NIELSEN EQUIVALENCE IN SMALL CANCELLATION GROUPS

ILYA KAPOVICH AND RICHARD WEIDMANN

Abstract. Let G be a group given by the presentation

$$\langle a_1,\dots,a_k,b_1,\dots b_k \,|\, a_i=u_i(\bar b), b_i=v_i(\bar a) \text{ for } 1\le i\le k\rangle,$$

where $k\ge 2$ and where the $u_i\in F(b_1,\dots,b_k)$ and $w_i\in F(a_1,\dots,a_k)$ are random words. Generically such a group is a small cancellation group and it is clear that $(a_1,\dots,a_k)$ and $(b_1,\dots,b_k)$ are generating n-tuples for G. We prove for generic choices of $u_1,\dots,u_k$ and $v_1,\dots,v_k$ the "once-stabilized" tuples $(a_1,\dots,a_k,1)$ and $(b_1,\dots,b_k,1)$ are not Nielsen equivalent in G. This provides a counter-example for a Wiegold-type conjecture in the setting of word-hyperbolic groups. We conjecture that in the above construction at least k stabilizations are needed to make the tuples $(a_1,\dots,a_k)$ and $(b_1,\dots,b_k)$ Nielsen equivalent.

1. Introduction

Let G be a group. Nielsen equivalence is an equivalence relation for k-tuples of elements of G. Let $\mathcal{T}=(g_1,\dots,g_n)\in G^k$ and $\mathcal{T}'=(g_1',\dots,g_n')\in G^k$ be two k-tuples. Then $\mathcal{T}$ and $\mathcal{T}'$ are called *elementary equivalent* (write $\mathcal{T}\sim_e\mathcal{T}'$) if one of the following holds:

1. There exists some $\sigma\in S_n$ such that $g_i'=g_{\sigma(i)}$ for $1\le i\le n$.
2. $g_i'=g_i^{-1}$ for some $i=1,\dots,n$ and $g_j'=g_j$ for $j\ne i$.
3. $g_i'=g_ig_j$ for some $i\ne j$ and $g_k'=g_k$ for $k\ne i$.

Any of these transformations is called a *Nielsen transformation* or a *Nielsen move*. Two tuples $\mathcal{T}$ and $\mathcal{T}'$ are called *Nielsen equivalent* or simply *equivalent* if there exists some finite sequence $\mathcal{T}_0,\mathcal{T}_1,\dots,\mathcal{T}_l$ such that

$$\mathcal{T}=\mathcal{T}_0\sim_e\mathcal{T}_1\sim_e\dots\sim_e\mathcal{T}_l=\mathcal{T}'.$$

Nielsen introduced this equivalence relation to study subgroups of free groups. Among other things he showed that in the free groups $F_n=F(x_1,\dots,x_n)$ any generating k-tuple is Nielsen equivalent to the tuple $(x_1,\dots,x_n,1,\dots,1)$, in particular any two generating k-tuples are Nielsen equivalent. This together with the fact that any Nielsen move on a basis of F_n induces an automorphism implies the following alternative definition of Nielsen equivalence:

Two k-tuples $\mathcal{T}=(g_1,\dots,g_k)\in G^k$ and $\mathcal{T}'=(g_1',\dots,g_k')\in G^k$ are Nielsen equivalent if and only if there exists a homomorphism $\phi:F_k\to G$ and an automorphism α of F_k such that the following hold:

1. $g_i=\phi(x_i)$ for $1\le i\le n$.
2. $g_i'=\phi\circ\alpha(x_i)$ for $1\le i\le n$.

Let $F_k=F(x_1,\dots,x_k)$ be a free group of rank k with a fixed free basis $(x_1,\dots,x_k)$. There is a natural identification between the set $Hom(F_k,G)$ of homomorphisms from F_k to G and the set G^k of k-tuples of elements of G. There is also a natural left action of $Aut(F_k)$ on $Hom(F_k,G)$ by pre-composition. In view of the above remark, two k-tuples of elements of G are Nielsen equivalent if and only if the corresponding elements of $Hom(F_k,G)$ lie in the same $Aut(F_k)$-orbit.

In general it is very difficult to decide if two k-tuples are Nielsen equivalent in a given group.

If two tuples $\mathcal{T} = (g_1, \ldots, g_k) \in G^k$ and $\mathcal{T}' = (g_1', \ldots, g_k') \in G^k$ are Nielsen equivalent then they generate the same subgroup of G, that is $\langle \mathcal{T} \rangle = \langle \mathcal{T}' \rangle \leq G$. Thus if two tuples generate different subgroups of G, the tuples are not Nielsen equivalent. However, this observation does not help in distinguishing Nielsen equivalence classes of tuples generating the same subgroup, in particular those tuples that generate the entire group G (*generating* tuples). Under the identification of the set of k-tuples in G with $Hom(F_k, G)$ discussed above, the set of generating k-tuples of G corresponds to the set $Epi(F_k, G)$ of epimorphisms from F_k to G.

The only exception is the case $k = 2$. A basic fact due to Nielsen shows that if 2-tuples (g_1, g_2) and (h_1, h_2) are Nielsen equivalent in G then $[g_1, g_2]$ is conjugate to $[h_1, h_2]^{\pm 1}$ in G. No such criteria exist for $k \geq 3$ and there are very few known results distinguishing Nielsen-equivalence classes of generating k-tuples for $k \geq 3$.

Let us note here that even in the algorithmically nice setting of torsion-free word-hyperbolic groups the problem of deciding if two tuples are Nielsen equivalent is algorithmically undecidable.

Indeed the subgroup membership problem is a special case of this problem since two tuples $(g_1, \ldots, g_n, h)$ and $(g_1, \ldots, g_n, 1)$ are Nielsen equivalent if and only if $h \in \langle g_1, \ldots, g_n \rangle$. This implies in particular that Nielsen equivalence is not decidable for finitely presented torsion-free small cancellation groups as they do not have decidable subgroup membership problem as shown by Rips [R].

As noted above, understanding Nielsen equivalence of generating k-tuples is particularly difficult and the problem becomes even harder if $k > rank(G)$, where $rank(G)$ is the smallest size of a generating set of G.

Of particular interest here is the so-called *Wiegold conjecture* about generating tuples of finite simple groups. We say that a generating k-tuple is *redundant* if it contains a proper subtuple that still generates G. We say that a generating tuple is *weakly redundant* if it is Nielsen equivalent to a redundant tuple. Note that a redundant tuple $(g_1, \ldots, g_k)$ is always Nielsen equivalent to a k-tuple of the form $(h_1, \ldots, h_{k-1}, 1)$. Thus a generating tuple if weakly redundant if and only if it is Nielsen equivalent to a tuple containing a trivial entry. It is well-known, as a consequence of classification, that every finite simple group is two-generated, so that $rank(G) \leq 2$. The Wiegold conjecture says that if G is a finite simple group and $k \geq 3$ then any two generating k-tuples of G are Nielsen equivalent; in other words the action of $Aut(F_k)$ on $Epi(F_k, G)$ is transitive in this case. Since G is two-generated and has a generating k-tuple of the form $(a, b, 1, \ldots, 1)$, this implies that any generating k-tuple of G is redundant. The Wiegold conjecture is closely related to the so-called "product replacement algorithm" in finite groups and there is substantial experimental evidence and a number of partial theoretical results in favor of the validity of the Wiegold conjecture. We refer the reader to [Pa, LGM, LP] for a more extensive discussion of this topic.

One obvious way of producing redundant tuples is via the so-called "stabilization" moves. A stabilization move on a k-tuple $(g_1, \ldots, g_k)$ gives a $(k+1)$-tuple $(g_1, \ldots, g_k, 1)$. It is easy to see that for any generating k-tuples $(g_1, \ldots, g_k)$ and $(h_1, \ldots, h_k)$ of a group G, applying k stabilization moves to each of them produces Nielsen equivalent $2k$-tuples $(g_1, \ldots, g_k, 1, \ldots, 1)$ and $(h_1, \ldots, h_k, 1, \ldots, 1)$. The Wiegold conjecture implies that for any two generating pairs (a, b) and (a_1, b_1) of a finite simple group G, the once-stabilized tuples $(a, b, 1)$ and $(a_1, b_1, 1)$ are Nielsen equivalent.

Let us mention here the (few) known results on distinguishing Nielsen equivalence of generating k-tuples for infinite groups. Apart from the special and much easier case of $k = 2$, these can be mostly separated into two distinct approaches.

The first one is algebraic (K-theoretic) and has been very successfully applied to study and distinguish Nielsen equivalence classes. The earliest work is due to Noskov [No] who showed that there exist non-minimal generating tuples that are not Nielsen equivalent to a tuple containing the trivial element and thereby giving a negative answer to a question of Waldhausen, these results where then generalized by Evans [E1].

Lustig and Moriah [LM1], [LM2], [LM3] used algebraic methods to distinguish Nielsen equivalence classes of Fuchsian groups and other groups with appropriate presentations. This enabled them to distinguish isotopy classes of vertical Heegaard splittings of Seifert manifolds.

Recently Evans [E2, E3] has found for any given number N large generating tuples of metabelian groups that do not become Nielsen equivalent after adding the trivial element to the tuples N times, making those the first examples of this type in the literature even for the case $N = 1$. The generating tuples however are much bigger than the rank of the group.

The second approach is combinatorial-geometric and is closer in spirit to Nielsen's original work, it relies mostly on using cancellation methods. First in line is Grushko's theorem [G] which states that any generating tuple of a free product is Nielsen equivalent to a tuple of elements that lie in the union of the factors. Together with recent work of the second author [W] this implies that Nielsen equivalence of irreducible generating tuples in a free product is decidable iff it is decidable in the factors.

Zieschang [Z] proves that any minimal generating tuple of a surface group is Nielsen equivalent to the standard generating tuple and proves a similar result for Fuchsian groups that lead to the solution of the rank problem [PRZ]. Nielsen equivalence in Fuchsian groups has been studied by many authors.

The finiteness of Nielsen equivalence classes of k-tuples for torsion-free locally quasiconvex-hyperbolic groups has been established by the authors [KW] generalizing a result of Delzant [D] who studied 2-generated groups. The first author and Schupp [KS] have recently established uniqueness of the Nielsen equivalence class of minimal generating tuples for a class of groups closely related to the one studied in the present article.

The main result of this paper is the following result, which implies in particular that there exist 2-generated torsion-free word-hyperbolic groups that have generating pairs (a_1, a_2) and (b_1, b_2) such that $(a_1, a_2, 1)$ and $(b_1, b_2, 1)$ are not Nielsen equivalent. See Section 3 for precise definitions related to genericity.

Theorem 1.1. *Let $k \geq m \geq 2$ be arbitrary and let G be a group given by the presentation*

$$(\dagger) \qquad \langle a_1, \ldots, a_k, b_1, \ldots b_m \mid a_i = u_i(\bar{b}), b_j = v_j(\bar{a}) \text{ for } 1 \leq i \leq k, 1 \leq j \leq m \rangle.$$

There exist generic sets $\mathcal{U}_k$ of k-tuples of cyclically reduced words in $F(b_1, \ldots, b_m)$ and $\mathcal{V}_m$ of generic m-tuples of cyclically reduced words in $F(a_1, \ldots, a_k)$ such that for any $(u_1, \ldots, u_k) \in \mathcal{U}_k$ and any $(v_1, \ldots, v_m) \in \mathcal{V}_m$ the group G given by the above presentation has the following property:

For any $g_1, \ldots, g_{k-1} \in G$ the generating $(k+1)$-tuple $(a_1, \ldots, a_k, 1)$ of G is not Nielsen equivalent to a tuple $(b_1, b_2, g_1, \ldots, g_{k-1})$.

The theorem immediately implies that the the generating $(k+1)$-tuples $(a_1,\dots,a_k,1)$ and $(b_1,\dots,b_m,1,\dots,1)$ are not Nielsen equivalent (where $k+1-m$ trivial entries are present in the second tuple). However, we do believe that much more is true, namely that following holds:

Conjecture 1.2. *Let $k\ge m\ge 2$. There exist generic sets $\widetilde{\mathcal{U}}_k$ of k-tuples of cyclically reduced words in $F(b_1,\dots,b_m)$ and $\widetilde{\mathcal{V}}_m$ of generic m-tuples of cyclically reduced words in $F(a_1,\dots,a_k)$ such that for any $(u_1,\dots,u_k)\in\widetilde{\mathcal{U}}_k$ and any $(v_1,\dots,v_m)\in\widetilde{\mathcal{V}}_m$ the group G given by presentation (†) above has the following property:*

Then for any $t<m$ the generating $(k+t)$-tuple $(a_1,\dots,a_k,1,\dots,1)$ is not equivalent to a $(k+t)$-tuple of type $(b_1,\dots,b_{k+1},*,\dots,*)$.

In particular, if $k=m$, the conjecture would imply that at least k stabilizations are needed in G in order to make the generating k-tuples $(a_1,\dots,a_k)$ and $(b_1,\dots,b_k)$ Nielsen equivalent.

The conclusion of Theorem 1.1 implies that for G as in the assumption of the theorem with generic u_i and v_j, the generating k-tuples $(a_1,\dots,a_k)$ and $(b_1,b_2,\dots,b_m,1,\dots,1)$ (where $(k-m)\ge 0$ trivial entries are present in the second tuple) are not Nielsen equivalent in G. While the proof of Theorem 1.1 is very complicated, we also give a simple proof (see Theorem 4.4) that in this situation $(a_1,\dots,a_k)$ is not Nielsen equivalent to a k-tuple of the form $(b_1,,*,\dots,*)$.

Although we do not prove it in this paper, for the case $k=m\ge 2$ and G as in Theorem 1.1 one can use the methods of [KS] to show that G has *exactly* two Nielsen-equivalence classes of generating k-tuples, namely $(a_1,\dots,a_k)$ and $(b_1,\dots,b_k)$.

2. Small cancellation theory

Recall that a set R of cyclically reduced words in $F=F(a_1,\dots,a_k)$ is *symmetrized* if for every $r\in R$ all cyclic permutations of $r^{\pm1}$ belong to R. For a symmetrized set $R\subseteq F(a_1,\dots,a_k)$, a freely reduced word $v\in F(a_1,\dots,a_k)$ is a *piece with respect to* R if there exist $r_1,r_2\in R$, such that $r_1\ne r_2$ and such that v is an initial segment of each of r_1,r_2.

Definition 2.1 (Small Cancellation Condition)**.** Let $R\subseteq F(a_1,\dots,a_k)$ be a symmetrized set of cyclically reduced words. Let $0<\lambda<1$. We say that R satisfies the $C'(\lambda)$-*small cancellation condition* if, whenever v is a piece with respect to R and v is a subword of some $r\in R$, then $|v|<\lambda|r|$.

We say that a presentation $G=\langle a_1,\dots,a_k|R\rangle$ satisfies the $C'(\lambda)$-*small cancellation condition* if $R\subseteq F(a_1,\dots,a_k)$ is a $C'(\lambda)$ set.

The following fact is a well-known basic property of small cancellation groups [LS]:

Proposition 2.2. *Let $G=\langle a_1,\dots,a_k|R\rangle$ be a $C'(\lambda)$-presentation, where $\lambda\le 1/6$. Let $w\in F(a_1,\dots,a_k)$ be a nontrivial freely reduced word such that $w=_G 1$. Then w has a subword u such that for some $r\in R$ the word u is a subword of r satisfying $|u|>(1-3\lambda)|r|$.*

Definition 2.3. Let $G=\langle a_1,\dots,a_k|R\rangle$ be a $C'(\lambda)$-presentation, where $\lambda\le 1/100$. For a freely reduced word $w\in F(a_1,\dots,a_k)$ we say that w is *Dehn reduced with respect to* R if w does not contain a subword u such that u is also a subword of some $r\in R$ with $|u|>|r|/2$.

We say that a freely reduced word $w\in F(a_1,\dots,a_k)$ is λ-*reduced with respect to* R if w does not contain a subword u such that u is also a subword of some $r\in R$ with $|u|>(1-3\lambda)|r|$.

Similarly, we say that a cyclically reduced word $w \in F(a_1, \dots, a_k)$ is λ-*cyclically reduced with respect to* R if every cyclic permutation of w is λ-reduced with respect to R. We also say that a cyclically reduced word $w \in F(a_1, \dots, a_k)$ is *cyclically Dehn-reduced with respect to* R if every cyclic permutation of w is Dehn-reduced with respect to R.

Note that if $\lambda \le 1/6$ then any Denh-reduced word is λ-reduced. Proposition 2.2 says that for a $C'(\lambda)$-presentation with $\lambda \le 1/6$ a nontrivial freely reduced and λ-reduced word in $F(a_1, \dots, a_k)$ represents a nontrivial element of G.

The following statement follows from the basic results of small cancellation theory, established in Ch. V, Sections 3-5 of [LS] (see also [Str]).

Proposition 2.4. *[Equality and Conjugacy diagrams in $C'(\lambda)$-groups]*
Let

$$(*) \qquad G = \langle a_1, \dots, a_k | R \rangle$$

be a $C'(\lambda)$-presentation, where $\lambda \le 1/100$.

(1) *Let $w_1, w_2 \in F(a_1, \dots, a_k)$ be freely reduced and λ-reduced words such that $w_1 =_G w_2$. Then any reduced equality diagram D over $(*)$, realizing the equality $w_1 =_G w_2$, has the form as shown in Figure 1. Specifically, any region Q of D labelled by $r \in R$ intersects both the upper boundary of D (labelled by w_1) and the lower boundary of D (labelled by w_2) in simple segments α_1, α_2 accordingly, satisfying*

$$\lambda |r| \le |\alpha_j| \le (1 - 3\lambda)|r|.$$

Moreover, if two regions Q, Q' in D, labelled by $r, r' \in R$, have a common edge, then they intersect in closed simple segment γ joining a point of the upper boundary of D with a point of the lower boundary of D and labelled by a piece with respect to R. In particular $|\gamma| < \lambda |r|$ and $|\gamma| < \lambda |r'|$. Further, if both w_1 and w_2 are also Dehn-reduced, then for the segments α_j above we have $|\alpha_j| \ge (1/2 - 2\lambda)|r| \ge |r|/3$.

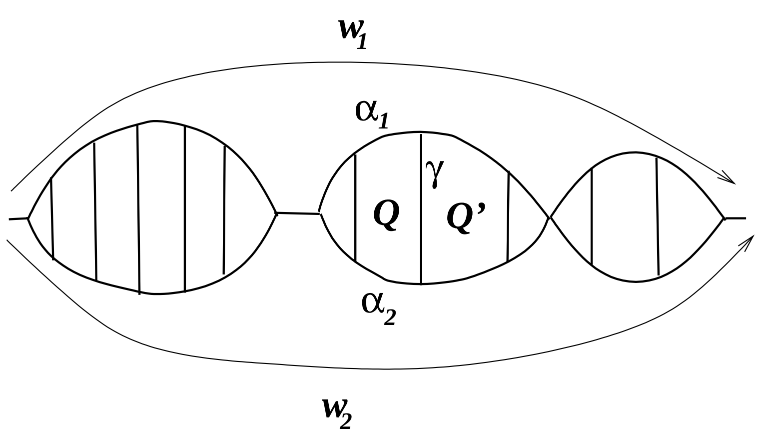

FIGURE 1. Equality diagram in a small cancellation group

(2) *Let $w_1, w_2 \in F(a_1, \dots, a_k)$ be nontrivial cyclically reduced and strongly λ-cyclically reduced words representing conjugate elements of G. Then there exists a reduced conjugacy diagram D over $(*)$ with the inner cycle boundary labelled by a cyclic permutation of w_2 and the outer cycle boundary labelled by a cyclic permutation of w_1, of the form shown in Figure 2. Specifically, any region Q of D labelled by $r \in R$ intersects both the outer boundary of D*

(labelled by a cyclic permutation of w_1) and the inner boundary of D (labelled by a cyclic permutation of w_2) in simple segments α_1, α_2 accordingly, satisfying

$$\lambda|r| \le |\alpha_j| \le (1-3\lambda)|r|.$$

Moreover, if two regions Q, Q' in D, labelled by $r, r' \in R$, have a common edge, then they intersect in closed simple segment γ joining a point of the inner boundary of D with a point of the outer boundary of D and labelled by a piece with respect to R. In particular $|\gamma| < \lambda|r|$ and $|\gamma| < \lambda|r'|$. Further, if both w_1 and w_2 are also Dehn-reduced, then for the segments α_j above we have $|\alpha_j| \ge (1/2 - 2\lambda)|r| \ge |r|/3$.

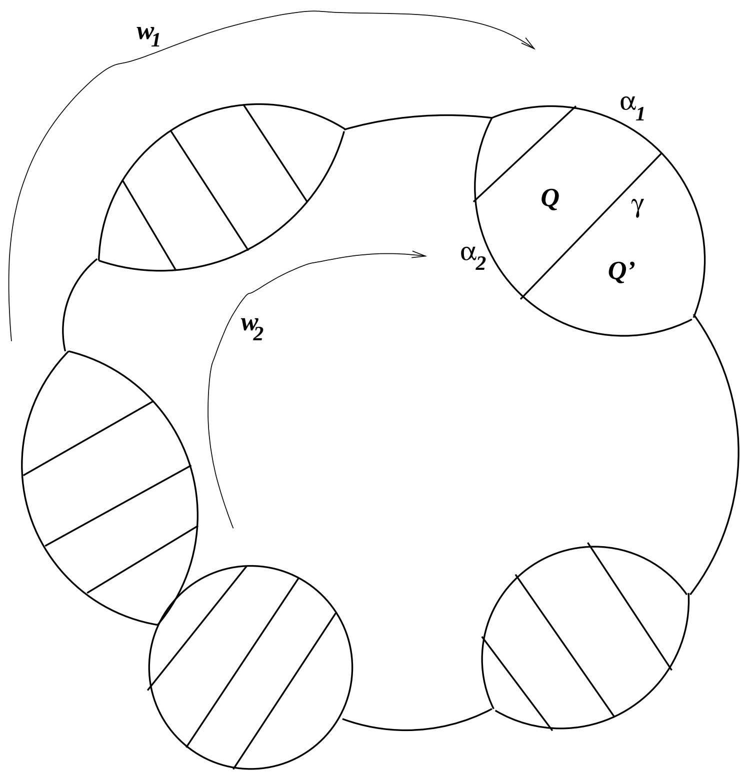

FIGURE 2. Conjugacy diagram in a small cancellation group

Remark 2.5. Suppose that w_1 and w_2 are as in part (1) of Proposition 2.4 and suppose that w_1 and w_2 are Dehn-reduced (which is a stronger assumption than being λ-reduced). Let Q be a region of D intersecting the upper and the lower boundaries of D in segments α_1 and α_2. Let $r \in R$ be the label of Q. Since the overlaps of Q with the neighboring regions in D have length at most $\lambda|r|$ each and since $|\alpha_j| \le |r|/2$ by Dehn-reduceness of w_1 and w_2, it follows that

$$|\alpha_j| \ge |r| - 2\lambda|r| - |r|/2 = |r|(1/2 - 2\lambda) \ge |r|\frac{48}{100} \ge |r|/3,$$

since $\lambda \le 1/100$. Thus in this case

$$|r|/3 \le \alpha_j \le |r|/2.$$

Similar conclusions apply to case (2) of Proposition 2.4 if we assume w_1 and w_2 to be cyclically Dehn reduced there.

Corollary 2.6. *Let $G = \langle a_1, \dots, a_k | R\rangle$ be a $C'(\lambda)$ presentation where $\lambda \le 1/100$. Suppose that for every $r \in R$ we have $|r| \ge 2/\lambda + 1$. Then:*

1. *For $i \ne j$ the elements $a_i^{\pm 1}$ and $a_j^{\pm 1}$ are not conjugate in G.*
2. *Suppose that wherever a_s^p is a subword of some $r \in R$, where $1 \le s \le k$, then $|p| < \lambda|r|$. Then for any $i \ne j$ and any $m \ne 0, n \ne 0$, the elements a_i^m and a_j^n are not conjugate in G.*
3. *Suppose that w is a cyclically reduced word in $F(A)$ that is conjugate in G to a_1. Then either $w = a_1$ in $F(A)$ or w is not cyclically λ-reduced.*

Lemma 2.7. *Let $G = \langle a_1, \dots, a_k | R\rangle$ be a $C'(\lambda)$ presentation where $\lambda \le 1/100$. Let $g \in G, g \ne 1$ be arbitrary. Then there exist freely reduced words $u, v \in F(a_1, \dots, a_k)$ such that:*

1. *We have $g =_G vuv^{-1}$.*
2. *The word u is cyclically reduced and cyclically Dehn-reduced. Moreover, u is d_G-geodesic and the element of G represented by u is of shortest possible length among all elements conjugate to g in G.*
3. *The word v is Dehn-reduced.*
4. *The word vuv^{-1} is freely reduced, as written, and is λ-reduced.*
5. *If z is any λ-reduced word in $F(a_1, \dots, a_k)$ representing g, then either $z = vuv^{-1}$ in $F(a_1, \dots, a_k)$ or there is some $r \in R$ such that both z and vuv^{-1} contain subwords of r representing of length $\ge \lambda|r|$.*

Proof. Consider all representations of g as $g = hg_0h^{-1}$ where $g_0 \in G$ is shortest in the conjugacy class of g. Once g_0 is fixed, among all such representations of g as $g = hg_0h^{-1}$, choose one, $g = hg_0h^{-1}$, where $h \in G$ is the shortest possible. Let u be a G-geodesic representative of g_0 and let v be a G-geodesic representative of h. Note that the minimality in the choice of g_0 implies that u is cyclically reduced and cyclically Dehn-reduced. Also, the minimality in the choice of h implies that the word vuv^{-1} is freely reduced as written. Since v is a geodesic word, it is Dehn-reduced. We claim that the word vuv^{-1} is λ-reduced. Indeed, suppose not. Then vuv^{-1} contains a subword w such that w is also a subword of some $r \in R$ with $|w| \ge (1 - 3\lambda)|r|$.

Note that since u and v are Dehn-reduced, the subword w overlaps at least two of the subwords v, u, v^{-1} in vuv^{-1}.

Case 1. The word w is a subword of vu or of uv^{-1}. We assume that w is a subword of vu, as the other case is symmetric. Then $w = w_1w_2$, where w_1 is a terminal segment of v and w_2 is an initial segment of u. We write v and u as $v = v'w_1$ and $u = w_2u'$. Let $y \in F(A)$ be such that $r = wy = w_1w_2y$ in $F(A)$, so that $|y| < 3\lambda|r|$. Note that since u and v are Dehn-reduced and $|w| \ge (1 - 3\lambda)|r|$, it follows that $|w_1|, |w_2| \ge (\frac{1}{2} - 3\lambda)|r| \ge |r|/3 > 3\lambda|r|$. Observe also that $v = v'w_1 =_G v'y^{-1}w_2^{-1}$. Therefore

$$g =_G vuv^{-1} = (v'w_1)(w_2u')(w_1^{-1}(v')^{-1}) =_G$$
$$=_G v'y^{-1}w_2^{-1}w_2u'w_2y(v')^{-1} = (v'y^{-1})(u'w_2)(y(v')^{-1}).$$

Since $|y| < 3\lambda|r|$ and $|w_1| > 3\lambda|r|$, it follows that $|v'y^{-1}|_A < |v|$, contradicting the minimality in the choice of h.

Case 2. The subword w overlaps both v and v^{-1} in vuv^{-1}

If the overlap of w with one of v, v^{-1} has length $\le \lambda|r|$, then either vu or uv^{-1} contains a subword of r of length $\ge (1-5\lambda)|r|$, and we get a contradiction similarly to Case 1.

If the overlaps of w with each of v, v^{-1} have length $> \lambda|r|$, we get a contradiction with the $C'(\lambda)$-small cancellation condition.

Thus part (1), (2), (3) and (4) of the lemma are established. Since the words vuv^{-1} and z are both λ-reduced, part (5) of the lemma now follows from Proposition 2.4. $\square$

3. Genericity

In this paper we work with the Arzhantseva-Ol'shanskii model of genericity in free groups based on the asymptotic density considerations.

Convention 3.1. Let $F = F(A)$, where $A = \{a_1, \ldots, a_k\}$ and $k \geq 2$. Let $m \geq 1$ and let $U \subseteq F^m$ be a subset of F^m. For $n \geq 0$ we denote by $\gamma_A(n, U)$ the number of all m-tuples $(u_1, \ldots, u_m) \in U$ such that $|u_i|_A = n$ for $i = 1, \ldots, m$. Note that for $n \geq 1$ we have $\gamma_A(n, F^m) = \big(2k(2k-1)^{n-1}\big)^m$. We say that $U \subseteq F^m$ is *spherically homogeneous* if for every $(u_1, \ldots, u_m) \in U$ we have $|u_1|_A = \cdots = |u_m|_A$. Note that this restriction is vacuous if $m = 1$. Let $S_m = S_{m,A}$ denote the set of all m-tuples $(u_1, \ldots, u_m) \in F^m$ such that $|u_1|_A = \cdots = |u_m|_A$.

Let $\mathcal{C} = \mathcal{C}_A$ be the set of all cyclically reduced words in $F(A)$. Let $\mathcal{C}_m = \mathcal{C}_{m,A} = \mathcal{C}^m \cap S_m$. Thus $\mathcal{C}_m$ consists of all m-tuples $(u_1, \ldots, u_m)$ of cyclically reduced words in $F(A)$ such that $|u_1| = \cdots = |u_m|$.

Although the notion of genericity make sense for arbitrary subsets of F^m, for reasons of simplicity we will restrict ourselves to spherically homogeneous subsets in this paper. Moreover, in applications we will only be concerned with tuples of cyclically reduced words.

Definition 3.2. Let $F = F(A)$, where $A = \{a_1, \ldots, a_k\}$ and $k \geq 2$. Let $m \geq 1$. Let $U \subseteq S_m$ be a spherically homogeneous subset. Let $U' \subseteq U$.

(1) We say that U' *is generic in* U if
$$\lim_{n\to\infty} \frac{\gamma_A(n, U')}{\gamma_A(n, U)} = 1.$$
If, in addition, the convergence to 1 in the above limit is exponentially fast, we say that U' is *exponentially generic in* U.

(2) We say that a subset $U' \subseteq U$ is *negligible in* U (correspondingly *exponentially negligible in* U) if $U - U'$ is generic (correspondingly exponentially generic) in U.

We stress that the above notions of genericity and negligibility are highly dependent on the choice of a free basis A of F. Therefore in all our discussions regarding genericity such a free basis is assumed to be fixed.

The following is a straightforward corollary of the definitions:

Lemma 3.3. *Let $U \subseteq \mathcal{C}_m$. Then the following are equivalent:*

(1) *The set U is exponentially negligible in $\mathcal{C}_m$.*

(2) *We have*
$$\lim_{n\to\infty} \frac{\gamma_A(n, U)}{(2k-1)^{nm}} = 0,$$
with exponentially fast convergence.

(3) *We have*
$$\limsup_{n\to\infty} \frac{\log \gamma_A(n, U)}{nm} < \log(2k-1).$$

Similarly, the definitions imply:

Lemma 3.4. *Let $m \geq 1$.*

(1) *The union of a finite number of (exponentially) negligible subsets of $\mathcal{C}_m$ is (exponentially) negligible in $\mathcal{C}_m$.*

(2) *The intersection of a finite number of (exponentially) generic subsets of* $\mathcal{C}_m$ *is (exponentially) generic in* $\mathcal{C}_m$.
(3) *Let* $U \subseteq \mathcal{C}$ *be an exponentially generic subset. Then* $U^m \cap S_m$ *is exponentially generic in* $\mathcal{C}_m$.

Convention 3.5. We say that a certain property of m-tuples of cyclically reduced elements of $F = F(A)$ is generic (correspondingly, exponentially generic) if there exists a generic (correspondingly, exponentially generic) subset $U \subseteq \mathcal{C}_m$ such that ever m-tuple in U has the property in question.

We list some properties of m-tuples of cyclically reduced elements of F are known to be exponentially generic.

Proposition 3.6. *Let* $F = F(A)$*, where* $A = \{a_1, \ldots, a_k\}$ *and* $k \geq 2$*. Let* $m \geq 1$*. Then:*
(1) *The property that no element of an* m*-tuple is a proper power in* F *is exponentially generic in* $\mathcal{C}_m$.
(2) *Let* $0 < \lambda < 1$ *be arbitrary. Then the property that an* m*-tuple* $(u_1, \ldots, u_m)$*, after cyclic reduction and symmetrization, satisfies the* $C'(\lambda)$ *small cancellation condition, is exponentially generic in* $\mathcal{C}_m$.
(3) *The property that for an* m*-tuple* $(u_1, \ldots, u_m)$ *for every* $i \neq j$ *the element* u_i *is not conjugate to* $u_j^{\pm 1}$ *in* F*, is exponentially generic in* $\mathcal{C}_m$.
(4) *Let* $K \geq 1$ *be an integer and let* $0 < \lambda < 1$*. Then the property of an* m*-tuple* $(u_1, \ldots, u_m)$ *that every subword* u *of some* u_i *of length* $\geq \lambda|u_i|$ *contains as a subword every freely reduced word of length* $\leq K$ *in* $F(A)$*, is exponentially generic in* $\mathcal{C}_m$.

4. Stallings folds and Nielsen equivalence

In this section we review Stallings folds and prove a weak version of our main theorem, we do the latter as the proof of this special case is quite easy but still illustrates some of the ingredients of the proof of the general case.

In a beautiful article [Sta] Stallings used graphs to represent subgroups of free groups and described how to use simple operations called folds to transform the graph into a graph at which a basis of the subgroup can be read off. We discuss Stallings folds for the free group $F = F(a_1, \ldots, a_k)$. Put $A = \{a_1, \ldots, a_k\}$.

Let now R_A be the directed labeled graph consisting of a single vertex v_0 and k loop-edges with labels $a_1, \ldots, a_k$. There is an obvious isomorphism

$$\phi : \pi_1(R_A, v_0) \to F$$

that maps the homotopy class represented by the ith loop edge travelled in positive direction to a_i.

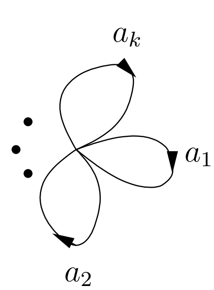

Figure 3. The graph R_A

For any directed labeled graph Γ with edge-labels from A there exists a unique label preserving graph map $p : \Gamma \to R_A$. After choosing a base point x of Γ the morphism p induces a homomorphism $p_* : \pi_1(\Gamma, x) \to \pi_1(R_A, v_0)$.

Given a tuple $T = (g_1, \ldots, g_s)$ of elements from $\pi_1(R_A, v_0) = F$ we can construct a graph S_T with base vertex x_0 such that $p_*(\pi_1(S_T, x_0)) = \langle g_1, \ldots, g_s \rangle$ as follows. We assume that $g_i \neq 1$ for $1 \leq i \leq l \leq s$ and $g_i = 1$ for $l < i \leq s$.

(1) S_T is a wedge of l circles with base vertex x_0 where the ith circle is of simplicial length $|g_i|_A$.
(2) The label of the ith circle is the reduced word representing g_i.

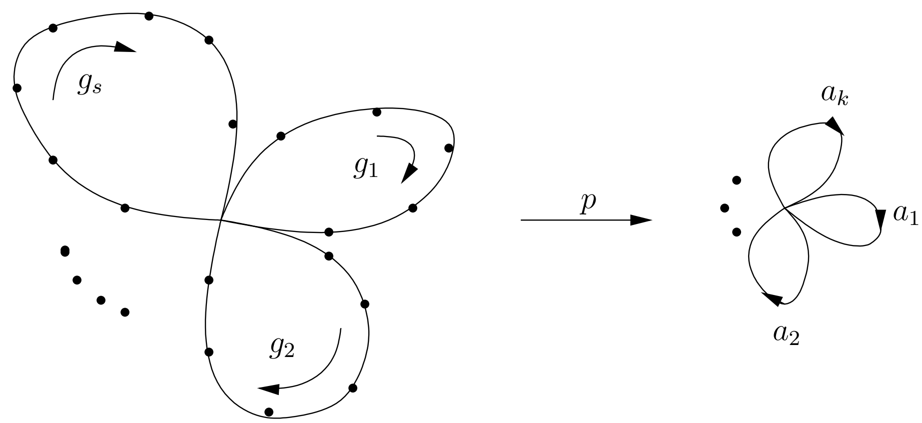

FIGURE 4. The map p from S_T to R_A

Stallings observed that any such marked graph can be folded without changing the image of the induced homomorphism. A fold here is the identification of two edges that have the same marking and the same initial or terminal vertex. Note that one or both vertices can be loop edges. In particular we have the following.

Lemma 4.1. *Let Γ be a marked graph such that $p_* : \pi_1(\Gamma, x_0) \to \pi_1(R_A, v_0)$ is surjective. Then there exists a finite sequence of labeled graphs*

$$\Gamma = \Gamma_0, \Gamma_1, \ldots, \Gamma_n = R_A$$

such that Γ_i can be obtained from Γ_{i-1} by a fold for $1 \leq i \leq n$.

Definition 4.2. We say that a finite connected graph Γ is a *core graph*, if it does not have any degree-one vertices. For a finite connected graph Γ with a nontrivial fundamental group let $Core(\Gamma)$ be the unique smallest subgraph of Γ whose inclusion into Γ is a homotopy equivalence.

Thus $Core(\Gamma)$ is obtained from Γ by cutting off a (possibly empty) collection of tree-branches. Note that $Core(\Gamma)$ is a core graph. We need the following simple observation for some graphs that fold onto R_A with a single fold.

Lemma 4.3. *Let Γ be a labeled directed core graph of rank k.*

(1) *If Γ folds onto R_A with a single fold then there exists a reduced word of length 2 that cannot be read as the label of an edge-path in Γ.*
(2) *If Γ folds onto R_A with a single edge sticking out then there exists a reduced word of length 4 that cannot be read as the label of an edge-path in Γ.*

Proof. (1) Note that the fold must identify a loop edge with a non-loop edge as identifying two loop edges decreases the rank and identifying two non-loop edges yields a graph with a non-loop edge. Thus w.l.o.g. we can assume that Γ has two vertices x and y and the fold identifies a loop edge at x and an edge from x to y both with label a_1. Γ has $k-1$ more edges that are labelled by $a_2, \ldots, a_k$.

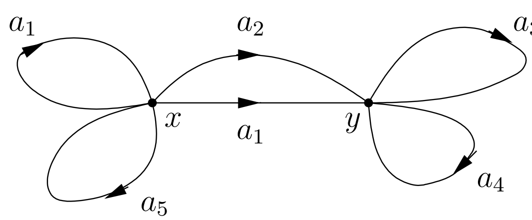

FIGURE 5. A core graph that folds onto R_A with a single fold

As we assume Γ to be a core graph there must either exist a loop edge at y or a second edge from x and y. W.l.o.g. we can assume that this edge is labelled by a_2, see Figure 5. In either case we cannot read the word a_2a_1 in Γ as the path is at the vertex y after reading a_2; this proves the assertion.

(2) In this case the fold must identify two non-loop edge with distinct endpoints, i.e. we must be in the situation of Figure 6. If a reduced word is read by a path in Γ then the vertex at which the fold is based can only occur as the initial and the terminal vertex of this path. As there is clearly a word of length 2 that cannot be read by the graph spanned by the remaining vertices (the word is a_1a_1 in the example) the claim follows. □

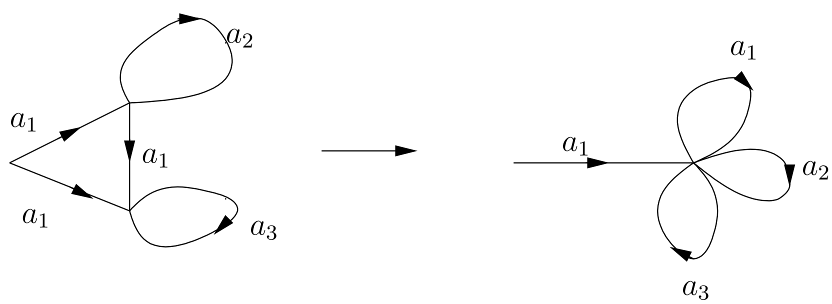

FIGURE 6. A core graph that folds onto R_A with an edge sticking out

We are now able to prove a weak version of our main theorem. It is a trivial consequence of Stalling folds and the above observations. One of the main observations used in the proof of the main theorem already shows up.

Theorem 4.4. *Let $k \geq m \geq 2$ and let $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_m\}$ be free bases of $F(A)$ and $F(B)$. Then there exist exponentially generic subsets $V \subseteq \mathcal{C}_{m,A}$ and $U \subseteq \mathcal{C}_{k,B}$ with the following properties.*

Let $(u_1(\bar{b}), \ldots, u_k(\bar{b})) \in U$ and $(v_1(\bar{a}), \ldots, v_m(\bar{a})) \in V$ be arbitrary and let G be given by the presentation

(!) $\quad G = \langle a_1, \ldots, a_k, b_1, \ldots b_m \mid a_i = u_i(\bar{b}), b_j = v_i(\bar{a})$ *for* $1 \leq i \leq k, 1 \leq j \leq m\rangle$.

Then the k-tuple $(a_1, \ldots, a_k)$ is not Nielsen equivalent in G to a k-tuple of form $(b_1, *, \ldots, *)$.

Proof. We can choose exponentially generic sets $V \subseteq \mathcal{C}_{m,A}$ and $U \subseteq \mathcal{C}_{k,B}$ so that the following hold:

(1) For every $(u_1(\bar{b}),\dots,u_k(\bar{b}))\in U$ and $(v_1(\bar{a}),\dots,v_m(\bar{a}))\in V$ the presentation (!) satisfies the $C'(1/100)$ small cancellation condition, after symmetrization.
(2) For every $(u_1(\bar{b}),\dots,u_k(\bar{b}))\in U$ and $(v_1(\bar{a}),\dots,v_m(\bar{a}))\in V$ we have $|u_i|_B\ge 10^{10}$ and $|v_j|_A\ge 10^{10}$.
(3) For every $(v_1(\bar{a}),\dots,v_m(\bar{a}))\in V$ and every $j=1,\dots,m$, every subword of v_j of length $\ge |v_j|/100$ contains all freely reduced words of length two in $F(A)$ as subwords.

We now argue by contradiction. Assume that for some G as in Theorem 4.4, the k-tuple $(a_1,\dots,a_k)$ is Nielsen equivalent in G to a k-tuple $(b_1,*,\dots,*)$. This implies that $(a_1,\dots,a_k)$ is in $F(A)$ Nielsen equivalent to a tuple $T=(w_1(\bar{a}),\dots,w_k(\bar{a}))$ such that $w_1(\bar{a})=b_1$ in G. Note that we might have $w_1(\bar{a})\neq v_1(\bar{a})$ in $F(A)$. We do not distinguish between the $w_i(\bar{a})$ and freely reduced words in $F(A)$ representing them. After a possible conjugation of T in $F(A)$, we may assume that w_1 is cyclically reduced in $F(A)$ and that w_1 is conjugate to b_1 in G.

By Corollary 2.6, $w_1(\bar{a})$ must contain at least one fourth of a cyclic permutation of one of the defining relations $\left(v_j(\bar{a})b_j^{-1}\right)^{\pm 1}$, since $w_1(\bar{a})$ is cyclically reduced and is conjugate to b_1 in G. This implies in particular that w_1 contains every freely reduced word of length two in $F(A)$ as a subword. Let now S_T be the wedge as above and choose a sequence

$$S_T=\Gamma_0,\Gamma_1,\dots,\Gamma_n=R_A$$

such that Γ_i can be obtained from Γ_{i-1} by a fold. Note that each Γ_i is a core graph. This holds as Γ_i is the image of loops labeled with freely reduced word and the base point must lie in the core as the cyclically reduced word w_i can be read in each Γ_i by a closed path based at the base vertex. Thus Γ_{n-1} is a core graph that folds onto R_A with a single fold and w_1 can be read by a closed path in Γ_i. This contradicts Lemma 4.3. □

Remark 4.5. The proof of Theorem 4.4 recovers the well-known fact that for any cyclically reduced primitive element g in $F(a_1,\dots,a_k)$ there exists some reduced word w of length two such that neither w nor w^{-1} occur as a subword of g.

5. The proof of Theorem 1.1

This section is dedicated to the proof of the main theorem. While the strategy of the proof is similar to the strategy of the proof of Theorem 4.4 the argument turns out to be significantly more involved.

We will prove the following theorem which is the precise formulation of Theorem 1.1 from the Introduction.

Theorem 5.1. *Let $k\ge m\ge 2$ and let $A=\{a_1,\dots,a_k\}$ and $B=\{b_1,\dots,b_m\}$ be free bases of $F(A)$ and $F(B)$. Then there exist exponentially generic subsets $V\subseteq \mathcal{C}_{m,A}$ and $U\subseteq \mathcal{C}_{k,B}$ with the following properties.*

Let $(u_1(\bar{b}),\dots,u_k(\bar{b}))\in U$ and $(v_1(\bar{a}),\dots,v_m(\bar{a}))\in V$ be arbitrary such that $|u_i|=|v_j|$ for $1\le i\le k$ and $1\le j\le m$. Let G be given by the presentation

(!!) $\quad G=\langle a_1,\dots,a_k,b_1,\dots b_m \mid a_i=u_i(\bar{b}), b_j=v_j(\bar{a}) \text{ for } 1\le i\le k, 1\le j\le m\rangle.$

Then the $(k+1)$-tuple $(a_1,\dots,a_k,1)$ is not Nielsen equivalent in G to a $(k+1)$-tuple of form $(b_1,b_2,*,\dots,*)$.

We will not specify at the outset an explicit list of conditions for V and U that are exponentially generic and that will ensure that the statement of Theorem 5.1 holds. Rather we will make the several required choices for V and U along the way in the proof, and each of U, V will be constructed as the intersection of a finite number of exponentially generic sets.

However, at the start we do require V and U to satisfy the following conditions. We choose $0 \le \lambda \le 10^{-100}$ and $0 < \lambda_1 \le 10^{-100}\lambda$.

Condition 5.2. The following hold for U and V:

(1) There are exponentially generic subsets $\mathcal{V} \subseteq \mathcal{C}_A$ and $\mathcal{U} \subseteq \mathcal{C}_B$ such that $U \subseteq \mathcal{U}^k$ and $V \subseteq \mathcal{V}^m$.
(2) For every $(u_1(\bar{b}), \ldots, u_k(\bar{b})) \in U$ and $(v_1(\bar{a}), \ldots, v_m(\bar{a})) \in V$ the presentation (!!) satisfies the $C'(\lambda_1)$ small cancellation condition, after symmetrization.
(3) For every $u \in \mathcal{U}$ and $v \in \mathcal{V}$ we have $|u|_B \ge 10^{100000}/\lambda_1$ and $|v|_A \ge 10^{100000}/\lambda_1$.
(4) For every $v \in \mathcal{V}$, every subword of v of length $\ge \lambda_1(|v| + 1)/1000$ contains all freely reduced words of length 1000 in $F(A)$ as subwords.
(5) For every $u \in \mathcal{U}$, every subword of u of length $\ge \lambda_1(|u| + 1)/1000$ contains all freely reduced words of length 1000 in $F(B)$ as subwords.
(6) If some defining relation r of (!!) contains two distinct occurrences of a subword y then $|y| \le \lambda_1|r|/1000$.

Proposition 3.6 implies that exponentially generic subsets $V \subseteq \mathcal{C}_{m,A}$ and $U \subseteq \mathcal{C}_{k,B}$ with the above properties exist.

The proof of Theorem 5.1 is again by contradiction. Thus we assume that the $(k+1)$-tuple $(a_1, \ldots, a_k, 1)$ is Nielsen equivalent in G to a tuple of type $(b_1, b_2, *, \ldots, *)$. It follows that in $F(A)$ the tuple $(a_1, \ldots, a_k, 1)$ is Nielsen equivalent to a tuple of type $(w_1(\bar{a}), \ldots, w_k(\bar{a}), w_{k+1}(\bar{a}))$ such that $b_i = w_i(\bar{a})$ in G for $i = 1, 2$.

Let $T = (w_1(\bar{a}), \ldots, w_k(\bar{a}), w_{k+1}(\bar{a}))$ be a tuple with the above properties such that $|w_1|_A + |w_2|_A$ is minimal among all such tuples.

Let S_T be the labeled graph as in Section 4 and choose a sequence

$$S_T = \Gamma_0, \Gamma_1, \ldots, \Gamma_l = R_A$$

such that Γ_i can be obtained from Γ_{i-1} by a Stallings fold for $1 \le i \le l$. Choose q maximal such that Γ_q does not contain a subgraph R'_A that is isomorphic to R_A as a labeled directed graph. Put $\Delta := \Gamma_q$. Δ contains a subgraph Ψ of rank k with at most $k+2$ edges that folds onto a subgraph of Γ_{q+1} that is either isomorphic to R_A or to R_A with a single edge sticking out, in particular Ψ contains at most $k+2$ edges.

Observe also that, since Ψ contains at most $k+2$ edges, there are only finitely many possibilities for Ψ. Hence by Proposition 3.6 and Lemma 4.3, the set of all freely reduced words that can be read along some paths in such graphs Ψ is an exponentially negligible subset $\mathcal{S}$ of $F(A)$. Hence we may assume in our choice of $V \subseteq S_{m,A}$ that the following holds:

Condition 5.3. If some freely reduced word y in $F(A)$ is readable along an edge-path in Ψ and y is a subword of some defining relation r then $|y| \le \lambda_1|r|/1000$.

Note that $Core(\Delta)$ cannot coincide with Ψ by the same argument used in the proof of Theorem 4.4. Therefore the following observation is obvious, here a lollipop is loop (candy) with an edge (the stick) attached. We will allow the stick to be degenerate, i.e. the lollipop to be a loop.

Lemma 5.4. *The graph Δ has the following properties:*

(1) *The subgraph $\bar{\Delta} = Core(\Delta)$ carries the fundamental group of Δ and one of the following holds:*

(a) *$\bar{\Delta}$ can be obtained from Ψ by attaching a (subdivided) segment γ along both endpoints to distinct vertices of Ψ.*
(b) *$\bar{\Delta}$ can be obtained from Ψ by attaching a (subdivided) lollipop along its unique valence one vertex if the stick is non-degenerate and an arbitrary vertex otherwise.*

(2) *Either the base-vertex of Δ lies in $\bar{\Delta}$ in which case $\Delta = \bar{\Delta}$, or Δ is obtained from $\bar{\Delta}$ by attaching a segment to $\bar{\Delta}$ along one of its endpoints. The other endpoint of the segment is then the base vertex x_0 of Δ.*

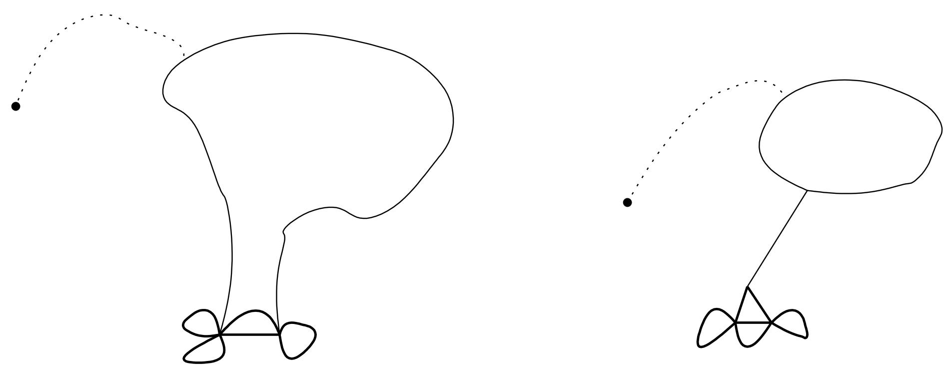

FIGURE 7. The two options for the graph Δ

Note that in Figure 7 the graph Ψ consists of the fat lines. A trivial but important fact is that any of the words $w_i(\bar{a})$ can be read by a path in Δ based at the base vertex as the map p from S_T to R_A factors through Δ. This path is clearly reduced as the words $w_i(\bar{a})$ are reduced.

The following lemma implies in particular that there exists a reduced word that is not the label of a path in Δ, it is a generalisation of Lemma 4.3.

Lemma 5.5. *Let Γ be a finite labeled graph of rank $k+1$ that folds onto R_A but does not contain a isomoprhic copy of R_A. Then there exists a reduced word that is not a label of a path in Γ.*

Proof. It suffices to deal with the case that Γ is such that any applicable fold produces a subgraph that is isomorphic to R_A as we could otherwise just apply a fold that does not produce such a subgraph. As any word that was readable before the fold is also readable after the fold it follows that the assertion for the new graph implies the assertion for the original one. We can further assume that Γ is a core graph as a word that is not readable in $Core(\Gamma)$ clearly has a power that is not readable in Γ.

Thus Γ contains a graph Ψ which either folds onto R_A or onto R_A with a single edge sticking out. Furthermore Γ is obtained from Ψ by attaching an edge or an lollipop as in described in Lemma 5.4. There are two cases: Case 1, where Ψ folds onto R_A, and Case 2, where Ψ folds onto R_A with an edge sticking out. We will give all details in Case 1 and leave Case 2 to the reader. No new arguments are needed to complete the proof.

Case 1: Ψ folds onto R_A with a single fold. Recall (see proof of Lemma 4.3) that w.l.o.g. Ψ has two vertices x and y, a loop edge at x with label a_1, an edge with label a_1 from x to y and $k-1$ more edge with labels $a_2, \ldots, a_k$, see Figure 5. Γ is obtained from Ψ by attaching a segment, resp. lollipop, with, say l, edges.

In the following we will denote the attached segment, resp. lollipop, by s. By the label of s we mean the label of the segment if s is a segment and the label of the path that walks once around the lollipop otherwise. It is easily verified that we can assume that the label of s is freely reduced as we are only looking at freely reduced words that can be read in Γ. We distinguish three subcases depending on the size of l, the number of edges of s.

Case 1A: Suppose first that $l \geq 3$. Then we cannot fold an edge of s onto Ψ as such a fold could not produce a copy of R_A contradicting our assumption. Note next that this assumption implies that the label of s cannot be a power of a_1 and cannot, if it starts (ends) at x start (end) with an edge labelled by $a_1^{\pm 1}$ as we could otherwise fold this edge onto Ψ.

It is now easily seen that any path that reads a_1^{-l} as a subword must be either at the vertex x after having read a_1^{-l} or must have traveled exclusively on the candy of the lollipop which itself is labeled by a power of a_1 and has a non-degenerate stick, see Figure 8.

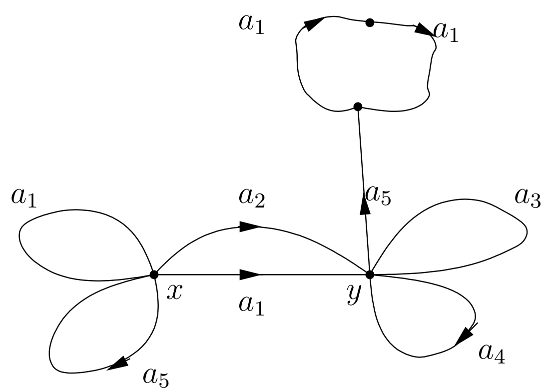

FIGURE 8. A graph Γ where a_i^l can be read in different places

If we now pick $a_j^\eta \neq a_1^{\pm 1}$ such that a_j^η is not the last edge on the stick of lollipop if the lollipop has a non-degenerate stick, i.e. choose $a_j^\eta \neq a_5$ in the example drawn in Figure 8. Then any path that reads $a_j^\eta a_1^{-l}$ must by at vertex x after having read $a_j^\eta a_1^{-l}$.

Now there is a word $a_j^\eta a_1^{-l} a_k^\varepsilon$ that can only be read by entering s (or the first edge of s^{-1}) to read a_k^ε. This implies in particular that s is attached to Ψ at x. It follows that the word $a_j^\eta a_1^{-l} a_k^\varepsilon a_l^\xi$ cannot be read in Γ if a_l^ξ is such that $a_k^\varepsilon a_l^\xi$ is not the initial word of the label of s (if s starts at x) and not the initial word of the label of s^{-1} if s ends at x. This proves the claim as it implies that $a_j^\eta a_1^{-l} a_k^\varepsilon a_l^\xi$ is not readable in Γ.

Case 1B: Suppose now that $l = 2$. If s does not fold onto Ψ then we argue as in the case $l \geq 3$. Thus we can assume that one of the two edges of s folds onto an edge of Ψ. As such a fold does not identify x and y it follows that this fold must produce a new loop edge attached to either x or y and that there must have been already a wedge of $k-1$ circles at x or y. We distinguish three different configurations.

Suppose first that s is a segment of length 2 from x to y. In this case an edge of s must be folded onto an edge from x to y to produce a loop.

If the fold produces a copy of R_A based at y then Ψ must have consisted of a single loop with label a_1 at x, a single edge with label a_1 from x to y and $k-1$ loop edges with labels $a_2, \ldots, a_k$ at y. The fold must add a loop with label a_1 which implies that the segment s from x to y has label $a_1 a_1$. This clearly means that the path $a_1^{-1}a_2$ is not readable in Γ as no path is at the vertex y after reading a^{-1}, see Figure 9.

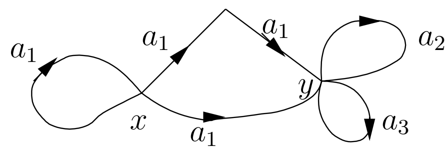

FIGURE 9. A copy of R_A emerges at y

If the fold produces a copy of R_A based at x then Ψ had w.l.o.g. $k-1$ loop edges based at x with labels $a_1, a_3, \ldots, a_k$, an edge with label a_1 from x to y and either (a) a loop edge with label a_2 based at y or (b) an edge with label a_2^η from x to y. The fold must produce a loop edge with label a_2 at x which implies that the segment s from x to y must have label $a_2^\eta a_1$ in case (a) or $a_2^{2\eta}$ or $a_2^\eta a_1$ in case (b). In all of these cases it is apparant that $a_2^2 a_1^{-1}$ is not readable, see Figure 10.

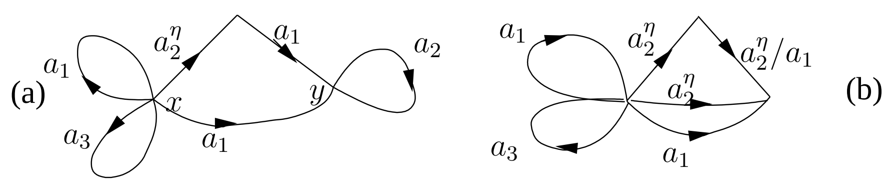

FIGURE 10. A copy of R_A emerges at x

Suppose now that s is a loop of length 2. As the fold must produce a loop edge at x or y it must fold an edge of s onto a loop edge. The descriptions of Ψ are as in the case before.

If a loop edge is added at y then the added loop edge must have label a_1, thus s must have been a loop based at y with label $a_i^{\pm 1}a_1^{\pm 1}$ with $i \geq 2$. In this case the word $a_1^{-2}a_2$ is not readable in Γ.

If a loop edge is added at x then the added loop edge must have label a_2 which implies that s must be a loop based at x with label $a_i^{\pm 1}a_2^{\pm 1}$ with $i \neq 2$. It follows that $a_2^2 a_1^{-1}$ is not readable in Γ.

Suppose lastly that s is a lollipop whose stick and loop both have a single edge. In this case we can argue precisely as in case 1A.

Case 1C: Suppose now that $l = 1$. We distinguish the cases were s is a loop edge and were it is an edge from x to y.

Suppose first that s is a loop edge. If we add a loop edge with label a_j to x then there must still be a path $a_1^{-1}a_k^{\pm 1}$ that cannot be read in Γ as a path must be at x after reading a_1^{-1} and we assume that Γ contains no copy of R_A meaning there is not a loop edge at x for every $i = 1, \ldots, k$. The same argument works if we add a loop edge with label different from a_1 at y. If we add a loop edge with label a_1 to y then inspecting the two possibilities reveals that there always exists a word of type $a_2^\eta a_j^\epsilon$ that is not readable as we assume that not all loop edges are present at y.

Suppose lastly that s is an edge from x to y. If the edge is labelled by $a_1^{\pm 1}$ then there is a word of type $a_2^\eta a_j^\varepsilon$ that is not readable in Γ. Thus we can assume that the edge is labelled by an element different from $a_1^{\pm 1}$. Note that this implies that any path that starts with a_1^{-1} is at the vertex x after reading a_1^{-1}, thus we can assume that any $a_i^{\pm 1}$ can be read by an edge emanating at x. For at least one a_j^ε

the corresponding edge leads to y. This implies that the word $a_1^{-1}a_j^{\varepsilon}a_1$ cannot be read in Γ. $\square$

Let now z be the word defined as follows: In the case (1)(a) of Lemma 5.4 z is the label of the segment attached to Ψ. In the case (1)(b) of Lemma 5.4 z is the label of the closed reduced edge-path in $\bar{\Delta}$ based at the attaching point of the lollipop that walks once around the lollipop, i.e. it is a word of type uwu^{-1} where u is the label of the stick of the lollipop.

Let now $\bar{z}$ be a geodesic word in G (i.e. a word in the a_i, b_i and their inverses) such that in G we have the identity $z =_G \bar{z}$. We will distinguish two cases, namely that $\bar{z}$ is a word in the a_i containing no subword of a defining relation r of length greater than $\frac{1}{100}|r|$ or that it does contain either a subword of length $\frac{1}{100}|r|$ of some defining relation r or some letter $b_i^{\pm 1}$.

Case A: The word $\bar{z}$ contains either some letter $b_i^{\pm 1}$ or a subword of length $\geq |r|/100$ of some defining relation r.

Let now x_1 be a base point in Ψ that is incident to the edge or lollipop with label z. We study elements represented by closed reduced paths in Δ based at x_1. It follows from the construction that there exist such elements for $i = 1, 2$ that represent elements that are conjugate to b_i in G. Clearly such elements are of type

$$w_0 z^{\varepsilon_1} w_1 \cdot \ldots \cdot w_{k-1} z^{\varepsilon_k} w_k$$

where the w_i are reduced words read by reduced paths in in Ψ and $\varepsilon \in \{-1, 1\}$ for all i. As in G we have $z = \bar{z}$ it follows that there are also products of type

$$w_0 \bar{z}^{\varepsilon_1} w_1 \cdot \ldots \cdot w_{k-1} \bar{z}^{\varepsilon_k} w_k$$

with the w_i as above that represent elements conjugate to b_1 and b_2. Note that the new word might no longer be freely reduced even if the initial one was.

Recall that all relations are assumed to be of the same length, in the following we will denote this length by N.

We will now study how long portions of some relation can occur in such products after free reduction. The following lemma implies in particular that in words of type $\bar{z}^{\varepsilon} w \bar{z}^{\varepsilon}$ with $\varepsilon \in \{-1, 1\}$ this can happen for at most one choice of w.

Lemma 5.6. *Let h be a reduced word that either contains some $b_i^{\pm 1}$ or a subword of a relation of length at least $\geq N/100$.*

For any freely reduced word t that can be read by a path in Ψ let W^t be the word obtained from hth by free reduction. We write W^t as a reduced product $W^t = W_1^t W_2^t W_3^t$, where W_1^t, W_2^t and W_3^t are the portions of h, t, h accordingly that survive a free reduction of the product hth.

There is a reduced word w readable in Ψ such that for each reduced $\bar{w} \neq w$ readable in Ψ the following holds:

1. *$W^{\bar{w}}$ does not contain a subword y of one of the defining relators r of the symmetrized form of presentation (!!) such that y has overlaps with both $W_1^{\bar{w}}$ and $W_3^{\bar{w}}$ in subwords of length at least $100\lambda_1 N$;*
2. *If in the free reduction of $h\bar{w}h$ some letter b_i^{ε} cancels then $h = ub_i^{\varepsilon} r_1 \hat{g} r_2 b_i^{-\varepsilon} v^{-1}$ such that the following hold:*
   (a) *u and v are words in the $a_i^{\pm 1}$.*
   (b) *r_1 and r_2 are words in the $b_i^{\pm 1}$ and $|r_1|, |r_2| \geq 100\lambda_1 N - 1$.*
   (c) *$W^w = ub_i^{\varepsilon} r_1 \hat{g} r_2 b_i^{-\varepsilon} a_j^{\eta} b_i^{\varepsilon} r_1 \hat{g} r_2 b_i^{-\varepsilon} v^{-1}$ and $r_2 b_i^{-\varepsilon} a_j^{\eta} b_i^{\varepsilon} r_1$ is a subword of a relation, i.e. situation (1) occurs for the word w.*
   (d) *$W^{\bar{w}} = ub_i^{\varepsilon} r_1 \hat{g} \bar{r}_2 \bar{r}_1 \hat{g} r_2 b_i^{-\varepsilon} v^{-1}$ where for $i = 1, 2$ $\bar{r}_i$ is the remnant of r_i and $|\bar{r}_i| \geq 90\lambda_1 N$. Furthermore $\bar{r}_2 \bar{r}_1$ is not a subword of a relation.*

(3) *In the free reduction of* $h\bar{w}h$ *no subword* y *of one of the two occurrences of* h *cancels such that* y *is also a subword of some word* $P \in \mathcal{U} \cup \mathcal{V}$ *of length* $|P| = N$ *with the property that* $|y| \geq 100\lambda_1 N$.

Moreover there is at most one word $\bar{w}$ *such that (2) occurs.*

Proof. We first show that there is at most one word w such that the conclusion of either (1) or (3) fails. We can assume that h contains a subword of a relation of length at least $100\lambda_1 N$ as there is nothing to show otherwise.

Let Y be the right-most subword in h such that $|Y| = 99\lambda_1 N$ and such that $Y \in \mathcal{V} \cup \mathcal{U}$. Similarly, let Y' be the left-most subword in h such that $|Y'| = 99\lambda_1 N$ and such that $Y' \in \mathcal{V} \cup \mathcal{U}$. If Y, Y' as above do not exist, then cases (1) and (3) of the lemma do not occur and there is nothing left to prove.

In what follows, when referring to Y we think of it as a subword in the first h in hwh and when referring to Y' we think of it as a subword in the second h in hwh. Recall that a freely reduced word w that is readable in Ψ can contain at most $\lambda_1|r|/100$ portion of a defining relator r.

Thus we assume that Y and Y' do exist. If case (3) fails, then a subword of Y of length at least $96\lambda_1 N$ must cancel with the subword of Y' of the same length in the free reduction of hwh. Therefore, if (3) fails for two distinct $w_1 \neq w_2$, then either Y has the form $Y = Y_1 y_2 = y_1 Y_1$ with $0 < |y_1| = |y_2| \leq 4\lambda N$ or Y' has the form $Y' = Y_1' y_2' = y_1' Y_1'$ with $0 < |y_1'| = |y_2'| \leq 4\lambda N$. It follows that Y or Y' is a periodic word, contrary to our assumptions about generic properties of $\mathcal{U}$ and $\mathcal{V}$. Thus (3) can fail for at most one w.

Suppose there exist distinct w_1 and w_2 as in the statement of the lemma such that (1) fails for w_1 and (3) fails for w_2. Again, in the free reduction of hw_2h, a subword of Y of length at least $96\lambda_1 N$ must cancel with the subword of Y'. If in the free reduction of hw_1h a subword of Y or Y' of length $\geq 70\lambda_1 N$ cancels, then we get a contradiction similarly to the above proof that (3) can fail for at most one w. Hence an initial segment s of Y of length $|s| \geq 29N\lambda_1$ and a terminal segment s' of Y' of length $|s'| \geq 29N\lambda_1$ survive the free reduction of hw_1h. The choice of Y and Y' now implies that each of s, s' is a subword of the word y, where y is a subword of some defining relator r. Recall however that since (3) fails for hw_2h the words Y and Y' almost cancel with each other and hence they are almost inverse. More precisely, a subword of Y of length at least $96\lambda_1 N$ must cancel with the subword of Y' of the same length. This means that there is a subword t of s of length $|q| \geq 10\lambda_1 N$ such that t^{-1} is a subword of s'. Thus both t and t^{-1} are subwords of r which contradicts the small cancellation assumption about (!!).

Suppose now that (1) fails for two different w_1 and w_2 and let y_1, y_2, r_1, r_2 be the corresponding subwords and relations as spelt out in (1). By the assumptions in Condition 5.3 the word $W_2^{w_i}$ can have an overlap with r_i of length at most $\lambda_1|r_i|/1000$. Hence $|W_2^{w_i}| \leq \lambda_1|r_i|/1000$ for $i = 1, 2$.

We claim that $||W_i^{w_1}| - |W_i^{w_2}|| \leq \lambda_1 N/10$ for $i = 1, 3$. Indeed, if that were not the case, then (3) would fail for either w_1 or w_2. But we have already established that it is not possible for (1) to fail for one w and for (3) to fail for a different w. Thus indeed $||W_i^{w_1}| - |W_i^{w_2}|| \leq \lambda_1 N/10$ for $i = 1, 3$. The fact that the subword y_i of W^{w_i} overlaps each of $W_i^{w_1}$ and $W_3^{w_i}$ in subwords of length $\geq N/10$, together with the small cancellation assumptions, implies that r_1 and r_2 are cyclic permutations of the same relator r. We now claim that $W_j^{w_1} = W_j^{w_2}$ for $j = 1, 2, 3$. If this were not the case, then r would contain subwords of the form $\alpha\beta\gamma$ and $\alpha\beta'\gamma$ where $|\alpha|, |\gamma| \geq \lambda|r|/100$, where $|\beta|, |\beta'| \leq \lambda_1|r_i|/1000$ and where $\beta \neq \beta'$. This contradicts the small cancellation condition. This shows that (1) can fail for at most one w as in the lemma.

Thus we have shown that there exists most one word w such that either (1) or (3) fails. Note further that there is at most one word $\bar{w}$ such that in $h\bar{w}h$ some letter $b_i^{\pm 1}$ cancels. If $w = \bar{w}$ or if w or $\bar{w}$ with the above properties does not exist then there is nothing to show. Thus we can assume that there is a word w such that either (1) or (3) fails and that for $\bar{w}$ some letter $b_i^{\pm 1}$ cancels. We have to show that this puts us into situation (2).

Suppose that (3) fails for w. Let b be the right-most occurrence of a letter $b_i^{\pm}1$ in the first h. Then the left-most occurrence of a letter from $B^{\pm 1}$ in the second h must be b^{-1} since in $h\bar{w}h$ these occurrences must cancel each other. Suppose first that b is to the left of the mid-point of Y in h. Then $Y, Y' \in F(A)$ and $Y, Y' \in \mathcal{V}$. Then all of Y, Y' have to cancel in both hwh and in $h\bar{w}h$. This is impossible by the same argument as in the proof that (3) can happen for at most one w. Thus b is to the right of the midpoint of Y in the first h. Hence b has to cancel both in hwh and $h\bar{w}h$, which is impossible, as noted above.

Suppose now that (1) fails for w. Again, let b be the right-most occurrence of a letter $b_i^{\pm}1$ in the first h. Then the left-most occurrence of a letter from $B^{\pm 1}$ in the second h must be b^{-1} since in $h\bar{w}h$ these occurrences must cancel each other. Since $w \neq \bar{w}$, these occurrences of b and b^{-1} do not cancel in hwh. Recall that the freely reduced form $W^w = W_1^w W_2^w W_3^w$ of hwh contains a subword y such that y is also a subword of a defining relation r and that y overlaps with both W_1^w and W_3^w in subwords of length at least $100\lambda_1 N$.

Suppose first that, up to cyclic permutation and inversion, r has the form $a_i = u_i(\tilde{b})$, where $u_i \in \mathcal{U}$. It follows that exactly one letter, namely $a_i^{\pm 1}$ survives in the free reduction of the segment between b and b^{-1} in hwh and, moreover, both the occurrences b and b^{-1} are inside of y. It follows that W_1^w ends with a subword c_1 and W_3^w begins with a subword c_2 which are of length at least $100\lambda_1 N$ and which are the terminal (initial) word of some $u_i(\tilde{b})^{\pm 1}$. Now genericity of u_i implies that not more than a subword of length $10\lambda_1 N$ can cancel in $c_1 c_2$ which puts us into (2)

Suppose lastly that up to a cyclic permutation and inversion, r has the form $b_i = v_i(\tilde{a})$. If b occurs to the left of the mid-point of Y in h then more than half of Y cancels in $h\bar{w}h$, which yields a contradiction as in the above argument that (3) and (1) cannot happen for different values of w. Hence b occurs after the midpoint of Y in the first h and, similarly, b^{-1} occurs before the mid-point of Y' in the second h. It follows that in order for (1) to occur at least one and therefore both of b and b^{-1} must cancel contradicting the fact that letters $b_i^{\pm 1}$ cannot cancel for different words. The last part of the proof clearly implies the more detailed description of (2).

This completes the proof of the lemma. □

Lemma 5.7. *Let h and w be as in the conclusion of Lemma 5.6. There exists a reduced word $\bar{h}$ in the a_i and b_i such that $h = \bar{h}$ in G such that the following hold:*

1. *$\bar{h}$ contains some $b_i^{\pm 1}$ or a subword of some a relation of length $\geq |v_j|/100$.*
2. *The conclusion of Lemma 5.6 applied to $\bar{h}$ holds for the same w as before.*
3. *$\bar{h}$ does not contain contain a subword of a relation r whose length is more than $\frac{6}{10}$ of the length of r.*
4. *For any $n \in \mathbb{N}$ the word obtained by free reduction of $(\bar{h}w)^n\bar{h}$ does not contain a subword of a relation r whose length is more than $\frac{99}{100}$ of the length of r.*
5. *If $\bar{h}$ contains a subword of some relation r of length $\geq |r|/100$ then in $(\bar{h}w)^n\bar{h}$ both the first and the last occurrence of $\bar{h}$ have subwords of length $\geq |r|/500$ of some relation that survive the free cancellation (of $(\bar{h}w)^n\bar{h}$).*

Proof We start with h and modify the word preserving properties (1), (2) and (3) which clearly hold for the initial word h. The final word obtained in this process is then $\bar{h}$.

Choose u maximal such that $h = u\hat{h}u^{-1}$ (as words). Note now that if (4) does not hold then we can already find a subword in $u\hat{h}\hat{h}u^{-1}$ that is a subword of a relation r whose length is more than $\frac{99}{100}$ of the length of r. This is true as such subword of $u\hat{h}^n u^{-1}$ that cannot already seen as a subword of $u\hat{h}^2u^{-1}$ would either contain a significant suffix of u and a significant prefix of u^{-1} or would contain a significant subword of $\hat{h}^n$ which is not a subword of $\hat{h}^2$; in both cases it is easy to deduce a contradiction to genericity of the relators.

Thus we can assume that $u\hat{h}\hat{h}u^{-1}$ contains a subword that is a subword of a relation r whose length is more than $\frac{99}{100}$ of the length of r. Genericity and (2) now implies that $\hat{h} = W_1W_2W_3$ such that W_3W_1 is a subword of r that is of length at least than $\frac{98}{100}$ of the length of r and that $|W_1| = |W_3|$, in particular $r = W_3W_1\bar{W}$ with $\bar{W}$ of length at most $\frac{2}{100}$ of the length of r. We replace $\hat{h}$ with $\tilde{h} = W_1W_2\bar{W}^{-1}W_1^{-1}$ which clearly does not change the element of G represented. After free reduction of $u\tilde{h}u^{-1}$ we obtain a new word that by construction satisfies conditions (1) and (3). It also satisfies condition (2) as the modification was only applied if the conclusion of (1) of Lemma 5.6 did not hold for $w = 1$ and in that case the conclusion of (3) of Lemma 5.6 does not hold for $w = 1$ after the modification.

Assertion (5) follows immediately from the above proof. □

Lemma 5.8. *Let h be as in the hypothesis of Lemma 5.6 and let w be a nontrivial freely reduced word read along a closed edge-path in Ψ.*

Let further W be the word obtained by freely reducing the product hwh^{-1}. Write W as a reduced product $W = W_1W_2W_3$, where W_1, W_2 and W_3 are the portions of h, w and h^{-1} accordingly that survive a free reduction of the product hwh^{-1}.

Then the following hold

1. *The free cancellation does not involve a letter $b_i^{\pm1}$ of $h^{\pm1}$ or a subword y of $h^{\pm r}$ of length $|y| > \lambda_1|r|/10$ where r is a defining relation and y is a subword of r.*
2. *Nontrivial portions of h and h^{-1} survive the free reduction.*
3. *Let y be a subword of W such that y is also a subword of some defining relation r that overlaps both W_1 and W_3 in W. Then the shorter of these two overlaps has length $\le \lambda_1|r|/10$ and $|W_2| \le \lambda_1|r|/1000$.*

Proof. Let $N = |r|$, where r is any defining relator from (!!), (recall that by assumption all the defining relators have the same length).

(1) The first claim is obvious while the later follows from the assumptions in Condition 5.2 and Condition 5.3.

Part (2) follows from (1) and the assumption the h either contains a letter $b_i^{\pm1}$ or a long subword of some relation.

For part (3), suppose that y overlaps both W_1 and W_3. Hence W_2 is a subword of y and hence of r. Therefore Condition 5.2 implies that $|W_2| \le \lambda_1|r|/1000$. Suppose that (2) fails and that the overlaps of y with each of W_1, W_3 have lengths $> \lambda_1|r|/10$. Thus $y = y_1W_2y_3$ where y_1 is a terminal segment of W_1, where y_3 is an initial segment of W_3 and where $|y_1|, |y_3| > \lambda_1|r|/10$. Recall that $W = W_1W_2W_3$ is the freely reduced form of hwh^{-1}. Since $|W_2| \le \lambda_1|r|/1000$, the definitions of W_1, W_2, W_3 then imply that there is a subword y_0 of y_1 of length $|y_0| \ge \lambda_1|r|/30$ such that y_0^{-1} is also a subword of y_3. Hence r contains two disjoint occurrences of y_0 and y_0^{-1}, where $|y_0| \ge \lambda_1|r|/30$. This contradicts the genericity assumptions in Condition 5.2. □

Lemma 5.9. *Any word that can be read by a closed loop in $\bar{\Delta} = Core\ \Delta$ and which represents an element of G that is conjugate to b_i cyclically reduces to b_i for $i = 1, 2$.*

Proof. Let w such that the conclusion of Lemma 5.6 holds for $h = \bar{z}$ and the word w. Possibly after replacing $\bar{z}$ with another word preserving the hypothesis on $\bar{z}$ for case A (see Lemma 5.7) we can assume the conclusions of Lemma 5.6, Lemma 5.7 and Lemma 5.8 hold for $\bar{z}$ and w.

Suppose now that p is a word that can be read by a closed loop in $\bar{\Delta}$ and which represents an element conjugate to b_i in G. If p cyclically reduces to b_i then there is nothing to show. Thus we can assume that it does not. We denote the word obtained from p by cyclic reduction by $\tilde{p}$. It now follows from Lemma 2.7 (3) that $\tilde{p}\tilde{p}$ contains a subword that is $\frac{9999}{10000}$ of some relation.

We will show that this is not possible, we distinguish two cases, namely that $\bar{z}$ contains a subword of length $\frac{1}{100}N$ that is also a subword of a relation or that it does not contains such a subword and contains a letter of type $b_i^{\pm 1}$.

Case 1: Suppose first that $\bar{z}$ contains a subword of length $\frac{1}{100}N$ that is also a subword of a relation. Recall that w is chosen as in Lemma 5.6. In the following we will denote the word obtained from $(\bar{z}w)^{n-1}\bar{z}$ by free reduction by z_n.

We can clearly assume that p is of the form

$$p = z_{j_1}^{\varepsilon_1} w_1 \cdot \ldots \cdot w_{k-1} z_{j_k}^{\varepsilon_k} w_k$$

where $\varepsilon_i = \pm 1$ for $1 \le i \le k$ and the w_i are reduced words read off by paths in Ψ such that $w_i \neq w^{\varepsilon_i}$ if $\varepsilon_i = \varepsilon_{i+1}$, otherwise we could replace the subword $z_{j_i}^{\varepsilon_i} w_i z_{j_{i+1}}^{\varepsilon_{i+1}} = z_{j_i}^{\varepsilon_i} w^{\varepsilon_i} z_{j_{i+1}}^{\varepsilon_i}$ by $z_{j_i + j_{i+1}}$. We can further assume that this product is freely reduced, i.e. that $w_i \neq 1$ if $\varepsilon_i = -\varepsilon_{i+1}$, and that it is cyclically reduced, i.e. that either $\varepsilon_1 = \varepsilon_k$ or that $\varepsilon_1 = -\varepsilon_k$ and $w_k \neq 1$. Note that the element represented by p is not necessarily conjugate to b_1 by a word in the a_i.

By Lemma 5.7 an element $z_n = (\bar{z}w)^{n-1}\bar{z}$ does not contain a subword that represents more than $\frac{99}{100}$ of a relation. By Lemma 5.7 (5) any of the words z_n contain two subwords of relations of length at least $|r|/500$ inherited from the first and last occurrence of $\bar{z}$ by the free cancellation of the product $(\bar{z}w)^{n-1}\bar{z}$. In particular we can choose words U and V that are the leftmost, respectively rightmost, subwords of z_n that represent $\frac{1}{500}$ of a relation, clearly these words are the same for all z_n with $n \ge 1$.

It now follows from Lemma 5.6 and Lemma 5.8 that none of the occurrences of $U^{\pm 1}$ and $V^{\pm 1}$ cancels completely in the free and cyclic reduction of p^2 to $\tilde{p}^2$. It also follows that the remnants of the $U^{\pm 1}$ and $V^{\pm 1}$ of two adjacent $z_{j_i}^{\varepsilon_i}$ do not join up to subwords of a relation. As the w_i do not contain subwords that are of length greater than $\frac{1}{10000}$ of a relation and are subwords of of a relation this implies that any subword of relation that occurs in $\tilde{p}\tilde{p}$ is of length at most $\frac{99}{100} + 2\frac{1}{500} + 2\frac{1}{10000}$ of a relation. As this number is smaller than $\frac{9999}{10000}$ it follows that p cannot have represented an element conjugate to b_1.

Case 2: Assume now that $\bar{z}$ contains a letter $b_i^{\pm 1}$and does not contain a subword of length $\frac{1}{100}N$ that is also a subword of a relation.

Write $\bar{z}$ as a product $ub_i^{\varepsilon}\tilde{w}b_j^{\eta}v^{-1}$ such that u and v are words in the a_i, this decomposition always exists unless $\bar{z}$ is conjugate to an element of type $b_i^{\pm 1}$ in which case the following arguments still apply with minor modifications. Now there is at most one word $\hat{w}$ in the a_i such that in the free cancellation of $\bar{z}\hat{w}\bar{z}$ the subword $v^{-1}\hat{w}u$ reduces to the trivial word.

Now in the word Z_n obtained by free reduction from product of type $(\bar{z}\hat{w})^n\bar{z}$ not more than $\frac{1}{10}$ of a relation can occur as such a word would need to be periodic because of the assumption that $\bar{z}$ does not contain more than $\frac{1}{100}$ of a relation.

It is further clear that Z_n with prefix ub_i^{ε} and suffix b_j^{η}, in particular it contains a letter of type $b_k^{\pm 1}$. We can clearly assume that p is of the form

$$p = Z_{j_1}^{\varepsilon_1} w_1 \cdot \ldots \cdot w_{k-1} Z_{j_k}^{\varepsilon_k} w_k$$

where $\varepsilon_i = \pm 1$ for $1 \leq i \leq k$ and the w_i are reduced words read off by paths in Ψ such that $w_i \neq \hat{w}^{\varepsilon_i}$ if $\varepsilon_i = \varepsilon_{i+1}$, otherwise we could replace the subword $Z_{j_i}^{\varepsilon_i} w_i Z_{j_{i+1}}^{\varepsilon_{i+1}} = Z_{j_i}^{\varepsilon_i} \hat{w}^{\varepsilon_i} Z_{j_{i+1}}^{\varepsilon_i}$ by $Z_{j_i + j_{i+1}}$.

Now by the assumptions made it follows at least one letter of type $b_k^{\pm 1}$ of each $Z_{j_i}^{\varepsilon_i}$ and one letter $a_i^{\pm 1}$ of each w_i does not cancel in the free reduction of p. Thus, as each of this factors does not contain more than $\frac{1}{10}$ of a relation the freely reduced word obtained from p cannot contain $\frac{1}{2}$ of a relation. Thus Case 2 is ruled out. □

We can now conclude case A. It follows from Lemma 5.9 that if a word is readable in Δ that represents an element conjugate to b_i then this word freely reduces to b_i. The proof of Case 2 of Lemma 5.9 implies that of each factor $\bar{z}^{\pm 1}$ at least one letter $b_i^{\pm 1}$ does not cancel. This implies that $w_1 \bar{z}^{\pm 1} \bar{w}_1$ must cyclically reduce to b_1 where w_i and $\bar{w}_1$ are words in the $a_i^{\pm 1}$ which in turn implies that $\bar{z}$ contains a single letter of type $b_i^{\pm 1}$, namely $b_1^{\pm 1}$. The same argument applied to b_2 implies that $\bar{z}$ contains $b_2^{\pm 1}$ which yields a contradiction. This concludes the proof of case A.

Case B: $\bar{z}$ is a word in the a_i containing no subword of some defining relation r of length $\geq |r|/100$. There are 2 subcases, namely $z = \bar{z}$ and $z \neq \bar{z}$.

Subcase B1: Suppose first that $z \neq \bar{z}$. The aim in this case is to get a contradiction to the minimality assumption on the sum of the lengths of w_1 and w_2. The underlying idea is simple: we just replace in Δ the edge (respectively lollipop) with label z by an edge (respectively lollipop) with label $\bar{z}$ and then do the corresponding replacement in the loops of S_T upstairs. Since in this case the word $\bar{z}$ is considerably shorter than z, this yields a contradiction with the minimality assumption on the pair (w_1, w_2). However taking into account the basepoint makes the proof more technical.

The following lemma describes the relationship between z and $\bar{z}$, note that all identities are identities of words and not only of group elements.

Lemma 5.10. *Let z and and $\bar{z}$ be as above. Then the following hold:*

(1) *There exist words $\alpha_0, \ldots, \alpha_t$, $\beta_1, \ldots, \beta_t$ (nontrivial all except possibly α_0 and α_t) and words $\bar{\beta}_1, \ldots, \bar{\beta}_t$ such that the following hold:*
  (a) $z = \alpha_0 \beta_1 \alpha_1 \ldots \beta_t \alpha_t$.
  (b) $\bar{z} = \alpha_0 \bar{\beta}_1 \alpha_1 \ldots \bar{\beta}_t \alpha_t$.
  (c) $|\bar{\beta}_i| \leq \frac{1}{100}|\beta_i|$ *and* $|\beta_i| \geq 1000N$ *for all i, where N is the length of the defining relations of G.*

(2) *If u and $\bar{u}$ are maximal words such that $z = uwu^{-1}$ and $\bar{z} = \bar{u}\bar{w}\bar{u}^{-1}$ for some words w and $\bar{w}$ then $|\bar{w}| \leq |w|$.*

(3) *If a_i^{η} is a letter of u such that neither a_i^{η} nor the the corresponding letter $a_i^{-\eta}$ of u^{-1} is a letter of a subword β_i then in $\bar{z}$ a_i^{η} is a letter of $\bar{u}$ and $a_i^{-\eta}$ is a letter of $\bar{u}^{-1}$.*

Proof. Recall that the presentation for G satisfies the $C'(\lambda_0)$ condition with $\lambda_0 = 1/1000$. By a λ_0*-reduction* on a word in the generators of G we mean replacing a subword of that word, that is also a subword of one of the defining relations r of length $\geq (1 - 3\lambda_0)|r|$ by the complementary portion of r.

First note that if $\tilde{z}$ is any other freely reduced and λ_0-reduced representative of z then in fact $\tilde{z} = \bar{z}$ as words. Indeed, if $\tilde{z} \neq \bar{z}$ then the equality diagram for $\tilde{z} =_G \bar{z}$ has the form provided by Proposition 2.4 and, also by Proposition 2.4 the word $\overline{z}$ contains a subword of a defining relation of length $\geq 1/100$ of that relation. This is impossible by the assumption on $\bar{z}$ in Case B.

Thus any maximal sequence of free and λ_0-reductions applied to z must terminate with the word $\bar{z}$ and $\bar{z}$ is the unique freely reduced and λ_0-reduced word representing the same element as z in G.

Let $\bar{z} =_G \bar{u}\bar{w}\bar{u}^{-1}$ be the representation of the element $z \in G$ provided by Lemma 2.7. Recall that by Lemma 2.7 the element of G represented by the word $\bar{w}$ is the shortest (in G) representative of the conjugacy class of $\bar{z}$, the word $\bar{u}\bar{w}\bar{u}^{-1}$ is freely reduced and Dehn reduced (in particular, λ_0-reduced). Since both $\bar{z}$ and $\bar{u}\bar{w}\bar{u}^{-1}$ are freely reduced and λ_0-reduced, it follows from the above remark that $\bar{z} = \bar{u}\bar{w}\bar{u}^{-1}$ as words. Moreover, we can obtain $\bar{z}$ from z by applying λ_0-reductions and free cancellations. Note that since the element represented by the geodesic word $\bar{w}$ is shortest in the G-conjugacy class of z, part (2) of Lemma 5.10 is now established.

Every λ_0-reduction replaces a subword Q by a word of length $\leq |Q|/100$. A free reduction replaces a subword of positive even length by an empty word. Recall that $\bar{z}$ is the unique freely reduced and λ_0-reduced representative of z. Moreover, since $z \neq \bar{z}$, at least one λ_0-reduction is necessary to get to $\bar{z}$ from z. Hence, by an inductive argument, we can show that z and $\bar{z}$ can be represented as $z = \alpha_0\beta_1\alpha_1\ldots\beta_t\alpha_t$, $\bar{z} = \alpha_0\bar{\beta}_1\alpha_1\ldots\bar{\beta}_t\alpha_t$, where $|\bar{\beta}_i| \leq \frac{1}{100}|\beta_i|$, where $\beta_i =_G \bar{\beta}_i$ and where the process of λ_0-reducing z to $\bar{z}$ consists in λ_0-reducing each of β_i to $\bar{\beta}_i$ separately. Note that this implies part (3) of Lemma 5.10.

The fact that $|\beta_i| \geq 1000N$ requires a more involved argument whose precise details we omit. We give a sketch of the argument here:

The process of λ_0-reducing z to $\bar{z}$ can be conducted in stages. At the first stage we find a maximal collection of non-overlapping subwords y_q, $q = 1, 2, ..., M$, of z such that each of them is also a subword of some defining relator $b_j = v_j(\overline{a})$ (or a cyclic permutation of this relator or its inverse) of length comprising $\geq (1 - 3\lambda_0)$ fraction of that relator. Since z is a word in a_i, each y_q is also a word in a_i. We then replace each of these subwords y_q by the missing portions of the corresponding defining relators, which introduces M letter from $\{b_1, \ldots, b_m\}^{\pm 1}$ into the word and denote the result by z_0'. We then freely reduce z_0' to get a word z_1 which is the result of the first stage. Note that no $b_j^{\pm}1$ from z_0' get cancelled in freely reducing z_0' to z_1 since otherwise one could show that the word z was not freely reduced. Also, we have $t \leq M$ and each of y_q is a subword of some β_i from the conclusion of the lemma. Recall that the λ_0-reduced form $\bar{z}$ of z does not involve any $b_j^{\pm}1$. In the word z_1 obtained after the first stage the only applicable λ_0-reductions are those that involve large portions of the defining relators $a_j = u_j(\overline{b})$ (technically this in not entirely correct, since in z we may have had a portion of the relation $b_j = v_j(\overline{a})$ of length slightly less than $(1 - 3\lambda_0)$ fraction of that relation which, in z_1 acquired a small extra portion and became $(1 - 3\lambda_0)$ fraction of that relation; however, we can disregard this technicality in our sketch). At the second stage, we apply a similar process to z_1 by finding a maximal collection of disjoint subwords x_s of z_1 such that each x_s is also a subword of some relation $a_j = b_j(\overline{b})$ comprising $\geq (1 - 3\lambda_0)$ fraction of that relation. Thus each x_s is, except for a single letter, a word in $\{b_1, \ldots, b_m\}^{\pm}1$ of length $\geq 0.99N$. When tracked back to z, this causes, for each x_s, a conglomeration of $\geq 0.99N$ of the original y_qs into a single β_j of length $\geq (1 - 3\lambda_0)N \cdot 0.99N \geq 1000N$ for sufficiently large N. Some of the b_j introduced in z_1 at stage one may not be contained in any of the x_s at stage two. However, eventually all of the b_j that were introduced in z_1 at stage one have to disappear in the λ_0-reduction process (since the word z involves only $\{a_1, \ldots, a_k\}^{\pm 1}$) and once can show that the condition $|\beta_i| \geq 1000N$ holds for every $i = 1, \ldots, t$.

□

Note that Lemma 5.10 (1c) implies that $\bar{z}$ is significantly shorter than z, namely by at least $\frac{99}{100} \sum |\beta_i| \geq 990 \cdot k \cdot |v_i|$. In the following we denote $\sum |\beta_i|$ by L.

Note further that any vertex except the initial and terminal vertex of the edge, resp. lollipop, correspond to a subword of length two of z which is read of the edge, resp. lollipop when reading z. In the case of the lollipop and the vertex lying on the stick there are two such words. We say that a vertex on the edge, resp. lollipop, is inessential if none of these at most two subwords is a subword of one the β_i. Note that in particular the initial and the terminal vertices are inessential.

We perform some changes to S_T and Δ, these changes do not alter the following facts

(1) S_T folds onto Δ and Δ folds onto R_A. Thus S_T folds onto R_A.
(2) The first two loops of S_T are labeled by words in the $a_i^{\pm 1}$ that represent b_1 and b_2.

The conclusion then follows from the fact that after the final modification the length of the first two loops of S_T has decreased.

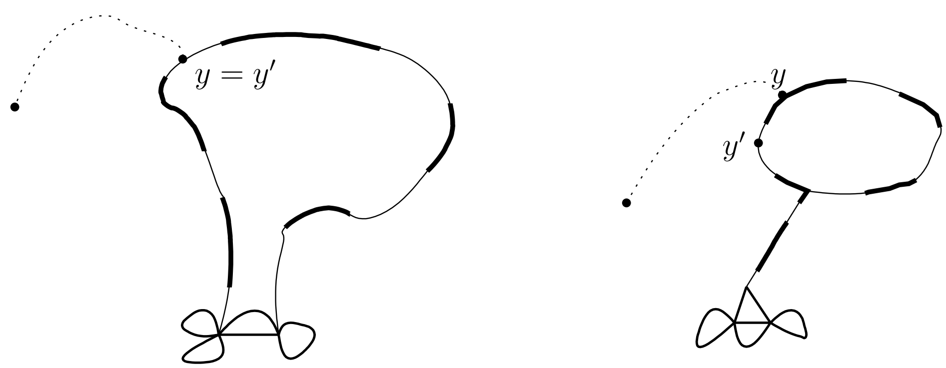

FIGURE 11. Fat parts correspond to paths read by the β_i

Let y be the vertex of $\bar{\Delta} = Core(\Delta)$ such that $d(v_0, y) = d(v_0, \bar{\Delta})$. Thus $y = v_0$ if $v_0 \in \bar{\Delta}$ and y is the attaching point of the segment joining $\bar{\Delta}$ and v_0 otherwise. We now choose a new vertex $y' \in \Delta$ such that the following hold (see Figure 11):

(1) y' is inessential.
(2) $d(y, y') \leq 2L$
(3) If y is in distance less than L from the stick of the lollipop then y' is on the stick of the lollipop (if we are in the lollipop situation).

The existence of such a vertex y' follows immediately from the definition of L. Note from now on we only provide figures for the lollipop situation as it is the more subtle one.

Let $[y', y]$ be an arbitrary geodesic in $\bar{\Delta}$. We now unfold along $[y', y]$ (starting at y). In the end v_0 is attached to $\bar{\Delta}$ by a segment of length $d(v_0, y) + d(y, y')$ which intersects $\bar{\Delta}$ only in y'. We denote the just constructed unfolded
delta by Δ_+. Note that we have not changed Ψ or $Core(\Delta)$ as $Core(\Delta_+) = Core(\Delta) = \bar{\Delta}$.

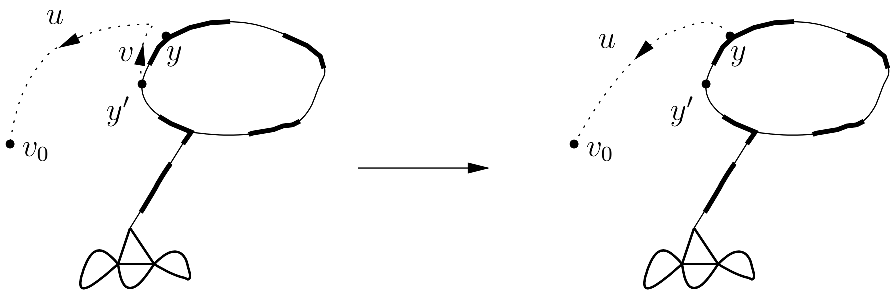

Figure 12. Unfolding Δ to Δ_+

Now S_T does not necessarily map onto Δ_+ but it does if we introduce at the beginning and at the end of each loop the trivial word that corresponds to the unfolding when going from Δ to Δ_+. More precisely:

If u is the word corresponding to $[y, v_0]$ in Δ and v the word corresponding to $[y', y]$ in Δ then the i-th loop in S_T has a label of type $u^{-1}d_iu$ for some word d_i as it gets mapped onto a path in Δ that starts and ends at v_0, after it starts (before it ends) at v_0 it has no other option than going to y first (or must come from y).

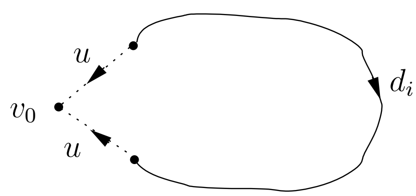

Figure 13. A loop of S_T

We replace the i-th loop of S_T with a loop with label $u^{-1}v^{-1}vd_iv^{-1}vu$. These new loops represent the same group element and now map to Δ_+ as the word $u^{-1}v^{-1}$ maps to the edge path $[v_0, y']$ in Δ_+ and the word $u^{-1}v^{-1}v$ onto $[v_0, y'] \cup [y', y]$ in Δ_+. We further modify the newly constructed graph such that we perform the free reduction on the subword vd_iv^{-1} that corresponds to backtracking on the edge, respectively lollipop. Thus we are left with labels of type $u^{-1}v^{-1}p_i\bar{d}_iq_i^{-1}vu$ where p_i is the prefix of v and q_i^{-1} is the suffix of v^{-1} of vd_iv^{-1} that have not been cancelled. We call the new graph S_T^+.

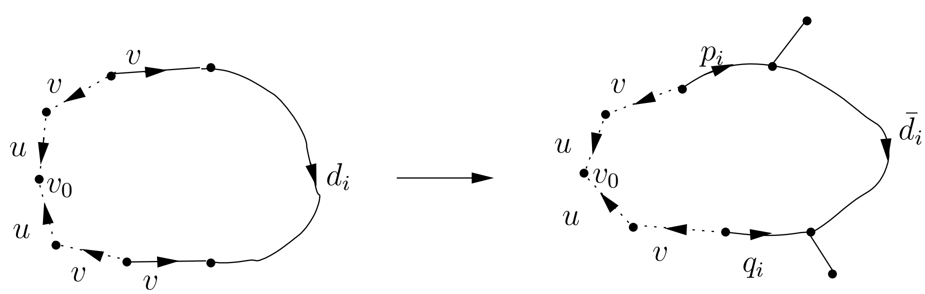

Figure 14. Constructing $S_T^+ +$ (only one loop drawn)

Note that the i-th loop of S_T^+ still represents the element b_i in G for $i = 1, 2$. By construction and as $d(y, y') = |v| \leq 2L$ we have increased the length of each loop by at most $8L$.

Now the subwords $p_i\bar{d}_i q_i^{-1}$ correspond to closed paths in $\bar{\Delta}$ based at y' and are therefore of type

$$\theta_1\gamma_0 uw^{r_1}u^{-1}\gamma_1\cdots uw^{r_l}u^{-1}\gamma_l\theta_2^{-1}$$

where

(1) θ_1 and θ_2 are words corresponding to reduced paths from y' to Ψ. In the case of a lollipop these paths might go around the lollipop a number of times.
(2) The words γ_i are read in Ψ and are non-trivial for $1 \le i \le l-1$.
(3) $r_i \in \mathbb{Z} - 0$ with $r_i \in \{-1, 1\}$ if z is the label of an edge rather than a lollipop.

We claim that $l > 20$. Indeed after conjugating the above word by θ_2^{-1} we get

$$\theta_2^{-1}\theta_1\gamma_0 uw^{r_1}u^{-1}\gamma_1\cdots uw^{r_l}u^{-1}\gamma_l = uw^{r_0}u^{-1}\gamma_0 uw^{r_1}u^{-1}\gamma_1\cdots uw^{r_l}u^{-1}\gamma_l$$

with $r_0 \in \mathbb{Z}$. Replacing $z^{r_i} = uw^{r_i}u^{-1}$ by $\bar{z}^{r_i} = \bar{u}\bar{w}^{r_i}\bar{u}^{-1}$ and free reduction produces a word in the $a_i^{\pm 1}$ that still represents a conjugate of y_1 (or y_2) and must therefore contain half a relation after free reduction. As this word consists of the remnants of the γ_i and the $\bar{u}\bar{w}^{r_i}\bar{u}^{-1}$, none of which contains more than $\frac{1}{50}$th of a relation (this follows from the assumption on $\bar{z}$ for the $\bar{u}\bar{w}^{r_i}\bar{u}^{-1}$ and by non-genericity for words readable in Ψ for the γ_i) it follows that $l > 20$.

We now construct a new graph Δ_* from Δ_+. We first replace $\bar{\Delta} = Core(\Delta_+)$ by replacing the segment, respectively lollipop, with label z as follows. In the case of a segment we replace it with a segment with label $\bar{z}$ and the same attaching points and in the case of a lollipop we replace the lollipop with a lollipop whose stick is labeled by $\bar{u}$ and whose loop based at the stick is labeled by $\bar{w}$. In both cases walking along the segment or around the lollipop reads $\bar{z} = \bar{u}\bar{w}\bar{u}^{-1}$. Let $\bar{\Delta}_*$ be the modified $\bar{\Delta}$.

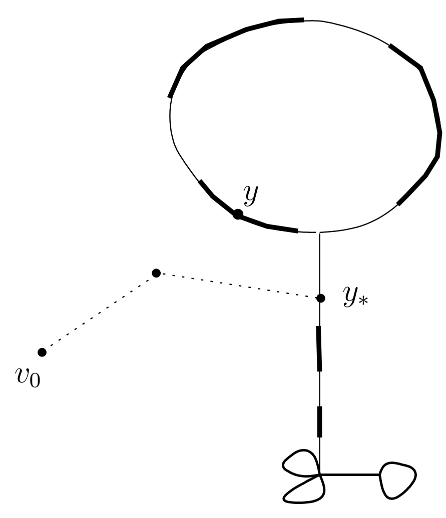

FIGURE 15. In Δ_* the point $y' = y_*$ could move to stick

Note that as the vertex y' was chosen to be inessential it corresponds to a unique vertex of $\bar{\Delta}_*$ which we call y_*. We now define Δ_* to be the graph obtained from $\bar{\Delta}_*$ by attaching the segment $[v_0, y']$ (from Δ) to $\bar{\Delta}_*$ by identifying y' and y_*. Note that in the lollipop case if y' was on the stick of the original lollipop then by Lemma 5.10 (3) y_* is on the stick of the new one but the converse is not necessarily true. Clearly Γ_* still folds onto R_A as it still contains Ψ.

Next we need to adjust S_T^+ such that it maps onto Δ_*. Recall that the labels of the loops are of type $u^{-1}v^{-1}p_i\bar{d}_iq_i^{-1}vu$ where $p_i\bar{d}_iq_i^{-1}$ is of the form

$$\theta_1\gamma_0 uw^{r_1}u^{-1}\gamma_1\cdots uw^{r_l}u^{-1}\gamma_l\theta_2^{-1}$$

with θ_1 and θ_2 words read by paths that start at y' and end in Ψ. We will not change the initial subword $u^{-1}v^{-1}$ or the terminal subword vu, nor will we change the subwords γ_i. We first replace the subwords $uw^{r_i}u^{-1}$ by $\bar{u}\bar{w}^{r_i}\bar{u}^{-1}$. This reduces the length of the word by at least $20\cdot\frac{99}{100}L$.

We will now replace θ_1 and θ_2 by words $\bar{\theta}_1$ and $\bar{\theta}_2$ that correspond to paths in Δ_* whose initial vertex is y_*, who have the same terminal vertices in Ψ as θ_1 and θ_2 and who are at most as long as θ_1 and θ_2, respectively. This clearly implies the assertion as now the sum of the lengths of the first two loops has decreased as we first increased their lengths by at most $8L$ and then decreased their lengths by at least $20\cdot\frac{99}{100}L$. This yields a contradiction to the minimality assumption as it obvious from the construction that the first two loops still represent b_1 and b_2 in G.

We give the construction for θ_1, the case for θ_2 is analogous. As θ_1 is freely reduced it is a suffix of the word uw^ru^{-1} for some $r\in\mathbb{Z}$ with $r\in\{-1,1\}$ in the edge case. We choose r such that $|r|$ is minimal, w.l.o.g. we can assume that $r>0$. If $|r|=1$ then θ_1 is a suffix of z and the conclusion follows after replacing all subwords β_i of θ_1 by β_i', our choice of y' guarantees that that each β_i either is a subword of θ_1 or does not overlap with θ_1. This case deals in particular with the non-lollipop case.

If $\theta_1=tw^ru^{-1}$, i.e. if y' lies on the stick of the lollipop, then y_* also lies on the stick of the lollipop corresponding to $\bar{u}\bar{w}\bar{u}^{-1}$ by Lemma 5.10 (3) and we replace θ_1 by $\bar{t}\bar{w}^r\bar{u}^{-1}$ where $\bar{t}\bar{w}\bar{u}^{-1}$ is the word obtained from twu^{-1} as in the case $r=1$. The claim now follows as $|\bar{t}\bar{w}\bar{u}^{-1}|\le|twu^{-1}|$ and as $|\bar{w}|\le|w|$ by Lemma 5.10 (2).

The remaining case is that $\theta_1=sw^{r-1}u^{-1}$ where s is a suffix of w. Thus $\theta_1=\hat{w}^{r-1}su^{-1}$ where $\hat{w}$ is the word corresponding to the loop of the lollipop based at y'. Let now $\hat{\theta}$ be the word obtaind from su^{-1} as in the case $r=1$ and $\hat{w}'$ be the path corresponding the closed path at y_* in the new lollipop that goes once around the lollipop. Note it follows from the fact that y_* is at least in distance L from the stick of the lollipop that $|\hat{w}'|\le|\hat{w}|$. We put $\bar{\theta}_1$ to be the word obtained from $(\hat{w}')^{r-1}\hat{\theta}$ by free reduction and the assertion is immediate.

Subcase B2: $z=\bar{z}$. If the length of z is less than $100k^2$ (recall that k is the number of the a_i) then $\bar{\Delta}=Core(\Delta)$ is one of finitely many graphs $\Gamma_1,\ldots,\Gamma_k$ and for each Γ_i an exponentially generic set of words cannot be read in Γ_i by Lemma 5.5. As the intersection of finitely many exponentially generic sets is exponentially generic it follows that an exponentially generic set of words cannot be read in any Γ_i and therefore cannot be read in $\bar{\Delta}$. Thus the assertion follows as in the proof of Theorem 4.4.

If the length of z is at least $100k^2$ then most subwords of length greater than $100k^2$ of a generic word that can be read in Γ must intersect a subword that is in Γ read by walking the edge or lollipop corresponding to z as they can't be read in ψ (they contain all reduced words of type $a_i^{\pm1}a_j^{\pm1}$). This means that at least half of these words consists of copies of z or powers of z making them non-generic, a contradiction.

This concludes the proof of Theorem 1.1.

References

[D] T. Delzant *Sous-groupes a deux generateurs des groupes hyperboliques.* Ghys, ?. (ed.) et al., Group theory from a geometrical viewpoint. Proceedings of a workshop, held at the International Centre for Theoretical Physics in Trieste, Italy, 26 March to 6 April 1990. Singapore: World Scientific., 1991, 177–189.

[E1] M.J. Evans *Presentations of Groups involving more Generators than are neceessary*, Proc. LMS **67**, 1993, 106-126.

[E2] M.J. Evans *Presentations of Groups involving more Generators than are neceessary II*, Cont. Math. **421**, 2006, 101–112.

[E3] M.J. Evans *Nielsen equivalence classes and stability graphs of finitely generated groups*, Ischia Group theory 2006, World Scientific, 2007.

[G] I. Grushko *On the bases of a free product of groups* Mat. Sbornik **8**, 1940, 169-182.

[KS] I. Kapovich and P. Schupp, *Genericity, the Arzhantseva-Ol'shanskii method and the Isomorphism Problem for One-Relator Groups*, Math. Ann. **331**, 2005, 1–19.

[KW] I. Kapovich and R. Weidmann, *Freely indecomposable groups acting on hyperbolic spaces*, IJAC 14 (2004), 115-171.

[LGM] C. Leedham-Green and S. Murray, *Variants of product replacement.* Computational and statistical group theory (Las Vegas, NV/Hoboken, NJ, 2001), 97–104, Contemp. Math., 298, Amer. Math. Soc., Providence, RI, 2002

[LP] A. Lubotzky and I. Pak, *The product replacement algorithm and Kazhdan's property (T).* J. Amer. Math. Soc. **14** (2001), no. 2, 347–363

[LM1] M. Lustig and Y. Moriah, *Nielsen equivalence and simple homotopy type*, Proc. LMS **62**, 1991, 537–562.

[LM2] M. Lustig and Y. Moriah, *Nielsen equivalence in Fuchsian groups and Seifert fibered spaces*, Topology **30**, 1991, 191–204.

[LM3] M. Lustig and Y. Moriah, *Generating systems of groups and Reidemeister-Whitehead torsion*, J. Algebra **157**, 1993, 170-198.

[LS] R. Lyndon and P. Schupp, *Cominatorial Group Theory*, Springer-Verlag, 1977

[No] G.A. Noskov, *Primitive elements in a free group*, Math. Zametki **30** (4), 1981, 497-500.

[Pa] I. Pak, *What do we know about the product replacement algorithm?*, Groups and computation, III (Columbus, OH, 1999), 301–347, Ohio State Univ. Math. Res. Inst. Publ., 8, de Gruyter, Berlin, 2001

[PRZ] N. Peczynski, G. Rosenberger and H. Zieschang *Über Erzeugende ebener diskontinuierlicher Gruppen* Invent. Math. **29**, 1975, 161–180.

[R] E. Rips, *Subgroups of small cancellation groups*, Bull. Lond. Math. Soc. **14**, 45-47, 1982.

[Sta] J.R. Stallings *Topology of Finite Graphs* Invent. Math. **71**, 1983, 551-565.

[Str] Ralph Strebel, *Small cancellation groups*, in: "Sur les groupes hyperboliques d'aprs Mikhael Gromov." Papers from the Swiss Seminar on Hyperbolic Groups held in Bern, 1988. Edited by . Ghys and P. de la Harpe. Appendix, 227–273. Progress in Mathematics, 83. Birkhuser Boston, 1990.

[W] R. Weidmann, *Generating tuples of free products*, Bull. LMS, **39**, 2007, 393–403.

[Z] H. Zieschang, *Über der Nielsensche Kürzungsmethode in freien Produkten mit Amalgam*, Inv. Math. **10**, 4–37, 1970.