $N_{\mathcal{Q}}(\Delta)$ the number of $\mathcal{Q}$-cells of Δ. Obviously given a word W in $X \cup \mathcal{H}$ that represents 1 in G, we have

$$\mathrm{Area}^{\mathrm{rel}}(W) = \min_{\mathbf{Lab}\,(\partial \Delta) \equiv W} \{N_{\mathcal{Q}}(\Delta)\},$$

where the minimum is taken over all van Kampen diagrams with boundary label W.

H_λ-components. Finally we are going to recall an auxiliary terminology introduced in [28], which plays an important role in our paper. Let G be a group, $\{H_\lambda\}_{\lambda \in \Lambda}$ a collection of subgroups of G, X a finite generating set of G with respect to $\{H_\lambda\}_{\lambda \in \Lambda}$, q a path in the Cayley graph $\Gamma(G, X \cup \mathcal{H})$. A subpath p of q is called an H_λ-*component* for some $\lambda \in \Lambda$ (or simply a *component*) of q, if the label of p is a word in the alphabet $H_\lambda \setminus \{1\}$ and p is not contained in a bigger subpath of q with this property. Two H_λ-components p_1, p_2 of a path q in $\Gamma(G, X \cup \mathcal{H})$ are called *connected* if there exists a path c in $\Gamma(G, X \cup \mathcal{H})$ that connects some vertex of p_1 to some vertex of p_2 and $\mathbf{Lab}\,(c)$ is a word consisting of letters from $H_\lambda \setminus \{1\}$. In algebraic terms this means that all vertices of p_1 and p_2 belong to the same coset gH_λ for a certain $g \in G$. Note that we can always assume c to have length at most 1, as every nontrivial element of H_λ is included in the set of generators. An H_λ-component p of a path q is called *isolated* if no distinct H_λ-component of q is connected to p.

The lemma below is a simplification of [28, Lemma 2.27].

LEMMA 3.10. *Suppose that* G *is a group that is hyperbolic relative to a collection of subgroups* $\{H_\lambda\}_{\lambda \in \Lambda}$, X *a finite generating set of* G *with respect to* $\{H_\lambda\}_{\lambda \in \Lambda}$. *Then there exists a constant* $K > 0$ *and a finite subset* $\Omega \subseteq G$ *such that the following condition holds. Let* q *be a cycle in* $\Gamma(G, X \cup \mathcal{H})$, $p_1, \ldots, p_k$ *a set of isolated* H_λ-*components of* q *for some* $\lambda \in \Lambda$, $g_1, \ldots, g_k$ *the elements of* G *represented by the labels of* $p_1, \ldots, p_k$ *respectively. Then for any* $i = 1, \ldots, k$, g_i *belongs to the subgroup* $\langle \Omega \rangle \leq G$ *and the lengths of* g_i *with respect to* Ω *satisfy the inequality*

$$\sum_{i=1}^{k} |g_i|_\Omega \leq Kl(q).$$

4. Small cancellation conditions

The main aim of this and the following four sections is to generalize the small cancellation theory over hyperbolic groups developed by Olshanskii in [23] to relatively hyperbolic groups. The fact that the Cayley graph $\Gamma(G, X)$ of a hyperbolic group G generated by a finite set X is a hyperbolic metric space plays the key role in [23]. Lemma 3.9 allows to extend this theory to the case of relatively hyperbolic groups. However this extension is not straightforward as the Cayley graph $\Gamma(G, X \cup \mathcal{H})$ defined in the previous section is not necessary locally finite.

Roughly speaking, one can divide results about small cancellation conditions from [23] into three classes. The first class consists of results about diagrams over presentations satisfying small cancellation conditions, which do not use local finiteness of Cayley graphs at all. They can be stated and proved in our settings without any essential changes. The main result of this kind is Lemma 4.4 stated below.

Proofs of results from the second class do not use local finiteness of the Cayley graph either, but they do employ certain facts concerning geometric and algebraic properties of ordinary hyperbolic groups. These results can also be reproved with minor changes modulo the paper [28], where the corresponding facts about relatively hyperbolic groups are contained. For convenience of the reader we provide self-contained proofs in Sections 5 and 6.

Finally results of the third type explain how to choose words of some specific form satisfying small cancellation conditions. Unlike in ordinary small cancellation theory over a free group, verifying small cancellation conditions over hyperbolic groups is much harder and local finiteness of Cayley graphs is essentially used in [23]. Our approach here is different and is explained in Section 7.

Given a set of words $\mathcal{R}$ in an alphabet $\mathcal{A}$, we say that $\mathcal{R}$ is *symmetrized* if for any $R \in \mathcal{R}$, $\mathcal{R}$ contains all cyclic shifts of $R^{\pm 1}$. Further let G be a group generated by a set $\mathcal{A}$. We say that a word R is (λ, c)-*quasi-geodesic in G*, if any path in the Cayley graph $\Gamma(G, \mathcal{A})$ labeled R is (λ, c)-quasi-geodesic.

Definition 4.1. Let G be a group generated by a set $\mathcal{A}$, $\mathcal{R}$ a symmetrized set of words in $\mathcal{A}$. For $\varepsilon > 0$, a subword U of a word $R \in \mathcal{R}$ is called an ε-*piece* if there exists a word $R' \in \mathcal{R}$ such that:

(1) $R \equiv UV$, $R' \equiv U'V'$, for some V, U', V';

(2) $U' = YUZ$ in G for some words Y, Z in $\mathcal{A}$ such that $\max\{\|Y\|, \|Z\|\} \leq \varepsilon$;

(3) $YRY^{-1} \neq R'$ in the group G.

Similarly, a subword U of $R \in \mathcal{R}$ is called *an ε'-piece* if:

(1') $R \equiv UVU'V'$ for some V, U', V';

(2') $U' = YU^{\pm 1}Z$ in the group G for some Y, Z satisfying $\max\{\|Y\|, \|Z\|\} \leq \varepsilon$.
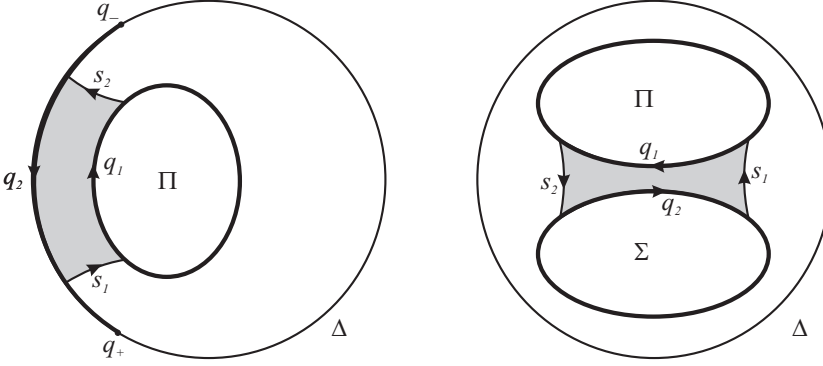
Figure 2. Contiguity subdiagrams.

Recall that a word R in $\mathscr{A}$ is said to be (λ, c)-*quasi-geodesic*, if some (equivalently, any) path labeled by R in $\Gamma(G, \mathscr{A})$ is (λ, c)-quasi-geodesic.

Definition 4.2. We say that the set $\mathscr{R}$ satisfies the $C(\varepsilon, \mu, \lambda, c, \rho)$-*condition* for some $\varepsilon \geq 0$, $\mu > 0$, $\lambda > 0$, $c \geq 0$, $\rho > 0$, if

(1) $\|R\| \geq \rho$ for any $R \in \mathscr{R}$;

(2) any word $R \in \mathscr{R}$ is (λ, c)-quasi-geodesic;

(3) for any ε-piece of any word $R \in \mathscr{R}$, the inequality $\max\{\|U\|, \|U'\|\} < \mu\|R\|$ holds (using the notation of Definition 4.1).

Further the set $\mathscr{R}$ satisfies the $C_1(\varepsilon, \mu, \lambda, c, \rho)$-*condition* if in addition the condition (3) holds for any ε'-piece of any word $R \in \mathscr{R}$.

Suppose that G is a group defined by (4). Given a set of words $\mathscr{R}$, we consider the quotient group of G represented by

$$(10) \qquad\qquad G_1 = \langle \mathscr{A} \mid \mathbb{O} \cup \mathscr{R} \rangle.$$

A cell in a van Kampen diagram over (10) is called an $\mathscr{R}$-*cell* if its boundary label is a word from $\mathscr{R}$.

Let Δ be a van Kampen diagram over (10), q a subpath of $\partial\Delta$. Let also Π, Σ be $\mathscr{R}$-cells of Δ. Suppose that there is a simple closed paths

$$(11) \qquad\qquad p = s_1 q_1 s_2 q_2$$

in Δ, where q_1 is a subpath of $\partial\Pi$, q_2 is a subpath of q (or $\partial\Sigma$), and

$$(12) \qquad\qquad \max\{l(s_1), l(s_2)\} \leq \varepsilon$$

for some constant $\varepsilon > 0$. By Γ we denote the subdiagram of Δ bounded by p (see Figure 2). If Γ contains no $\mathscr{R}$-cells, we say that Γ is an ε-*contiguity subdiagram* (or simply a contiguity subdiagram if ε is fixed) of Π to the subpath q of $\partial\Delta$ (or to Σ, respectively) and q_1 is the *contiguity arc* of Π to q (respectively, Σ). The

ratio $l(q_1)/l(\partial\Pi)$ is called the *contiguity degree* of Π to q (respectively, Σ) and is denoted by (Π, Γ, q) (respectively, (Π, Γ, Σ)). In case $q = \partial\Delta$, we talk about contiguity subdiagrams, etc., of Π to $\partial\Delta$.

A van Kampen diagram Δ over (10) is said to be *reduced* if Δ has minimal number of $\mathcal{R}$-cells among all diagrams over (10) having the same boundary label.

The lemma below may be seen as a geometric reformulation of the small cancellation condition (see [23, Lemma 5.2]). We provide the proof here for the sake of completeness.

LEMMA 4.3. *Suppose that $\mathcal{R}$ satisfies the $C(\varepsilon, \mu, \lambda, c, \rho)$-condition for some values of the parameters. Let Δ be a reduced diagram over (10). Then for every ε-contiguity subdiagram Γ of a cell Π to another cell Σ, we have*

$$(13) \qquad\qquad \max\{(\Pi, \Gamma, \Sigma), (\Sigma, \Gamma, \Pi)\} < \mu.$$

Proof. Let $\partial\Gamma = s_1 q_1 s_2 q_2$ as above (see Figure 2). Let $\partial\Pi = q_1 a$ and $\partial\Sigma = q_2 b$. Then the word $U \equiv \mathbf{Lab}(q_1)$ is an ε-piece. Indeed the first two conditions in Definition 4.2 are satisfied. If the third condition is not, then

$$\mathbf{Lab}(s_1)\mathbf{Lab}(\partial\Pi)\mathbf{Lab}(s_1)^{-1} = \mathbf{Lab}(\partial\Sigma).$$

That is, $\mathbf{Lab}(s_1 q_1 a s_1^{-1} b q_2) = 1$ in G. Hence we can cut the subdiagram of Δ bounded by $s_1 q_1 a s_1^{-1} b q_2$ and fill the obtained hole with a diagram over (4) without $\mathcal{R}$-cells reducing the number of $\mathcal{R}$-cells by 2. This contradicts the assumption that Δ is reduced. Thus U is an ε-piece, and (13) follows from the $C(\varepsilon, \mu, \lambda, c, \rho)$-condition. $\qquad\square$

When dealing with a diagram Δ over (10), it is convenient to consider the following transformations. Let Σ be a subdiagram of Δ which contains no $\mathcal{R}$-cells, Σ' another diagram over (4) with $\mathbf{Lab}(\partial\Sigma) = \mathbf{Lab}(\partial\Sigma')$. Then we can cut off Σ and fill the obtained hole with Σ'. Note that this transformation does not affect $\mathbf{Lab}(\partial\Delta)$ and the number of $\mathcal{R}$-cells in Δ. If two diagrams over (10) can be obtained from each other by a sequence of such transformations, we call them $\mathbb{0}$-*equivalent*.

The following is an analogue of the well-known Greendlinger lemma. Recall that a path p in Δ is called (λ, c)-*quasi-geodesic*, if its label is (λ, c)-quasi-geodesic, i.e., some (equivalently, any) path in $\Gamma(G, \mathscr{A})$ with the same label is (λ, c)-quasi-geodesic.

LEMMA 4.4. *Let G be a group with a presentation $G = \langle \mathscr{A} \,|\, \mathbb{0} \rangle$. Suppose that the Cayley graph $\Gamma(G, \mathscr{A})$ of G is hyperbolic. Then for any $\lambda \in (0, 1]$, $c \geq 0$, and $\mu \in (0, 1/16]$, there exist $\varepsilon \geq 0$ and $\rho > 0$ with the following property. Let $\mathcal{R}$ be a symmetrized set of words in $\mathscr{A}$ satisfying the $C(\varepsilon, \mu, \lambda, c, \rho)$-condition, Δ a reduced van Kampen diagram over the presentation (10) whose boundary is (λ, c)-quasi-geodesic. Assume that Δ has at least one $\mathcal{R}$-cell. Then there exists a diagram*

Δ' which is $\mathbb{O}$-equivalent to Δ, an $\mathscr{R}$-cell Π in Δ', and an ε-contiguity subdiagram Γ of Π to $\partial\Delta'$ such that

$$(\Pi, \Gamma, \partial\Delta') > 1 - 13\mu.$$

The proof of this lemma is actually given in [23, Lemma 6.6] (see also the addendum in [24]) under the additional assumption that $\mathscr{A}$ is finite (i.e., the group G is hyperbolic). However the finiteness of $\mathscr{A}$ was not used in the proof, so the same proof works in our case without any changes. Following the referee's recommendation, we provide a self-contained proof of Lemma 4.4 in Appendix.

5. **Relative hyperbolicity of the quotient**

Our next goal is to show that adding relations satisfying sufficiently strong small cancellation conditions preserves relative hyperbolicity. Throughout the rest of the paper (except the appendix), let G be a group hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda\in\Lambda}$, X a finite relative generating set of G with respect to $\{H_\lambda\}_{\lambda\in\Lambda}$. We set $\mathscr{A} = X \cup \mathscr{H}$ and $\mathbb{O} = \mathscr{S} \cup \mathscr{Q}$, where $\mathscr{S}$ and $\mathscr{Q}$ are defined as in (9). In the proof of the lemma below we follow the idea from [23, Lemma 6.7] with little changes.

LEMMA 5.1. *For any* $\lambda \in (0,1]$, $c \geq 0$, $N > 0$, *there exist* $\mu > 0$, $\varepsilon \geq 0$, *and* $\rho > 0$ *such that for any finite symmetrized set of words* $\mathscr{R}$ *satisfying the* $C(\varepsilon, \mu, \lambda, c, \rho)$-*condition, the following hold*:

1. *The group* G_1 *defined by* (10) *is hyperbolic relative to the collection of images of subgroups* $H_\lambda, \lambda \in \Lambda$, *under the natural homomorphism* $G \to G_1$. *In particular, the Cayley graph of* G_1 *with respect to the generating set* $\mathscr{A}$ *is hyperbolic*.

2. *The restriction of the natural homomorphism* $\gamma\colon G \to G_1$ *to the subset of elements of length at most* N *with respect to the generating set* $\mathscr{A}$ *is injective*.

Proof. Let us fix $\lambda, c > 0$ and sufficiently small $\mu < 1/16$. The exact value of μ required can be easily extracted from the proof. In what follows, we assume that μ is small enough comparably to λ. Further we choose the constants ε and ρ according to Lemma 4.4. Since the $C(\varepsilon, \mu, \lambda, c, \rho)$-condition becomes stronger as ρ increases, we may increase ρ if necessary without violating the conclusion of Lemma 4.4. The exact lower bound can be easily extracted from our proof.

For a word W in the alphabet $X \cup \mathscr{H}$ that represents 1 in G_1, we denote by $\mathrm{Area}_1^{\mathrm{rel}}(W)$ its relative area, that is the minimal number k in a representation of type

$$W =_F \prod_{i=1}^{k} f_i R_i^{\pm 1} f_i^{-1}$$

with the equality in the group F defined by (5), where $R_i \in \mathscr{R} \cup \mathscr{Q}$. Similarly, by $\mathrm{Area}^{\mathrm{rel}}(W)$ we denote the relative area of a word W representing 1 in G, i.e., the minimal number k in the above decomposition, where $R_i \in \mathscr{R}$.

As G is hyperbolic relative to $\{H_\lambda\}_{\lambda \in \Lambda}$, there exists a constant $L > 0$ such that for any word W in $X \cup \mathcal{H}$ representing 1 in G, we have $\mathrm{Area}^{\mathrm{rel}}(W) \leq L\|W\|$. To prove the lemma it suffices to show that there is a constant $\alpha > 0$ such that for any word W representing 1 in G_1, $\mathrm{Area}_1^{\mathrm{rel}}(W) \leq \alpha\|W\|$. We are going to prove this inequality for

$$\alpha = \frac{3L}{\lambda - 27\mu}.$$

We proceed by induction on the length of W. Let p be a path labeled by W in $\Gamma(G, X \cup \mathcal{H})$. Suppose first that p is not $(1/2, 0)$-quasi-geodesic. Passing to a cyclic shift of W if necessary, we may assume that $p = p_0 p_1$, where p_0 is a subpath of p such that $\mathrm{dist}_{X \cup \mathcal{H}}((p_0)_-, (p_0)_+) < l(p_0)/2$. Let q be a geodesic path in $\Gamma(G, X \cup \mathcal{H})$ that goes from $(p_0)_-$ to $(p_0)_+$. Thus $l(q) < l(p_0)/2$. We denote by U_0, U_1, and V the labels of p_0, p_1, and q respectively. Then

$$W =_F (U_0 V^{-1})(V U_1),$$

where $U_0 V^{-1}$ represents 1 in G and $V U_1$ represents 1 in G_1. Obviously we have

$$\|V U_1\| \leq l(p) - l(p_0) + l(q) < \|W\| - \frac{l(p_0)}{2}$$

and

$$\|U_0 V^{-1}\| \leq l(p_0) + l(q) < \frac{3l(p_0)}{2}.$$

Using the inductive hypothesis, we obtain

$$\mathrm{Area}_1^{\mathrm{rel}}(W) \leq \mathrm{Area}_1^{\mathrm{rel}}(V U_1) + \mathrm{Area}^{\mathrm{rel}}(U_0 V^{-1})$$
$$< \alpha \left(\|W\| - \frac{1}{2} l(p_0) \right) + \frac{3}{2} L l(p_0) < \alpha \|W\|$$

as $\alpha > 3L$.

Now suppose that p is a $(1/2, 0)$-quasi-geodesic path in $\Gamma(G, X \cup \mathcal{H})$. Since $C(\varepsilon, \mu, \lambda, c, \rho)$-condition implies $C(\varepsilon, \mu, 1/2, c, \rho)$ whenever $\lambda > 1/2$, it suffices to prove the lemma for $\lambda \leq 1/2$. So we may assume that p is (λ, c)-quasi-geodesic as well. Let Δ be a reduced diagram over (10) such that $\partial \Delta$ is labeled W. Assume that Δ has at least one $\mathcal{R}$-cell. (Otherwise the lemma is obviously true.) Passing to an $\mathcal{O}$-equivalent diagram if necessary and using Lemma 4.4, we can find an $\mathcal{R}$-cell Π and a contiguity subdiagram Γ of Π to $\partial \Delta$ with $(\Pi, \Gamma, \partial \Delta) > 1 - 13\mu$. Let $\partial \Gamma = s_1 q_1 s_2 q_2$, where q_1 (respectively q_2) is a subpath of $\partial \Pi$ (respectively $\partial \Delta$) and $\max\{\|s_1\|, \|s_2\|\} \leq \varepsilon$. Let also $\partial \Pi = q_1 u$ and $\partial \Delta = w q_2$.

Note that perimeter of the subdiagram Ξ of Δ bounded by the path $s_2^{-1} u s_1^{-1} w$ is smaller than perimeter of Δ if ρ is large enough and μ is close to zero. Indeed as Γ contains no $\mathcal{R}$-cells, we can regard $s_1 q_1 s_2 q_2$ as a cycle in the Cayley graph $\Gamma(G, X \cup \mathcal{H})$. Thus,

(14)
$$l(q_2) \geq \mathrm{dist}_{X \cup \mathcal{H}}((q_2)_-, (q_2)_+) \geq \mathrm{dist}_{X \cup \mathcal{H}}((q_1)_-, (q_1)_+) - 2\varepsilon \geq \lambda l(q_1) - c - 2\varepsilon$$
$$\geq \lambda(1 - 13\mu) l(\partial \Pi) - c - 2\varepsilon \geq \lambda(1 - 13\mu)\rho - c - 2\varepsilon.$$

(Recall that $l(\partial\Pi) \geq \rho$ by the $C(\varepsilon, \mu, \lambda, c, \rho)$-condition.) On the other hand,

$$(15) \qquad\qquad l(s_2^{-1}us_1^{-1}) \leq 2\varepsilon + 13\mu l(\partial\Pi) \leq 2\varepsilon + 13\mu\rho.$$

If ρ is big enough and μ is small enough, the right side of (14) is greater than the right side of (15) and hence $l(\partial\Xi) < l(\partial\Delta)$.

Therefore, by induction the total number n_1 of $\mathcal{R}$- and $\mathcal{Q}$-cells in Ξ is at most

$$(16) \qquad\qquad n_1 \leq \alpha l(\partial\Xi) \leq \alpha\big(\|W\| - l(q_2) + l(s_2^{-1}us_1^{-1})\big)$$
$$\leq \alpha\big(\|W\| - l(\partial\Pi)(\lambda - 26\mu) + c + 4\varepsilon\big).$$

Furthermore, as q_2 is $(1/2, 0)$-quasi-geodesic in $\Gamma(G, X \cup \mathcal{H})$, we have

$$l(q_2) \leq 2\mathrm{dist}_{X\cup\mathcal{H}}((q_2)_-, (q_2)_+) \leq 2(\mathrm{dist}_{X\cup\mathcal{H}}((q_1)_-, (q_1)_+) + 2\varepsilon) \leq 2l(\partial\Pi) + 4\varepsilon.$$

Therefore, the perimeter of Γ satisfies

$$l(\partial\Gamma) \leq 2\varepsilon + l(q_1) + l(q_2) \leq 3l(\partial\Pi) + 6\varepsilon.$$

Hence we may assume that the number n_2 of $\mathcal{Q}$-cells of Γ is at most

$$(17) \qquad n_2 \leq Ll(\partial\Gamma) \leq 3L(l(\partial\Pi) + 2\varepsilon) \leq \alpha(l(\partial\Pi)(\lambda - 27\mu) + 2\varepsilon).$$

Finally, combining (16) and (17), we obtain

$$\mathrm{Area}_1^{\mathrm{rel}}(W) \leq n_1 + n_2 \leq \alpha(\|W\| - \mu l(\partial\Pi) + c + 6\varepsilon) \leq \alpha\|W\|$$

whenever ρ is big enough. This completes the proof of relative hyperbolicity of G_1. The hyperbolicity of the Cayley graph follows from Lemma 3.9.

Note that if $\mu < 1/30$, the inequality (14) implies

$$\|W\| \geq \frac{1}{2}\lambda\rho - c - 2\varepsilon$$

for every nontrivial word W which is geodesic in G and represents 1 in G_1. Therefore, the second statement of the lemma holds if we assume $\rho \geq 2(N + c + 2\varepsilon)/\lambda$.

$\square$

6. **Torsion in the quotient**

We keep all assumptions and notation introduced in the beginning of the previous section. Our next goal is to describe periodic elements in the quotient group (10) of a relatively hyperbolic group G. To this end we need an auxiliary result.

Recall that for an element $g \in G$, the *translation number* of g with respect to $\mathcal{A}$ is defined to be

$$\tau_{\mathcal{A}}(g) = \lim_{n\to\infty} \frac{|g^n|_{\mathcal{A}}}{n}.$$

This limit always exists and is equal to $\inf_n(|g^n|_{\mathcal{A}}/n)$ [9]. The lemma below can be found in [28, Th. 4.43].

LEMMA 6.1. *There exists $d > 0$ such that for any hyperbolic element of infinite order $g \in G$ we have $\tau_{X\cup\mathcal{H}}(g) \geq d$.*

For our goals, even a stronger result is necessary.

LEMMA 6.2. *There exist $1 \geq \alpha > 0$ and $a \geq 0$ with the following property. Suppose that g is a hyperbolic element of G of infinite order such that g has the smallest length among all elements of the conjugacy class g^G. Denote by U a shortest word in the alphabet $X \cup \mathcal{H}$ representing g. Then for any $n \in \mathbb{N}$, any path in $\Gamma(G, X \cup \mathcal{H})$ labeled U^n is (α, a)-quasi-geodesic.*

Proof. Recall that $\Gamma(G, X \cup \mathcal{H})$ is hyperbolic by Lemma 3.9. First suppose that $|g|_{X \cup \mathcal{H}} = \|U\| > 8\delta$, where δ is the hyperbolicity constant of $\Gamma(G, X \cup \mathcal{H})$. Since g is a shortest element in g^G and U is a shortest word representing g, there exists $k > 8\delta$ so that the path p labeled U^n is a k-local geodesic in $\Gamma(G, X \cup \mathcal{H})$ for any n. Therefore, by Lemma 3.6, p is $(1/3, 2\delta)$-quasi-geodesic.

Further if $|g|_{X \cup \mathcal{H}} = \|U\| \leq 8\delta$, then for any $n \in \mathbb{N}$, we have

$$|g^n|_{X \cup \mathcal{H}} \geq n \inf_i \left(\frac{1}{i} |g^i|_{X \cup \mathcal{H}} \right) \geq nd \geq \frac{d}{8\delta} n |g|_{X \cup \mathcal{H}},$$

where d is the constant provided by Lemma 6.1. Hence the path p labeled U^n is $\left(\frac{d}{8\delta}, 8\delta \right)$-quasi-geodesic. It remains to set $\alpha = \min \left\{ \frac{1}{3}, \frac{d}{8\delta} \right\}$ and $a = 8\delta$. $\qquad \square$

LEMMA 6.3. *For any $\lambda \in (0, 1]$, $c \geq 0$ there are $\mu > 0$, $\varepsilon \geq 0$, and $\rho > 0$ such that the following condition holds. Suppose that $\mathfrak{R}$ is a symmetrized set of words in $\mathcal{A}$ satisfying the $C_1(\varepsilon, \mu, \lambda, c, \rho)$-condition. Then every element of finite order in the group G_1 given by (10) is the image of an element of finite order of G.*

Proof. Let us fix $\lambda, c > 0$. Observe that the $C_1(\varepsilon, \mu, \lambda, c, \rho)$-condition becomes stronger as λ increases and c decreases. Hence it suffices to prove the lemma assuming that $\lambda < \alpha$ and $c > a$ for α and a provided by Lemma 6.2. Let us choose constants $1/16 > \mu > 0$, $\varepsilon > 0$, $\rho > 0$ such that

$(*)$ the conclusion of Lemma 4.4 holds.

Again as in the proof of Lemma 5.1, we note that the $C_1(\varepsilon, \mu, \lambda, c, \rho)$-condition becomes stronger as μ decreases and ε, ρ increase. Hence we may decrease μ and then increase ρ and ε if necessary without violating $(*)$. In particular, without loss of generality, we may assume that

$$(18) \qquad\qquad\qquad \varepsilon > 2\kappa + 8\delta,$$

where $\kappa = \kappa(\lambda, c)$ is the constant provided by Lemma 3.1 and δ is the hyperbolicity constant of $\Gamma(G, X \cup \mathcal{H})$. We fix ε from now on.

Suppose that g is an element of order $n > 0$ in G_1 such that its preimage has infinite order in G. Assume also that g has the smallest length among all elements from the conjugacy class g^{G_1}. Denote by U a shortest word in the alphabet $X \cup \mathcal{H}$ representing g in G_1. Then there exists a diagram Δ over (10) with boundary label U^n. By Lemma 6.2, the label of $\partial \Delta$ is (λ, c)-quasi-geodesic (in G) and Δ contains at least one $\mathfrak{R}$-cell. (For otherwise Δ is a diagram over (4) and $U^n = 1$ in G.) Note

that passing from Δ to an $\mathbb{O}$-equivalent diagram does not affect $\partial\Delta$ and the number of $\mathcal{R}$-cells. Thus by Lemma 4.4 we can assume that there exists an $\mathcal{R}$-cell Π in Δ and ε-contiguity subdiagram Γ of Π to $\partial\Delta$ such that $(\Pi, \Gamma, \partial\Delta) > 1 - 13\mu$.

Let $\partial\Gamma = s_1 q_1 s_2 q_2$ as on Fig. 1. Since Γ has no $\mathcal{R}$-cells, we may think of $s_1 q_1 s_2 q_2$ as a quadrangle in $\Gamma(G, X \cup \mathcal{H})$, where q_i is (λ, c)-quasi-geodesic, and $l(s_i) \le \varepsilon$ for $i = 1, 2$. Therefore we have

(19)
$$l(q_2) \ge \mathrm{dist}_{X \cup \mathcal{H}}((q_2)_-, (q_2)_+) \ge \mathrm{dist}_{X \cup \mathcal{H}}((q_1)_-, (q_1)_+) - l(s_1) - l(s_2)$$
$$\ge \lambda l(q_1) - c - 2\varepsilon \ge \lambda(1 - 13\mu)l(\partial\Pi) - c - 2\varepsilon \ge \lambda(1 - 13\mu)\rho - c - 2\varepsilon.$$

(Recall again that $l(\partial\Pi) \ge \rho$ by the $C_1(\varepsilon, \mu, \lambda, c, \rho)$-condition.) In particular, choosing ρ large enough we can make $l(q_2)$ as large as we want.

Passing from U to a cyclic shift of $U^{\pm 1}$ if necessary, we may assume that **Lab**(q_2) is a prefix of U^n. We now have three cases to consider. Our goal is to show that neither of them is possible whenever μ is small enough and ρ is large enough.

Case 1. Suppose that $l(q_2) \ge 4\|U\|/3$. This allows us to find two long disjoint equal subwords of **Lab**(q_2). More precisely, we decompose **Lab**(q_2) as **Lab**$(q_2) \equiv W V_1 W V_2$, where

$$\lambda^2 l(q_2)/5 \le \|W\| \le \lambda^2 l(q_2)/4$$

and

$$\|V_1\| > l(q_2)/3.$$

Let $q_2 = w_1 v_1 w_2 v_2$ be the corresponding decomposition of the path q_2. Corollary 3.3 applied to the quadrangle $s_1 q_1 s_2 q_2$ implies that there is a point $o \in q_1$ such that $\mathrm{dist}_{X \cup \mathcal{H}}(o, (w_1)_+) \le K + \varepsilon$, where K depends on λ, c, and δ only. Let $q_1^{-1} = r_1 t$, where $(r_1)_+ = o$. Thus we have

(20) $\mathrm{dist}_{X \cup \mathcal{H}}((r_1)_\pm, (w_1)_\pm) \le K + \varepsilon.$

Similarly one can find a subpath r_2^ξ, $\xi = \pm 1$, of q_1^{-1} such that

(21) $\mathrm{dist}_{X \cup \mathcal{H}}((r_2^\xi)_\pm, (w_2)_\pm) \le K + \varepsilon.$

We note that r_1 and r_2 are disjoint. Indeed, otherwise r_1 passes through $(r_2)_-$ or $(r_2)_+$. For definiteness, assume that $(r_2)_- \in r_1$. Then we have

$$l(r_1) \le \frac{1}{\lambda}(\mathrm{dist}_{X \cup \mathcal{H}}((r_1)_-, (r_1)_+) + c) \le \frac{1}{\lambda}(l(w_1) + 2\varepsilon + K + c)$$
$$= \frac{1}{\lambda}(\|W\| + 2\varepsilon + K + c) \le \lambda l(q_2)/4 + (2\varepsilon + K + c)/\lambda.$$

On the other hand,

$$l(r_1) \ge \mathrm{dist}_{X \cup \mathcal{H}}((r_1)_-, (r_2)_-) \ge \lambda l(w_1 v_1) - c - 2\varepsilon - K$$
$$\ge \lambda\|V_1\| - c - 2\varepsilon - K \ge \lambda l(q_2)/3 - c - 2\varepsilon - K.$$

These inequalities contradict each other if $l(q_2)$ is large enough. As explained before, the later condition can always be ensured by choosing sufficiently large ρ (see (19)).

Thus we have a decomposition $q_1^{-1} = r_1 t_1 r_2^{\xi} t_2$, $\xi = \pm 1$. Let $\mathbf{Lab}\,(q_1)^{-1} \equiv R_1 T_1 R_2 T_2$ be the corresponding decomposition of the label. By (20) and (21), we have $R_1 = Y_1 W Z_1$ and $R_2 = Y_2 W^{\pm 1} Z_2$ in G, where $\|Y_i\|, \|X_i\| \le K + \varepsilon$ for $i = 1, 2$. Hence $R_1 = Y R_2^{\pm 1} Z$ in G, where $\|Y\|, \|Z\| \le 2(K + \varepsilon)$. Without loss of generality, we may assume that words Y and Z are geodesic in G. Let $aybz$ be the corresponding rectangle in $\Gamma(G, X \cup \mathscr{H})$, where $\mathbf{Lab}\,(a) \equiv R_1^{-1}$, $\mathbf{Lab}\,(y) \equiv Y$, $\mathbf{Lab}\,(b) \equiv R_2^{\pm 1}$, $\mathbf{Lab}\,(z) \equiv Z$. Recall that R_1, R_2 are subwords of the (λ, c)-quasi-geodesic word $\mathbf{Lab}\,(\partial \Pi)$. Hence a, b are (λ, c)-quasi-geodesic. Using (19) and the triangle inequality, we obtain

$$(22) \qquad l(a) = \|R_1\| \ge \mathrm{dist}_{X \cup \mathscr{H}}((r_1)_-, (r_1)_-)$$
$$\ge \mathrm{dist}_{X \cup \mathscr{H}}((w_1)_-, (w_1)_+) - l(s_1) - \mathrm{dist}_{X \cup \mathscr{H}}((r_1)_+, (w_1)_+)$$
$$\ge \lambda l(w_1) - c - 2\varepsilon - K = \lambda \|W\| - c - 2\varepsilon - K$$
$$\ge \lambda^3 l(q_2)/5 - c - 2\varepsilon - K$$
$$\ge \lambda^3 [\lambda(1 - 13\mu)\rho - c - 2\varepsilon]/5 - c - 2\varepsilon - K.$$

Clearly we can ensure $l(a) > (4 \max\{l(y), l(z)\} + c)/\lambda$ by taking ρ large enough. This allows us to apply Corollary 3.5. Thus we obtain subsegments a', b' of the sides a and b, respectively, such that the distances between the corresponding ends of these subsegments is at most $8\delta + 2\kappa$. Let $A = \mathbf{Lab}\,(a')$, $B = \mathbf{Lab}\,(b')$. Then $A = CB^{\pm 1}D$ in G, where $\|C\|, \|D\| \le 8\delta + 2\kappa < \varepsilon$ by (18). Using this and (22), we obtain

$$\min\{\|A\|, \|B\|\} \ge \frac{7}{20}(\lambda l(a) - c) - 8\delta - 2\kappa \ge \mu\rho \ge \mu l(\partial \Pi)$$

if μ is small enough and ρ is large enough. Since A and B are disjoint subwords of $\mathbf{Lab}\,(\partial \Pi)$, this contradicts the $C_1(\varepsilon, \mu, \lambda, c, \rho)$-condition.

Case 2. Suppose that $\|U\| \le l(q_2) \le 4\|U\|/3$, i.e. $\mathbf{Lab}\,(q_2) \equiv UV$ for some (may be empty) word V, $\|V\| \le \|U\|/3$. Note that $\mathbf{Lab}\,(q_2) = \mathbf{Lab}\,(s_2^{-1} u s_1^{-1})$ in G_1, hence $g = \mathbf{Lab}\,(s_2^{-1} u s_1^{-1})V^{-1}$ in G_1. Since U is the shortest word representing g in G_1, we obtain

$$\|U\| \le 2\varepsilon + 13\mu l(\partial \Pi) + \|V\| \le 2\varepsilon + 13\mu l(\partial \Pi) + \|U\|/3.$$

Consequently,

$$(23) \qquad \|U\| \le 3(2\varepsilon + 13\mu l(\partial \Pi))/2 \le 3(2\varepsilon + 13\mu\rho)/2.$$

On the other hand, using (19) we obtain

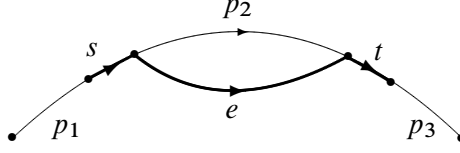$$\|U\| \ge 3l(q_2)/4 \ge 3(\lambda(1 - 13\mu)\rho - c - 2\varepsilon)/4.$$

Figure 3

The later inequality contradicts (23) whenever μ is small enough and ρ is large enough.

Case 3. Suppose that $l(q_2) < \|U\|$, i.e., **Lab** (q_2) is a subword of U. Again since U is the shortest word representing g in G_1, we have $\|\textbf{Lab}\,(q_2)\| \leq \|Q\|$ for every word Q such that $Q = \textbf{Lab}\,(q_2)$ in G_1. In particular, for $Q \equiv \textbf{Lab}\,(s_2^{-1} u s_1^{-1})$, we obtain

$$(24) \qquad \|Q\| \geq \|\textbf{Lab}\,(q_2)\| = l(q_2) \geq \lambda(1 - 13\mu)\rho - c - 2\varepsilon$$

by (19). On the other hand, we obviously have $\|Q\| \leq 2\varepsilon + 13\mu l(\partial\Pi) < 2\varepsilon + 13\mu\rho$, which contradicts (24) if μ is small enough and ρ is large enough. $\square$

7. **Words with small cancellations**

Recall that G denotes a group hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda}$, X denotes a finite relative generating set of G with respect to $\{H_\lambda\}_{\lambda \in \Lambda}$. As above we set $\mathscr{A} = X \cup \mathscr{H}$. Our main goal here is to show that a certain set of words over the alphabet $\mathscr{A}$ satisfies the small cancellation conditions described above.

More precisely, suppose that W is a word satisfying the following conditions.

(W1) $W \equiv x a_1 b_1 \ldots a_n b_n$ for some $n \geq 1$, where:

(W2) $x \in X \cup \{1\}$;

(W3) $a_1, \ldots, a_n$ (respectively $b_1, \ldots, b_n$) are elements of H_α (respectively H_β), where $H_\alpha \cap H_\beta = \{1\}$;

(W4) the elements $a_1, \ldots, a_n, b_1, \ldots, b_n$ do not belong to the set

$$(25) \qquad \mathscr{F} = \mathscr{F}(\varepsilon) = \{g \in \langle \Omega \rangle \; : \; |g|_\Omega \leq K(32\varepsilon + 70)\},$$

where ε is some nonnegative integer and the set Ω and the constant K are provided by Lemma 3.10.

Let $\mathscr{SW}$ denote the set of all subwords of cyclic shifts of $W^{\pm 1}$. As in [28], we say that a path p in $\Gamma(G, X \cup \mathscr{H})$ is a *path without backtracking* if all components of p are isolated.

LEMMA 7.1. *Suppose p is a path in $\Gamma(G, X \cup \mathscr{H})$ such that* **Lab** *$(p) \in \mathscr{SW}$.* *Then*

1) *p is a path without backtracking.*

2) *p is $(1/3, 2)$-quasi-geodesic.*

Proof. 1) Suppose that $p = p_1 s p_2 t p_3$, where s and t are two connected components. Passing to another pair of connected components of p if necessary, we may assume that p_2 is a path without backtracking. For definiteness, we also assume that s and t are H_α-components. Let e denote a path of length at most 1 in $\Gamma(G, X \cup \mathcal{H})$ connecting s_+ to t_- and labeled by an element of H_α (see Figure 3). It follows from our choice of W and the condition $H_\alpha \cap H_\beta = \{1\}$ that $l(p_2) \geq 2$. Thus p_2 contains $m \geq l(p_2)/2 \geq 1$ H_β-components, say $r_1, \ldots, r_m$, and all these components are isolated components of the cycle $d = e p_2^{-1}$. Let $g_1, \ldots, g_m$ be elements of G represented by the labels of $r_1, \ldots, r_m$. By Lemma 3.10, $g_i \in \langle \Omega \rangle$, $i = 1, \ldots, m$. According to (W4), we have

$$\sum_{i=1}^{m} |g_i|_\Omega \geq 70 K m \geq 35 K l(p_2) > K(l(p_2) + 1) = K l(d).$$

This contradicts Lemma 3.10.

2) Since the set $\mathcal{S}W$ is closed under taking subwords, it suffices to show that $\text{dist}_{X \cup \mathcal{H}}(p_-, p_+) \geq l(p)/3 - 2$. In case $l(p) \leq 6$ this is obvious, so we assume $l(p) > 6$. Suppose that $\text{dist}_{X \cup \mathcal{H}}(p_-, p_+) < l(p)/3 - 2$. Let c denote a geodesic in $\Gamma(G, X \cup \mathcal{H})$ such that $c_- = p_-$ and $c_+ = p_+$. Since p is a path without backtracking, any H_α-component of p is connected to at most one H_α-component of c. Obviously the path p contains at least $l(p)/2 - 1$ H_α-components. Therefore, at least

$$k = l(p)/2 - 1 - l(c) > l(p)/2 - 1 - (l(p)/3 - 2) > l(p)/6 > 1$$

of them are isolated H_α-components of the cycle $p c^{-1}$. Let $f_1, \ldots, f_k$ be the elements of G represented by these components. Then as above we have $f_i \in \langle \Omega \rangle$, $i = 1, \ldots, k$. By (W4), we obtain

$$\sum_{i=1}^{k} |f_i|_\Omega \geq 70 K k > 2 K l(p) > K l(p c^{-1}).$$

This leads to a contradiction again.   □

LEMMA 7.2. *Suppose $u p v^{-1} q^{-1}$ is an an arbitrary quadrangle in $\Gamma(G, X \cup \mathcal{H})$ satisfying the following conditions:*

(a) **Lab** $(p) \equiv$ **Lab** $(q^{-1}) \in \mathcal{S}W$;

(b) $\max\{l(u), l(v)\} \leq \varepsilon$;

(c) $l(p) = l(q) \geq 6\varepsilon + 22.$

Then the paths p and q have a common edge.

The proof of Lemma 7.2 is based on the following two results. (We keep the notation of Lemma 7.2 there.)

LEMMA 7.3. *Let us divide p into three parts $p = p_1 p_0 p_2$ such that*

(26) $$l(p_1) = l(p_2) = 3\varepsilon + 6.$$

Suppose that s is a component of p_0. Then s can not be connected to a component of paths u or v.

Proof. Suppose that a component s of p_0 is connected to a component t of u. Then $\mathrm{dist}_{X \cup \mathcal{H}}(s_+, t_+) \le 1$. Recall that the segment $[p_-, s_+]$ of p is $(1/3, 2)$-quasi-geodesic by Lemma 7.1. Hence,

$$l(p_1) \; < l([p_-, s_+]) \le 3(\mathrm{dist}_{X \cup \mathcal{H}}(p_-, s_+) + 2)$$

$$\le 3(\mathrm{dist}_{X \cup \mathcal{H}}(p_-, t_+) + \mathrm{dist}_{X \cup \mathcal{H}}(t_+, s_+) + 2)$$

$$\le 3(l(u) - 1 + 1 + 2) \le 3\varepsilon + 6.$$

However, this contradicts (26). Similarly s can not be connected to a component of v.   □

LEMMA 7.4. *Let $s_1, \ldots, s_k$ be consecutive components of p_0. Then q can be decomposed as $q = q_1 t_1 \ldots q_k t_k q_{k+1}$, where*
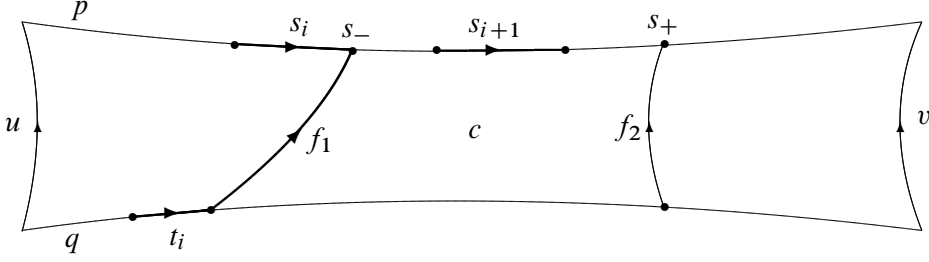
1) t_i *is a component of q connected to s_i, $i = 1, \ldots, k$;*

2) q_i *contains no components for $i = 2, \ldots, k$, i.e. either $\mathbf{Lab}\,(q_i) \equiv x$ or q_i is trivial.*

Proof. To prove the first assertion of the lemma we proceed by induction. First let us show that s_1 is not isolated in $d = u p v^{-1} q^{-1}$.

Indeed assume s_1 is isolated in d. Suppose for definiteness that s_1 is an H_α-component. We consider the maximal subpath s of p_0 such that s contains s_1 and all H_α-components of s are isolated in d. By maximality of s, either $s_- = (p_0)_-$, or $s_- = r_+$ for a certain H_α-component r of p_0 such that r is not isolated in d. (According to Lemma 7.3 this means that r is connected to an H_α-component of q.) In the first case we denote by f_1 the path $u p_1$. In the second case, let f_1 be a path of length ≤ 1 that connects an H_α-component of q to r. In both cases we have $l(f_1) \le 4\varepsilon + 6$. It follows from the choice of s that no H_α-component of s is connected to an H_α-component of f_1. Similarly we construct a path f_2 such that $(f_2)_- \in q$, $(f_2)_+ = s_+$, $l(f_2) \le 4\varepsilon + 6$, and no H_α-component of s is connected to an H_α-component of f_2.

Clearly all H_α-components of s are isolated in the cycle

$$c = f_1 s f_2^{-1} [(f_2)_-, (f_1)_-],$$

Figure 4

where $[(f_2)_-, (f_1)_-]$ is a segment of $q^{\pm 1}$. We have

$$(27) \qquad \mathrm{dist}_{X \cup \mathcal{H}}((f_1)_-, (f_2)_-) \leq l(f_1) + l(s) + l(f_2) \leq 8\varepsilon + 12 + l(s).$$

Consequently,

$$(28) \qquad l([(f_1)_-, (f_2)_-]) \leq 3\mathrm{dist}_{X \cup \mathcal{H}}((f_1)_-, (f_2)_-) + 2) \leq 24\varepsilon + 42 + 3l(s).$$

Finally,

$$(29) \qquad l(c) \leq l(f_1) + l(s) + l(f_2) + l([(f_1)_-, (f_2)_-]) \leq 32\varepsilon + 54 + 4l(s).$$

Let $g_1, \ldots, g_m$ denote the elements represented by H_α-components of s. Note that $l(s) \leq 2m + 2$. Applying Lemma 3.10, we obtain $g_i \in \langle \Omega \rangle$, $i = 1, \ldots, m$, and

$$(30) \qquad \sum_{i=1}^{m} |g_i|_\Omega \leq K l(c) \leq K(32\varepsilon + 54 + 4l(s)) \leq K(32\varepsilon + 62 + 8m).$$

Therefore, at least one of the elements $g_1, \ldots, g_m$ has length at least

$$(31) \qquad \frac{1}{m} K(32\varepsilon + 62 + 8m) \leq K(32\varepsilon + 70)$$

that contradicts (W4). Thus s_1 is not isolated in d. By Lemma 7.3 this means that s_1 is connected to an H_α-component t_1 of q.

Now assume that we have already found components $t_1, \ldots, t_i$ of q, $1 \leq i < k$, that are connected to $s_1, \ldots, s_i$ respectively. The inductive step is similar to the above considerations. For definiteness, we assume that s_i is an H_α-component. Then s_{i+1} is an H_β-component by the choice of W. We denote by f_1 a path of length ≤ 1 labeled by an element of H_α that connects $(t_i)_+$ to $(s_i)_+$ (Figure 4). If s_{i+1} is isolated in the cycle $c = f_1[(s_i)_+, p_+]v^{-1}[q_+, (t_i)_+]$, we denote by s the maximal initial subpath of the segment $[(s_i)_+, (p_0)_+]$ of p_0 such that s contains s_{i+1} and all H_α-components of s are isolated in c. As above, we can find a path f_2 such that $(f_2)_- \in q$, $(f_2)_+ = s_+$, $l(f_2) \leq 4\varepsilon + 6$, and no H_α-component of s is connected to an H_α-component of f_2. The inequalities (27)–(31) remain valid and we arrive at a contradiction in the same way. Thus s_{i+1} is not isolated in c, i.e., it is connected to a component t_{i+1} of the segment $[(t_i)_+, q_+]$ of q. This completes the inductive step.
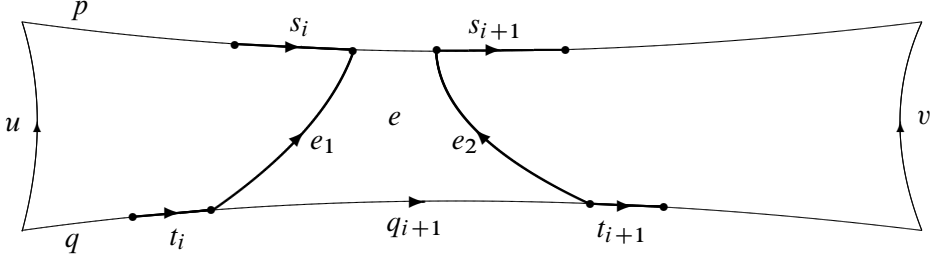
Figure 5

Let us prove the second assertion of the lemma. Suppose that s_i, s_{i+1} are two subsequent components of p_0 and such that q_{i+1} contains at least one component (say, an H_α-component). As above for definiteness we assume that s_i (respectively s_{i+1}) is an H_α-component (respectively H_β-component). Let e_1, e_2 be the paths of lengths ≤ 1 labeled by elements of H_α and H_β respectively such that $(e_1)_+ = (s_i)_+$, $(e_1)_- = (q_{i+1})_-$, $(e_2)_+ = (s_{i+1})_-$, $(e_2)_- = (q_{i+1})_+$ (see Figure 5). As q is a path without backtracking, each H_α-component of q_{i+1} is isolated in the cycle $e = q_{i+1}e_2[(s_{i+1})_-, (s_i)_+]e_1^{-1}$, where $[(s_{i+1})_-, (s_i)_+]$ is a segment of p^{-1}. Notice that $l([(s_{i+1})_-, (s_i)_+]) \leq 1$ as s_i and s_{i+1} are subsequent components of p_0. We denote by $f_1, \ldots, f_m$ the elements represented by H_α-components of q_{i+1}. By Lemma 3.10, we have $f_j \in \langle \Omega \rangle$, $j = 1, \ldots, m$, and

$$\sum_{j=1}^{m} |f_j|_\Omega \leq Kl(e) \leq K(3 + l(q_{i+1})) \leq K(3 + 2m + 2) \leq 7mK.$$

This contradicts (W4) again. □

Proof of Lemma 7.2. We keep the notation introduced in Lemmas 5.3 and 5.4. Let also

$$p_0 = p_1 s_1 \ldots p_k s_k p_{k+1}$$

for some (may be trivial) subpaths $p_1, \ldots p_{k+1}$ of p_0.

According to (26) and condition c) of Lemma 7.2, we have

$$l(p_0) \geq 6\varepsilon + 22 - l(p_1) - l(p_2) \geq 10.$$

Since $s_1, \ldots, s_k$ are subsequent components and **Lab** (p_0) is a subword of a cyclic shift of $W^{\pm 1}$, at most one of the paths $p_2, \ldots, p_k$ is nontrivial. Therefore, there are at least 5 subsequent components $s_i, \ldots, s_{i+4}$, such that $p_{i+1}, \ldots, p_{i+4}$ are trivial. Without loss of generality we may assume $i = 1$. Similarly, by Lemma 7.4, we can find at least three subsequent components among $t_1, \ldots t_5$, say t_1, t_2, t_3, such that $(t_1)_+ = (t_2)_-$ and $(t_2)_+ = (t_3)_-$. Let w be an element represented by the label of any path that goes from $(t_1)_+ = (t_2)_-$ to $(s_1)_+ = (s_2)_-$. For definiteness we assume that t_1 and s_1 are H_α-components. Since t_1 and s_1 are connected, we have $w \in H_\alpha$. On the other hand, the H_β-components t_2 and s_2 are also connected. Hence $w \in H_\beta$. Thus $w \in H_\alpha \cap H_\beta = \{1\}$, i.e. the vertices $(t_2)_-$ and $(s_2)_-$ coincide.

Similarly the vertices $(t_2)_+$ and $(s_2)_+$ coincide. In particular, t_2 and s_2 are edges labeled by the same element of H_β, i.e., t_2 and s_2 coincide. $\square$

Now we are ready to prove the main result of this section.

THEOREM 7.5. *Suppose that W is a word in $\mathcal{A}$ satisfying the conditions* (W1)–(W4) *and, in addition, $a_i \neq a_j^{\pm 1}$, $b_i \neq b_j^{\pm 1}$ whenever $i \neq j$ and $a_i \neq a_i^{-1}$, $b_i \neq b_i^{-1}$, $i, j \in \{1, \ldots, n\}$. Then the set $\mathcal{W}$ of all cyclic shifts of $W^{\pm 1}$ satisfies the $C_1(\varepsilon, \frac{3\varepsilon + 11}{n}, \frac{1}{3}, 2, 2n + 1)$ small cancellation condition.*

Proof. The first two conditions from Definition 4.2 follow from the choice of W and Lemma 7.1. Suppose that U is an ε-piece of a word $R \in \mathcal{W}$. Assume that $\max\{\|U\|, \|U'\|\} \geq \mu \|R\|$ for $\mu = \frac{3\varepsilon + 11}{n}$, that is,

$$\max\{\|U\|, \|U'\|\} \geq \frac{3\varepsilon + 11}{n}(2n + 1) > 6\varepsilon + 22.$$

(Here and below we use the notation of Definitions 4.2 and 4.1.) Without loss of generality we may assume that $\|U\| \geq 6\varepsilon + 22$. By the definition of an ε-piece, there is a quadrangle $upv^{-1}q^{-1}$ in $\Gamma(G, X \cup \mathcal{H})$ satisfying conditions (a)–(c) of Lemma 7.2 and such that labels of p and q are U and U' respectively.

Let e be the common edge of p and q. Then we have

$$R \equiv U_1 \mathbf{Lab}\,(e) U_2 V$$

and

$$R' \equiv U_1' \mathbf{Lab}\,(e) U_2' V',$$

where $U_1 \mathbf{Lab}\,(e) U_2 \equiv U$ and $U' \equiv U_1' \mathbf{Lab}\,(e) U_2'$. Since $\mathbf{Lab}\,(e)$ appears in $W^{\pm 1}$ only once, R, R' are cyclic shifts of the same word $W^{\pm 1}$ and

$$U_2 V U_1 \equiv U_2' V' U_1'.$$

Note also that

$$Y = U_1' U_1^{-1}$$

in G as $Y^{-1} U_1' U_1^{-1}$ is a label of a cycle in $\Gamma(G, X \cup \mathcal{H})$. Therefore, the following equalities hold in the group G:

$$YRY^{-1} = U_1' U_1^{-1} U_1 \mathbf{Lab}\,(e) U_2 V U_1 (U_1')^{-1} = U_1' \mathbf{Lab}\,(e) U_2' V' U_1' (U_1')^{-1} = R'$$

that contradicts the third condition from Definition 4.2.

Similarly, if U is an ε'-piece, then $R \equiv UVU'V'$ for some U, U', V, V', where both subwords U and U' contain a certain letter from $X \cup \mathcal{H}$. In this case we arrive at a contradiction again as any letter $a \in X \cup \mathcal{H}$ appears in R only once, and if a appears in R, then a^{-1} does not. $\square$

Finally we note that the condition $x \in X \cup \{1\}$ in Theorem 7.5 is not really restrictive since we can always add any element $x \in G$ to the set X without violating relative hyperbolicity.

8. **Suitable subgroups and quotients**

Throughout this section, we keep the assumption that G is hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda}$. The proof of Lemma 2.3 is based on the following auxiliary result.

LEMMA 8.1. *Suppose that for some* $\lambda, \mu \in \Lambda$, $\lambda \neq \mu$, H_λ *and* H_μ *contain elements of infinite order and* $H_\lambda \cap H_\mu = \{1\}$. *Then there are* $f \in H_\lambda$, $g \in H_\mu$ *such that* fg *is a hyperbolic element of infinite order and* $E_G(fg) = \langle fg \rangle$.

Proof. We set $\varepsilon = 2(\kappa + \delta)$, where $\kappa = \kappa(\delta, 1/3, 2)$ is the constant provided by Lemma 3.1 and δ is the hyperbolicity constant of $\Gamma(G, X \cup \mathcal{H})$. It is convenient to assume that $\kappa, \delta \in \mathbb{N}$.

Let $\mathcal{F} = \mathcal{F}(\varepsilon)$ be the set defined by (25). Since Ω is finite, $\mathcal{F}$ is finite, and hence there are elements $f \in H_\lambda \setminus \mathcal{F}$ and $g \in H_\mu \setminus \mathcal{F}$ of infinite order. In particular,

$$(32) \qquad\qquad\qquad f^2 \neq 1, \quad g^2 \neq 1.$$

Note that the word $W = (fg)^m$ satisfies conditions (W1)–(W4). (Here f and g are regarded as letters of $\mathcal{H}$.) By Lemma 7.1, for any $m \in \mathbb{N}$, the word $(fg)^m$ is $(1/3, 2)$-quasi-geodesic in $\Gamma(G, X \cup \mathcal{H})$. In particular, fg is a hyperbolic element. Indeed otherwise the length $|(fg)^m|_{X \cup \mathcal{H}}$ would be bounded uniformly on m.

Now suppose that $a \in E_G(fg)$. Then by Theorem 2.1, $a(fg)^m a^{-1} = (fg)^{\pm m}$ for some $m \in \mathbb{N}$. Passing to a multiple of m if necessary, we may assume that

$$(33) \qquad\qquad m \geq 3|a|_{X \cup \mathcal{H}} + 12(\kappa + \delta) + 20.$$

Let $a_1 b_1 a_2 b_2$ be a quadrangle in $\Gamma(G, X \cup \mathcal{H})$ such that a_1, a_2 are geodesic, the labels $\mathbf{Lab}\,(a_1) = \mathbf{Lab}\,(a_2^{-1})$ represent a in G, and $\mathbf{Lab}\,(b_1) = \mathbf{Lab}\,(b_2^{-1}) \equiv (fg)^{\pm m}$. Let also $b_1 = b_1' p b_1''$ Figure 6, where

$$(34) \qquad\qquad l(b_1') = l(b_1'') = 3(l(a_1) + 2(\kappa + \delta) + 3).$$
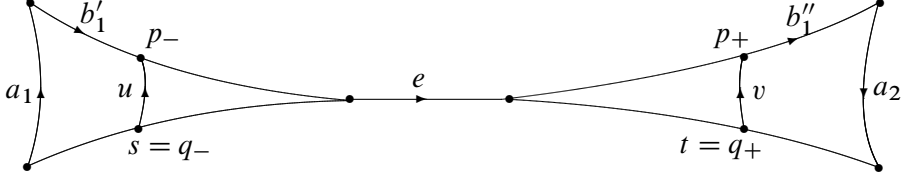
As b_1 is $(1/3, 2)$-quasi-geodesic, we have

$$(35) \qquad \mathrm{dist}_{X \cup \mathcal{H}}(p_\pm, (b_1)_\pm) \geq \frac{1}{3} l(b_1') - 2 \geq l(a_1) + 2(\kappa + \delta) + 1.$$

Further by Corollary 3.3, there is a point $s \in a_1 \cup b_2 \cup a_2$ such that $\mathrm{dist}_{X \cup \mathcal{H}}(s, p_-) \leq 2(\kappa + \delta)$. If $s \in a_1$, then we have

$$\mathrm{dist}_{X \cup \mathcal{H}}(p_-, b_-) \leq \mathrm{dist}_{X \cup \mathcal{H}}(p_-, s) + \mathrm{dist}_{X \cup \mathcal{H}}(s, b_-) \leq 2(\kappa + \delta) + l(a_1)$$

that contradicts (35). For the same reason s can not belong to a_2. Thus $s \in b_2$. Without loss of generality, we may assume that s is a vertex of $\Gamma(G, X \cup \mathcal{H})$. Similarly there exists a vertex $t \in b_2$ such that $\mathrm{dist}_{X \cup \mathcal{H}}(t, p_+) \leq 2(\kappa + \delta)$. Let u, v be geodesics in $\Gamma(G, X \cup \mathcal{H})$ connecting s to p_- and t to p_+ respectively and let q denote the segment $[s, t]$ of b_2^{-1}.

Figure 6

According to (33) and (34), we have

$$l(p) \geq 2m - 6(l(a_1) + 2(\kappa + \delta) + 3) \geq 12(\kappa + \delta) + 22.$$

Hence we may apply Lemma 7.2 for the quadrangle $upv^{-1}q^{-1}$ and $\varepsilon = 2(\kappa + \delta)$. Thus there exists a common edge e of p and q. In particular, this and (32) imply that $a(fg)^m a^{-1} = (fg)^m$ (not $(fg)^{-m}$).

There are two options for labels of the segments $[(b_1)_-, e_-]$ and $[(b_2)_+, e_-]$ of b_1 and b_2^{-1} respectively. Namely both these labels are either of the form $(fg)^n$ (possibly for different n) or of the form $(fg)^k f$. In both cases $a = (fg)^l$ for a certain l, as labels of a_1 and $[(b_2)_+, e_-][(b_1)_-, e_-]^{-1}$ represent the same element of G. Thus $a \in \langle fg \rangle$ and $E_G(fg) = \langle fg \rangle$. $\qquad\square$

Proof of Lemma 2.3. Let $f_1, f_2 \in H^0$ be noncommensurable elements of H such that $E_G(f_1) \cap E_G(f_2) = \{1\}$. By Theorem 2.1, G is hyperbolic relative to the collection $\{H_\lambda\}_{\lambda \in \Lambda} \cup E_G(f_1) \cup E_G(f_2)$.

We construct a sequence of desired elements $h_1, h_2, \ldots$ by induction. By Lemma 8.1, there are $f \in E_G(f_1)$, $g \in E_G(f_2)$ such that the element $h_1 = fg$ is hyperbolic (with respect to the collection $\{H_\lambda\}_{\lambda \in \Lambda} \cup E_G(f_1) \cup E_G(f_2)$) and $E_G(h_1) = \langle h_1 \rangle$. Theorem 2.1 implies that G is hyperbolic relative to the collection $\{H_\lambda\}_{\lambda \in \Lambda} \cup E_G(f_1) \cup E_G(f_2) \cup E_G(h_1)$. Further we construct a hyperbolic (with respect to $\{H_\lambda\}_{\lambda \in \Lambda} \cup E_G(f_1) \cup E_G(f_2) \cup E_G(h_1)$) element h_2 as a product of an element of $E_G(f_1)$ and an element of $E_G(f_2)$ as above. As h_2 is hyperbolic, it is not commensurable with h_1. Applying Theorem 2.1 again we join $E_G(h_2)$ to the collection of subgroups with respect to which G is hyperbolic and so on. Continuing this procedure, we get what we need. $\qquad\square$

To prove Theorem 2.4, we need the following two observations. The first one is a particular case of Theorem 2.40 from [28].

LEMMA 8.2. *Suppose that a group G is hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda} \cup \{S_1, \ldots, S_m\}$, where $S_1, \ldots, S_m$ are finitely generated and hyperbolic in the ordinary (nonrelative) sense. Then G is hyperbolic relative to $\{H_\lambda\}_{\lambda \in \Lambda}$.*

The next lemma is a particular case of Theorem 1.4 from [28].

LEMMA 8.3. *Suppose that a group G is hyperbolic relative to a collection of subgroups $\{H_\lambda\}_{\lambda \in \Lambda}$. Then*

(a) *For any $g \in G$ and any $\lambda, \mu \in \Lambda$, $\lambda \neq \mu$, the intersection $H_\lambda^g \cap H_\mu$ is finite.*

(b) *For any $\lambda \in \Lambda$ and any $g \notin H_\lambda$, the intersection $H_\lambda^g \cap H_\lambda$ is finite.*

Proof of Theorem 2.4. Obviously it suffices to deal with the case $m = 1$. The general case will follow if we apply the theorem for $m = 1$ several times.

Let $t \in G$ be an arbitrary element. Passing to a new relative generating set $X' = X \cup \{t\}$ if necessary, we may assume that $t \in X$. By Lemma 2.3 there are noncommensurable elements $h_1, h_2 \in H^0$ such that $E_G(h_1)$ and $E_G(h_2)$ are cyclic. According to Theorem 2.1, G is hyperbolic relative to $\{H_\lambda\}_{\lambda \in \Lambda} \cup E_G(h_1) \cup E_G(h_2)$.

Let μ, ε, ρ be constants such that the conclusions of Lemmas 5.1 and 6.3 are satisfied for $\lambda = 1/3$, $c = 2$, $N = 1$. By Theorem 7.5, there are n and $m_1, \dots, m_n$ such that the set $\mathcal{R}$ of all cyclic shifts and their inverses of the word

$$R \equiv t h_1^{m_1} h_2^{m_1} \dots h_1^{m_n} h_2^{m_n}$$

in the alphabet $\mathcal{A} = X \cup \mathcal{H} \cup (E_G(h_1) \setminus \{1\}) \cup (E_G(h_2) \setminus \{1\})$ satisfies the $C(1/3, 2, \varepsilon, \mu, \rho)$-condition (here $h_i^{m_j}$ is regarded as a letter in $E_G(h_i) \setminus \{1\}$, $i = 1, 2$, $j = 1, \dots, n$). Indeed it suffices to choose large enough n and $m_1, \dots, m_n$ satisfying $m_i \neq \pm m_j$ whenever $i \neq j$. Let G_1 be the quotient of G obtained by imposing the relation $R = 1$ and η the corresponding natural homomorphism.

By Lemma 5.1, G_1 is hyperbolic relative to the images of $H_\lambda, \lambda \in \Lambda$ and $E_G(h_1)$, $E_G(h_2)$. As any elementary group is hyperbolic, G_1 is also hyperbolic relative to $\{\eta(H_\lambda)_{\lambda \in \Lambda}\}$ according to Lemma 8.2. The inclusion $\eta(t) \in \eta(H)$ follows immediately from the equality $R = 1$ in G_1. The third assertion of the theorem follows from Lemma 5.1 b) as any element from the union $\bigcup_{\lambda \in \Lambda} H_\lambda$ has length 1.

Similarly as η is injective on $E_G(h_1) \cup E_G(h_2)$, $\eta(h_1)$ and $\eta(h_2)$ are elements of infinite order. Note also that $\eta(h_1)$ and $\eta(h_2)$ are not commensurable in G_1. Indeed otherwise the intersection $\big(\eta(E_G(h_1))\big)^g \cap \eta(E_G(h_2))$ is infinite for some $g \in G$ contradictory to the first assertion of Lemma 8.3. Assume now that $g \in E_{G_1}(\eta(h_1))$, where $E_{G_1}(\eta(h_1))$ is the maximal elementary subgroup of G_1 containing $\eta(h_1)$. By the first assertion of Theorem 2.1, $\big(\eta(h_1^m)\big)^g = \eta(h_1^{\pm m})$ for a certain $m \neq 0$. Therefore, $\big(\eta(E_G(h_1))\big)^g \cap \eta(E_G(h_1))$ contains $\eta(h_1^m)$ and in particular this intersection is infinite. By the second assertion of Lemma 8.3, this means that $g \in \eta(E_G(h_1))$. Thus we proved that $\eta(E_G(h_1)) = E_{G_1}(\eta(h_1))$. The same is true for h_2. Finally, using injectivity of η on $E_G(h_1) \cup E_G(h_2)$ again, we obtain

$$E_{G_1}(\eta(h_1)) \cap E_{G_1}(\eta(h_2)) = \eta(E_G(h_1)) \cap \eta(E_G(h_2))$$
$$= \eta\big(E_G(h_1) \cap E_G(h_2)\big) = \{1\}.$$

This means that the image of H is a suitable subgroup of G_1. To complete the proof it remains to note that the last assertion of the theorem follows from Lemma 6.3.

$\square$

9. Appendix. The proof of Lemma 4.4

Following the referee's recommendation, we provide here the proof of Lemma 4.4, which is a particular case of Lemma 9.7 below. As we mentioned in Section 4, Lemma 4.4 (as well as Lemma 9.7) is, in fact, proved in [23], although it is stated in a slightly different way there. The proof below follows [23] with little improvements. We stress that all results of this section should be credited to Olshanskii.

Throughout the appendix, let G denote a group with a presentation (4). Suppose that the Cayley graph $\Gamma(G, \mathscr{A})$ of G is hyperbolic.

We start with an auxiliary result. Recall that a graph is *simple* if it has no loops and multiple edges.

LEMMA 9.1. *Let Ξ be a simple planar graph. Suppose that some nonnegative weights $\nu(o)$, $\nu(e)$ are assigned to each vertex o and each edge e of Ξ. Assume also that there exists a constant a such that $\nu(e) \le a\nu(o)$ for any incident edge e and vertex o. Then the sums ν_0, ν_1 of the weights of all vertices and edges of Φ, respectively, satisfy $\nu_1 \le 5a\nu_0$.*

Proof. By the well-known consequence of the Euler Formula, every simple planar graph contains a vertex of degree at most 5. The statement of the lemma easily follows from this by induction. $\square$

The proof of Lemma 4.4 is divided into a sequence of lemmas. In what follows, let G be a group given by (4) such that the Cayley graph $\Gamma(G, \mathscr{A})$ of G is hyperbolic. We fix $\lambda \in (0, 1]$, $c \ge 0$, $\mu \in (0, 1/16)$ and take

(36) $$\varepsilon > c_1 + 2\kappa,$$

where $c_1 = c_1(\delta)$ and $\kappa = \kappa(\delta, \lambda, c)$ are the constants provided by Lemmas 3.7 and 3.1 respectively. Let (10) be a presentation that satisfies the $C(\varepsilon, \mu, \lambda, c, \rho)$-condition, where ρ is sufficiently large. (The exact lower bound for ρ, which ensures that all arguments below are correct, can be easily extracted from the proofs.)

Let also Δ be a diagram over (10). Below we prove Lemma 4.4 by induction on the number of $\mathscr{R}$-cells in Δ. To complete the inductive step, we will have to deal with the case when $\partial\Delta$ consists of at most 4 quasi-geodesic segments. More precisely, let $\partial\Delta = q_1 \dots q_r$ for some $1 \le r \le 4$, where $q_1, \dots, q_r$ are (λ, c)-quasi-geodesic. We call the subpaths $q_1, \dots, q_r$ *sections* of $\partial\Delta$.

Definition 9.2. A set $\mathcal{M}$ of ε-contiguity subdiagrams of cells to cells or cells to sections of $\partial\Delta$ is called *distinguished*, if

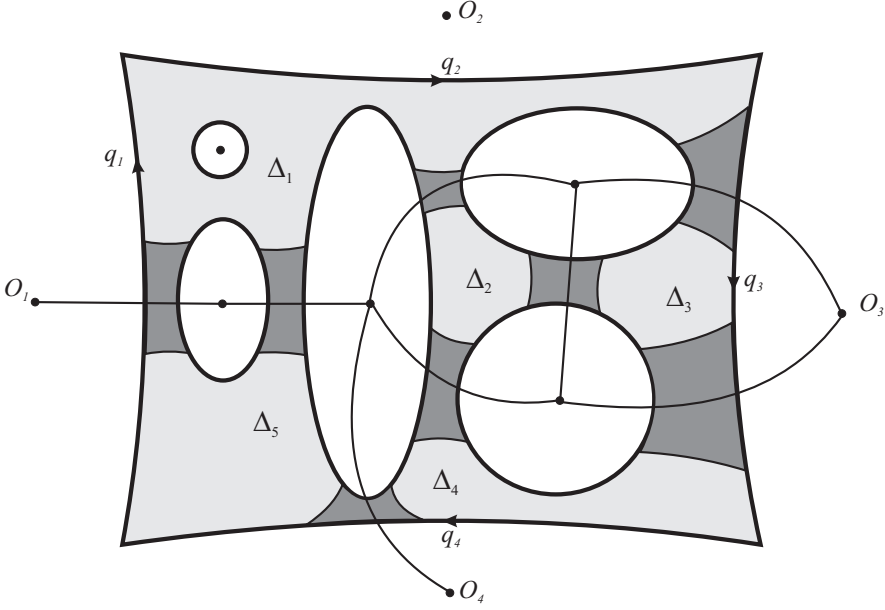(a) Distinct subdiagrams in $\mathcal{M}$ are disjoint.

Figure 7. A distinguished set of ε-contiguity subdiagrams (dark gray) in a diagram Δ with five $\mathcal{R}$-cells (white) and the corresponding graph $\Phi'_{\mathcal{M}}$.

(b) The sum of lengths of contiguity arcs of all subdiagrams from $\mathcal{M}$ is not less than the same sum for any other collections satisfying (a).

(c) $\mathcal{M}$ contains minimal number of subdiagrams among all collections satisfying (a), (b).

Further let $\Gamma \in \mathcal{M}$ be a ε-contiguity subdiagram whose boundary is decomposed according to (11) and (12). The paths s_1, s_2 are called *side arcs* of Γ. If Γ is an ε-contiguity subdiagram of a cell Π_1 to a section of $\partial\Delta$, the contiguity arc q_1 of $\partial\Pi$ (and all its edges) is called *outer*. If Γ is an ε-contiguity subdiagram of a cell Π_1 to a cell Π_2, q_1 (and all its edges) is called *inner*. The edges of Π_1 that are neither inner nor outer are called *unbound*. Every maximal subpath of $\partial\Pi_1$ consisting of unbound edges is called an *unbound arc* of Π_1. The notion of an unbound arc of a section of $\partial\Delta$ is defined in the same way.

To each distinguished set $\mathcal{M}$ of ε-contiguity diagrams, we associate a planar graph $\Phi_{\mathcal{M}}$ as follows. Choose a point $o(\Pi)$ inside every $\mathcal{R}$-cell Π of Δ. The set of vertices of $\Phi_{\mathcal{M}}$ consists of all such $o(\Pi)$. If there exists an ε-contiguity subdiagram $\Gamma \in \mathcal{M}$ of a cell Π_1 to a cell Π_2, the two vertices $o(\Pi_1)$ and $o(\Pi_2)$ are connected by an (unoriented) edge through Γ. Further let $\Phi'_{\mathcal{M}}$ be the graph obtained from $\Phi_{\mathcal{M}}$ in the following way. For each $i = 1, \ldots, r$, we add to $\Phi_{\mathcal{M}}$ a vertex O_i outside of Δ and for each ε-contiguity subdiagram Γ of a cell Π to q_i, we connect $o(\Pi)$ to O_i by an edge passing through Γ. Clearly $\Phi'_{\mathcal{M}}$ is also planar (see Figure 7).

Let now Δ be a reduced diagram over (10), which has $n \geq 1$ $\mathcal{R}$-cells. The next four results are proved under the following additional assumption. It will be eliminated later in Corollary 9.6.

(∗) *For any distinguished system of ε-contiguity subdiagrams $\mathcal{M}$ in Δ, the graph $\Phi_{\mathcal{M}}$ is simple and inside every 2-gon of $\Phi'_{\mathcal{M}}$, there is a vertex of $\Phi_{\mathcal{M}}$.*

Let also $\mathcal{M}$ be a distinguished system of ε-contiguity subdiagrams of Δ. Cutting off all $\mathcal{R}$-cells and subdiagrams $\Gamma \in \mathcal{M}$, we obtain a set of diagrams $\Delta_1, \ldots, \Delta_d$ over (4) (see Figure 7). Each of them may have holes. The boundary $\partial \Delta_i$ of each diagram may be thought of as a union of n_i arcs, where each arc is of one of the following types:

(A1) An unbound arc of an $\mathcal{R}$-face.

(A2) An unbound arc of $\partial \Delta$.

(A3) A side arc of some ε-contiguity subdiagram from $\mathcal{M}$.

LEMMA 9.3. *Suppose that Δ satisfies (∗). In the notation introduced above, we have $\sum_{i=1}^{d} n_i \leq 53n$.*

Proof. Let v, e, f denote the number of vertices, edges, and regions of (the planar realization of) $\Phi'_{\mathcal{M}}$, respectively. By (∗), every region of it (except possibly for the outer one) has degree at least 3. Hence $f \leq 2e/3 + 1$. By the Euler formula, we have $e \leq v + f - 2 \leq v + 2e/3 - 1$. This implies $e \leq 3(v-1) \leq 3(n+3) \leq 12n$.

Since unbound arcs of boundaries of $\mathcal{R}$-cells and sections of $\partial \Delta$ are separated by contiguity arcs, the total number of arcs of type (A1) or (A2) in Δ is not greater than

$$2|\mathcal{M}| + n + r \leq 2e + n + 4 \leq 24n + n + 4 \leq 29n.$$

Clearly the number of arcs of type (A3) is at most $2|\mathcal{M}| = 2e \leq 24n$. So the total number of arcs of types (A1)–(A3) is at most $53n$. $\qquad\square$

All results below are proved up to passing from Δ to an $\mathcal{O}$-equivalent diagram (see the definition before Lemma 4.4).

LEMMA 9.4. *Suppose that Δ satisfies (∗). Let S denote the sum of lengths of all unbound arcs of type (A1) in Δ. Then $S < n\sqrt{\rho}$.*

Proof. Let S_i denote the sum of lengths of all arcs of type (A1) in $\partial \Delta_i$, $i = 1, \ldots, d$. Assume that $S \geq n\sqrt{\rho}$. Then

$$(37) \qquad\qquad S_i \geq n_i \sqrt{\rho}/60$$

for some i. Indeed, otherwise we have

$$S = \sum_{i=1}^{d} S_i \leq \frac{\sqrt{\rho}}{60} \sum_{i=1}^{d} n_i < n\sqrt{\rho}$$

by Lemma 9.3.

For every $i = 1, \ldots, d$, the diagram Δ_i has at most n_i holes. Hence we can cut it by $l \leq n_i$ paths $t_1, \ldots, t_l$ into a simply connected diagram $\widetilde{\Delta_i}$. We assume that the collection of cutting paths $t_1, \ldots, t_l$ is chosen so that the sum $\sum\limits_{i=1}^{l} l(t_i)$ is minimal. The boundary of $\widetilde{\Delta_i}$ decomposes into k_i subpaths, each of which is either $t_j^{\pm 1}$ for some $j = 1, \ldots, l$, or an arc of one of the types (A1)–(A3) (or a part of such an arc arising after cutting along $t_1, \ldots, t_l$). Therefore,

$$(38) \qquad\qquad\qquad k_i \leq 4n_i.$$

Note also that passing to a diagram which is $\mathbb{O}$-equivalent to Δ if necessary, we may assume that the labels of t_i's are geodesic in G without loss of generality. Indeed if w_i is a geodesic word representing the same element as $\mathbf{Lab}\,(t_i)$ in G, let Σ_i be a diagram over (4) with boundary label $w_i (\mathbf{Lab}\,(t_i))^{-1}$. Further let Ξ_i be the diagram obtained by gluing Σ_i and its mirror copy along w_i. Clearly $\mathbf{Lab}\,\partial(\Xi_i) = \mathbf{Lab}\,(t_i)(\mathbf{Lab}\,(t_i))^{-1}$. We can use 0-refinement (see the subsection "van Kampen diagrams" in Section 3) to create a copy t_i' of the paths t_i in Δ with the same label and endpoints such that $t_i (t_i')^{-1}$ bounds a subdiagram consisting of 0-cells. Further we cut this subdiagram and fill in the obtained hole with Ξ_i. After this transformation, the vertices $(t_i)_-$ and $(t_i)_+$ are connected by a path whose label w_i is geodesic in G, and we can replace t_i with that paths. Note that this transformation does not affect $\partial\Delta$, $\mathfrak{R}$-cells, and distinguished ε-contiguity subdiagrams.

Similarly we may assume that for every (sub)arc p of an arc of type (A1)–(A3) in $\partial\widetilde{\Delta_i}$, the diagram $\widetilde{\Delta_i}$ contains a path $\bar{p}$ with the same endpoints as p such that $\mathbf{Lab}\,(\bar{p})$ is geodesic in G and $p(\bar{p})^{-1}$ bounds a subdiagram over (4). Since $\widetilde{\Delta_i}$ is a simply connected diagram over (4), its 1-skeleton can be naturally mapped to $\Gamma(G, \mathcal{A})$ by a map preserving labels and orientation. In what follows we keep the same notation $t_i^{\pm 1}$, p, etc., for the images of the paths $t_i^{\pm 1}$, p, etc., in $\Gamma(G, \mathcal{A})$ (although t_i^{-1} is not the inverse path of t_i there). In particular, p and $\bar{p}$ belong to the closed κ-neighborhoods of each other in $\Gamma(G, \mathcal{A})$ by Lemma 3.1.

In the notation of Lemma 3.7, let N_2 consist of images of segments $t_i^{\pm 1}$, $i = 1, \ldots, l$, in $\Gamma(G, \mathcal{A})$ and images of $\bar{p}$ for (sub)arcs p of type (A2). Further let N_1 and N_3 consist of images of $\bar{p}$ for (sub)arcs p of type (A1) and (A3), respectively.

Clearly $\sigma_3 \leq \varepsilon k_i$. On the other hand, by (37) and (38) we obtain

$$\sigma_1 = \sum_{\bar{p} \in N_1} l(p_1) \geq \sum_{\bar{p} \in N_1} (\lambda l(p) - c) \geq \lambda S_i - k_i c > \lambda n_i \sqrt{\rho}/60 - k_i c$$
$$\geq \lambda k_i \sqrt{\rho}/240 - k_i c = k_i (\lambda \sqrt{\rho}/240 - c).$$

Taking ρ large enough, we may assume that $a = \lambda \sqrt{\rho}/240 - c > \max\{1000\varepsilon, c_2\}$. Hence by Lemma 3.7, there exist subsegments $\bar{q}_j$, $j = 1, 2$, of some $\bar{p}_1 \in N_1$ and $\bar{p}_2 \in N_1 \cup N_2$, respectively, such that $l(p_j) \geq 10^{-3}a$, $j = 1, 2$, and the distance
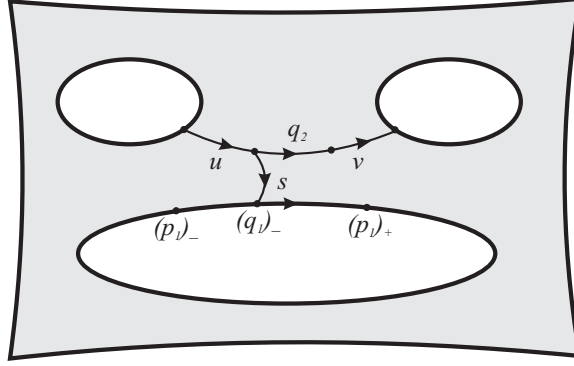
Figure 8. Replacing a cutting path with a shorter one.

between the respective endpoints of p_1 and p_2 is at most c_1. This means that there exist subsegments q_j of p_j, $j = 1, 2$, of lengths at least $10^{-3}a - 2\kappa$ such that

$$(39) \qquad \max\{\mathrm{dist}_{X \cup \mathscr{H}}((q_1)_-, (q_2)_-), \ \mathrm{dist}_{X \cup \mathscr{H}}((q_1)_+, (q_2)_+)\} \le 2\kappa + c_1 < \varepsilon.$$

The latter inequality uses (36). Now there are two cases to consider.

Case 1. $\bar{p}_2 \in N_1$. That is, q_1 and q_2 are subpaths of boundaries of some $\mathscr{R}$-cells in Δ. After passing to an $\mathbb{O}$-equivalent diagram if necessary, we may assume that there exists a subdiagram Γ of Δ such that $\partial\Gamma = q_1 s_1 q_2^{-1} s_2$, where $l(s_j) < \varepsilon$, $j = 1, 2$, and Γ does not intersect $\mathscr{R}$-cells and distinguished ε-contiguity subdiagrams. This contradicts the maximality of $\mathcal{M}$.

Case 2. $\bar{p}_2 \in N_2$. If q_2 is a subpath of $\partial\Delta$, we get a contradiction as above by constructing a new ε-contiguity subdiagram of an $\mathscr{R}$-cell to $\partial\Delta$. Thus we only need to deal with the case $p_2 = t_j^{\pm 1}$. Without loss of generality, we may assume that $p_2 = t_j$. Let $t_j = u q_2 v$. By (39), after passing to an $\mathbb{O}$-equivalent diagram if necessary, we can find a paths s of lengths at most ε in Δ connecting $(q_2)_-$ to $(q_1)_-$ (see Figure 8). Note that

$$l(us) \le l(t_j) - l(q_2) + \varepsilon \le l(t_j) - (10^{-3}(\lambda\sqrt{\rho}/240 - c) - 2\kappa) + \varepsilon < l(t_j)$$

if ρ is large enough. This contradicts our assumption that $\sum_{i=1}^{l} l(t_i)$ is minimal.

Thus in both cases we obtain a contradiction. Hence $S < n\sqrt{\rho}$. □

LEMMA 9.5. *Suppose that Δ satisfies* (∗). *Let Σ_0 and Σ denote the sum of lengths of all outer arcs and the sum of perimeters of all $\mathscr{R}$-cells in Δ, respectively. Then $\Sigma_0 > (1 - 11\mu)\Sigma$.*

Proof. For each $\mathscr{R}$-cell Π of Δ, we assign the weight $\nu(o) = l(\partial\Pi)$ to the corresponding vertex o of the estimating graph $\Phi_{\mathcal{M}}$. Clearly the sum of weights of all vertices in $\Phi_{\mathcal{M}}$ equals Σ. Further let $\Gamma \in \mathcal{M}$ be a distinguished ε-contiguity subdiagram with $\partial\Gamma = s_1 q_1 s_2 q_2$, where q_1, q_2 are the contiguity arcs. We assign

the weight $\nu(e) = l(q_1) + l(q_2)$ to the edge e corresponding to Γ. Note that the sum of weights of all edges in $\Gamma \in \mathcal{M}$ is equal to the sum of lengths Σ_{inn} of all inner arcs in Δ.

By Lemma 4.3, $\nu(e) \le 2\mu\nu(o)$ whenever o and e are incident. Hence $\Sigma_{inn} \le 10\mu\Sigma$ by Lemma 9.1 and (∗). By Lemma 9.4, we have $S < \Sigma/\sqrt{\rho}$, where S is the sum of lengths of all unbound arcs, since the perimeter of every $\mathcal{R}$-cell in Δ is at least ρ by the $C(\varepsilon, \mu, \lambda, c, \rho)$-condition. Therefore,

$$\Sigma_0 = \Sigma - \Sigma_{int} - S > \Sigma - 10\mu\Sigma - \Sigma/\sqrt{\rho} > (1 - 11\mu)\Sigma$$

if $\rho > \mu^{-2}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

The following corollary is immediate.

COROLLARY 9.6. *Suppose that* Δ *satisfies* (∗). *Passing to an* $\mathbb{O}$-*equivalent diagram if necessary, we can find an* $\mathcal{R}$-*cell* Π *of* Δ *and disjoint* ε-*contiguity subdiagrams* $\Gamma_{i,j}$ *of* Π *to sections* q_j, $j = 1, \ldots, r$ *of* $\partial\Delta$ *such that*

$$\sum_{i,j}(\Pi, \Gamma_{i,j}, q_j) > 1 - 11\mu.$$

We are now ready to eliminate assumption (∗) and prove the main result of the appendix, of which Lemma 4.4 is a particular case corresponding to $r = 1$. For convenience, we recall all assumptions here.

LEMMA 9.7. *Let* G *be a group with a presentation* (4). *Suppose that the Cayley graph* $\Gamma(G, \mathcal{A})$ *of* G *is hyperbolic. Then for any* $\lambda \in (0, 1]$, $c \ge 0$, *and* $\mu \in (0, 1/16]$, *there exist* $\varepsilon \ge 0$ *and* $\rho > 0$ *with the following property. Let* $\mathcal{R}$ *be a symmetrized set of words in* $\mathcal{A}$ *satisfying the* $C(\varepsilon, \mu, \lambda, c, \rho)$-*condition,* Δ *a reduced van Kampen diagram over the presentation* (10) *such that* $\partial\Delta = q_1 \ldots q_r$ *for some* $1 \le r \le 4$, *where* $q_1, \ldots, q_r$ *are* (λ, c)-*quasi-geodesic. Assume that* Δ *has at least one* $\mathcal{R}$-*cell. Then up to passing to an* $\mathbb{O}$-*equivalent diagram, the following conditions hold.*

(a) *The diagram* Δ *satisfies* (∗).

(b) *There is an* $\mathcal{R}$-*cell* Π *of* Δ *and disjoint* ε-*contiguity subdiagrams* Γ_j *of* Π *to sections* q_j, $j = 1, \ldots, r$, *of* $\partial\Delta$ *(some of them may be absent) such that*

$$\sum_{j=1}^{r}(\Pi, \Gamma_j, q_j) > 1 - 13\mu.$$

Proof. We proceed by induction on n, the number of $\mathcal{R}$-cells in Δ.

(a) Suppose that there are multiple edges in $\Phi_{\mathcal{M}}$. That is, there are two distinguished ε-contiguity subdiagrams Θ_1 and Θ_2 between some $\mathcal{R}$-cells Π_1 and Π_2. Consider a subdiagram Ξ in Δ such that: (i) $\partial\Xi = s_1 t_1 s_2 t_2$, where s_j is a side arc of Θ_j and t_j is a subpaths of $\partial\Pi_j$, $j = 1, 2$; (ii) Ξ does not contain Π_1 and Π_2. By the definition of $\mathcal{M}$, we can not include Θ_1 and Θ_2 into a single ε-contiguity subdiagram of Π_1 to Π_2. This means that Ξ contains at least one $\mathcal{R}$-cell. On the

other hand, the number of $\mathcal{R}$-cells in Ξ is smaller than n according to (ii). Hence by the inductive assumption, there is an $\mathcal{R}$-cell Π in Ξ and ε-contiguity subdiagrams $\Gamma_1, \ldots, \Gamma_4$ of Π to s_1, t_1, s_2, t_2, respectively, such that

$$(\Pi, \Gamma_1, s_1) + \cdots + (\Pi, \Gamma_4, t_2) > 1 - 13\mu.$$

On the other hand $(\Pi, \Gamma_2, t_1) + (\Pi, \Gamma_4, t_2) < 2\mu$ by Lemma 4.3 since Γ_2 and Γ_4 are ε-contiguity subdiagrams of Π to Π_1 and Π_2, respectively. Further if u is a contiguity arc of Π to s_1, then $l(u) \le \lambda^{-1}(3\varepsilon + c)$ since u is (λ, c) quasi-geodesic and $l(s_1) \le \varepsilon$. Therefore, $(\Pi, \Gamma_1, s_1) = l(u)/l(\partial\Pi) < l(u)/\rho < \mu/2$ if ρ is large enough. Similarly $(\Pi, \Gamma_3, s_2) < \mu/2$. Thus

$$(\Pi, \Gamma_1, s_1) + \cdots + (\Pi, \Gamma_4, t_2) < 3\mu.$$

We obtain a contradiction since $1 - 13\mu > 3\mu$ for $\mu < 1/16$.

Thus the graph $\Phi_{\mathcal{M}}$ does not have multiple edges. Arguing as above, it is easy to show that $\Phi_{\mathcal{M}}$ can not contain loops either. The only difference is that the subdiagram Ξ will be bounded by st, where $l(s) < \varepsilon$ and t is a subpaths of $\partial\Pi_1$. (Here Π_1 is the $\mathcal{R}$-cell of Δ such that there is a lop incident to the corresponding vertex in $\Phi_{\mathcal{M}}$.)

Finally inside every 2-gon ef of $\Phi'_{\mathcal{M}}$ there is a vertex of $\Phi_{\mathcal{M}}$ since otherwise one can include the ε-contiguity subdiagrams corresponding to the edges e and f of $\Phi'_{\mathcal{M}}$ into a single ε-contiguity subdiagram in the obvious way, contrary to the definition of $\mathcal{M}$.

(b) By (a), we can choose an $\mathcal{R}$-cell Π in Δ (passing to an $\mathbb{O}$-equivalent diagram if necessary) and subdiagrams $\Gamma_{i,j}$ satisfying the inequality in Corollary 9.6. Let us consider the subdiagram Γ^1 of Δ such that: (i) $\partial\Gamma^1 = s_1 t_1 s_2 t_2$, where s_1, s_2 are side arcs of some of $\Gamma_{i,1}$'s and t_1, t_2 are subpaths of $\partial\Pi$ and the section q_1 of $\partial\Delta$, respectively; (ii) Γ^1 contains all $\Gamma_{i,1}$'s. Let m_1 be the number of $\mathcal{R}$-cells in Γ^1. We similarly construct $\Gamma^2, \ldots, \Gamma^r$ and define $m_2, \ldots, m_r$.

Suppose that the cell Π is chosen in such a way that the sum $m(\Pi) = m_1 + \cdots + m_r$ is minimal among all cells satisfying the condition

$$\sum_{i,j}(\Pi, \Gamma_{i,j}, q_j) > 1 - 11\mu.$$

If $m_1 = \cdots = m_r = 0$, then each of the sets $\{\Gamma_i, 1\}, \ldots, \{\Gamma_i, 1\}$ consists of at most one diagram. Indeed otherwise one could include at least 2 of them into a single ε-contiguity subdiagram, which contradicts the definition of $\mathcal{M}$.

Thus we may assume that $m_1 > 0$. Then by Corollary 9.6 and the inductive assumption, the subdiagram Γ^1 contains an $\mathcal{R}$-cell Π' and ε-contiguity subdiagrams $\Gamma'_{i,1}, \ldots, \Gamma'_{i,4}$ of Π' to s_1, t_1, s_2, t_2, respectively, such that

$$(40) \qquad \sum_{i=1}^{k_1}(\Pi', \Gamma'_{i,1}, s_1) + \cdots + \sum_{i=1}^{k_4}(\Pi', \Gamma'_{i,4}, t_2) > 1 - 11\mu.$$

Note that $k_2 \leq 1$ since all $\Gamma'_{i,2}$'s are ε-contiguity subdiagrams of Π' to Π and Δ satisfies $(*)$ by (a). If $k_2 = 1$, we have $(\Pi', \Gamma'_{1,2}, t_1) < \mu$ by Lemma 4.3. Further if $k_1 \geq 1$, we can construct an ε-contiguity subdiagram Γ of Π' to t_1 such that: (i) $\partial\Gamma = st$, where $l(s) \leq 3\varepsilon$ and t is a subpaths of $\partial\Pi'$; (ii) Γ contains at least one $\mathcal{R}$-cell. This leads to a contradiction as in part (a). Thus $k_1 \leq 1$ and if $k_1 = 1$ we obtain $(\Pi', \Gamma'_{i,1}, s_1) < \mu/2$ as in part (a). Similarly $k_3 \leq 1$ and if $k_3 = 1$, we have $(\Pi', \Gamma'_{i,3}, s_2) < \mu/2$. These inequalities together with (40) imply $\sum_{i=1}^{k_4} (\Pi', \Gamma'_{i,4}, t_2) > 1 - 13\mu$. However $m(\Pi') < m(\Pi)$ since the cell Π' counts in $m(\Pi)$ but not in $m(\Pi')$. This contradicts the choice of Π. $\qquad\square$

References

[1] G. ARZHANTSEVA, A. MINASYAN, and D. OSIN, The SQ-universality and residual properties of relatively hyperbolic groups, *J. Algebra* **315** (2007), 165–177. MR 2008g:20095 Zbl 1132.20022

[2] G. ARZHANTSEVA, M. R. BRIDSON, T. JANUSZKIEWICZ, I. J. LEARY, A. MINASYAN, and J. ŚWIĄTKOWSKI, Infinite groups with fixed point properties, *Geom. Topol.* **13** (2009), 1229–1263. MR 2496045 Zbl 05530204

[3] G. BAUMSLAG, A. G. MYASNIKOV, and V. SHPILRAIN, Open problems in combinatorial group theory, in *Combinatorial and Geometric Group Theory* (*New York*, 2000/*Hoboken, NJ*, 2001), *Contemp. Math.* **296**, Amer. Math. Soc., Providence, RI, second ed., 2002, pp. 1–38. MR 2003h:20001 Zbl 1065.20042

[4] I. BELEGRADEK and D. OSIN, Rips construction and Kazhdan property (T), *Groups Geom. Dyn.* **2** (2008), 1–12. MR 2009c:20076 Zbl 1152.20039

[5] B. H. BOWDITCH, Relatively hyperbolic groups, preprint, 1999.

[6] M. R. BRIDSON and A. HAEFLIGER, *Metric Spaces of Non-Positive Curvature*, Grundl. Math. Wissen. **319**, Springer-Verlag, New York, 1999. MR 2000k:53038 Zbl 0988.53001

[7] F. DAHMANI, Combination of convergence groups, *Geom. Topol.* **7** (2003), 933–963. MR 2005g:20063 Zbl 1037.20042

[8] B. FARB, Relatively hyperbolic groups, *Geom. Funct. Anal.* **8** (1998), 810–840. MR 99j:20043 Zbl 0985.20027

[9] S. M. GERSTEN and H. B. SHORT, Rational subgroups of biautomatic groups, *Ann. of Math.* **134** (1991), 125–158. MR 92g:20092 Zbl 0744.20035

[10] E. GHYS and P. DE LA HARPE (eds.), *Sur les Groupes Hyperboliques d'après Mikhael Gromov*, Progr. Math. **83**, Birkhäuser, Boston, MA, 1990. MR 92f:53050

[11] M. GROMOV, Hyperbolic groups, in *Essays in Group Theory*, Math. Sci. Res. Inst. Publ. **8**, Springer-Verlag, New York, 1987, pp. 75–263. MR 89e:20070 Zbl 0634.20015

[12] V. S. GUBA, A finitely generated complete group, *Izv. Akad. Nauk* SSSR *Ser. Mat.* **50** (1986), 883–924, In Russian; translated in *Math.* USSR-*Izv.* **29** (1987), 233–277. MR 88e:20034 Zbl 0631.20025

[13] G. HIGMAN, B. H. NEUMANN, and H. NEUMANN, Embedding theorems for groups, *J. London Math. Soc.* **24** (1949), 247–254. MR 11,322d Zbl 0034.30101

[14] S. V. IVANOV, On some finiteness conditions in semigroup and group theory, *Semigroup Forum* **48** (1994), 28–36. MR 94k:20102 Zbl 0808.20046

[15] S. V. IVANOV and A. YU. OL′SHANSKII, Some applications of graded diagrams in combinatorial group theory, in *Groups—St. Andrews* 1989, Vol. 2, *London Math. Soc. Lecture Note Ser.* **160**, Cambridge Univ. Press, Cambridge, 1991, pp. 258–308. MR 92j:20022

[16] R. C. LYNDON and P. E. SCHUPP, *Combinatorial Group Theory*, Springer-Verlag, New York, 1977. MR 58 #28182 Zbl 0368.20023

[17] V. D. MAZUROV and E. I. KHUKHRO (eds.), *The Kourovka Notebook. Unsolved Problems in Group Theory*, fifteenth augmented ed., Rossiĭskaya Akademiya Nauk Sibirskoe Otdelenie, Institut Matematiki im. S. L. Soboleva, Novosibirsk, 2002. MR 2004f:20001a

[18] K. V. MIKHAJLOVSKII and A. YU. OL′SHANSKII, Some constructions relating to hyperbolic groups, in *Geometry and Cohomology in Group Theory* (Durham, 1994), *London Math. Soc. Lecture Note Ser.* **252**, Cambridge Univ. Press, Cambridge, 1998, pp. 263–290. MR 2000h:20083

[19] A. MINASYAN, Groups with finitely many conjugacy classes and their automorphisms, *Comment. Math. Helv.* **84** (2009), 259–296. MR 2495795 Zbl 05545771

[20] A. MINASYAN and D. OSIN, Normal automorphisms of hyperbolic groups, preprint, 2008, *Trans. Amer. Math. Soc.*, to appear. arXiv 0809.2408

[21] A. YU. OL′SHANSKIĬ, *Geometry of Defining Relations in Groups*, Mathematics and its Applications (*Soviet Series*) **70**, Kluwer Academic Publishers Group, Dordrecht, 1991. MR 93g:20071 Zbl 0830.20053

[22] _____ , Periodic quotient groups of hyperbolic groups, *Mat. Sb.* **182** (1991), 543–567, In Russian; translated in *Math*. USSR *Sbornik* **72** (1992), 519–541. MR 92d:20050

[23] _____ , On residualing homomorphisms and *G*-subgroups of hyperbolic groups, *Internat. J. Algebra Comput.* **3** (1993), 365–409. MR 94i:20069

[24] _____ , On the Bass-Lubotzky question about quotients of hyperbolic groups, *J. Algebra* **226** (2000), 807–817. MR 2001i:20069 Zbl 0957.20026

[25] D. V. OSIN, Factorizable groups and finiteness conditions, in preparation.

[26] _____ , Elementary subgroups of relatively hyperbolic groups and bounded generation, *Internat. J. Algebra Comput.* **16** (2006), 99–118. MR 2007b:20092 Zbl 1100.20033

[27] _____ , Relative Dehn functions of amalgamated products and HNN-extensions, in *Topological and Asymptotic Aspects of Group Theory*, *Contemp. Math.* **394**, Amer. Math. Soc., Providence, RI, 2006, pp. 209–220. MR 2007g:20042 Zbl 1111.20036

[28] _____ , Relatively hyperbolic groups: intrinsic geometry, algebraic properties, and algorithmic problems, *Mem. Amer. Math. Soc.* **179** (2006), vi+100. MR 2006i:20047

E-mail address: denis.osin@gmail.com

STEVENSON CENTER 1326, DEPARTMENT OF MATHEMATICS, VANDERBILT UNIVERSITY, NASHVILLE, TN 37240, UNITED STATES

Editorial correspondence

Papers submitted for publication and editorial correspondence should be addressed to Maureen Schupsky, Annals of Mathematics, Fine Hall-Washington Road, Princeton University, Princeton, NJ, 08544-1000 U.S.A. The e-mail address is annals@math.princeton.edu.

Preparing and submitting papers

The Annals requests that all papers include an abstract of about 150 words which explains to the nonspecialist mathematician what the paper is about. It should not make any reference to the bibliography. Authors are encouraged to initially submit their papers electronically and in PDF format. Please send the file to: annals@math.princeton.edu or to the Mathematics e-print arXiv: front.math.ucdavis.edu/submissions. If a paper is submitted through the arXiv, then please e-mail us with the arXiv number of the paper.

Proofs

A PDF file of the galley proof will be sent to the corresponding author for correction. If requested, a paper copy will also be sent to the author.

Offprints

Authors of single-authored papers will receive 30 offprints. (Authors of papers with one co-author will receive 15 offprints, and authors of papers with two or more co-authors will receive 10 offprints.) Extra offprints may be purchased through the editorial office.

Subscriptions

The price for a print and online subscription, or an online-only subscription, is $390 per year for institutions. In addition, there is a postage surcharge of $40 for print subscriptions that are mailed to countries outside of the United States. Individuals interested in subscriptions for their own personal use should contact the publisher at the address below. Subscriptions and changes of address should be sent to Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840 (e-mail: contact@mathscipub.org; phone: 1-510-643-8638; fax: 1-510-295-2608). (Checks should be made payable to "Mathematical Sciences Publishers".)

Back issues and reprints

Orders for missing issues and back issues should be sent to Mathematical Sciences Publishers at the above address. Claims for missing issues must be made within 12 months of the publication date. Online versions of papers published five or more years ago are available through JSTOR (www.jstor.org).

Microfilm

Beginning with Volume 1, microfilm may be purchased from NA Publishing, Inc., 4750 Venture Drive, Suite 400, PO Box 998, Ann Arbor, MI 48106-0998; phone: 1-800-420-6272 or 734-302-6500; email: info@napubco.com, website: www.napubco.com/contact.html.

TABLE OF CONTENTS