

# Cohomology of group theoretic Dehn fillings I: Cohen-Lyndon type theorems

Bin Sun

## Abstract

This is the first paper of three papers in a row aiming to study cohomology of group theoretic Dehn fillings. In the present paper, we prove a particular free product structure, which is termed the Cohen-Lyndon property, of Dehn filling kernels. As an application, we describe the structure of relative relation modules of Dehn fillings.

## 1 Introduction

**1.1 Dehn surgery of 3-manifolds.** In 3-dimensional topology, Dehn surgery is an operation of modifying a 3-manifold by cutting off a solid torus and then gluing it back in a different way. The Lickorish-Wallace theorem, which states that every closed connected orientable 3-manifold can be obtained from the 3-dimensional sphere by performing finitely many Dehn surgeries, serves as a motivation of the study of Dehn surgeries.

The second step of the surgery, called *Dehn filling*, can be formalized as follows. Let  $M$  be a 3-manifold with toral boundary. Topologically distinct ways of gluing a solid torus to  $M$  are parametrized by free homotopy classes of essential simple closed curves on  $\partial M$ , called *slopes*. For a slope  $s$ , the Dehn filling  $M(s)$  is obtained by attaching a solid torus  $S^1 \times D^2$  to  $\partial M$  such that  $\partial D^2$  is mapped to a curve of the slope  $s$ . The following is a special case of Thurston's hyperbolic Dehn filling theorem.

**Theorem 1.1** ([Thu82, Theorem [TH1]]). *Let  $M$  be a compact orientable 3-manifold with toral boundary such that  $M \setminus \partial M$  admits a complete finite-volume hyperbolic structure. Then  $M(s)$  is hyperbolic for all but finitely many slopes  $s$ .*

**1.2 Group theoretic Dehn fillings.** In group theoretic settings, Dehn filling can be generalized as follows. Let  $G$  be a group, let  $H$  be a subgroup of  $G$ , and let  $N$  be a normal subgroup of  $H$ . The *group theoretic Dehn filling* associated with the data  $(G, H, N)$  is the process of forming the quotient group  $G/\langle\langle N \rangle\rangle$ , where  $\langle\langle N \rangle\rangle$  is the normal closure of  $N$  in  $G$ .

Under the assumptions of Theorem 1.1, let  $G = \pi_1(M)$ . The natural map  $\pi_1(\partial M) \rightarrow \pi_1(M)$  is injective. We think of  $\pi_1(\partial M)$  as a subgroup of  $\pi_1(M)$  and let  $H = \pi_1(\partial M)$ . Let  $N \triangleleft H$  be the subgroup generated by the slope  $s$ . Then  $G/\langle\langle N \rangle\rangle = \pi_1(M(s))$  by the Seifert-van Kampen theorem.

Dehn filling is a fundamental tool in group theory. The solution of the virtually Haken conjecture uses Dehn fillings of hyperbolic groups [AGM13]. For a large number of relatively hyperbolic groups, Dehn fillings are used to prove the Farrell-Jones conjecture [ACG18] and solve the isomorphism problem [DG18]. By considering Dehn fillings of hyperbolically embedded subgroups, [DGO17] constructs purely pseudo-Anosov normal subgroups of mapping class groups. Other applications of Dehn fillings can be found in [AGM16, GMS16].

In group theoretic settings, Thurston's theorem was first generalized by Osin [Osi07], and independently by Groves-Manning [GM08] to Dehn fillings of peripheral subgroups of relatively hyperbolic groups. More recently, Dahmani-Guirardel-Osin [DGO17] proved an analog of Thurston's theorem in the more general settings of groups with hyperbolically embedded subgroups (see Theorem 1.4 below and the discussion afterwards). We discuss here some examples and refer to Section 3.2 for the definition. We use  $H \hookrightarrow_h G$  to indicate that  $H$  is a hyperbolically embedded subgroup of  $G$ .

**Example 1.2.** If  $H$  is a peripheral subgroup of a relatively hyperbolic group  $G$ , then  $H \hookrightarrow_h G$  [DGO17, Proposition 2.4]. For example,

- (a) if a group  $G$  decomposes as a free product  $G = A * B$ , then we have  $A \hookrightarrow_h G$  and  $B \hookrightarrow_h G$  [Osi06];
- (b) under the assumptions of Theorem 1.1, we have  $\pi_1(\partial M) \hookrightarrow_h \pi_1(M)$  [Bow12, Far98].

**Example 1.3.** Let  $G$  be a group acting acylindrically on a Gromov hyperbolic space and let  $g$  be a loxodromic element of  $G$ . Then there exists a maximal virtually cyclic subgroup  $E(g) \leq G$  containing  $g$  such that  $E(g) \hookrightarrow_h G$  [DGO17, Corollary 2.9]. In particular, if  $G$  is a hyperbolic group (resp. the mapping class group of a punctured closed orientable surface [DGO17, Theorem 2.19], outer automorphism group of a finite rank non-abelian free group [DGO17, Theorem 2.20]) and  $g$  is a loxodromic (resp. pseudo-Anosov, fully irreducible) element, then  $E(g) \hookrightarrow_h G$ .

Other examples of hyperbolically embedded subgroups can be found in [DGO17].

**Theorem 1.4** ([DGO17, Theorem 2.27]). *Let  $G$  be a group with a subgroup  $H \hookrightarrow_h G$ . Then there exists a finite set  $\mathcal{F} \subset H \setminus \{1\}$  such that if  $N \triangleleft H$  and  $N \cap \mathcal{F} = \emptyset$ , then the natural homomorphism  $H/N \rightarrow G/\langle\langle N \rangle\rangle$  maps  $H/N$  injectively onto a hyperbolically embedded subgroup of  $G/\langle\langle N \rangle\rangle$ .*

Under the assumptions of Theorem 1.1, the above theorem, together with some basic facts about relatively hyperbolic groups, implies that  $\pi_1(M(s))$  is non-virtually-cyclic and word-hyperbolic for all but finitely many slopes  $s$ . Thurston's geometrization conjecture, proved by Perelman, implies that this algebraic statement about  $\pi_1(M(s))$  is equivalent to the hyperbolicity of  $M(s)$ . Thus, the above theorem indeed provides a generalization of Theorem 1.1.

**1.3 Motivation: a question on group cohomology.** Note that in the settings of Thurston's theorem, i.e., if  $G = \pi_1(M)$ ,  $H = \pi_1(\partial M)$ , and  $M(s)$  admits a hyperbolic structure, we have

$$H^*(G/\langle\langle N \rangle\rangle; \cdot) \cong H^*(\pi_1(M(s)); \cdot),$$

which can be computed via  $M(s)$ . Indeed, as  $M(s)$  admits a hyperbolic structure, the universal cover of  $M(s)$  is  $\mathbb{H}^3$ , which is contractible, and thus  $M(s)$  is a model of  $K(G/\langle\langle N \rangle\rangle, 1)$ .

However, there are no analogous methods for Dehn fillings of hyperbolically embedded subgroups. The main question motivating our research is the following.

**Question A.** *For a group  $G$  with a subgroup  $H \hookrightarrow_h G$  and a normal subgroup  $N \triangleleft H$ , what can be said about  $H^*(G/\langle\langle N \rangle\rangle; \cdot)$ ?*

In this series of three papers, we answer this question and discuss some applications. The first task is to understand the structure of  $\langle\langle N \rangle\rangle$ , which is solved by the present paper. In the second paper [Sun19a], we will combine structural results obtained in this paper and the Lyndon-Hochschild-Serre spectral sequence to compute  $H^*(G/\langle\langle N \rangle\rangle; \cdot)$ . In the third paper [Sun19b], we will study cohomological properties of  $G/\langle\langle N \rangle\rangle$  and discuss some applications to acylindrically hyperbolic groups.

## 2 Main results

**2.1 Cohen-Lyndon type theorems for  $\langle\langle N \rangle\rangle$ .** In general,  $\langle\langle N \rangle\rangle$  does not need to have any particular structure. Nevertheless, it turns out that if  $N$  avoids a finite set of bad elements, then  $\langle\langle N \rangle\rangle$  enjoys a nice free product structure. In order to state our main results, we introduce the following terminology.

**Definition 2.1.** Let  $G$  be a group with a subgroup  $H \hookrightarrow_h G$ . We say that a property  $P$  holds *for all sufficiently deep* normal subgroups  $N \triangleleft H$  if there exists a finite set  $\mathcal{F} \subset H \setminus \{1\}$  such that  $P$  holds for all normal subgroups  $N \triangleleft H$  with  $N \cap \mathcal{F} = \emptyset$ .

**Definition 2.2.** Let  $G$  be a group with a subgroup  $H$  and let  $N \triangleleft H$ . We say that the triple  $(G, H, N)$  has the *Cohen-Lyndon property* if there exists a left transversal  $T$  of  $H\langle\langle N \rangle\rangle$  in  $G$  such that  $\langle\langle N \rangle\rangle$  is the free product of its subgroups  $N^t = tNt^{-1}$  for  $t \in T$ , denoted as

$$\langle\langle N \rangle\rangle = \prod_{t \in T}^* N^t.$$

The latter definition is motivated by the following result [CL63, Theorem 4.1], which was later generalized by [EH87, Theorem 1.1] to free products of locally indicable groups.

**Theorem 2.3** (Cohen-Lyndon). *Let  $F$  be a free group and let  $C$  be a maximal cyclic subgroup of  $F$ . Then for all  $f \in C \setminus \{1\}$ , the triple  $(F, C, \langle f \rangle)$  has the Cohen-Lyndon property.*

Note that  $F$  is hyperbolic and  $f$  has infinite order. By Example 1.3, we have  $C = E(f) \hookrightarrow_h F$  and thus the above theorem fits in the general framework of group theoretic Dehn fillings. For general hyperbolically embedded subgroups, a weak version of the Cohen-Lyndon property is given in [DGO17, Theorem 2.27].

**Theorem 2.4** (Dahmani-Guirardel-Osin). *Let  $G$  be a group with a subgroup  $H \hookrightarrow_h G$ . Then for all sufficiently deep  $N \triangleleft H$ , we have*

$$\langle\langle N \rangle\rangle = \prod_{t \in T}^* N^t$$

for some subset  $T \subset G$ .

The main difference between Theorems 2.4 and 2.3 is that in Theorem 2.4,  $T$  is just some subset of  $G$ , instead of being a left transversal of  $H\langle\langle N \rangle\rangle$  in  $G$ . Our result improves Theorem 2.4.

**Theorem 2.5.** *Suppose that  $G$  is a group with a subgroup  $H \hookrightarrow_h G$ . Then  $(G, H, N)$  has the Cohen-Lyndon property for all sufficiently deep  $N \triangleleft H$ .*

In the special case where  $G$  and  $H$  are finitely generated and  $G$  is hyperbolic relative to  $H$ , Theorem 2.5 is proved in [GMS16, Theorem 4.8]. The proofs of [DGO17, Theorem 7.15] and [GMS16, Theorem 4.8] use windmills, very rotating families, and spiderwebs. It is possible to prove Theorem 2.5 with these notions. However, the present paper provides a proof of a different flavor, using surgery on van Kampen diagrams and geodesic polygons of Cayley graphs.

**Remark 2.6.** In fact, we prove Theorem 2.5 in more general settings of a group with a family of *weakly hyperbolically embedded subgroups* (see Definition 3.8). As an application, we obtain Cohen-Lyndon type theorems for graphs of groups, e.g., amalgamated free products and HNN-extensions (see Corollaries 6.8, 6.9, and 6.10).

Combining Theorem 2.5 and Example 1.3, we obtain:

**Corollary 2.7.** *Let  $G$  be a group acting acylindrically on a Gromov hyperbolic space, and let  $g \in G$  be a loxodromic element. Then  $(G, E(g), N)$  has the Cohen-Lyndon property for all sufficiently deep  $N \triangleleft E(g)$ .*

In the case where  $G = F$  and  $H = C$ , we recover Theorem 2.3 for sufficiently deep (but not all)  $\langle f \rangle \triangleleft C$ . In the case where  $G$  is a free product of locally indicable groups, by considering the action of  $G$  on the corresponding Bass-Serre tree, we recover [EH87, Theorem 1.1] for sufficiently deep normal subgroups.

**2.2 Structure of relative relation modules.** Let  $Rel(G, \langle\langle N \rangle\rangle)$  and  $Rel(H, N)$  be the relative relation modules of the exact sequences

$$1 \rightarrow \langle\langle N \rangle\rangle \rightarrow G \rightarrow Q \rightarrow 1$$

and

$$1 \rightarrow N \rightarrow H \rightarrow R \rightarrow 1,$$

respectively, i.e.  $Rel(G, \langle\langle N \rangle\rangle)$  (resp.  $Rel(H, N)$ ) is the  $\mathbb{Z}Q$ -module (resp.  $\mathbb{Z}R$ -module) whose base set is the abelianization of  $\langle\langle N \rangle\rangle$  (resp.  $N$ ) and the  $Q$ -action (resp.  $R$ -action) is induced by conjugation. If  $G$  is free, then  $Rel(G, \langle\langle N \rangle\rangle)$  is called a *relation module*. For sufficiently deep  $N$ , it follows immediately from Theorem 1.4 that the natural map identifies  $R$  with a subgroup of  $Q$ . We can then further identify  $\mathbb{Z}R$  with a subring of  $\mathbb{Z}Q$ . Thus, given any  $\mathbb{Z}R$ -module  $A$ , it makes sense to talk about the *induced module* of  $A$  from  $\mathbb{Z}R$  to  $\mathbb{Z}Q$ , which is denoted by  $Ind_R^Q A = \mathbb{Z}Q \otimes_{\mathbb{Z}R} A$ .

If  $G = F$  and  $H = C$ , Theorem 2.3 directly implies

$$Rel(F, \langle\langle f \rangle\rangle) \cong \mathbb{Z}[F/C\langle\langle f \rangle\rangle] \cong Ind_R^Q \mathbb{Z} \cong Ind_R^Q Rel(C, \langle f \rangle)$$

as  $\mathbb{Z}Q$ -modules. In general, we have the following corollary to Theorem 2.5.

**Corollary 2.8.** *Let  $G$  be a group with a subgroup  $H \hookrightarrow_h G$ . Then for all sufficiently deep  $N \triangleleft H$ , there is an isomorphism of  $\mathbb{Z}Q$ -modules*

$$Rel(G, \langle\langle N \rangle\rangle) \cong Ind_R^Q Rel(H, N). \quad (1)$$

**Remark 2.9.** Merely knowing that  $\langle\langle N \rangle\rangle = \prod_{t \in T}^* N^t$  for some subset  $T \subset G$  is not enough to guarantee (1). For example, let  $G$  be any abelian group and let  $H$  be a proper subgroup of  $G$ . Then for any subgroup  $N$  of  $H$ ,  $\langle\langle N \rangle\rangle = N = \prod_{t \in \{1\}}^* N^t$ . But  $Rel(G, \langle\langle N \rangle\rangle)$  (resp.  $Rel(H, N)$ ) is a  $\mathbb{Z}Q$ -module (resp.  $\mathbb{Z}R$ -module) with the trivial  $Q$ -action (resp.  $R$ -action) and thus  $Rel(G, \langle\langle N \rangle\rangle) \not\cong Ind_R^Q Rel(H, N)$ .

**2.3 Organization of the paper.** The strategy of this paper is to deal with the more general setting where  $G$  has a family of weakly hyperbolically embedded subgroups and prove the Cohen-Lyndon property in this case. Theorem 2.5 and Corollary 2.8 are simple consequences of the general results. Necessary preliminaries are provided in Section 3, which is divided into four subsections. Section 3.1 recalls the definition and basic properties of van Kampen diagrams. Section 3.2 surveys the notion of (weakly) hyperbolically embedded subgroups. The proof of the Cohen-Lyndon property relies heavily on surgeries on van Kampen diagrams and geodesic polygons of Cayley graphs. Section 3.3 is devoted to a concept called isolated components, which is vital to surgeries, and Section 3.4 collects some results about surgeries on van Kampen diagrams. In Section 4, we construct particular transversals (with the aid of Zorn's lemma). Section 5 states and proves the main theorem of this paper in the general case by using the transversals constructed in Section 4. Finally, Section 6 analyzes the structure of the relative relation modules of Dehn fillings.

**Acknowledgement.** This work was done when I was a graduate student at Vanderbilt University. I would like to thank my supervisor, Professor Denis Osin, for the priceless discussions. This paper would not have been written without his help. I would also like to thank the anonymous referee for the useful comments on an early version of this paper and for showing me an alternative proof of Theorem 2.5 using very rotating families.

### 3 Preliminaries

This section contains a brief discussion of concepts and tools used to prove the main theorem. Our main reference is [DGO17].

#### 3.1 Van Kampen diagrams

Let  $G$  be a group given by the presentation

$$G = \langle \mathcal{A} \mid \mathcal{R} \rangle, \quad (2)$$

where  $\mathcal{A}$  is a symmetric set of letters and  $\mathcal{R}$  is a symmetric set of words in  $\mathcal{A}$  (i.e., for every  $w \in \mathcal{R}$ , every cyclic shift of  $w$  or  $w^{-1}$  belongs to  $\mathcal{R}$ ).

A *van Kampen diagram*  $\Delta$  over (2) is a finite oriented connected planar 2-complex with labels on its oriented edges such that

- (a) each oriented edge of  $\Delta$  is labeled by a letter in  $\mathcal{A} \cup \{1\}$ ;
- (b) if an oriented edge  $e$  of  $\Delta$  has label  $a \in \mathcal{A} \cup \{1\}$ , then  $e^{-1}$  has label  $a^{-1}$ , where  $e^{-1}$  (resp.  $a^{-1}$ ) is the inverse of  $e$  (resp.  $a$ ).

Here, 1 is identified with the empty word over  $\mathcal{A}$  and thus  $1 = 1^{-1}$ . By convention, the empty word of  $\mathcal{A}$  represents the identity of  $G$ .

Let  $p = e_1 \cdots e_k$  be a path in a van Kampen diagram over (2) or in the Cayley graph  $\Gamma(G, \mathcal{A})$ . The initial vertex (resp. terminal vertex) of  $p$  is denoted as  $p^-$  (resp.  $p^+$ ). The *label* of  $p$ , denoted as  $Lab(p)$ , is obtained by first concatenating the labels of the edges  $e_1, \dots, e_k$  and then removing all 1's, as 1 is identified with the empty word. Therefore, the label of a path in a van Kampen diagram is a word over  $\mathcal{A}$ . If  $w$  is a word over  $\mathcal{A}$ , then the notation  $Lab(p) \equiv w$  will indicate a letter-by-letter equality between  $Lab(p)$  and  $w$ .

**Remark 3.1.** Suppose that  $p$  is a path in a van Kampen diagram over (2) with  $Lab(p) \equiv w_1 \cdots w_k$ . Then we can decompose  $p$  in the following way: Let  $p_{w_1}$  be the maximal subpath of  $p$  such that  $p_{w_1}^- = p^-$  and  $Lab(p_{w_1}) \equiv w_1$ . For  $i = 2, \dots, k$ , let  $p_{w_i}$  be the maximal subpath of  $p$  such that  $p_{w_i}^- = p_{w_{i-1}}^+$  and  $Lab(p_{w_i}) \equiv w_i$ .

Edges labeled by letters from  $\mathcal{A}$  are called *essential edges*, while edges labeled by the letter 1 are called *non-essential edges*. A *face* of  $\Delta$  is a 2-cell of  $\Delta$ . Let  $\Pi$  be a face of  $\Delta$ , the boundary of  $\Pi$  is denoted as  $\partial\Pi$ . Likewise, the boundary of  $\Delta$  is denoted by  $\partial\Delta$ . Note that if we choose a base point for  $\partial\Pi$  (resp.  $\partial\Delta$ ), then  $\partial\Pi$  (resp.  $\partial\Delta$ ) becomes a path in  $\Delta$ . For a word  $w$  over  $\mathcal{A}$ , we use the notation  $Lab(\partial\Pi) \equiv w$  (resp.  $Lab(\partial\Delta) \equiv w$ ) to indicate that one can pick a base point to turn  $\partial\Pi$  (resp.  $\partial\Delta$ ) into a path  $p$  so that  $Lab(p) \equiv w$ .

**Remark 3.2.** Suppose that  $\Delta$  is a diagram with  $Lab(\partial\Delta) \equiv w_1 \cdots w_k$ . Then we can decompose  $\partial\Delta$  in the following way: Let  $p_b$  be vertex of  $\partial\Delta$  such that when we use  $p_b$  as the base point of  $\partial\Delta$ , we can turn  $\partial\Delta$  into a path  $p$  with  $Lab(p) \equiv w_1 \cdots w_k$ . And then we use Remark 3.1 to decompose  $p$  and thus decompose  $\partial\Delta$ .

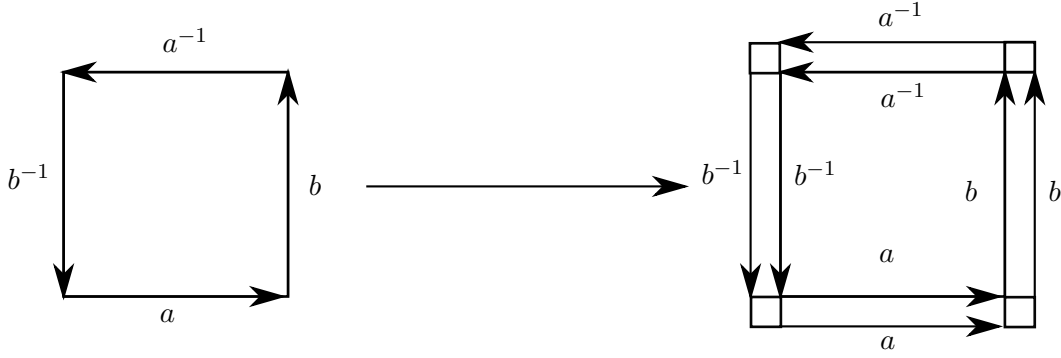


Figure 1: A refinement of a van Kampen diagram over the presentation  $G = \langle a, b \mid aba^{-1}b^{-1} = 1 \rangle$

Consider the following additional assumption on van Kampen diagrams:

- (c) For every face  $\Pi$  of a van Kampen diagram  $\Delta$  over the presentation (2), at least one of the following conditions (c1) and (c2) holds:
  - (c1)  $Lab(\partial\Pi)$  is equal to an element of  $\mathcal{R}$ .
  - (c2)  $\partial\Pi$  either consists entirely of non-essential edges or consists of exactly two essential edges with mutually inverse labels (in addition to non-essential edges).

A face satisfying (c<sub>2</sub>) is called a *non-essential face*. All other faces are called *essential faces*. The process of adding non-essential faces to a van Kampen diagram is called a *refinement*. Figure 1 illustrates a refinement on a van Kampen diagram, where the unlabeled edges are labeled by 1. The interested readers are referred to [Ol'91] for a formal discussion. By using refinements, we can ensure that

- (d) Every face is homeomorphic to a disc, i.e., its boundary has no self-intersection.

**Assumption 3.3.** In the sequel, the above assumptions (c) and (d) will be imposed on van Kampen diagrams.

The well-known van Kampen lemma states that a word  $w$  over  $\mathcal{A}$  represents 1 in  $G$  if and only if there is a van Kampen diagram  $\Delta$  over (2) such that  $\Delta$  is homeomorphic to a disc (such diagrams are called *disk diagrams*), and that  $Lab(\partial\Delta) \equiv w$ .

**Remark 3.4.** If a van Kampen diagram  $\Delta$  is homeomorphic to a disc, and  $O$  is a vertex of  $\Delta$ , then there exists a unique continuous map  $\mu$  from the 1-skeleton of  $\Delta$  to  $\Gamma(G, \mathcal{A})$  sending  $O$  to the identity vertex, preserving the labels of the essential edges and collapsing non-essential edges to points.

### 3.2 Hyperbolically embedded subgroups and group theoretic Dehn fillings

Let  $G$  be a group, let  $\{H_\lambda\}_{\lambda \in \Lambda}$  be a family of subgroups of  $G$ , let  $X$  be a subset of  $G$  such that  $G$  is generated by  $X$  together with the union of all  $H_\lambda, \lambda \in \Lambda$ , and let  $\mathcal{H} = \bigsqcup_{\lambda \in \Lambda} H_\lambda$ . Consider the Cayley graph  $\Gamma(G, X \sqcup \mathcal{H})$ .

**Remark 3.5.** It is possible that  $X$  and  $H_\lambda, \lambda \in \Lambda$ , as subsets of  $G$ , have non-empty intersections with each other. As a consequence, several letters of  $X \sqcup \mathcal{H}$  might represent the same element of  $G$ . If this is the case, the Cayley graph  $\Gamma(G, X \sqcup \mathcal{H})$  will have multiple edges corresponding to those letters.

**Notation.** Let  $w$  be a word over the alphabet  $X \sqcup \mathcal{H}$ . The inverse of  $w$  is denoted by  $w^{-1}$ . The *length* of  $w$ , denoted as  $\|w\|$ , is the number of letters in  $w$ . We identify  $w$  with the element of  $G$  represented by  $w$ . Thus, if  $S$  is a subset of  $G$ , then we write  $w \in S$  to indicate that  $w$  represents an element of  $S$ .

There are two types of equalities for words over  $X \sqcup \mathcal{H}$ . Given two words  $u$  and  $v$  over  $X \sqcup \mathcal{H}$ , the notation  $u \equiv v$  indicates the letter-by-letter equality between  $u$  and  $v$  and the notation  $u =_G v$  indicates that  $u$  and  $v$  represent the same element of  $G$ .

If  $g$  is an element of  $G$ , then  $|g|$  denotes the length of the shortest word in  $X \sqcup \mathcal{H}$  representing  $g$ .

Note that, for each  $\lambda \in \Lambda$ , the Cayley graph  $\Gamma(H_\lambda, H_\lambda)$  can be identified as the complete subgraph of  $\Gamma(G, X \sqcup \mathcal{H})$  whose vertex set is  $H_\lambda$ , and edges are the ones labeled by letters from  $H_\lambda$ .

**Definition 3.6.** Fix  $\lambda \in \Lambda$ . A (combinatorial) path  $p$  in  $\Gamma(G, X \sqcup \mathcal{H})$  between vertices of  $\Gamma(H_\lambda, H_\lambda)$  is called  $H_\lambda$ -*admissible* if it does not contain any edge of  $\Gamma(H_\lambda, H_\lambda)$ . Note that a  $H_\lambda$ -admissible path  $p$  is allowed to pass through vertices of  $\Gamma(H_\lambda, H_\lambda)$ . For every pair of elements  $h, k \in H_\lambda$ , let  $\widehat{d}_\lambda(h, k) \in [0, +\infty]$  be the length of a shortest  $H_\lambda$ -admissible path connecting  $h, k$ . If no such path exists, set  $\widehat{d}_\lambda(h, k) = +\infty$ . The laws of summation on  $[0, +\infty)$  extend naturally to  $[0, +\infty]$  and it is easy to verify that  $\widehat{d}_\lambda : H_\lambda \times H_\lambda \rightarrow [0, +\infty]$  defines a metric on  $\Gamma(H_\lambda, H_\lambda)$  called the *relative metric on  $\Gamma(H_\lambda, H_\lambda)$  with respect to  $X$* .

**Remark 3.7.** Let  $p$  be a path in  $\Gamma(G, X \sqcup \mathcal{H})$  with  $\text{Lab}(p) \equiv h \in H_\lambda$ , for some  $\lambda \in \Lambda$ . For simplicity, we denote  $\widehat{d}_\lambda(1, h)$  by  $\widehat{\ell}_\lambda(p)$ .

**Definition 3.8.** Let  $G$  be a group, let  $\{H_\lambda\}_{\lambda \in \Lambda}$  be a family of subgroups of  $G$ , let  $X$  be a subset of  $G$ , and let  $\mathcal{H} = \bigsqcup_{\lambda \in \Lambda} H_\lambda$ . We say that  $\{H_\lambda\}_{\lambda \in \Lambda}$  *weakly hyperbolically embeds into  $(G, X)$*  (denoted as  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$ ) if  $G$  is generated by the set  $X$  together with the union of all  $H_\lambda, \lambda \in \Lambda$ , and the Cayley graph  $\Gamma(G, X \sqcup \mathcal{H})$  is a Gromov hyperbolic space.

If the collection  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  and for each  $\lambda \in \Lambda$ , the metric space  $(H_\lambda, \widehat{d}_\lambda)$  is proper, i.e., every ball of finite radius contains only finitely many elements, then we say that  $\{H_\lambda\}_{\lambda \in \Lambda}$  *hyperbolically embeds into  $(G, X)$*  (denoted as  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_h (G, X)$ ).

Further, the collection  $\{H_\lambda\}_{\lambda \in \Lambda}$  *hyperbolically embeds into  $G$* , denoted as  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_h G$ , if there exists some subset  $X \subset G$  such that  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_h (G, X)$ .



**Remark 3.9.** Note that if the family  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  for some subset  $X \subset G$  and  $Y = X \cup X^{-1}$ , then we also have  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, Y)$ . In the sequel, we will always assume that the relative generating set  $X$  is symmetric, i.e.,  $X = X^{-1}$ .

**Definition 3.10.** Suppose  $G$  is a group with a family of subgroups  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  for some subset  $X \subset G$ . For  $\lambda \in \Lambda$ , let  $\widehat{d}_\lambda$  be the relative metric on  $\Gamma(H_\lambda, H_\lambda)$  with respect to  $X$ . We say that a property  $P$  holds for all sufficiently deep Dehn fillings of  $\{H_\lambda\}_{\lambda \in \Lambda}$  (or for all sufficiently deep  $N_\lambda \triangleleft H_\lambda, \lambda \in \Lambda$ ), if there exists a number  $C > 0$  such that if  $N_\lambda \triangleleft H_\lambda$  and  $\widehat{d}_\lambda(1, n) > C$  for all  $n \in N_\lambda \setminus \{1\}, \lambda \in \Lambda$ , then  $P$  holds.

One remarkable property of weakly hyperbolically embedded subgroups is the following group theoretic Dehn filling theorem.

**Theorem 3.11** ([DGO17, Theorem 7.15]). *Let  $G$  be a group with a family of subgroups  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  for some subset  $X \subset G$ . Then for all sufficiently deep  $N_\lambda \triangleleft H_\lambda, \lambda \in \Lambda$ , we have:*

- (a) *For each  $\lambda \in \Lambda$ , the natural homomorphism  $i_\lambda : H_\lambda/N_\lambda \rightarrow G/\langle\langle \mathcal{N} \rangle\rangle$  is injective (i.e.,  $H_\lambda \cap \langle\langle \mathcal{N} \rangle\rangle = N_\lambda$ ), where  $\mathcal{N} = \bigcup_{\lambda \in \Lambda} N_\lambda$ .*
- (b)  *$\{i_\lambda(H_\lambda/N_\lambda)\}_{\lambda \in \Lambda} \hookrightarrow_{wh} G/\langle\langle \mathcal{N} \rangle\rangle$ .*
- (c) *There exist subsets  $T_\lambda \subset G, \lambda \in \Lambda$ , such that  $\langle\langle \mathcal{N} \rangle\rangle = \prod_{\lambda \in \Lambda, t \in T_\lambda}^* N_\lambda^t$ .*

As mentioned in the introduction, we are going to improve part (c) of the above theorem. To simplify statements, we introduce the following terminologies.

**Definition 3.12.** Let  $G$  be a group with a subgroup  $H$ . The collection of left transversals of  $H$  in  $G$  is denoted by  $LT(H, G)$ , i.e., every element of  $LT(H, G)$  is a left transversal of  $H$  in  $G$ .

**Definition 3.13.** Let  $G$  be a group, let  $\{H_\lambda\}_{\lambda \in \Lambda}$  be a family of subgroups of  $G$ , and let  $N_\lambda$  be a normal subgroup of  $H_\lambda$  for every  $\lambda \in \Lambda$ . We say that the triple  $(G, \{H_\lambda\}_{\lambda \in \Lambda}, \{N_\lambda\}_{\lambda \in \Lambda})$  has the *Cohen-Lyndon property* if there exists a left transversal  $T_\lambda \in LT(H_\lambda \langle\langle \mathcal{N} \rangle\rangle, G)$  for every  $\lambda \in \Lambda$  such that  $\langle\langle \mathcal{N} \rangle\rangle = \prod_{\lambda \in \Lambda, t \in T_\lambda}^* N_\lambda^t$ , where  $\mathcal{N} = \bigcup_{\lambda \in \Lambda} N_\lambda$ .

### 3.3 Isolated components

Let us assume, until the end of Section 5, that  $G$  is a group with a family of subgroups  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  for some symmetric subset  $X \subset G$ . For each  $\lambda \in \Lambda$ , let  $\widehat{d}_\lambda$  be the relative metric on  $\Gamma(H_\lambda, H_\lambda)$  with respect to  $X$ , and let  $\mathcal{H} = \bigsqcup_{\lambda \in \Lambda} H_\lambda$ . The following terminology goes back to [Osi06].

**Definition 3.14.** Let  $p$  be a path in  $\Gamma(G, X \sqcup \mathcal{H})$ . Fix  $\lambda \in \Lambda$ . An  $H_\lambda$ -subpath  $q$  of  $p$  is a nontrivial subpath of  $p$  labeled by a word over the alphabet  $H_\lambda$  (if  $p$  is a cycle, we allow  $q$  to be a subpath of some cyclic shift of  $p$ ). An  $H_\lambda$ -subpath  $q$  of  $p$  is called an  $H_\lambda$ -component if  $q$  is not properly contained in any other  $H_\lambda$ -subpath. Two  $H_\lambda$ -components  $q_1, q_2$  of  $p$  are

called *connected* if there exists a path  $c$  in  $\Gamma(G, X \sqcup \mathcal{H})$  such that  $c$  connects a vertex of  $q_1$  to a vertex of  $q_2$ , and that  $\text{Lab}(c)$  is a letter from  $H_\lambda$ . An  $H_\lambda$ -component  $q$  of  $p$  is called *isolated* if it is not connected to any other  $H_\lambda$ -component of  $p$ .

The key property of isolated components is that, in a *geodesic polygon* (i.e., a polygon in  $\Gamma(G, X \sqcup \mathcal{H})$  with geodesic sides)  $p$ , the total  $\widehat{\ell}$ -length of isolated components is uniformly bounded above by a linear function of the number of sides. The following result is proved in [DGO17, Proposition 4.14], which is a straightforward generalization of [Osi07, Proposition 3.2].

**Lemma 3.15** (Dahmani-Guirardel-Osin). *There exists a positive number  $D$  satisfying the following property: Let  $p$  be an  $n$ -gon in  $\Gamma(G, X \sqcup \mathcal{H})$  with geodesic sides  $p_1, \dots, p_n$  and let  $I$  be a subset of the set of sides of  $p$  such that every side  $p_i \in I$  is an isolated  $H_{\lambda_i}$ -component of  $p$  for some  $\lambda_i \in \Lambda$ . Then*

$$\sum_{p_i \in I} \widehat{\ell}_{\lambda_i}(p_i) \leq Dn.$$

**Remark 3.16.** Theorem 3.11 asserts the existence of a constant  $C$  such that if  $\widehat{d}_\lambda(1, n) \geq C$  for every  $n \in N_\lambda \setminus \{1\}$  and  $\lambda \in \Lambda$ , then  $H_\lambda \cap \langle\langle \mathcal{N} \rangle\rangle = N_\lambda$  for all  $\lambda \in \Lambda$ . In fact, one can let  $C = 4D$ , where  $D$  is the constant provided by Lemma 3.15 (see [DGO17]).

### 3.4 Diagram surgery

The diagram surgery surveyed in this section was first introduced by Osin in [Osi07], where he proved a group theoretic Dehn filling theorem for relatively hyperbolic groups. Later, Dahmani et al. generalized this technique to deal with weakly hyperbolically embedded subgroups [DGO17].

Consider a symmetric set  $\mathcal{R}$  of words over the alphabet  $X \sqcup \mathcal{H}$  such that  $G$  has the presentation

$$G = \langle X \sqcup \mathcal{H} \mid \mathcal{R} \rangle, \tag{3}$$

and that for all  $\lambda \in \Lambda$ ,  $\mathcal{R}$  contains all words over the alphabet  $H_\lambda$  which represent the identity.

Suppose that  $N_\lambda$  is a normal subgroup of  $H_\lambda$  for each  $\lambda \in \Lambda$ . Denote the union of  $N_\lambda, \lambda \in \Lambda$ , by  $\mathcal{N}$ . Killing  $\langle\langle \mathcal{N} \rangle\rangle$  in  $G$  is equivalent to adding, to  $\mathcal{R}$ , all words over  $H_\lambda$  which represent elements of  $N_\lambda$ , for all  $\lambda \in \Lambda$ , to form a new presentation

$$Q = G / \langle\langle \mathcal{N} \rangle\rangle = \langle X \sqcup \mathcal{H}, \mathcal{R} \cup \mathcal{S} \rangle, \tag{4}$$

where  $\mathcal{S} = \bigcup_{\lambda \in \Lambda} \mathcal{S}_\lambda$  and  $\mathcal{S}_\lambda$  consists of all words over  $H_\lambda$  representing elements of  $N_\lambda$  in  $G$ .

In the sequel, let  $\mathcal{D}$  be the set of all van Kampen diagrams  $\Delta$  over (4) satisfying the following.

- (D1) Topologically  $\Delta$  is a disc with  $k \geq 0$  holes. The boundary of  $\Delta$  can be decomposed as  $\partial\Delta = \partial_{\text{ext}}\Delta \cup \partial_{\text{int}}\Delta$ , where  $\partial_{\text{ext}}\Delta$  is the boundary of the disc, and  $\partial_{\text{int}}\Delta$  consists of disjoint cycles (connected components)  $c_1, \dots, c_k$  that bound the holes.

- (D2) For  $i = 1, \dots, k$ ,  $c_i$  is labeled by a word from  $\mathcal{S}$ .
- (D3) Each diagram  $\Delta$  is equipped with a *cut system* that is a collection  $T = \{t_1, \dots, t_k\}$  of disjoint paths (*cuts*)  $t_1, \dots, t_k$  in  $\Delta$  without self-intersections such that, for  $i = 1, \dots, k$ , the two endpoints of  $t_i$  belong to  $\partial\Delta$ , and that after cutting  $\Delta$  along  $t_i$  for all  $i = 1, \dots, k$ , one gets a disc van Kampen diagram  $\tilde{\Delta}$  over (3).

See Figure 2 for an illustration of a diagram in  $\mathcal{D}$ .

**Lemma 3.17.** *A word  $w$  over  $X \sqcup \mathcal{H}$  represents 1 in  $Q$  if and only if there is a diagram  $\Delta \in \mathcal{D}$  such that  $\text{Lab}(\partial_{\text{ext}}\Delta) \equiv w$ .*

*Proof.* Let  $w$  be a word over  $X \sqcup \mathcal{H}$ . If there is a diagram  $\Delta \in \mathcal{D}$  such that  $\partial_{\text{ext}}\Delta \equiv w$ , by filling the holes of  $\Delta$  with faces whose boundaries are labeled by words from  $\mathcal{S}$ , one creates a disc van Kampen diagram over (3), whose boundary is labeled by  $w$ . Conversely, if  $w$  represents 1 in  $Q$ , then there exists a disc van Kampen diagram  $\bar{\Delta}$  over (3) with  $\text{Lab}(\partial\bar{\Delta}) \equiv w$ . By removing all faces of  $\bar{\Delta}$  labeled by words from  $\mathcal{S}$ , we obtain a diagram  $\Delta'$  satisfying (D1) and (D2). To produce a cut system, choose a vertex  $O$  in  $\partial_{\text{ext}}\Delta'$ . Connect  $O$  with each component of  $\partial_{\text{int}}\Delta'$  by a path so that these paths do not cross each other (although they do intersect each other). By passing to a refinement of  $\Delta'$ , one can separate these paths so that they no longer intersect each other and thus creates a diagram  $\Delta$  satisfying (D1), (D2), and (D3) with  $\text{Lab}(\partial_{\text{ext}}\Delta) \equiv w$ .  $\square$

Figure 2 illustrates the last step of the above proof. The left half shows the diagram  $\Delta'$  where a dash path and a dot-dash path connect  $O$  with two components of  $\partial_{\text{int}}\Delta'$ . By thickening these paths with a refinement, we obtain the right half. The regions shaded by horizontal and vertical lines consist of non-essential faces, while the paths from  $O_1$  to  $U_1$  and from  $O_3$  to  $V_2$  form a cut system.

Let  $\Delta$  be a diagram in  $\mathcal{D}$  and let  $\tilde{\Delta}$  be the disc van Kampen diagram resulting from cutting  $\Delta$  along its set of cuts. Define  $\kappa : \tilde{\Delta} \rightarrow \Delta$  to be the map that “sews” the cuts. Fix an arbitrary vertex  $O$  in  $\tilde{\Delta}$  and let  $\mu$  be a map sending the 1-skeleton of  $\Delta$  to  $\Gamma(G, X \sqcup \mathcal{H})$ , as described by Remark 3.4.

**Definition 3.18.** Let  $\Delta_1$  and  $\Delta_2$  be two diagrams of  $\mathcal{D}$  and let  $\Gamma_1$  (resp.  $\Gamma_2$ ) be the subgraph of the 1-skeleton of  $\Delta_1$  (resp.  $\Delta_2$ ) consisting of  $\partial\Delta_1$  (resp.  $\partial\Delta_2$ ) and all cuts of  $\Delta_1$  (resp.  $\Delta_2$ ). We say that  $\Delta_1$  and  $\Delta_2$  are *equivalent* if there exists a graph isomorphism  $\Gamma_1 \rightarrow \Gamma_2$  which preserves labels and orientations of edges, and maps the cuts and boundary of  $\Delta_1$  to the cuts and boundary of  $\Delta_2$ , respectively.

The following Lemmas 3.19 and 3.25 are results from [DGO17], which are straightforward generalizations of results of [Osi07]. Note that the authors of [DGO17] assume that the presentation (3) has a linear relative isoperimetric function, but this assumption is not used in the proofs of those lemmas.

**Lemma 3.19** ([DGO17, Lemma 7.11]). *Let  $a, b$  be two vertices on  $\partial\Delta$  and let  $\tilde{a}, \tilde{b}$  be two vertices on  $\partial\tilde{\Delta}$  such that  $\kappa(\tilde{a}) = a, \kappa(\tilde{b}) = b$ . Then for any path  $p$  in  $\Gamma(G, X \sqcup \mathcal{H})$  connecting  $\mu(\tilde{a})$  to  $\mu(\tilde{b})$ , there is a diagram  $\Delta_1 \in \mathcal{D}$  with the following properties:*

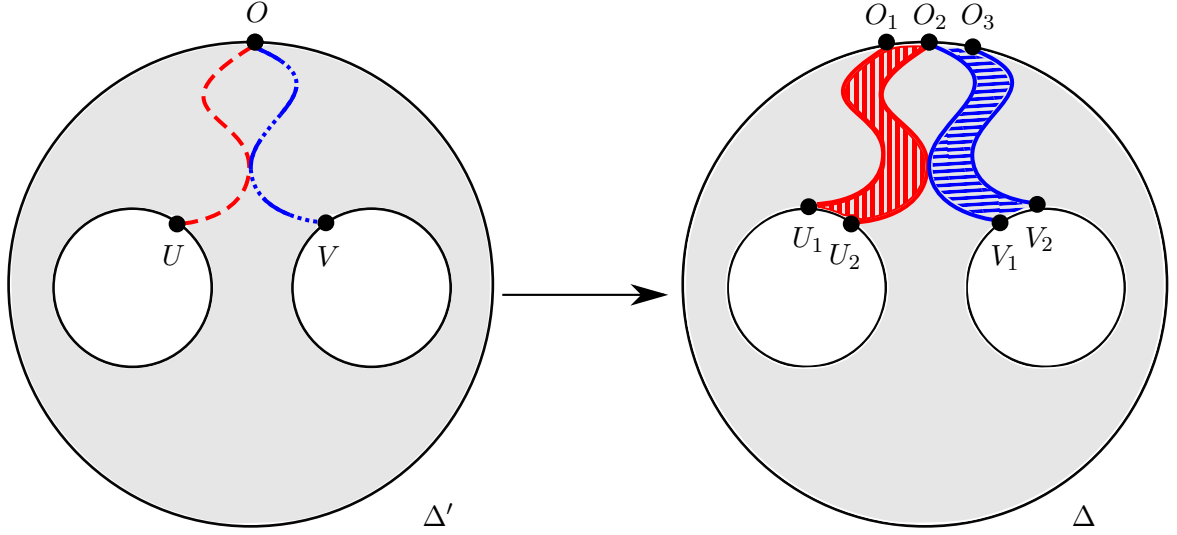


Figure 2: How to produce a cut system

- (a)  $\Delta$  and  $\Delta_1$  are equivalent.
- (b) There is a path  $q$  in  $\Delta_1$  without self-intersections such that (1)  $q$  connects  $a$  and  $b$ , (2)  $q$  has no common vertices with the cuts of  $\Delta_1$  except possibly for  $a, b$ , and (3)  $Lab(q) \equiv Lab(p)$ .

**Definition 3.20.** Fix  $\lambda \in \Lambda$ . An  $H_\lambda$ -subpath in  $\partial\Delta$  (resp.  $\partial\tilde{\Delta}$ ) for some  $\Delta \in \mathcal{D}$  is a path labeled by a nontrivial word over  $H_\lambda$ . An  $H_\lambda$ -subpath  $p$  of  $\partial\Delta$  (resp.  $\partial\tilde{\Delta}$ ) is called an  $H_\lambda$ -component if  $p$  is not properly contained in any other  $H_\lambda$ -subpath. Two  $H_\lambda$ -components  $p, q$  of  $\partial\Delta$  are *connected* if there exist  $H_\lambda$ -components  $a, b$  in  $\partial\tilde{\Delta}$  such that  $\kappa(a)$  (resp.  $\kappa(b)$ ) is a subpath of  $p$  (resp.  $q$ ), and that  $\mu(a), \mu(b)$  are connected in  $\Gamma(G, X \sqcup \mathcal{H})$  (in the sense of Definition 3.14).

**Remark 3.21.** The definitions of  $H_\lambda$ -subpaths,  $H_\lambda$ -components, and connected  $H_\lambda$ -components in  $\partial\Delta$  for a van Kampen diagram  $\Delta \in \mathcal{D}$  or  $\partial\tilde{\Delta}$  do not depend on the pre-chosen vertex  $O$ .

**Definition 3.22.** The *type* of  $\Delta$  is defined by the formula

$$\tau(\Delta) = (k, \sum_{i=1}^k \|Lab(t_i)\|),$$

where  $k$  is the number of holes in  $\Delta$  and  $t_1, \dots, t_k$  are the cuts. We order the types of diagrams in  $\mathcal{D}$  lexicographically:  $(k_1, \ell_1) < (k_2, \ell_2)$  if and only if either  $k_1 < k_2$  or  $k_1 = k_2$  and  $\ell_1 < \ell_2$ .

**Definition 3.23.** A word  $w$  over  $X \sqcup \mathcal{H}$  is called *geodesic* if it labels a geodesic in  $\Gamma(G, X \sqcup \mathcal{H})$ .

**Definition 3.24.** For any word  $w$  over  $X \sqcup \mathcal{H}$ , let  $\mathcal{D}(w)$  be the set of diagrams  $\Delta \in \mathcal{D}$  such that  $\text{Lab}(\partial_{\text{ext}}\Delta) \equiv w$ .

**Lemma 3.25** ([DGO17, Lemma 7.17] (see also [Osi07, Lemma 5.2])). *Suppose that for every  $\lambda \in \Lambda$  and  $n \in N_\lambda \setminus \{1\}$ , we have  $\widehat{d}_\lambda(1, n) > 4D$ , where  $D$  is the constant given by Lemma 3.15. Let  $w$  be a geodesic word over  $X \sqcup \mathcal{H}$  representing 1 in  $Q$ , and let  $\Delta$  be a diagram in  $\mathcal{D}(w)$  of minimal type. Then there exists  $\lambda \in \Lambda$  and a connected component  $c$  of  $\partial_{\text{int}}\Delta$  such that  $c$  is connected to an  $H_\lambda$ -component of  $\partial_{\text{ext}}\Delta$ .*

## 4 Construction of the transversals

The proof of Theorem 5.1 relies on constructing a particular left transversal  $T_\lambda \in LT(H_\lambda \langle\langle \mathcal{N} \rangle\rangle, G)$  for each  $\lambda \in \Lambda$ . It is convenient to construct a collection  $\{T_\lambda\}_{\lambda \in \Lambda}$  of sets of words over  $X \sqcup \mathcal{H}$  satisfying the following properties (P1) through (P3), and think of  $T_\lambda$  as a transversal in  $LT(H_\lambda \langle\langle \mathcal{N} \rangle\rangle, G)$  (identifying words over  $X \sqcup \mathcal{H}$  and the elements of  $G$  represented by those words) for  $\lambda \in \Lambda$ . Recall that  $\|w\|$  is the length of  $w$  for a word  $w$  over  $X \sqcup \mathcal{H}$ , and that  $|g|$  denotes the length of a geodesic word over  $X \sqcup \mathcal{H}$  representing an element  $g \in G$ .

- (P1) [ $\{T_\lambda\}_{\lambda \in \Lambda}$  is transversal] For each  $\lambda \in \Lambda$ ,  $T_\lambda \in LT(H_\lambda \langle\langle \mathcal{N} \rangle\rangle, G)$ .
- (P2) [ $\{T_\lambda\}_{\lambda \in \Lambda}$  is geodesic] If  $w \in T_\lambda$  for some  $\lambda \in \Lambda$ , and  $gH_\lambda \langle\langle \mathcal{N} \rangle\rangle = wH_\lambda \langle\langle \mathcal{N} \rangle\rangle$  for some  $g \in G$ , then  $\|w\| \leq |g|$ . This implies that, for all  $\lambda \in \Lambda$ , every  $w \in T_\lambda$  is a geodesic word over  $X \sqcup \mathcal{H}$ .
- (P3) [ $\{T_\lambda\}_{\lambda \in \Lambda}$  is prefix closed] Let  $\lambda, \mu \in \Lambda$ . If a word  $w \in T_\lambda$  can be decomposed as  $w \equiv uhv$  with  $h \in H_\mu \setminus \{1\}$  ( $u, v$  are allowed to be empty words), then  $u \in T_\mu$  and  $\widehat{d}_\mu(1, h) \leq \widehat{d}_\mu(1, h')$  for all  $h' \in hN_\mu$ .

**Lemma 4.1.** *There exists a collection  $\{T_\lambda\}_{\lambda \in \Lambda}$  satisfying (P1), (P2), and (P3).*

*Proof.* Let  $\mathcal{W}$  be the poset of collections  $\{W_\lambda\}_{\lambda \in \Lambda}$  of words satisfying (P2) and (P3), while instead of (P1), we only demand that the words of  $W_\lambda$  represent a subset of a transversal in  $LT(H_\lambda \langle\langle \mathcal{N} \rangle\rangle, G)$  for every  $\lambda \in \Lambda$ . We order  $\mathcal{W}$  by index-wise inclusion, i.e.,  $\{U_\lambda\}_{\lambda \in \Lambda}$  is less than  $\{V_\lambda\}_{\lambda \in \Lambda}$  if and only if  $U_\lambda \subset V_\lambda$  for every  $\lambda \in \Lambda$ .  $\mathcal{W}$  is non-empty because the collection  $\{W_\lambda\}_{\lambda \in \Lambda}$  with each  $W_\lambda$  consisting of only the empty word is a member of  $\mathcal{W}$ . Moreover, the union of any chain of  $\mathcal{W}$  is again a member of  $\mathcal{W}$ . Therefore, Zorn's lemma implies that  $\mathcal{W}$  has a maximal member  $\{T_\lambda\}_{\lambda \in \Lambda}$ . Suppose that  $\{T_\lambda\}_{\lambda \in \Lambda}$  does not satisfy (P1), i.e., there exist  $\lambda_0 \in \Lambda$  and  $g \in G$  such that no element of the coset  $gH_{\lambda_0}M$  is represented by a word in  $T_{\lambda_0}$ . Without loss of generality, let us assume that if  $g'$  is an element of  $G$  such that  $|g'| < |g|$ , then for each  $\lambda \in \Lambda$ ,  $g'H_\lambda \langle\langle \mathcal{N} \rangle\rangle \cap T_\lambda \neq \emptyset$ .

Let  $w$  be a geodesic word over  $X \sqcup \mathcal{H}$  representing  $g$ . Consider the collection  $\{U_\lambda\}_{\lambda \in \Lambda}$  constructed as follows. For every  $\lambda \in \Lambda \setminus \{\lambda_0\}$ , let  $U_\lambda = T_\lambda$ , and construct  $U_{\lambda_0}$  by the following manner: If  $w$  contains no letter from  $\mathcal{H}$ , let  $U_{\lambda_0} = T_{\lambda_0} \cup \{w\}$ . If  $w$  contains at least one letter from  $\mathcal{H}$ , then  $w$  can be decomposed as  $w \equiv uhv$  such that  $h \in H_\lambda \setminus \{1\}$  for some  $\lambda \in \Lambda$  and  $v$  contains no letter from  $\mathcal{H}$  ( $u, v$  are allowed to be empty words). As  $\|u\| < \|w\| = |g|$ , there exists a word  $u' \in T_\lambda$  such that  $u' \in uH_\lambda \langle\langle \mathcal{N} \rangle\rangle$ . Let  $h'$  be an element of  $H_\lambda$  such that  $u \langle\langle \mathcal{N} \rangle\rangle = u'h' \langle\langle \mathcal{N} \rangle\rangle$  and let  $h''$  be an element of  $H_\lambda$  such that (a)  $h''N_\lambda = h'hN_\lambda$  and (b) if  $k \in h''N_\lambda$ , then  $\widehat{d}_\lambda(1, h'') \leq \widehat{d}_\lambda(1, k)$ . Set  $U_{\lambda_0} = T_{\lambda_0} \cup \{u'h''v\}$ .

It is straight-forward to verify that  $\{U_\lambda\}_{\lambda \in \Lambda}$  is an element of  $\mathcal{W}$ . There is a word in  $U_{\lambda_0}$  representing an element in  $gH_{\lambda_0} \langle\langle \mathcal{N} \rangle\rangle$ , while  $T_{\lambda_0}$  has no such words. It follows that  $\{U_\lambda\}_{\lambda \in \Lambda}$  is strictly greater than  $\{T_\lambda\}_{\lambda \in \Lambda}$ , contradicting the choice of  $\{T_\lambda\}_{\lambda \in \Lambda}$ .  $\square$

## 5 Main theorem and its proof

The following Theorem 5.1 is the main theorem of this paper. We first use it to prove Theorem 2.5, and then we prove Theorem 5.1.

**Theorem 5.1.** *Let  $G$  be a group with a family of subgroups  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  for some  $X \subset G$ . Then the Cohen-Lyndon property holds for all sufficiently deep Dehn fillings of  $\{H_\lambda\}_{\lambda \in \Lambda}$ .*

*Proof of Theorem 2.5.* By assumption,  $H \hookrightarrow_h (G, X)$  for some subset  $X \subset G$ . Let  $\widehat{d}$  be the relative metric on  $\Gamma(H, H)$  with respect to  $X$ . Theorem 5.1 provides a constant  $C$  such that if  $N \triangleleft H$  and  $\widehat{d}(n) > C$  for all  $n \in N \setminus \{1\}$ , then  $(G, H, N)$  possesses the Cohen-Lyndon property. As  $H \hookrightarrow_h (G, X)$ ,  $\widehat{d}$  is locally finite. In particular,

$$\mathcal{F} = \{h \in H \setminus \{1\} \mid \widehat{d}(h) \leq C\}$$

is a finite set. The desired result follows by noting that if  $N \triangleleft H$  and  $N \cap \mathcal{F} = \emptyset$ , then  $(G, H, N)$  possesses the Cohen-Lyndon property.  $\square$

Let us prove Theorem 5.1. Recall that Lemma 3.15 provides a number  $D > 0$  to estimate the total length of isolated components in a geodesic polygon, and that Theorem 3.11 and Remark 3.16 implies that if  $\widehat{d}_\lambda(1, n) \geq 4D$  for every  $n \in N_\lambda \setminus \{1\}$  and  $\lambda \in \Lambda$ , then  $H_\lambda \cap \langle\langle \mathcal{N} \rangle\rangle = N_\lambda$  for all  $\lambda \in \Lambda$ . We assume the following condition.

$$(24D) \quad \widehat{d}_\lambda(1, n) > 24D \text{ for all } n \in N_\lambda \setminus \{1\} \text{ and } \lambda \in \Lambda.$$

We prove that (24D) implies the Cohen-Lyndon property of  $(G, \{H_\lambda\}_{\lambda \in \Lambda}, \{N_\lambda\}_{\lambda \in \Lambda})$ . Let  $\{T_\lambda\}_{\lambda \in \Lambda}$  be a collection of words over  $X \sqcup \mathcal{H}$  satisfying (P1), (P2), and (P3) (by Lemma 4.1, such a collection exists) and think of each  $T_\lambda$  as a left transversal in  $LT(H_\lambda \langle\langle \mathcal{N} \rangle\rangle, G)$ . For every  $\lambda \in \Lambda$ , we extend  $T_\lambda$  to a set  $T_\lambda^{ex}$ . Roughly speaking,  $T_\lambda^{ex}$  is the set of words obtained from  $T_\lambda$  by replacing letters from  $H_\lambda$  with other letters from the same coset of  $N_\lambda$  in  $H_\lambda$ .

**Definition 5.2.** For every  $\lambda \in \Lambda$ , let  $T_\lambda^{ex}$  be the set of words with the following property: Every word  $w \in T_\lambda^{ex}$  admits a decomposition  $w \equiv w_1 h_1 \cdots w_k h_k w_{k+1}$  ( $w_1, \dots, w_{k+1}$  are allowed to be empty words) such that for every  $i \in \{1, \dots, k\}$ , there exists  $\lambda_i \in \Lambda$  with the following properties.

- (a) For  $i = 1, \dots, k$ ,  $h_i$  is an element of  $H_{\lambda_i}$  ( $h_i$  is allowed to equal 1).
- (b) There exists an element  $h'_i \in H_{\lambda_i} \setminus \{1\}$  such that  $h'_i N_{\lambda_i} = h_i N_{\lambda_i}$  for  $i = 1, \dots, k$ , and that the concatenation  $w_1 h'_1 \cdots w_k h'_k w_{k+1}$  is a word in  $T_\lambda$ .

**Remark 5.3.** If  $k = 0$  in the above definition, conditions (a) and (b) will be satisfied trivially. Thus,  $T_\lambda$  is a subset of  $T_\lambda^{ex}$  for every  $\lambda \in \Lambda$ .

**Definition 5.4.** Let  $w$  be a word over  $X \sqcup \mathcal{H}$  and let  $\lambda \in \Lambda$ . If  $w \in T_\lambda^{ex}$ , let  $rank_\lambda(w)$  be the minimal number  $k$  obtained from the decompositions  $w \equiv w_1 h_1 \cdots w_k h_k w_{k+1}$  satisfying Definition 5.2. If  $w \notin T_\lambda$ , let  $rank_\lambda(w) = \infty$ .

For every word  $w$  over  $X \sqcup \mathcal{H}$ , the *rank* of  $w$ , denoted as  $rank(w)$ , is the number  $\min_{\lambda \in \Lambda} \{rank_\lambda(w)\}$ .

**Lemma 5.5.** Let  $w$  be a word in  $T_\lambda^{ex}$  for some  $\lambda \in \Lambda$ . Suppose that  $w$  can be decomposed as  $w \equiv u h v$  with  $h \in H_\mu \setminus \{1\}$  for some  $\mu \in \Lambda$ . Let  $h''$  be an element of  $H_\mu$  such that  $h'' N_\mu = h N_\mu$ . Then  $u h'' v \in T_\lambda^{ex}$ .

*Proof.* Let  $w \equiv w_1 h_1 \cdots w_k h_k w_{k+1}$  be a decomposition satisfying Definition 5.2 and let  $h'_1, \dots, h'_k$  be as in (b) of Definition 5.2.

Without loss of generality, we may assume that  $h = h_i$  for some number  $i \in \{1, \dots, k\}$ . Then  $u h'' v$  can be decomposed as

$$u h'' v \equiv w_1 h_1 \cdots w_{i-1} h_{i-1} w_i h'' w_{i+1} h_{i+1} w_{i+2} h_{i+2} \cdots w_k h_k w_{k+1}.$$

By replacing  $h_j$  with  $h'_j$  for  $j \neq i$  and  $h''$  with  $h'_i$ , we obtain a word in  $T_\lambda$  and thus  $u h'' v \in T_\lambda^{ex}$ .  $\square$

**Lemma 5.6.** Let  $w$  be a word in  $T_\lambda^{ex}$  for some  $\lambda \in \Lambda$  with a decomposition  $w \equiv w_1 h_1 \cdots w_k h_k w_{k+1}$  satisfying Definition 5.2. Then  $w_1 \in T_{\lambda_1}$ .

*Proof.* Let  $h'_1, \dots, h'_k$  be as in (b) of Definition 5.2. Note that the word  $w_1 h'_1 \cdots w_k h'_k w_{k+1}$  can be decomposed as

$$w_1 h'_1 \cdots w_k h'_k w_{k+1} \equiv w_1 h'_1 (w_2 h'_2 \cdots w_k h'_k w_{k+1}).$$

By (P3),  $w_1 \in T_{\lambda_1}$ .  $\square$

It will be shown that  $\langle\langle \mathcal{N} \rangle\rangle = \prod_{\lambda \in \Lambda, t \in T_\lambda}^* N_\lambda^t$ . For the moment, let

$$K = \langle N_\lambda^t, t \in T_\lambda, \lambda \in \Lambda \rangle \leq G.$$

**Lemma 5.7.** *Let  $w$  be a word in  $\bigcup_{\lambda \in \Lambda} T_\lambda^{ex}$ , and let  $n$  be an element of  $N_{\lambda_0}$  for some  $\lambda_0 \in \Lambda$ . Then  $wnw^{-1} \in K$ .*

*Proof.* Let  $\mu$  be an element of  $\Lambda$  with  $\text{rank}(w) = \text{rank}_\mu(w)$ . Thus,  $w$  admits a decomposition  $w \equiv w_1 h_1 \cdots w_k h_k w_{k+1}$  satisfying Definition 5.2 with  $k = \text{rank}(w)$ . We perform induction on  $\text{rank}(w)$ . If  $\text{rank}(w) = 0$ , then  $w \in T_\mu$  and thus  $wnw^{-1} \in K$ .

Suppose that, for all  $w' \in \bigcup_{\lambda \in \Lambda} T_\lambda^{ex}$  with  $\text{rank}(w') < \text{rank}(w)$  and all  $n' \in \bigcup_{\lambda \in \Lambda} N_\lambda$ , we have  $w'^{-1} n' w' \in K$ . Let  $h'_1, \dots, h'_k$  be as in (b) of Definition 5.2. Thus, there exists  $n_1 \in N_{\lambda_1}$  such that  $n_1 h'_1 = h_1$  (note that  $N_{\lambda_1}$  is a normal subgroup of  $H_{\lambda_1}$ ). Notice that

$$w =_G (w_1 n_1 w_1^{-1}) (w_1 h'_1 w_2 h_2 \cdots w_k h_k w_{k+1})$$

and thus

$$wnw^{-1} =_G (w_1 n_1 w_1^{-1}) (w' n w'^{-1}) (w_1 n_1 w_1^{-1})^{-1}, \quad (5)$$

where  $w' \equiv w_1 h'_1 w_2 h_2 \cdots w_k h_k w_{k+1}$ .

By replacing  $h_j$  with  $h'_j$  for  $j = 2, \dots, k$ , we can turn  $w'$  into a word in  $T_\mu$ . Thus,  $w' \in T_\mu^{ex}$  and  $\text{rank}(w') \leq k - 1 < \text{rank}(w)$ . It follows from the induction hypothesis that  $w' n (w')^{-1} \in K$ . By Lemma 5.6,  $w_1 \in T_{\lambda_1}$  and thus  $w_1 n_1 w_1^{-1} \in K$ . By (5),  $wnw^{-1}$  represents a product of elements of  $K$ .  $\square$

For the next two lemmas, recall that  $\|w\|$  denotes the length of a word  $w$  over  $X \sqcup \mathcal{H}$ , and that  $|g|$  denotes the length of a geodesic word over  $X \sqcup \mathcal{H}$  representing an element  $g \in G$ .

**Lemma 5.8.** *Let  $\lambda$  be an element of  $\Lambda$ ,  $u$  be a word in  $T_\lambda^{ex}$ ,  $h$  be a letter of  $H_\lambda \setminus \{1\}$ , and  $v$  be a word over  $X \sqcup \mathcal{H}$  with  $\|v\| = \|u\|$ . Suppose that every element  $m' \in \langle\langle \mathcal{N} \rangle\rangle$  with  $|m'| < 2\|u\| + 1$  belongs to  $K$ . If the concatenation  $uhv \in \langle\langle \mathcal{N} \rangle\rangle$ , then  $uhv \in K$ .*

*Proof.* If  $uhv$  is not a geodesic word, the desired result will follow from the assumptions trivially. So let us assume that  $uhv$  is geodesic. Consider a diagram  $\Delta \in \mathcal{D}(w)$  of minimal type (see Definition 3.22).

We prove Lemma 5.8 by an induction on the number of holes in  $\Delta$ . If  $\Delta$  has no holes, then it will be a disk van Kampen diagram over (3) with boundary labeled by  $uhv$  and thus  $uhv$  represents  $1 \in K$ .

Suppose that  $\Delta$  has  $k \geq 1$  holes. By Lemma 3.25, there exists  $\mu \in \Lambda$  and a connected component  $c$  of  $\partial_{int} \Delta$  such that  $c$  is connected to an  $H_\mu$ -component of  $\partial_{ext} \Delta$ . Let  $w$  be the label of  $c$ . Then  $w$  is a word over  $H_\mu$  representing an element  $n \in N_\mu$ . As  $\text{Lab}(\partial_{ext} \Delta) \equiv uhv$ , we can use Remark 3.2 to decompose  $\partial_{ext} \Delta$  as the concatenation  $p_u p_h p_v$  of three paths  $p_u, p_h$ , and  $p_v$  with  $\text{Lab}(p_u) \equiv u, \text{Lab}(p_h) \equiv h, \text{Lab}(p_v) \equiv v$ . Depending on where  $c$  is connected to, there are three possible cases.

*Case 1:*  $c$  is connected to an  $H_\mu$ -component of  $p_u$ .

In other words,  $u$  can be decomposed as  $u \equiv u_1 h_1 u_2$  with  $h_1 \in H_\mu \setminus \{1\}$ , and  $p_u$  can be decomposed as a concatenation  $p_{u_1} p_{h_1} p_{u_2}$  of three paths  $p_{u_1}, p_{h_1}$ , and  $p_{u_2}$  such that



$Lab(p_{u_1}) \equiv u_1, Lab(p_{h_1}) \equiv h_1, Lab(p_{u_2}) \equiv u_2$  and  $c$  is connected to  $p_{h_1}$  (see Remark 3.1). By Lemma 3.19, passing to an equivalent diagram if necessary, we may assume that there exists a path  $p_{h_2}$  in  $\Delta$  with  $Lab(p_{h_2}) \equiv h_2 \in H_\mu$ , connecting the common vertex of  $p_{h_1}$  and  $p_{u_1}$  to a vertex of  $c$ . Note that the conjugate  $n_1 = h_2 n h_2^{-1} \in N_\mu$ . Let  $h_3$  be the letter from  $H_\mu$  such that  $h_3 =_G n_1 h_1$ . Then

$$uhv \equiv u_1 h_1 u_2 h v =_G (u_1 n_1^{-1} u_1^{-1})(u_1 h_3 u_2 h v). \quad (6)$$

As  $h_1 \neq 1$ , we have  $\|u_1\| \leq \|u\| - 1$  and thus  $\|u_1 n_1^{-1} u_1^{-1}\| \leq 2\|u_1\| - 1 < 2\|u\| + 1$ . Note that  $u_1 n_1^{-1} u_1^{-1} \in \langle\langle \mathcal{N} \rangle\rangle$ . By the induction hypothesis,  $u_1 n_1^{-1} u_1^{-1} \in K$ .

Let  $u_4 \equiv u_1 h_3 u_2$ . Note that  $\|u_4\| \leq \|u\|$ . As  $uhv, u_1 n_1^{-1} u_1^{-1} \in \langle\langle \mathcal{N} \rangle\rangle$ , it follows from (6) that  $u_4 h v \in \langle\langle \mathcal{N} \rangle\rangle$ . If  $\|u_4\| < \|u\|$ , then  $\|u_4 h v\| < 2\|u\| + 1$  and thus  $u_4 h v \in K$ , by assumption. So let us assume that  $\|u_4\| = \|u\|$ . By Lemma 5.5,  $u_4 \in T_\lambda^{ex}$ . Let  $\Sigma$  be a disc van Kampen diagram over (3) such that

$$Lab(\partial\Sigma) \equiv h_2 w h_2^{-1} h_1 h_3^{-1}.$$

Cut  $\Delta$  along the path  $p_{h_2}$  to produce a diagram  $\Delta_1 \in \mathcal{D}$  with

$$Lab(\partial_{ext}\Delta_1) \equiv u_1 h_2 w h_2^{-1} h_1 u_2 h v.$$

Glue  $\Sigma$  to  $\Delta_1$  by identifying the paths with label  $h_2 w h_2^{-1} h_1$  (perform refinements if the non-essential edges of the two paths do not match) to construct a diagram  $\Delta_2 \in \mathcal{D}$  with

$$Lab(\partial_{ext}\Delta_2) \equiv u_4 h v$$

(see Figure 3). Note that the number of holes in  $\Delta_2$  is strictly less than that of  $\Delta$ . By the induction hypothesis,  $u_4 h v \in K$ . By (6),  $uhv$  is a product of elements of  $K$ .

*Case 2:*  $c$  is connected to an  $H_\mu$ -component of  $p_v$ .

This case is symmetric to Case 1 and the proof is left to the reader.

*Case 3:*  $c$  is connected to  $p_h$ .

In other words,  $\mu = \lambda$  and  $h \in H_\lambda \setminus \{1\}$ . By Lemma 3.19 and passing to an equivalent diagram if necessary, we may assume that there exists a path in  $\Delta$ , labeled by a letter  $h_1 \in H_\lambda$ , connecting the common vertex of  $p_h$  and  $p_u$  to a vertex of  $c$ . Note that the conjugate  $n_1 = h_1 n h_1^{-1} \in N_\lambda$ . Let  $h_2$  be a letter from  $H_\lambda$  such that  $h_2 =_G n_1 h$ . Consider the equality

$$uhv =_G (u n_1^{-1} u^{-1})(u h_2 v). \quad (7)$$

As  $u \in T_\lambda^{ex}$ , Lemma 5.7 implies that  $u n_1^{-1} u^{-1} \in K$ . An analysis similar to the one in Case 1 (with  $u h_2 v$  in place of  $u_4 h v$ ) shows that  $u h_2 v \in K$ . By (7),  $uhv$  is a product of elements of  $K$ .  $\square$

**Definition 5.9.** Let  $w$  be a word representing an element of  $\langle\langle \mathcal{N} \rangle\rangle$ . Define the number  $k(w)$  to be the minimal number of holes of a diagram  $\Delta \in \mathcal{D}(w)$ . The *type* of  $w$  is the pair  $\tau(w) = (\|w\|, k(w))$ . We order the set of types lexicographically (see Definition 3.22).

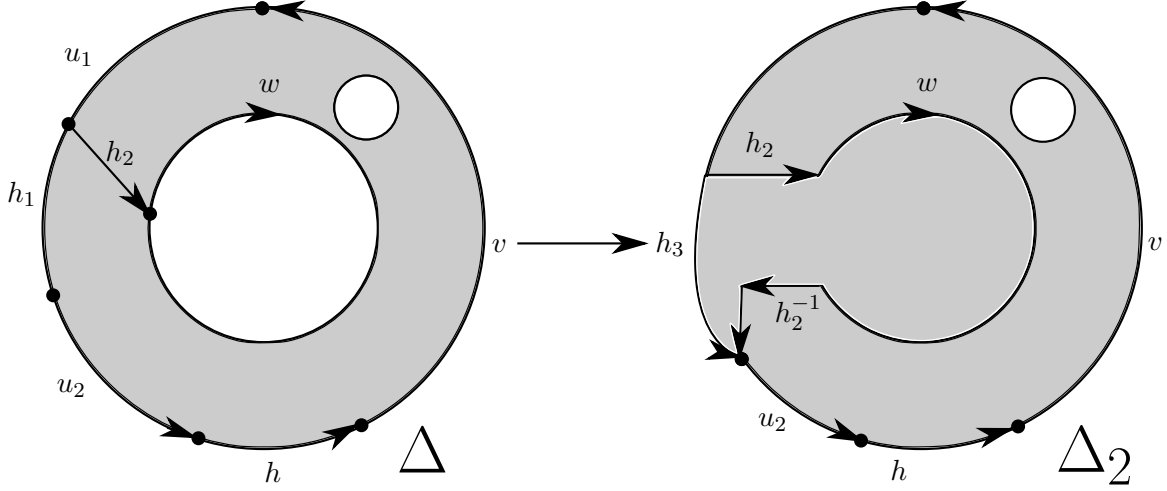


Figure 3: An illustration of Case 1 in the proof of Lemma 5.8

**Remark 5.10.** If  $w$  is a word representing an element of  $\langle\langle \mathcal{N} \rangle\rangle$  and  $\Delta$  is a diagram in  $\mathcal{D}(w)$  of minimal type, then  $\Delta$  necessarily has  $k(w)$  holes.

**Proposition 5.11.**  $\langle\langle \mathcal{N} \rangle\rangle = K$ .

*Proof.* Clearly, each of the groups  $N_\lambda^t, t \in T_\lambda, \lambda \in \Lambda$ , is contained in  $\langle\langle \mathcal{N} \rangle\rangle$  and thus  $K \leq \langle\langle \mathcal{N} \rangle\rangle$ . Let  $w$  be a word over  $X \sqcup \mathcal{H}$  such that  $w \in \langle\langle \mathcal{N} \rangle\rangle$ . Let us show that  $w \in K$  by performing induction on the type of  $w$ . Note that the base case  $\|w\| = k(w) = 0$  is trivial.

Suppose that, for every word  $w'$  over  $X \sqcup \mathcal{H}$  with  $w' \in \langle\langle \mathcal{N} \rangle\rangle$ ,  $\tau(w') < \tau(w)$  implies that  $w' \in K$ . If  $w$  is not a geodesic word, the induction hypothesis will imply  $w \in K$ . Thus, we may assume that  $w$  is geodesic. Consider a diagram  $\Delta \in \mathcal{D}(w)$  of minimal type.

By Lemma 3.25, there exist  $\lambda \in \Lambda$  and a connected component  $c$  of  $\partial_{int}\Delta$  connected to an  $H_\lambda$ -component of  $\partial_{ext}\Delta$ . In other words,  $w$  can be decomposed as  $uhv$  with  $h \in H_\lambda \setminus \{1\}$  ( $u, v$  are allowed to be empty words), and  $\partial_{ext}\Delta$  can be decomposed as a concatenation  $p_u p_h p_v$  of three paths  $p_u, p_h$ , and  $p_v$  such that  $Lab(p_u) = u, Lab(p_h) = h, Lab(p_v) = v$  and  $c$  is connected to  $p_h$  (see Remark 3.2). By Lemma 3.19 and passing to an equivalent diagram if necessary, we may assume that there exists a path  $p_{h_1}$  in  $\Delta$  with  $Lab(p_{h_1}) \equiv h_1 \in H_\lambda$ , connecting the common vertex of  $p_h$  and  $p_u$  to a vertex of  $c$ .

Note that, as  $h \neq 1$ , at least one of  $\|u\|$  and  $\|v\|$  is at most  $(\|w\| - 1)/2$ . Without loss of generality, we may assume that  $\|v\| \leq (\|w\| - 1)/2$ . The case  $\|u\| \leq (\|w\| - 1)/2$  can be analyzed in almost the same way (or just by considering  $w^{-1}$  and reversing every edge of  $\Delta$  if one wishes).

Let  $w_1 \equiv Lab(c)$ . Thus,  $w_1 \in N_\lambda$ . Let  $h_2$  be a letter from  $H_\lambda$  such that  $h_2 =_G h h_1 n h_1^{-1}$ . There exists  $t \in T_\lambda$  such that  $t$  and  $v^{-1}$  are in the same left  $H_\lambda \langle\langle \mathcal{N} \rangle\rangle$ -coset. In other words, there exists  $h_3 \in H_\lambda$  such that  $th_3v \in \langle\langle \mathcal{N} \rangle\rangle$ . Let  $n_1$  be a letter in  $N_\lambda$  such that

$$n_1 =_G h_3 h_1 n^{-1} h_1^{-1} h_3^{-1}.$$

Consider the equality

$$w \equiv uhv =_G (uh_2v)(v^{-1}h_3^{-1}t^{-1})(tn_1t^{-1})(th_3v). \quad (8)$$

Note that  $uh_2v \in \langle\langle \mathcal{N} \rangle\rangle$ , as all other brackets in (8) represents elements of  $\langle\langle \mathcal{N} \rangle\rangle$ . As in the proof of Lemma 5.8, let  $\Sigma$  be a disc van Kampen diagram over (3) with

$$Lab(\partial\Sigma) \equiv hh_1w_1h_1^{-1}h_2^{-1}.$$

Cut  $\Delta$  along  $p_{h_1}$  to produce a diagram  $\Delta_1 \in \mathcal{D}$  with  $Lab(\partial_{ext}\Delta_1) \equiv uhh_1w_1h_1^{-1}v$ . Glue  $\Delta_1$  to  $\Sigma$ , identifying the paths labeled by  $hh_1w_1h_1^{-1}$  (perform refinements if the non-essential edges of the two paths do not match). Denote the resulting diagram by  $\Delta_2$ . Clearly,  $\Delta_2 \in \mathcal{D}$  and  $Lab(\partial_{ext}\Delta_2) \equiv uh_2v$ . Note that the number of holes in  $\Delta_2$  is strictly less than that of  $\Delta$ , and that  $\|uh_2v\| \leq \|u\| + \|v\| + 1 = \|uhv\|$ , as  $uhv$  is a geodesic word. Thus,  $\tau(uh_2v) < \tau(w)$  and the induction hypothesis implies  $uh_2v \in K$ .

Clearly,  $tn_1t^{-1} \in K$ . Note also that  $th_3v \in K$ . Indeed, if either  $\|t\| < \|v\|$  or  $h_3 = 1$ , then  $\|th_3v\| < 2\|v\| + 1 = \|w\|$  and the induction hypothesis implies that  $th_3v \in K$ . If  $\|t\| = \|v\|$  and  $h_3 \neq 1$ , then Lemma 5.8 implies  $th_3v \in K$ .

As  $v^{-1}h_3^{-1}t^{-1} \equiv (th_3v)^{-1}$ , we also have  $v^{-1}h_3^{-1}t^{-1} \in K$ . By (8),  $w$  is a product of elements of  $K$ .  $\square$

The cutting process in the proof of Lemma 5.11 is exactly the same as the one for Lemma 5.8. See Figure 3 for an illustration.

The goal of the rest of this section is to prove the following.

**Proposition 5.12.**  $\langle\langle \mathcal{N} \rangle\rangle = \prod_{\lambda \in \Lambda, t \in T_\lambda}^* N_\lambda^t$ .

*Proof.* Assume, for the contrary, that there exists a word

$$z \equiv \prod_{i=1}^k t_i n_i t_i^{-1} \quad (9)$$

representing  $1 \in G$  such that

(Z1)  $k \geq 2$ ;

(Z2) for  $i = 1, \dots, k$ , there exists  $\lambda_i \in \Lambda$  such that  $n_i \in N_{\lambda_i} \setminus \{1\}$  and  $t_i \in T_{\lambda_i}$ ;

(Z3) for  $i = 1, \dots, k$ , either  $\lambda_i \neq \lambda_{i+1}$  or  $t_i \not\equiv t_{i+1}$  (subscripts are modulo  $k$ , i.e.,  $n_{k+1} = n_1$ ,  $t_0 = t_k$ , etc.).

Without loss of generality, we may also assume

(Z4)  $z$  is minimal, i.e., has the minimal  $k$  among all other words of the form (9) representing 1 in  $G$  and satisfying (Z1), (Z2), and (Z3).

The main idea of the proof of Lemma 5.12 is to show that the existence of such a word  $z$  contradicts Lemma 3.15. For this purpose, it is convenient to first cyclically permute  $z$  and consider the word

$$w \equiv t_k^{-1} \left( \prod_{i=1}^{k-1} t_i n_i t_i^{-1} \right) t_k n_k.$$

In what follows, subscripts are modulo  $k$ . Let  $p_w$  be the path in  $\Gamma(G, X \sqcup \mathcal{H})$  with  $\text{Lab}(p) \equiv w$  and  $p^- = 1$ . We use  $p_{n_i}, p_{t_i^{\pm 1}}$  to denote subpaths of  $p_w$  labeled by  $n_i, t_i^{\pm 1}$ , respectively. More precisely,  $p_{n_i}$  (resp.  $p_{t_i}, p_{t_i^{-1}}$ ) will denote the path in the Cayley graph  $\Gamma(G, X \sqcup \mathcal{H})$  with  $\text{Lab}(p_{n_i}) = n_i$  (resp.  $\text{Lab}(p_{t_i}) = t_i, \text{Lab}(p_{t_i^{-1}}) = t_i^{-1}$ ) and  $p_{n_i}^- = t_k^{-1}(\prod_{j=1}^{i-1} t_j n_j t_j^{-1}) t_i$  (resp.  $p_{t_i}^- = t_k^{-1}(\prod_{j=1}^{i-1} t_j n_j t_j^{-1}), p_{t_i^{-1}}^- = t_k^{-1}(\prod_{j=1}^{i-1} t_j n_j t_j^{-1}) t_i n_i$ ).

Recall that the collection  $\{T_\lambda\}_{\lambda \in \Lambda}$  satisfies (P1), (P2), and (P3). Note that, for every  $\lambda \in \Lambda$  and every word  $t \in T_\lambda$ , the word  $t$  does not end with a letter from  $H_\lambda$ , by (P2). It follows that  $p_{n_i}$  is an  $H_{\lambda_i}$ -component of  $p_w$  for  $i = 1, \dots, k$ . Being a cyclic permutation of  $z$ , the word  $w$  represents 1 in  $G$  and thus the terminal vertex of  $p_w$  is 1. Hence,  $p_w$  is a geodesic  $3k$ -gon. As  $\ell_{\lambda_i}(p_{n_i}) = \widehat{d}_{\lambda_i}(1, n_i)$  for  $i = 1, \dots, k$ , by Lemma 3.15 and (24D), there exists some  $i \in \{1, \dots, k\}$  such that  $p_{n_i}$  is not an isolated  $H_{\lambda_i}$ -component of  $p_w$ .

The rest of the proof is divided into several lemmas. All of them are stated under the assumptions (and using the notations) of Proposition 5.12.

**Lemma 5.13.** *If  $p_{n_i}$  is not an isolated  $H_{\lambda_i}$ -component of  $p_w$  for some  $i \in \{1, \dots, k\}$ , then there are only three possibilities:*

- (a)  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ , but not connected to any  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ .
- (b)  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ , but not connected to any  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ .
- (c)  $p_{n_i}$  is connected to both an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$  and an  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ .

*Proof.* Without loss of generality, let us assume that  $p_{n_1}$  is not isolated in  $p_w$ . There are six cases to consider (see Figure 4 for an illustration).

*Case 1:*  $p_{n_1}$  is connected to an  $H_{\lambda_1}$ -component of either  $p_{t_1}$  or  $p_{t_1^{-1}}$ . In this case, some terminal segment of  $t_1$  represents an element of  $H_{\lambda_1}$ , which contradicts (P2).

*Case 2:*  $p_{n_1}$  is connected to either  $p_{n_2}$  or  $p_{n_k}$ . If  $p_{n_1}$  is connected to  $p_{n_2}$ , then  $\lambda_1 = \lambda_2$ , which in turn implies  $t_1, t_2 \in T_{\lambda_1}$ . The assumption that  $p_{n_1}$  is connected to  $p_{n_2}$  also implies  $t_1^{-1} t_2 \in H_{\lambda_1}$ . By (P1),  $t_1 \equiv t_2$ , which contradicts (Z3) as  $\lambda_1 = \lambda_2$ . The analysis for the subcase where  $p_{n_1}$  is connected to  $p_{n_k}$  is similar.

*Case 3:*  $p_{n_1}$  is connected to  $p_{n_i}$  for some  $i \in \{3, \dots, k-1\}$ . In other words, there exists

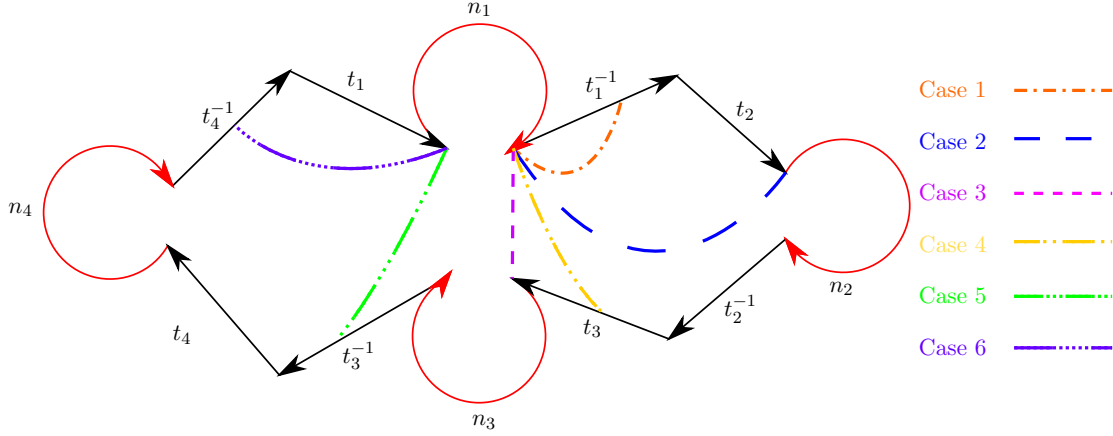


Figure 4: Cases 1 through 6 in the proof of Lemma 5.13

$h \in H_{\lambda_1}$  such that the word

$$u \equiv t_1^{-1} \left( \prod_{j=2}^{i-1} t_j n_j t_j^{-1} \right) t_i h$$

represents 1 in  $G$ . As  $\prod_{j=2}^{i-1} t_j n_j t_j^{-1} \in \langle\langle \mathcal{N} \rangle\rangle \triangleleft G$ , we have  $t_1^{-1} t_i \in H_{\lambda_1} \langle\langle \mathcal{N} \rangle\rangle$ . The assumption that  $p_{n_1}$  is connected to  $p_{n_i}$  also implies  $n_1, n_i \in N_{\lambda_1}$  and thus  $t_1, t_i \in T_{\lambda_1}$ . By (P1),  $t_1 \equiv t_i$ . Thus, the word

$$u' \equiv t_1 h t_1^{-1} \left( \prod_{j=2}^{i-1} t_j n_j t_j^{-1} \right)$$

is a cyclic permutation of  $u$  and represents 1 in  $G$ . It follows that  $t_1 h t_1^{-1} \in \langle\langle \mathcal{N} \rangle\rangle$ . By Theorem 3.11, Remark 3.16, and Condition (24D), we have  $h \in N_{\lambda_1}$ . Then the word  $t_1 h t_1^{-1} \left( \prod_{j=2}^{i-1} t_j n_j t_j^{-1} \right)$  represents 1 in  $G$ , contradicting (Z4).

*Case 4:*  $p_{n_1}$  is connected to an  $H_{\lambda_1}$ -component of  $p_{t_i}$  for some  $i \in \{3, \dots, k\}$ . Thus,  $t_i$  can be decomposed as  $t_i \equiv t'_i h' t''_i$  with  $h' \in H_{\lambda_1} \setminus \{1\}$  and there exists  $h \in H_{\lambda_1}$  such that the word

$$u \equiv t_1^{-1} \left( \prod_{j=2}^{i-1} t_j n_j t_j^{-1} \right) t'_i h$$

represents 1 in  $G$ . By (P3),  $t'_i$  belongs to  $T_{\lambda_1}$ . Arguing as in Case 3, we conclude that the word  $t_1 h t_1^{-1} \left( \prod_{j=2}^{i-1} t_j n_j t_j^{-1} \right)$  represents 1 in  $G$ , contradicting (Z4).

*Case 5:*  $p_{n_1}$  is connected to an  $H_{\lambda_1}$ -component of  $p_{t_i^{-1}}$  for some  $i \in \{2, \dots, k-1\}$ . This case can be reduced to Case 4 by considering  $w^{-1}$ .

Thus, the only possibilities left are (a), (b), and (c). □

**Lemma 5.14.** *If  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ , then  $t_{i+1}$  can be decomposed as  $t_{i+1} \equiv uhv$  with  $h \in H_{\lambda_i} \setminus \{1\}$  ( $u, v$  are allowed to be empty words),  $t_i \equiv u$ , and  $\widehat{d}_{\lambda_i}(1, n_i h) > 12D$ .*

*Proof.* By Definition 3.14,  $t_{i+1}$  can be decomposed as  $t_{i+1} \equiv uhv$  with  $h \in H_{\lambda_i} \setminus \{1\}$  such that  $p_{n_i}$  is connected to the path  $p_h$  in  $\Gamma(G, X \sqcup \mathcal{H})$  with  $\text{Lab}(p_h) \equiv h$  and  $p_h^- = t_k^{-1}(\prod_{j=1}^i t_j n_j t_j^{-1})u$ . By (P3),  $u \in T_{\lambda_i}$ . The assumption that  $p_{n_i}$  is connected to  $p_h$  also implies  $t_i^{-1}u \in H_{\lambda_i}$  and thus  $t_i \equiv u$ , by (P1). Another consequence of (P3) is

$$\widehat{d}_{\lambda_i}(1, h) \leq \widehat{d}_{\lambda_i}(1, h(h^{-1}n_i h)) = \widehat{d}_{\lambda_i}(1, n_i h).$$

Therefore, the triangle inequality implies

$$\widehat{d}_{\lambda_i}(1, n_i) \leq \widehat{d}_{\lambda_i}(1, n_i h) + \widehat{d}_{\lambda_i}(1, h^{-1}) = \widehat{d}_{\lambda_i}(1, n_i h) + \widehat{d}_{\lambda_i}(1, h) \leq 2\widehat{d}_{\lambda_i}(1, n_i h)$$

and thus

$$\widehat{d}_{\lambda_i}(1, n_i h) \geq \widehat{d}_{\lambda_i}(1, n_i)/2 > 12D,$$

by (24D). □

The next lemma follows from Lemma 5.14 by considering  $w^{-1}$ .

**Lemma 5.15.** *If  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ , then  $t_{i-1}$  can be decomposed as  $t_{i-1} \equiv uhv$  with  $h \in H_{\lambda_i} \setminus \{1\}$  ( $u, v$  are allowed to be empty words),  $t_i \equiv u$ , and  $\widehat{d}_{\lambda_i}(1, h^{-1}n_i) > 12D$ .*

**Lemma 5.16.** *If  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ , then  $p_{n_{i+1}}$  is not connected to any  $H_{\lambda_{i+1}}$ -component of  $p_{t_i^{-1}}$ . If  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ , then  $p_{n_{i-1}}$  is not connected to any  $H_{\lambda_{i-1}}$ -component of  $p_{t_i}$ .*

*Proof.* If  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ , then  $\|t_i\| < \|t_{i+1}\|$  by Lemma 5.14. If, in addition,  $p_{n_{i+1}}$  is connected to an  $H_{\lambda_{i+1}}$ -component of  $p_{t_i^{-1}}$ , then  $\|t_{i+1}\| < \|t_i\|$  by Lemma 5.15, a contradiction.

The second assertion of the Lemma can be proved by considering  $w^{-1}$ . □

Recall that we assume the existence of a word  $z$  satisfying (Z1) through (Z4) and construct  $w, p_w$  from  $z$ . The previous several lemmas reveal some properties of  $p_w$  and we are now ready to construct a geodesic polygon  $p$  from  $p_w$  so that  $p$  violates Lemma 3.15, and then we can conclude that  $z$  does not exist and prove Proposition 5.12. The idea is to merge all  $H_{\lambda_i}$ -components connected to  $p_{n_i}$  to form an isolated  $H_{\lambda_i}$ -component for  $i = 1, \dots, k-1$ . Of course, one can also merge  $p_{n_k}$  with the  $H_{\lambda_k}$ -components connected to it. We do not perform this merging only because it makes the construction more complicated. Pick elements  $h_1, \dots, h_{k-1} \in \mathcal{H}$  and  $g_{1,1}, g_{1,2}, g_{2,1}, g_{2,2}, \dots, g_{k-1,1}, g_{k-1,2} \in G$  by the following procedure.

**Procedure 5.17.** For  $i = 1, \dots, k-1$ , perform the following.

- (a) If  $p_{n_i}$  is an isolated  $H_{\lambda_i}$ -component in  $p_w$ , let  $g_{i,1} \in G$  (resp.  $g_{i,2} \in G$ ) be represented by the word  $t_k^{-1}(\prod_{j=1}^{i-1} t_j n_j t_j^{-1}) t_i$  (resp.  $t_k^{-1}(\prod_{j=1}^{i-1} t_j n_j t_j^{-1}) t_i n_i$ ), and let  $h_i = n_i$ .
- (b) If, in  $p_w$ ,  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ , but not connected to any  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ , then by Lemma 5.14,  $t_{i+1}$  can be decomposed as  $t_{i+1} \equiv u_i h'_i v_i$  with  $h'_i \in H_{\lambda_i} \setminus \{1\}$ ,  $t_i \equiv u_i$ , and  $\widehat{d}_{\lambda_i}(1, n_i h'_i) > 12D$ . Let  $h_i$  be a letter from  $H_{\lambda_i}$  such that  $h_i =_G n_i h'_i$ , and let  $g_{i,1} \in G$  (resp.  $g_{i,2} \in G$ ) be represented by the word  $t_k^{-1}(\prod_{j=1}^{i-1} t_j n_j t_j^{-1}) t_i$  (resp.  $t_k^{-1}(\prod_{j=1}^{i-1} t_j n_j t_j^{-1}) t_i h_i$ ).
- (c) If in  $p_w$ ,  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ , but not connected to any  $H_{\lambda_i}$ -component of  $t_{i+1}$ , then by Lemma 5.15,  $t_{i-1}$  can be decomposed as  $t_{i-1} \equiv u_i h'_i v_i$  with  $h'_i \in H_{\lambda_i} \setminus \{1\}$ ,  $t_i \equiv u_i$ , and  $\widehat{d}_{\lambda_i}(1, h_i'^{-1} n_i) > 12D$ . Let  $h_i$  be a letter from  $H_{\lambda_i}$  such that  $h_i =_G h_i'^{-1} n_i$ , and let  $g_{i,1} \in G$  (resp.  $g_{i,2} \in G$ ) be represented by the word  $t_k^{-1}(\prod_{j=1}^{i-2} t_j n_j t_j^{-1}) t_{i-1} n_{i-1} v_i^{-1}$  (resp.  $t_k^{-1}(\prod_{j=1}^{i-2} t_j n_j t_j^{-1}) t_{i-1} n_{i-1} v_i^{-1} h_i$ ).
- (d) If in  $p_w$ ,  $p_{n_i}$  is connected to both an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$  and an  $H_{\lambda_i}$ -component of  $p_{t_{i-1}^{-1}}$ , then by Lemmas 5.14 and 5.15,  $t_{i+1}$  (resp.  $t_{i-1}$ ) can be decomposed as  $t_{i+1} \equiv u_i h'_i v_i$  (resp.  $t_{i-1} \equiv u'_i h''_i v'_i$ ) with  $h'_i \in H_{\lambda_i} \setminus \{1\}$  (resp.  $h''_i \in H_{\lambda_i} \setminus \{1\}$ ),  $t_i \equiv u_i$  (resp.  $t_i \equiv u'_i$ ). Let  $h_i$  be a letter from  $H_{\lambda_i}$  such that  $h_i =_G h''_i^{-1} n_i h'_i$ , and let  $g_{i,1} \in G$  (resp.  $g_{i,2} \in G$ ) be represented by the word  $t_k^{-1}(\prod_{j=1}^{i-2} t_j n_j t_j^{-1}) t_{i-1} n_{i-1} (v'_i)^{-1}$  (resp.  $t_k^{-1}(\prod_{j=1}^{i-2} t_j n_j t_j^{-1}) t_{i-1} n_{i-1} (v'_i)^{-1} h_i$ ).

**Lemma 5.18.**  $g_{i,1}$  and  $g_{i,2}$  are vertices on  $p_w$  for  $i = 1, \dots, k-1$ . Moreover, the order in which  $p_w$  visits these vertices is  $g_{1,1}, g_{1,2}, g_{2,1}, g_{2,2}, \dots, g_{k-1,1}, g_{k-1,2}$ .

*Proof.* The first assertion follows directly from the choices of those vertices. Clearly, the path  $p_w$  visits  $g_{i,1}$  before visiting  $g_{i,2}$  for  $i = 1, \dots, k-1$ . Thus, the second assertion will be proved once we show that, for all  $i, j \in \{1, \dots, k-1\}$  with  $i < j$ , the path  $p_w$  visits  $g_{i,2}$  before visiting  $g_{j,1}$ .

Suppose, for the contrary, that for some  $i, j \in \{1, \dots, k-1\}$  with  $i < j$ , the path  $p_w$  visits  $g_{j,1}$  before visiting  $g_{i,2}$ . By Lemma 5.13, there is only one possibility for this case:  $j = i+1$ ,  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ , and  $p_{n_{i+1}}$  is connected to an  $H_{\lambda_{i+1}}$ -component of  $p_{t_i^{-1}}$ . By Lemma 5.16, if  $p_{n_i}$  is connected to an  $H_{\lambda_i}$ -component of  $p_{t_{i+1}}$ , then  $p_{n_{i+1}}$  is not connected to any  $H_{\lambda_{i+1}}$ -component of  $p_{t_i^{-1}}$ , a contradiction.  $\square$

**Lemma 5.19.** For  $i = 1, \dots, k-2$ , the subpath of  $p_w$  from  $g_{i,2}$  to  $g_{i+1,1}$  consists of at most two geodesic segments.

Lemma 5.19 follows immediately from the choices of the vertices  $g_{i,1}$  and  $g_{i,2}$ ,  $1 \leq i \leq k-1$ . We are now ready to construct a geodesic polygon  $p$  from  $p_w$ .

**Construction 5.20.** For  $i = 1, \dots, k-1$ , let  $p_{h_i}$  the edge of  $\Gamma(G, X \sqcup \mathcal{H})$  with  $\text{Lab}(p_{h_i}) = h_i$  and  $p_{h_i}^- = g_{i,1}$ . Let  $p$  be the path in  $\Gamma(G, X \sqcup \mathcal{H})$  satisfying:  $p^-$  is the identity vertex.  $p$  first follows the path of  $p_w$  (in the direction of  $p_w$ ) until  $p$  visits  $g_{1,1}$ , and then  $p$  travels along  $p_{h_1}$  and arrives at  $g_{1,2}$ . And then  $p$  follows the path  $p_w$  (in the direction of  $p_w$ ) until





*Proof.* First suppose  $\text{card}(I_1) > k/2$ . Then there exists a number  $i$  such that both  $i$  and  $i+1$  belong to  $I_1$ , contradicting Lemma 5.16. Thus,  $\text{card}(I_1) \leq k/2$ .

Suppose  $\text{card}(I_1) = k/2$ . Then  $k$  is even and  $I_1 = \{1, 3, \dots, k-3, k-1\}$ . For every even number  $i \in \{2, 4, \dots, k-2, k\}$ , Lemma 5.16 implies that  $p_{n_i}$  is an isolated  $H_{\lambda_i}$ -component of  $p_w$ . Note that  $\widehat{\ell}_{\lambda_i}(p_{n_i}) = \widehat{d}_{\lambda_i}(1, n_i) > 24D$  for  $i = 1, \dots, k$ , by (24D). Therefore, Lemma 3.15, applied to the geodesic  $3k$ -gon  $p_w$ , yields

$$\frac{24Dk}{2} < \widehat{\ell}_{\lambda_2}(p_{n_2}) + \widehat{\ell}_{\lambda_4}(p_{n_4}) + \dots + \widehat{\ell}_{\lambda_{k-2}}(p_{n_{k-2}}) + \widehat{\ell}_{\lambda_k}(p_{n_k}) < 3kD,$$

a contradiction.  $\square$

Thus,  $\text{card}(I_2) = k - 1 - \text{card}(I_1) \geq (k-1)/2$ . For each  $i \in I_2$ ,  $p_{h_i}$  is an isolated  $H_{\lambda_i}$ -component of  $p$  with  $\widehat{\ell}_{\lambda_i}(p_{h_i}) = \widehat{d}_{\lambda_i}(1, h_i) > 12D$ , by Procedure 5.17 and Construction 5.20. Lemma 3.15, applied to the geodesic polygon  $p$ , yields

$$6D(k-1) = 12D(k-1)/2 < \sum_{i \in I_2} \widehat{\ell}_{\lambda_i}(p_{h_i}) \leq 3kD. \quad (10)$$

In other words,  $k < 2$ , contradicting (Z1). Proposition 5.12 is proved.  $\square$

Finally, Theorem 5.1 follows from Proposition 5.11 and Proposition 5.12.

**Remark 5.23.** The proof of Theorem 5.1 implies that if  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  for some  $X \subset G$ ,  $N_\lambda \triangleleft H_\lambda$  for  $\lambda \in \Lambda$ , and (24D) holds, then for every collection  $\{T_\lambda\}_{\lambda \in \Lambda}$  satisfying (P1), (P2), and (P3), we have

$$\langle\langle \mathcal{N} \rangle\rangle = \prod_{\lambda \in \Lambda, t \in T_\lambda}^* N^t.$$

**Remark 5.24.** In fact, one can show that if  $\{H_\lambda\}_{\lambda \in \Lambda} \hookrightarrow_{wh} (G, X)$  for some  $X \subset G$ ,  $N_\lambda \triangleleft H_\lambda$  for  $\lambda \in \Lambda$ , and following condition

$$(4D) \quad \widehat{d}_\lambda(1, n) > 4D \text{ for all } n \in N_\lambda \setminus \{1\} \text{ and } \lambda \in \Lambda$$

holds, then the triple  $(G, \{H_\lambda\}_{\lambda \in \Lambda}, \{N_\lambda\}_{\lambda \in \Lambda})$  possesses the Cohen-Lyndon property. For the proof, one needs to merge  $p_{n_k}$  with the  $H_{\lambda_k}$ -components connected to it in the construction of  $p$ , and sharpen the coarse estimate (10).

## 6 Relative relation modules

Let  $H$  be a group with a normal subgroup  $N$  and  $R = H/N$ . The *relative relation module*  $\text{Rel}(H, N)$  of the exact sequence

$$1 \rightarrow N \rightarrow H \rightarrow R \rightarrow 1$$

is the abelianization  $\overline{N} = N/[N, N]$  equipped with the  $R$ -action by conjugation. More precisely, denote by  $\overline{n}$  the image of an element  $n \in N$  under the quotient map  $N \rightarrow \overline{N}$ . Then there is an action of  $H$  on  $\overline{N}$  given by  $h \circ \overline{n} = \overline{hnh^{-1}}$  for all  $h \in H, \overline{n} \in \overline{N}$ . Notice that if  $h$  belongs to  $N$ , then  $h \circ \overline{n} = \overline{hnh^{-1}} = \overline{h}\overline{n}\overline{h}^{-1} = \overline{n}$  for all  $\overline{n} \in \overline{N}$ , as  $\overline{h}$  commutes with  $\overline{n}$ . Hence, the action of  $H$  gives rises to an action of  $R$ , turning  $\overline{N}$  into a  $\mathbb{Z}R$ -module. If  $H$  is a free group, then  $Rel(H, N)$  is called a *relation module*.

The main goal of this section is to prove Proposition 6.1, which, together with Theorem 2.5, implies Corollary 2.8.

**Proposition 6.1.** *Suppose that  $G$  is a group, that  $\{H_\lambda\}_{\lambda \in \Lambda}$  is a family of subgroups of  $G$ , and that  $N_\lambda$  is a normal subgroup of  $H_\lambda$  for every  $\lambda \in \Lambda$ . Let  $\mathcal{N} = \bigcup_{\lambda \in \Lambda} N_\lambda$ ,  $Q = G/\langle\langle \mathcal{N} \rangle\rangle$ , and  $R_\lambda = H_\lambda/N_\lambda$  for  $\lambda \in \Lambda$ . If  $N_\lambda \neq \{1\}$  for every  $\lambda \in \Lambda$  and the triple  $(G, \{H_\lambda\}_{\lambda \in \Lambda}, \{N_\lambda\}_{\lambda \in \Lambda})$  has the Cohen-Lyndon property, then*

- (a) *for every  $\lambda \in \Lambda$ , the natural map  $R_\lambda \rightarrow Q$  is injective (i.e.,  $H_\lambda \cap \langle\langle \mathcal{N} \rangle\rangle = N_\lambda$ ), identifying  $R_\lambda$  with a subgroup of  $Q$ ;*
- (b)  *$Rel(G, \langle\langle \mathcal{N} \rangle\rangle) \cong \bigoplus_{\lambda \in \Lambda} Ind_{R_\lambda}^Q Rel(H_\lambda, N_\lambda)$  as  $\mathbb{Z}Q$ -modules.*

**Remark 6.2.** If  $N_{\lambda_0} = \{1\}$  for some  $\lambda_0 \in \Lambda$ , then we can consider the subset  $\Lambda'$  such that  $N_\lambda \neq \{1\}$  for every  $\lambda \in \Lambda'$ . Proposition 6.1 can then be applied to  $(G, \{H_\lambda\}_{\lambda \in \Lambda'}, \{N_\lambda\}_{\lambda \in \Lambda'})$ .

Suppose that the assumptions of Proposition 6.1 are satisfied. Let  $T_\lambda, \lambda \in \Lambda$ , be the transversals provided by Definition 3.13. Fix some  $\lambda \in \Lambda$  for the moment. Suppose  $h \in H_\lambda \cap \langle\langle \mathcal{N} \rangle\rangle$ . Then  $h \in N_{\langle\langle \mathcal{N} \rangle\rangle}(N_\lambda)$ , the normalizer of  $N_\lambda$  in  $\langle\langle \mathcal{N} \rangle\rangle$ . Note that

$$\langle\langle \mathcal{N} \rangle\rangle = \prod_{\mu \in \Lambda, t \in T_\mu}^* N_\mu^t = N_\lambda * \left( \prod_{t \in T_\lambda \setminus \{1\}}^* N_\lambda^t * \prod_{\mu \in \Lambda \setminus \{\lambda\}, t \in T_\mu}^* N_\mu^t \right)$$

and  $N_\lambda \neq \{1\}$ . Note also the following general fact.

**Lemma 6.3.** *Let  $A, B \neq \{1\}$  be groups. Then  $N_{A*B}(A) = A$ .*

*Proof.* Suppose that there exists  $a \in A \setminus \{1\}$  and  $g \in A * B \setminus A$  such that  $a^g \in A$ . Consider the Bass-Serre tree  $Tr$  corresponding to  $A * B$ . The vertex group  $A$  fixes a vertex  $v$  of  $Tr$  and thus  $a^g$  fixes  $v$ . Clearly, the vertex  $g \circ v$  is also fixed by  $a^g$ . As  $g \in A * B \setminus A$ ,  $g \circ v \neq v$  and thus  $a^g$  fixes a non-trivial path between  $v$  and  $g \circ v$ . In particular,  $a^g$  fixes an edge of  $Tr$  and thus conjugates into the unique edge subgroup  $\{1\}$  of  $A * B$ . It follows that  $a^g = 1$ , which is in contradiction with  $a \neq 1$ .  $\square$

Therefore,  $N_{\langle\langle \mathcal{N} \rangle\rangle}(N_\lambda) = N_\lambda$  and  $h \in N_\lambda$ . We conclude:

**Lemma 6.4.** *For every  $\lambda \in \Lambda$ ,  $H_\lambda \cap \langle\langle \mathcal{N} \rangle\rangle = N_\lambda$ .*

Let us consider the relative relation modules  $Rel(G, \langle\langle \mathcal{N} \rangle\rangle)$  and  $Rel(H_\lambda, N_\lambda)$ ,  $\lambda \in \Lambda$ . For every  $\lambda \in \Lambda$ , let  $M_\lambda$  be the subgroup of  $G$  generated by  $N_\lambda^t$ ,  $t \in T_\lambda$ . Note that  $M_\lambda = \prod_{t \in T_\lambda}^* N_\lambda^t$  for every  $\lambda \in \Lambda$ , as  $\langle\langle \mathcal{N} \rangle\rangle = \prod_{\lambda \in \Lambda, t \in T_\lambda}^* N_\lambda^t$ . Note also that  $\langle\langle \mathcal{N} \rangle\rangle = \prod_{\lambda \in \Lambda}^* M_\lambda$ .

For every  $\lambda \in \Lambda$ , the composition of natural maps  $M_\lambda \hookrightarrow \langle\langle \mathcal{N} \rangle\rangle \rightarrow \overline{\langle\langle \mathcal{N} \rangle\rangle}$  maps  $M_\lambda$  into the abelian group  $\overline{\langle\langle \mathcal{N} \rangle\rangle}$  and thus factors through

$$i_\lambda : \overline{M_\lambda} \rightarrow \overline{\langle\langle \mathcal{N} \rangle\rangle}.$$

The homomorphisms  $i_\lambda$ ,  $\lambda \in \Lambda$ , extend to an abelian group homomorphism

$$i : \bigoplus_{\lambda \in \Lambda} \overline{M_\lambda} \rightarrow \overline{\langle\langle \mathcal{N} \rangle\rangle}.$$

It is well-known that  $i$  is an abelian group isomorphism (for example, see [Rob96, Problem 4 of Exercise 6.2]). Thus, we identify  $\overline{M_\lambda}$  with its image  $i_\lambda(\overline{M_\lambda})$  for every  $\lambda \in \Lambda$  and write

$$Rel(G, \langle\langle \mathcal{N} \rangle\rangle) = \overline{\langle\langle \mathcal{N} \rangle\rangle} = \bigoplus_{\lambda \in \Lambda} \overline{M_\lambda}.$$

Fix  $\lambda \in \Lambda$  for the moment. By the same argument as the one above, we write

$$\overline{M_\lambda} = \bigoplus_{t \in T_\lambda} \overline{N_\lambda^t}.$$

**Lemma 6.5.**  *$\overline{M_\lambda}$  is a  $\mathbb{Z}Q$ -submodule of  $Rel(G, \langle\langle \mathcal{N} \rangle\rangle) = \bigoplus_{\lambda \in \Lambda} \overline{M_\lambda}$ . The  $Q$ -action on  $\overline{M_\lambda}$  transitively permutes the summands  $\overline{N_\lambda^t}$ ,  $t \in T_\lambda$ , and its isotropy group of  $\overline{N_\lambda}$  is  $R_\lambda$ , i.e., an element  $q \in Q$  satisfies  $q \circ \overline{n} \in \overline{N_\lambda}$  for all  $\overline{n} \in \overline{N_\lambda}$  if and only if  $q \in R_\lambda$ .*

*Proof.* Fix  $t_0 \in T_\lambda$  and  $g \in G$ . There exists  $t_1 \in T_\lambda$ ,  $h \in H_\lambda$ , and  $m \in \langle\langle \mathcal{N} \rangle\rangle$  such that

$$gt_0 = t_1hm. \tag{11}$$

Consider the summand  $\overline{N_\lambda^{t_0}}$ . For all  $n \in N_\lambda$ ,

$$g \circ \overline{t_0 n t_0^{-1}} = \overline{gt_0 n t_0^{-1} g^{-1}} = \overline{t_1 h m n m^{-1} h^{-1} t_1^{-1}} = \overline{t_1 h n h^{-1} t_1^{-1}} \in \overline{N_\lambda^{t_1}},$$

where the fact that the action of  $\langle\langle \mathcal{N} \rangle\rangle$  acts trivially on  $Rel(G, \langle\langle \mathcal{N} \rangle\rangle)$  is used in the second equality. Hence,  $g \circ (\overline{N_\lambda^{t_0}}) \subset \overline{N_\lambda^{t_1}}$ . As  $\overline{M_\lambda} = \bigoplus_{t \in T_\lambda} \overline{N_\lambda^t}$ , it follows that  $\overline{M_\lambda}$  is  $G$ -invariant and thus  $\overline{M_\lambda}$  is also  $Q$ -invariant.

The above paragraph shows that  $g$  maps  $\overline{N_\lambda^{t_0}}$  into  $\overline{N_\lambda^{t_1}}$ . Actually,  $g(\overline{N_\lambda^{t_0}}) = \overline{N_\lambda^{t_1}}$ . Indeed, given  $n \in N_\lambda$ , we find an element  $x$  of  $N_\lambda^{t_0}$  such that  $g \circ \overline{x} = \overline{n^{t_1}}$ . Let  $x = n^{t_0 h^{-1}}$ . Note that  $n^{h^{-1}} \in N_\lambda$ , as  $N_\lambda$  is normal in  $H_\lambda$ . Thus,  $x \in N_\lambda^{t_0}$ . Direct computation shows

$$g \circ \overline{x} = \overline{g x g^{-1}} = \overline{gt_0(h^{-1} n h) t_0^{-1} g^{-1}} = \overline{t_1 h m (h^{-1} n h) m^{-1} h^{-1} t_1^{-1}} = \overline{t_1 h (h^{-1} n h) h^{-1} t_1^{-1}} = \overline{n^{t_1}},$$

where the fact that the action of  $\langle\langle \mathcal{N} \rangle\rangle$  on  $\text{Rel}(G, \langle\langle N \rangle\rangle)$  is trivial is used in the second equality. Hence,  $g \circ \bar{x} = \overline{n^{t_1}}$ .

As a consequence,  $g \circ (\overline{N_\lambda^{t_0}}) = \overline{N_\lambda^{t_1}}$ , i.e., the action of  $G$  on  $\overline{M_\lambda}$  permutes the summands  $\overline{N_\lambda^t}, t \in T_\lambda$ . In fact, this permutation is transitive: Let  $t$  be any element of  $T_\lambda$ . We wish to find an element of  $G$  which maps  $\overline{N_\lambda^{t_0}}$  to  $\overline{N_\lambda^t}$ . This can be done by  $tt_0^{-1}$ :

$$tt_0^{-1} \circ \overline{N_\lambda^{t_0}} = \overline{N_\lambda^t}.$$

Thus, the action of  $G$  on  $\overline{M_\lambda}$  transitively permutes the summands  $\overline{N_\lambda^t}, t \in T_\lambda$ . The same is thus true for the action of  $Q$  on  $\overline{M_\lambda}$ .

Clearly, for the action of  $G$  on  $\overline{M_\lambda}$ , the isotropy group of  $\overline{N_\lambda}$  contains  $H_\lambda \langle\langle \mathcal{N} \rangle\rangle$ . Observe that in equation (11), if  $t_0 = 1$  and  $g \notin H_\lambda \langle\langle \mathcal{N} \rangle\rangle$ , then  $t_1 \neq 1$  as  $t_1^{-1}g \in H_\lambda \langle\langle \mathcal{N} \rangle\rangle$ . It follows that

$$g \circ \overline{N_\lambda} = \overline{N_\lambda^{t_1}} \neq \overline{N_\lambda},$$

i.e.,  $g$  does not fix  $\overline{N_\lambda}$  setwise. Therefore, for the action of  $G$  on  $\overline{M_\lambda}$ , the isotropy group of  $\overline{N_\lambda}$  is  $H_\lambda \langle\langle \mathcal{N} \rangle\rangle$ . As a consequence, for the action of  $Q$  on  $\overline{M_\lambda}$ , the isotropy group of  $\overline{N_\lambda}$  is  $R_\lambda$ .  $\square$

Recall that if  $\mathcal{O}$  is a ring,  $\mathcal{D}$  is a subring of  $\mathcal{O}$ , and  $A$  is a  $\mathcal{D}$ -module, the *induced module of  $A$  from  $\mathcal{D}$  to  $\mathcal{O}$* , denoted as  $\text{Ind}_{\mathcal{D}}^{\mathcal{O}} A$ , is the tensor product  $\mathcal{O} \otimes_{\mathcal{D}} A$ . If  $\mathcal{O}, \mathcal{D}$  are integral group rings, we simplify notations by dropping  $\mathbb{Z}$ , e.g., we write  $\text{Ind}_H^G$  instead of  $\text{Ind}_{\mathbb{Z}H}^{\mathbb{Z}G}$ . For  $\lambda \in \Lambda$ , Lemma 6.5, together with the following Proposition 6.6, which is a well-known characterization of induced modules (for example, see [Bro94, Proposition 5.3 of Chapter III]), implies that  $\overline{M_\lambda} \cong \text{Ind}_{R_\lambda}^Q \text{Rel}(H_\lambda, N_\lambda)$  as  $\mathbb{Z}Q$ -modules.

**Proposition 6.6.** *Let  $G$  be a group and  $A$  be a  $\mathbb{Z}G$ -module. Suppose that the underlying abelian group of  $A$  is a direct sum  $\bigoplus_{i \in I} A_i$  and that the  $G$ -action transitively permutes the summands. If  $H \leq G$  is the isotropy group of  $A_j$  for some  $j \in I$ . Then  $A_j$  is a  $\mathbb{Z}H$ -module and  $A \cong \text{Ind}_H^G A_j$  as  $\mathbb{Z}G$ -modules.*

*Proof of Proposition 6.1.* For every  $\lambda \in \Lambda$ ,  $\overline{M_\lambda} \cong \text{Ind}_{R_\lambda}^Q \text{Rel}(H_\lambda, N_\lambda)$  as  $\mathbb{Z}Q$ -modules. Thus,

$$\text{Rel}(G, \langle\langle \mathcal{N} \rangle\rangle) = \bigoplus_{\lambda \in \Lambda} \overline{M_\lambda} \cong \bigoplus_{\lambda \in \Lambda} \text{Ind}_{R_\lambda}^Q \text{Rel}(H_\lambda, N_\lambda)$$

as  $\mathbb{Z}Q$ -modules.  $\square$

**Example 6.7.** Let  $\mathcal{G}$  be a graph of groups,  $\pi_1(\mathcal{G})$  be the fundamental group of  $\mathcal{G}$ ,  $\{G_v\}_{v \in V\mathcal{G}}$  be the collection of vertex subgroups, and  $\{G_e\}_{e \in E\mathcal{G}}$  be the collection of edge subgroups. By [DGO17, Example 4.12],  $\{G_v\}_{v \in V\mathcal{G}} \hookrightarrow_{wh} (\pi_1(\mathcal{G}), X)$  provided that the subset  $X \subset \pi_1(\mathcal{G})$  consists of stable letters (i.e., generators corresponding to edges of  $\mathcal{G} \setminus T\mathcal{G}$ , where  $T\mathcal{G}$  is a spanning tree of  $\mathcal{G}$ ), and the corresponding relative metric on a vertex group  $G_v$  corresponding to a vertex  $v \in V\mathcal{G}$  is bi-Lipschitz equivalent to the word metric with respect to the union of the edge subgroups of  $G_v$  corresponding to edges incident to  $v$ . Thus, we have the following corollary to Theorems 3.11, 5.1 and Proposition 6.1.

**Corollary 6.8.** *Let  $\mathcal{G}$  be a graph of groups,  $\pi_1(\mathcal{G})$  be the fundamental group of  $\mathcal{G}$ ,  $\{G_v\}_{v \in V\mathcal{G}}$  be the collection of vertex subgroups, and  $\{G_e\}_{e \in E\mathcal{G}}$  be the collection of edge subgroups. Suppose that, for every  $v \in V\mathcal{G}$ ,  $N_v$  is normal subgroup of  $G_v$  with*

$$N_v \cap \langle G_e, v \in e \rangle = \emptyset.$$

*Then the triple  $(G, \{G_v\}_{v \in V\mathcal{G}}, \{N_v\}_{v \in V\mathcal{G}})$  has the Cohen-Lyndon property, and  $\text{Rel}(G, \langle\langle N \rangle\rangle) \cong \bigoplus \text{Ind}_{Q_v}^Q \text{Rel}(G_v, N_v)$  as  $\mathbb{Z}Q$ -modules, where  $\mathcal{N} = \bigcup_{v \in V\mathcal{G}} N_v$ ,  $Q = G/\langle\langle \mathcal{N} \rangle\rangle$ , and  $Q_v = G_v/N_v$  for  $v \in V\mathcal{G}$ .*

In particular,

**Corollary 6.9.** *Let  $G = A *_C B$  be an amalgamated free product. If  $N \triangleleft A$  and  $N \cap C = \{1\}$ , then  $(G, A, N)$  has the Cohen-Lyndon property, and  $\text{Rel}(G, \langle\langle N \rangle\rangle) \cong \text{Ind}_R^Q \text{Rel}(A, N)$  as  $\mathbb{Z}Q$ -modules, where  $Q = G/\langle\langle N \rangle\rangle$  and  $R = A/N$ .*

**Corollary 6.10.** *Let  $G = H *_t$  be an HNN-extension with associated subgroups  $A, B \leq H$ . If  $N \triangleleft H$  and  $N \cap (A \cup B) = \{1\}$ , then  $(G, H, N)$  has the Cohen-Lyndon property, and  $\text{Rel}(G, \langle\langle N \rangle\rangle) \cong \text{Ind}_R^Q \text{Rel}(H, N)$  as  $\mathbb{Z}Q$ -modules, where  $Q = G/\langle\langle N \rangle\rangle$  and  $R = H/N$ .*

Alternatively, Corollary 6.9 can be deduced from [KS70] and both of Corollaries 6.9, 6.10 can be deduced from the Bass-Serre theory.

## References

- [ACG18] Y. Antolín, R. Coulon, and G. Gandini. Farrell-Jones via Dehn fillings. *J. Topol. Anal.*, 10(4):873–895, 2018.
- [AGM13] I. Agol, D. Groves, and J. Manning. The virtual Haken conjecture. *Doc. Math.*, 18:1045–1087, 2013.
- [AGM16] I. Agol, D. Groves, and J. Manning. An alternate proof of Wise’s malnormal special quotient theorem. *Forum Math. Pi*, 4:e1, 54, 2016.
- [Bow12] B. Bowditch. Relatively hyperbolic groups. *Internat. J. Algebra Comput.*, 22(3):1250016, 66, 2012.
- [Bro94] K. Brown. *Cohomology of groups*, volume 87 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994. Corrected reprint of the 1982 original.
- [CL63] D. Cohen and R. Lyndon. Free bases for normal subgroups of free groups. *Trans. Amer. Math. Soc.*, 108:526–537, 1963.
- [DG18] F. Dahmani and V. Guirardel. Recognizing a relatively hyperbolic group by its Dehn fillings. *Duke Math. J.*, 167(12):2189–2241, 2018.

- [DGO17] F. Dahmani, V. Guirardel, and D. Osin. Hyperbolically embedded subgroups and rotating families in groups acting on hyperbolic spaces. *Mem. Amer. Math. Soc.*, 245(1156):v+152, 2017.
- [EH87] M. Edjvet and J. Howie. A Cohen-Lyndon theorem for free products of locally indicable groups. *J. Pure Appl. Algebra*, 45(1):41–44, 1987.
- [Far98] B. Farb. Relatively hyperbolic groups. *Geom. Funct. Anal.*, 8(5):810–840, 1998.
- [GM08] D. Groves and J. Manning. Dehn filling in relatively hyperbolic groups. *Israel J. Math.*, 168:317–429, 2008.
- [GMS16] D. Groves, J. Manning, and A. Sisto. Boundaries of Dehn fillings. *arXiv:1612.03497*, 2016.
- [KS70] A. Karrass and D. Solitar. The subgroups of a free product of two groups with an amalgamated subgroup. *Trans. Amer. Math. Soc.*, 150:227–255, 1970.
- [Ol’91] A. Ol’shanskii. *Geometry of defining relations in groups*, volume 70 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1991. Translated from the 1989 Russian original by Yu. A. Bakhturin.
- [Osi06] D. Osin. Relatively hyperbolic groups: intrinsic geometry, algebraic properties, and algorithmic problems. *Mem. Amer. Math. Soc.*, 179(843):vi+100, 2006.
- [Osi07] D. Osin. Peripheral fillings of relatively hyperbolic groups. *Invent. Math.*, 167(2):295–326, 2007.
- [Rob96] D. Robinson. *A course in the theory of groups*, volume 80 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996.
- [Sun19a] B. Sun. Cohomology of group theoretic Dehn fillings II: a spectral sequence. *arXiv:1907.12183*, 2019.
- [Sun19b] B. Sun. Cohomology of group theoretic Dehn fillings III: Applications. *arXiv:1908.01290*, 2019.
- [Thu82] W. Thurston. Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 6(3):357–381, 1982.

Bin Sun, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE, RIVERSIDE, CALIFORNIA, 92521

*E-mail address*, `bin.sun@ucr.edu`