Groups finitely presented in Burnside varieties

A.Yu. Olshanskii *

Contents

*The author was supported in part by the NSF grants DMS-1500180 and DMS-1901976.

Abstract

For all sufficiently large odd integers n, the following version of Higman's embedding theorem is proved in the variety $\mathcal{B}_n$ of all groups satisfying the identity $x^n = 1$. A finitely generated group G from $\mathcal{B}_n$ has a presentation $G = \langle A \mid R \rangle$ with a finite set of generators A and a recursively enumerable set R of defining relations if and only if it is a subgroup of a group H finitely presented in the variety $\mathcal{B}_n$. It follows that there is a 'universal' 2-generated finitely presented in $\mathcal{B}_n$ group containing isomorphic copies of all finitely presented in $\mathcal{B}_n$ groups as subgroups.

Key words: generators and relations in groups, finitely and recursively presented groups, Burnside variety, van Kampen diagram, Turing machine, S-machine

AMS Mathematical Subject Classification: 20F05, 20F06, 20F50, 20F65, 20E06, 20E10, 03D10, 20F10, 03D25

1 Introduction

The celebrated theorem of G. Higman [7] asserts that a finitely generated group is recursively presented if and only if it is a subgroup of a finitely presented group. This theorem detects a deep connection between the logic concept of recursiveness and properties of finitely presented groups.

The Higman embedding $G \hookrightarrow H$ of a recursively presented group G makes the finitely presented group H quite large, saturating it with free subgroups. Therefore there were no nontrivial analogies of Higman's theorem in proper subvarieties of the variety of all groups. (The statement is trivial if every finitely generated group of a variety $\mathcal{V}$ is finitely presented in $\mathcal{V}$.) O. Kharlampovich finished the paper [10] on the following question:

Is every finitely generated group with the identity $x^n = 1$ and recursively enumerable set of defining relations embeddable into a group given by this identity and a finite set of relations (n is a large number) ? In other words, is the analogy of Higman's embedding theorem true in the variety $\mathcal{B}_n$?

Even earlier the same problem was formulated in [11] (problem 12.32) by S.V. Ivanov :

Prove an analogue of Higmans theorem for the Burnside variety $\mathcal{B}_n$ of groups of odd exponent $n >> 1$, that is, prove that every recursively presented group of exponent n can be embedded in a finitely presented (in $\mathcal{B}_n$) group of exponent n.

Recall that the *Burnside variety* of exponent $n \geq 1$, denoted by $\mathcal{B}_n$, is the class of all groups satisfying the identical relation $x^n = 1$. So $\mathcal{B}_n$ is the class of groups having finite exponent n. Every m-generated group of this variety is isomorphic to a factor group of the m-generated free group $B(m,n)$ in $\mathcal{B}_n$, which in turn is isomorphic to the factor group of a free group $F(m)$ of rank m modulo the normal subgroup F_m^n generated by all powers g^n, where $g \in F_m$. (If two words from $F(m)$ have equal images in the *free Burnside group* $B(m,n)$, we say that these words are equal *modulo the Burnside relations.*)

A group G is called *finitely presented* in the variety $\mathcal{B}_n$ if $G \cong B(m,n)/N$ for some $m \geq 1$, where N is a normal closure of a finite subset $R \subset B(m,n)$, in other words, G satisfies the identity $x^n = 1$ and have a presentation $\langle a_1, \dots, a_m \mid R \cup S \rangle$ in the class of all groups, where R is a finite set of relators and every relator from S is an n-th power in the free group $F(a_1, \dots, a_m)$.

Theorem 1.1. *There is a constant C such that for every odd integer $n \geq C$, the following is true.*

A finitely generated group G satisfying the identity $x^n = 1$ has a presentation $G = \langle A \mid R \rangle$ with a finite set of generators A and a recursively enumerable set R of defining relations if and only if it is a subgroup of a group H finitely presented in the variety $\mathcal{B}_n$.

The embedding property of Theorem 1.1 holds for groups G with countable sets of generators as well. Furthermore, the group H can be chosen with two generators. More precisely, we have

Corollary 1.2. *There is a constant C such that for every odd integer $n \geq C$, the following is true.*

Let a group G from the variety $\mathcal{B}_n$ be given by a countable set of generators $x_1, x_2, \dots$ and a recursively enumerable set of defining relations R in these generators. Then G is a subgroup of a 2-generated group E finitely presented in $\mathcal{B}_n$.

Proof. The group G is isomorphic to the factor group $B(\infty, n)/N$, where $B(\infty, n)$ is the free Burnside group of exponent n with countable set of free generators $x_1, x_2, \dots$ and N is the normal closure of the set R in $B(\infty, n)$.

We will use a theorem of D. Sonkin (Theorem 2.2 [30]). According to this theorem the 2-generated free Burnside group $B(2,n)$ with free basis (a,b) has a subgroup K with the following properties.

The subgroup K is generated by a countable set of words $w_1, w_2, ...$ in the generators a, b such that this set is recursive, K is a free Burnside group of exponent n with free generators $w_1, w_2, \ldots$, and K has *Congruence Extension Property* (CEP). CEP means that every normal subgroup of K is the intersection of a normal subgroup of $B(2,n)$ with K. Equivalently, a normal closure in K of arbitrary subset $S \subset K$ is the intersection of the normal closure of S in $B(\infty, n)$ with K.

Now the group G can be presented as the factor group K/M, where M is the normal closure in K of a set S, where $S = \{r(w_1, w_2, \ldots) \mid r(x_1, x_2, \ldots) \in R\}$. So the set S is a recursively enumerable set of words in the generators a, b.

By CEP, we have $M = K \cap P$ for the normal closure P of S in $B(2,n)$. Hence we obtain that $G \cong K/M = K/K \cap P \cong KP/P$, where KP/P is a subgroup of the two generated group $B(2,n)/P$ and P is the normal closure in $B(2,n)$ of the recursively enumerable set of words S. By Theorem 1.1, the group $B(2,n)/P$ is, in turn, a subgroup of a finitely presented in $\mathcal{B}_n$ group H.

It remains to embed the group H in a 2-generated finitely presented in $\mathcal{B}_n$ group. If $H \cong B(m,n)/L$, then we repeat the above trick using that $B(m,n)$ is embeddable in $B(2,n)$ as a CEP-subgroup. Indeed, the group $B(m,n)$ is a retract of $B(\infty, n)$, and so a CEP-subgroup. Since the CEP property is transitive, $B(m,n)$ is isomorphic to a CEP subgroup K' of $B(2,n)$.

The normal subgroup L becomes a normal closure of a finite set R' in K'. Therefore L is the intersection of K' with the normal closure Q of R' in $B(2,n)$. Hence the group H is embeddable in the 2-generated finitely presented in $\mathcal{B}_n$ group $E = B(2,n)/Q$, as required. □

Examples. **1.** The direct product of all finite groups of exponent dividing n is embeddable in a 2-generated group finitely presented in the variety $\mathcal{B}_n$.

2. There are infinite 2-generated, simple groups of large odd exponent n, where all maximal subgroups have order n. (In particular, every proper subgroup has order n id n is prime.) Moreover, there exist such "monsters" with decidable word problem by Theorem 28.4[19]. By Corollary 1.2, they are embeddable in 2-generated groups finitely presented in the variety $\mathcal{B}_n$.

The next statement is similar to another Higman's theorem [7].

Corollary 1.3. *For every sufficiently large odd integer n, there exists a 2-generated finitely presented in $\mathcal{B}_n$ group E containing, as subgroups with pairwise trivial intersections, isomorphic copies of all recursively presented groups $\{G_i\}_{i=1}^{\infty}$ of exponent n.*

Proof. The set of finite presentations $G_i = \langle A_i \mid R_i \rangle$ in $\mathcal{B}_n$ is recursively enumerable. One can assume that the finite sets A_i-s are disjoint and obtain a recursive presentation $\langle X \mid \cup_i R_i \rangle$, where $X = \cup_i A_i = \{x_1, x_2, \ldots\}$. By Dyck's lemma the group G defined by this presentation in $\mathcal{B}_n$ admits a retraction ρ_i onto every G_i, where ρ_i is identical on A_i and maps all other generators from X to 1. Therefore G contains an isomorphic copy of every group G_i. So does the 2-generated finitely presented group $E \geq G$ given by Corollary 1.2. By Theorem 1.1, every recursively presented group from $\mathcal{B}_n$ is also embeddable in E. □

The next corollary shows that there are finitely presented in $\mathcal{B}_n$ groups with undecidable word problem. Recall that the first examples of a finitely presented groups with undecidable word problem were constructed by W.W. Boone [4] and P.S. Novikov [15]. M.Sapir [28] obtained the first example of a finitely presented in a variety group of finite exponent with undecidable word problem. Then O. Kharlampovich [10] found finitely presented in varieties $\mathcal{B}_n$ groups with undecidable word problem, provided $n = pr$, p was an odd prime, and either r had an odd divisor ≥ 665 or $r \geq 2^{48}$. Sapir and Kharlampovich referred to the property that non-cyclic free Burnside groups of large exponent n are infinite ([17, 1, 9]), and both of them simulated the work of Minskiy machine in a normal subgroup of smaller exponent. Therefore the exponent n was a composite number in their examples. We obtain new examples now, in particular, for all large prime exponents:

Corollary 1.4. *For every sufficiently large odd integer $n \geq C$, there exists a 2-generated finitely presented in $\mathcal{B}_n$ group with undecidable word problem.*

Proof. We will use the properties of the subgroup $K \leq B(2,n)$ defined in the proof of Corollary 1.2, and choose a recursively enumerable but non-recursive subset T in the set of free generators $W = \{w_1, w_2, \dots\}$. Let N be the normal closure of T in $B(2,n)$, and so the group $G = B(2,n)/N$ is recursively presented. A word $w_i \in W$ in the generators a, b of $B(2,n)$ is trivial in G iff it belongs to N. By CEP, this condition is equivalent to the property $w_i \in T$, because W is the free basis of K. Since the set T is not recursive, the word problem in G is undecidable. By Corollary 1.2, the finitely generated group G is a subgroup of a finitely presented in $\mathcal{B}_n$ group H with two generators. Hence the word problem in H is undecidable too. □

The following proposition will be considered at the end of this paper since the proof of it is based on a few lemmas needed for the main theorem. Recall that a subgroup $A \leq B$ is called *Frattini embedded* in the group B (see [32]) provided two elements of A are conjugate in B if and only if they are conjugate in A.

Proposition 1.5. *The embeddings $G \hookrightarrow H$, $G \hookrightarrow E$ and $G_i \hookrightarrow E$ given by Theorem 1.1, Corollary 1.2 and Corollary 1.3 enjoy the following additional properties.*

(1) The group G (every group G_i) is a CEP subgroup in H (resp., in E).

(2) The group G (every group G_i) is Frattini embedded in H (resp., in E).

W.W. Boone and G. Higman obtained a pure algebraic characterization of groups with decidable word problem. They proved in [5] that the word problem is decidable in a finitely generated group G iff G is embeddable in a simple subgroup of a finitely presented group. Similar characterization is obtained in the Burnside varieties $\mathcal{B}_n$ (n is a large odd)

Theorem 1.6. *There is a constant C such that for every odd $n \geq C$, the following is true. A finitely generated group G from the variety $\mathcal{B}_n$ has decidable word problem if and only if G is a embeddable in a simple subgroup of a group finitely presented in the variety $\mathcal{B}_n$.*

There are two main ingredients in the proof of Theorem 1.1: diagram analysis of the consequences of defining relations and programming of machines producing the set of relations of the group G. The work of a Turing machine or of an S-machine has finite program; a proper interpretation of their commands in terms of finite sets of group relations leads to the proof of the classical Higman's theorem. In the present paper, we should make the "machine relations" compatible with Burnside relations so that their union does not kill nontrivial elements of G.

Recall that the first finitely generated infinite groups satisfying the Burnside identity $x^n = 1$ for a large odd exponent n were constructed by P.S. Novikov and S.I. Adian (see [16], [17], [1]). Much shorter proof, although with a worse estimate for the exponent n, was given in [18] (see [19] for other results). Among the extensions and the applications of our approach, we emphasize the development presented in the joint paper of Olshanskii and Sapir [23], which is the important tool of the current exposition too.

As in [19] and [23], the concept of A-map is permanently used in the diagram part of this paper. To define it, we use so called contiguity submaps Γ between two cells (or faces, regions) Π_1 and Π_2 of a planar map, i.e. very thin submaps in comparison with the perimeters $|\partial\Pi_1|$ and $|\partial\Pi_2|$. Loosely speaking, the A-property says that $|\partial\Pi_1| << |\partial\Pi_2|$ if Γ is not too short in comparison with $|\partial\Pi_1|$. (See accurate definitions and basic properties of A-maps in Section 2). Note that there are no groups or van Kampen diagrams in Chapter 5 of [19]; the properties of A-maps obtained there used global estimates for surface maps on regarded as metric spaces. Lemmas 2.3 and 2.4 borrowed from [19] make redundant the local analysis of cancellations and inductive classification of periodic words typical for [17], [1]. In particular, Lemma 2.4 helps to prove that the free Burnside group $B(m, n)$ is infinite. This follows from well-known Prochet - Thue - Morse examples [26, 31, 14] of infinite sets of aperiodic words in 2-letter alphabet.

We start with a version of Higman's embedding of a recursively presented group G in a finitely presented group $\tilde{G}$. Then we impose Burnside relations $w^n = 1$ on $\tilde{G}$: at first we consider the words in so called a-generators of $\tilde{G}$, then inductively introduce the periods w depending on a- and θ-generators, finally the periods depend on all the generators of $\tilde{G}$. At each of these steps, the van Kampen diagram with standard metric do not satisfy the hyperbolic properties of A-map since the inductively defined groups $G(i)$ have many abelian subgroups. However they becomes A-maps with respect of special metrics. In particular, as in [23], we prescribe length 0 to every a-letter when introducing periods w depending on a- and θ-letters only, and even θ-letters get length 0 when all the generators are involved.

It was proved in [23] that the inductive transitions $\mathbf{G}(0) \to \ldots \to \mathbf{G}(i) \to \ldots \to \mathbf{G}(\infty)$ with surjective homomorphism are successful, i.e. the diagrams over $\mathbf{G}(\infty)$ are A-maps, if the original presentations of $\mathbf{G}(0)$ and of $\tilde{G} = \mathbf{G}(1/2)$ have Properties (Z1), (Z2), (Z3) (see Subsection 3.1). Property (Z1) imposes some restrictions on the relations of $\mathbf{G}(0)$ implying that the corresponding cells can be further regarded having rank 0 in diagrams. Property (Z2) is imposed on hub relations used for the Higman's embedding. (This property is empty, if the set of hubs is empty.) Property (Z3) is related to the known property of periodic words (see [6]) saying that simple periods of the same long periodic word are equal up to cyclic permutation. (Various inductive generalizations of such properties of periodic words were obtained in [17, 1]; for more direct generalizations formulated in different terms, see [19, 23].) Property (Z3) reduces the information we need on periodic words to the following condition.

(*) *Let W be a word of positive length which is not a conjugate of a shorter word in* $\mathbf{G}(1/2)$*, and X be a* 0*-word (i.e. every letter of X has length* 0*). Assume that the word* $W^{-4}XW^4$ *is also equal in* $\mathbf{G}(0)$ *to a* 0*-word Y. Then* $(WX)^n = W^n$ *in* $\mathbf{G}(0)$.

(The exponent 4 can be replaced with 3, but here we use the exponent from [23].)

To obtain Condition (*), we simplify the diagram Δ corresponding to the equality $W^{-4}XW^4 = Y$. Changing the pair (W, X) by conjugate pairs in $\mathbf{G}(0)$, we step-by-step remove unwanted cells until Δ becomes a diagram over the group M whose relations simulate the work of the machine $\mathbf{M}$ recognizing the relations of G. After further transformations one can find trapezia in Δ corresponding to some computations of $\mathbf{M}$. The labels of "horizontal" lines of a trapezium are the words W_i-s in a computation $W_0 \to W_1 \to \ldots \to W_k$, where W_i is obtained from W_{i-1} by an application of some rule θ_i, and the product $\theta_1 \ldots \theta_k$ is called the history h of this computation. It turns out that Condition (*) is a consequence of the properties of M-computations. For example:

(**) *If two computations with histories h and ghg^{-1} both start and end with the same word W (under some restrictions), then g is equal, modulo the Burnside relations, to a word g' such that there is a computation with history g' starting and ending with the word W.*

The designing of a machine $\mathbf{M}$ recognizing the recursively enumerable set of defining relations of given group G and satisfying conditions like (**), is the central point of this paper. The crucial issue and the important difference in comparison with [23] is the following. In the beginning, we know nothing about the Turing machine enumerating the defining relations of G, except for the *existence* of such a machine, whereas the S-machine $\mathcal{M}$ from [23] was explicitly defined to produce the words of the form w^n in some alphabet. (Note that S-machines invented by M. Sapir in [29] are more suitable for many group-theoretic applications than classical Turing machines working with positive words only, see [22].)

Now we start with a Turing machine M_0 having a few additional properties (e.g., no configuration can appear in a computation of M_0 two times). Then each command of M_0 should be transformed in a set of rules of an S-machine. Unfortunately, known constructions become insecure when one interprets the computations of S-machine by trapezia, because the cells corresponding to the relations of G and to the Burnside relations can be inserted between the horizontal bands, which distorts the computation. In [23], we had very concrete S-machine which produced arbitrary word w, and a after number of properly defined rewritings one obtained the required relator w^n. Those rewritings were organized and controlled so that arbitrary interference of unwanted cells replaced w^n with some word $(w')^n$, which was also a relator. But now M_0 looks like a black box device without information about its internal workings.

Therefore we suggest a new way of replacement of a Turing machine M_0 with an equivalent S-machine M_1. For every command θ of M_0, we construct an S-machine $M(\theta)$, and for every configuration $\mathcal{W}$ of M_0, we define a configuration $W = F(\mathcal{W})$ of M_1, where the inductive definition of the function F is based on aperiodic endomorphisms ψ_j of 2-generated free semigroup. For example, if θ inserts a letter a_j at the end of some tape, then the canonical word of $M(\theta)$ replaces the whole tape subword w of W with a word $\phi_j(w) = \mathbf{w_j}\psi_j(w)$, where $\mathbf{w}_j$ is an aperiodic word with a small cancellation property. The S-machine M_1 is a union of all $M(\theta)$-s.

The machine M_1 has a number of helpful properties. For instance, if some tape has a word w in the beginning of a computation and the word w' at the end, then either w' and w are equal in the free group or they have different canonical images in the free Burnside group. This and other features make M_1 interference-free, i.e. arbitrary replacements of the tape words with equal words modulo the Burnside relations, during a computation changes the output with an equal word modulo the Burnside relations too.

Many copies of M_1 gives us M_2, since one needs this modification to obtain Property (Z2) later. Then we should add a machine T which translates the language recognized by M_0 into the language recognized by M_1. Such a modification adds one more tape which does not belong to M_2, but T just cleans up this tape, and so it is easy to control possible unwanted "noise" during a computation.

Note that M_1 and the main machine $\mathbf{M}$ do not satisfy the usual definition of S-machine from [29], [22], [25]. So we have to give a more general definition, where the tape alphabets of θ-admissible words are replaced with finitely generated subgroups (depending on the rule θ) of free groups. As a consequence, some rules cannot be written in "'inserting"' form now, and a rule θ may insert a word in sectors locked by θ.

In sections 10 and 11, we get back to groups and diagrams and obtain the required embedding $G \hookrightarrow H$ starting with the Higman-type embedding $G \hookrightarrow \tilde{G}$, where the finite set of defining relations of $\tilde{G}$ is based on the rules of the machine $\mathbf{M}$. It turns out that the canonical mapping of G to the Burnside quotient H of $\tilde{G}$ is injective, which completes the proof of Theorem 1.1. Section 11 is closer to Sections 8 - 10 of [23], although some proofs are different and shorter now.

Theorem 1.6 is obtained in Subsection 12.2 as a corollary of Theorem 1.1. In addition, we must prove there that every finitely generated group of exponent n with decidable word problem is embeddable in a recursively presented simple group of exponent n. Embeddings of countable groups in simple groups of exponent n are defined in Chapter 11 of [19]. They were obtained by V.N. Obraztsov who modified the constructions of "monsters" from Chapters 7 - 9 [19]. By induction, the relations are imposed on free products in Chapter 11, while in Chapter 8 we start with a free group. Now we have to explain that the relations can be recursively enumerated if the free factors have decidable word problem. We do this using many references to lemmas from [19], but without long definitions such as "Condition R", or "B-map", because otherwise the exposition in Subsection 12.2 would be considerably longer. The motivated reader can find all the details in [19].

2 Maps and diagrams

Planar maps and van Kampen diagrams are standard tools to study the presentations of complicated groups (see, for example, [13] and [19]). Here we present the main concepts related to maps and diagrams. A van Kampen diagram is a labeled map, so we start with discussing maps.

2.1 Graphs and maps

We are using the standard definition of a *graph*. In particular, every edge has a direction, and every edge $\mathbf{e}$ has an inverse edge $\mathbf{e}^{-1}$ (having the opposite direction). For every edge $\mathbf{e}$, $\mathbf{e}_-$ and $\mathbf{e}_+$ are the beginning and the end vertices of the edge.

A *map* on a surface X (in this paper, on a disk or on an annulus) is simply a finite connected graph drawn on X which subdivides this surface into polygonal 2-cells (= cells). Edges do not have labels, so no group presentations are involved in studying maps. On the other hand the properties of maps help finding the structure of van Kampen diagrams because every diagram becomes a map after we remove all the labels.

Let us recall the necessary definitions from [19].

A map is called *graded* if every cell Π is assigned a non-negative number, $r(\Pi)$, its *rank*. A map Δ is called a *map of rank* i if all its cells have ranks $\le i$. Every graded map Δ has a *type* $\tau(\Delta)$ which is the vector whose first coordinate is the maximal rank of a cell in the map, say, r, the second coordinate is the number of cells of rank r in the map, the third coordinate is the number of cells of maximal rank $< r$ and so on. We compare types lexicographically.

Let Δ be a graded map on a surface X. The cells of rank 0 in Δ are called 0-*cells*. Some of the edges in Δ are called 0-*edges*. Other edges and cells will be called *positive*. For every path $\mathbf{p}$ in Δ, there are the beginning and the end vertices $\mathbf{p}_-$ and $\mathbf{p}_+$. We define the *length of a path* $\mathbf{p}$ in a map as the number of positive edges in it. The length of a path $\mathbf{p}$ is denoted by $|\mathbf{p}|$, in particular, $|\partial(\Pi)|$ is the *perimeter* perimeter of a cell Π and $|\partial(\Delta)|$ is the perimeter of a disk map Δ.

We assume that

(M1) the inverse edge of a 0-edge is also a 0-edge,

(M2) if Π is a 0-cell, then we have either $|\partial\Pi| = 0$ or $|\partial\Pi| = 2$,

(M3) for every cell Π of rank $r(\Pi) > 0$ the length $|\partial(\Pi)|$ of its contour is positive.

2.2 Bands

A van Kampen *diagram* over a presentation $\langle \mathcal{X} \mid \mathcal{R} \rangle$ is a planar map with edges labeled by elements from $\mathcal{X}^{\pm 1}$ (and therefore paths labeled by words), such that the label of the contour of each cell belongs to $\mathcal{R}$ up to taking cyclic shifts and inverses [19]. In §11 of [19], 0-edges of van Kampen diagrams have label 1, but in [23] and in this paper, 0-edges will be labeled also by the so called *0-letters*. All diagrams which we shall consider in this paper, when considered as planar maps, will obviously satisfy conditions (M1), (M2), (M3) from Section 2.1. By van Kampen Lemma, a word w over $\mathcal{X}$ is equal to 1 modulo $\mathcal{R}$ if and only if there exists a van Kampen disk diagram over $\mathcal{R}$ with boundary label w. Similarly (see [13], [19]) two words w_1 and w_2 are conjugate modulo R if and only if there exists an annular diagram with the internal counterclockwise contour labeled by w_1 and external counterclockwise contour labeled by w_2.

There may be many diagrams with the same boundary label. Sometimes we can reduce the number of cells in a diagram or ranks of cells by replacing a 2-cell subdiagram by another subdiagram with the same boundary label. The simplest such situation is when there are two positive cells in a diagram which have a common edge and are mirror images of each other, i.e. the labels of the boundary paths of these cells starting with this edge are equal. In this case one can *cancel* these two cells (see [13]). In order strictly define this cancellation (to preserve the topological type of the diagram), one needs to use the so called 0-refinement. The 0-refinement consists of adding 0-cells with (freely trivial) boundary labels of the form $1aa^{-1}$ or 111 (see [19], §11.5). A diagram having no such mirror pairs of positive cells is called *reduced*.

As in the book [27] and our the papers [29], [3] and [20], one of the main tools to study van Kampen diagrams are bands and annuli.

Let S be a subset of $\mathcal{X}$. An *S-band* $\mathcal{B}$ is a sequence of cells $\pi_1, ..., \pi_n$, for some $n \geq 0$ in a van Kampen diagram such that

- Each two consecutive cells in this sequence have a common edge labeled by a letter from S.
- Each cell π_i, $i = 1, ..., n$ has exactly two S-edges (i.e. edges labeled by a letter from S).
- If $n = 0$, then the boundary of S has form $\mathbf{e}\mathbf{e}^{-1}$ for an S-edge $\mathbf{e}$.

Figure 1 illustrates this concept. In this Figure, edges $\mathbf{e}, \mathbf{e}_1, ..., \mathbf{e}_{n-1}, \mathbf{f}$ are S-edges, the lines $l(\pi_i, \mathbf{e}_i), l(\pi_i, \mathbf{e}_{i-1})$ connect fixed points in the cells with fixed points of the corresponding edges.

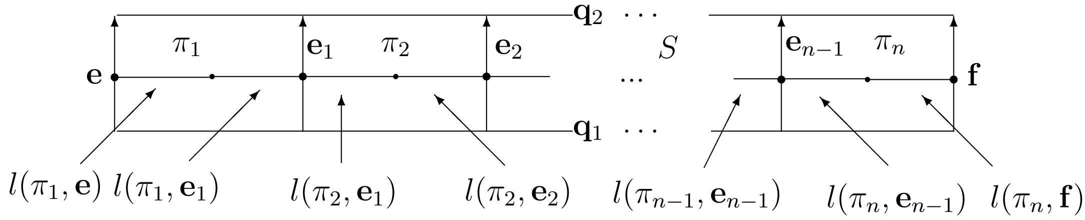

Fig. 1.

The broken line formed by the lines $l(\pi_i, \mathbf{e}_i)$, $l(\pi_i, \mathbf{e}_{i-1})$ connecting points inside neighboring cells is called the *connecting line* $\mathcal{B}$. The S-edges $\mathbf{e}$ and $\mathbf{f}$ are called the *start* and *end* edges of the band.

The counterclockwise boundary of the subdiagram formed by the cells $\pi_1, ..., \pi_n$ of $\mathcal{B}$ has the form $\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f}\mathbf{q}_2^{-1}$. We call $\mathbf{q}_1$ the *bottom* of $\mathcal{B}$ and $\mathbf{q}_2$ the *top* of $\mathcal{B}$, denoted $\mathbf{bot}(\mathcal{B})$ and $\mathbf{top}(\mathcal{B})$. (Both of them are the *sides* of $\mathcal{B}$.)

2.3 Bonds and contiguity submaps

Now let us introduce the main concepts from Chapter 5 of [19]. Consider a map Δ. A path in Δ is called *reduced* if it does not contain subpaths of length 2 which are homotopic to paths of length 0 by homotopies involving only 0-cells. We shall usually suppose that the contour of Δ is subdivided into several subpaths called *sections*. We usually assume that sections are reduced paths.

By property (M2), if a 0-cell has a positive edge, it has exactly two positive edges. Thus we can consider bands of 0-cells having positive edges. The start and end edges of such bands are called *adjacent*. (We obtain a simple generalization of the cancellation of a pair of positive cells in a diagram replacing the common edge of their boundaries with two adjacent edges.)

Let $\mathbf{e}$ and $\mathbf{f}$ be adjacent edges of Δ belonging to two positive cells Π_1 and Π_2 or to sections of the contour of Δ. Then there exists a band of 0-cells connecting these edges. The union of the cells of this band is called a 0-*bond* between the cells Π_1 and Π_2 (or between a cell and a section of the contour of Δ, or between two sections of the contour). The contour of the 0-bond has the form $\mathbf{p}^{-1}\mathbf{e}\mathbf{s}\mathbf{f}^{-1}$, where $\mathbf{p}$ and $\mathbf{s}$ are paths of length 0 because every 0-cell has at most two positive edges.

Now suppose we have chosen two pairs $(\mathbf{e}_1, \mathbf{f}_1)$ and $(\mathbf{e}_2, \mathbf{f}_2)$ of adjacent edges such that $\mathbf{e}_1, \mathbf{e}_2$ belong to the contour of a cell Π_1 and $\mathbf{f}_1^{-1}, \mathbf{f}_2^{-1}$ to the contour of a cell Π_2. Then we have two bonds E_1 and E_2. If $E_1 = E_2$ (that is $\mathbf{e}_1 = \mathbf{e}_2, \mathbf{f}_1 = \mathbf{f}_2$) then let $\Gamma = E_1$. Now let $E_1 \neq E_2$ and $\mathbf{z}_1\mathbf{e}_1\mathbf{w}_1\mathbf{f}_1^{-1}$ and $\mathbf{z}_2\mathbf{e}_2\mathbf{w}_2\mathbf{f}_2^{-1}$ be the contours of these bonds. Further let $\mathbf{y}_1$ and $\mathbf{y}_2$ be subpaths of the contours of Π_1 and Π_2 (of Π_1 and $\mathbf{q}$, or $\mathbf{q}_1$ and $\mathbf{q}_2$, where $\mathbf{q}$, $\mathbf{q}_1$ and $\mathbf{q}_2$ are some segments of the boundary contours) where $\mathbf{y}_1$ (or $\mathbf{y}_2$) has the form $\mathbf{e}_1\mathbf{p}\mathbf{e}_2$ or $\mathbf{e}_2\mathbf{p}\mathbf{e}_1$ (or $(\mathbf{f}_1\mathbf{u}\mathbf{f}_2)^{-1}$ or $(\mathbf{f}_2\mathbf{u}\mathbf{f}_1)^{-1}$). If $\mathbf{z}_1\mathbf{y}_1\mathbf{w}_2\mathbf{y}_2$ (or $\mathbf{z}_2\mathbf{y}_1\mathbf{w}_1\mathbf{y}_2$) is a contour of a disk submap Γ which does not contain Π_1 or Π_2, then Γ is called a 0-*contiguity submap* of Π_1 to Π_2 (or Π_1 to $\mathbf{q}$ or of $\mathbf{q}_1$ to $\mathbf{q}_2$). The contour of Γ is naturally subdivided into four parts. The paths $\mathbf{y}_1$ and $\mathbf{y}_2$ are called the *contiguity arcs*. We write $\mathbf{y}_1 = \Gamma \bigwedge \Pi_1$, $\mathbf{y}_2 = \Gamma \bigwedge \Pi_2$ or $\mathbf{y}_2 = \Gamma \bigwedge \mathbf{q}$. The other two paths are called the *side arcs* of the contiguity submap.

The ratio $|\mathbf{y}_1|/|\partial(\Pi_1)|$ (or $|\mathbf{y}_2|/|\partial(\Pi_2)|$ is called the *degree of contiguity* of the cell Π_1 to the cell Π_2 or to $\mathbf{q}$ (or of Π_2 to Π_1). We denote the degree of contiguity of Π_1 to Π_2 (or Π_1 to $\mathbf{q}$) by (Π_1, Γ, Π_2) (or $(\Pi_1, \Gamma, \mathbf{q})$). Notice that this definition is not symmetric and when $\Pi_1 = \Pi_2 = \Pi$, for example, then (Π, Γ, Π) is a pair of numbers.

We say that two contiguity submaps are *disjoint* if none of them has a common point with the interior of the other one, and their contiguity arcs do not have common edges.

Now we are going to define k-bonds and k-contiguity submaps for $k > 0$. In these definitions we need a fixed number ε, $0 < \varepsilon < 1$.

Let $k > 0$ and suppose that we have defined the concepts of j-bond and j-contiguity submap for all $j < k$. Consider three cells π, Π_1, Π_2 (possibly with $\Pi_1 = \Pi_2$) satisfying the following conditions:

1. $r(\pi) = k, r(\Pi_1) > k, r(\Pi_2) > k$,

2. there are disjoint submaps Γ_1 and Γ_2 of j_1-contiguity of π to Π_1 and of j_2-contiguity of π to Π_2, respectively, with $j_1 < k, j_2 < k$, such that Π_1 is not contained in Γ_2 and Π_2 is not contained in Γ_1,

3. $(\pi, \Gamma_1, \Pi_1) \geq \varepsilon$, $(\pi, \Gamma_2, \Pi_2) \geq \varepsilon$.

Then there is a minimal submap E in Δ containing π, Γ_1, Γ_2. This submap is called a *k-bond between Π_1 and Π_2 defined by the contiguity submaps Γ_1 and Γ_2 with principal cell π* (see Figure 2).

The *contiguity arc* q_1 of the bond E to Π_1 is $\Gamma_1 \bigwedge \Pi_1$. It will be denoted by $E \bigwedge \Pi_1$. Similarly $E \bigwedge \Pi_2$ is by definition the arc $\mathbf{q}_2 = \Gamma_2 \bigwedge \Pi_2$. The contour $\partial(E)$ can be written in the form $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$ where $\mathbf{p}_1$, $\mathbf{p}_2$ are called the *side arcs* of the bond E.

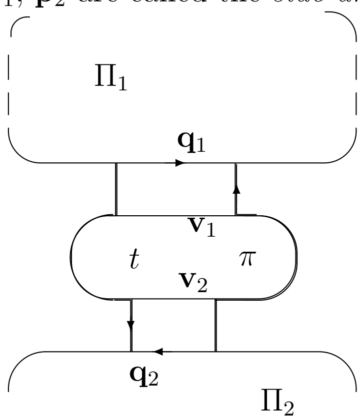

Fig. 2.

Bonds between a cell and a section of the contour of Δ or between two sections of the contour are defined in a similar way.

Now let E_1 be a k-bond and E_2 be a j-bond between two cells Π_1 and Π_2, $j \leq k$ and either $E_1 = E_2$ or these bonds are disjoint. If $E_1 = E_2$ then $\Gamma = E_1 = E_2$ is called the k-contiguity submap between Π_1 and Π_2 determined by the bond $E_1 = E_2$. If E_1 and E_2 are disjoint then the corresponding *contiguity submap* Γ is defined as the smallest submap containing E_1 and E_2, bounded by these submaps and segments of $\partial\Pi_1$ and $\partial\Pi_2$, and not containing Π_1 and Π_2 (see Figure 3).

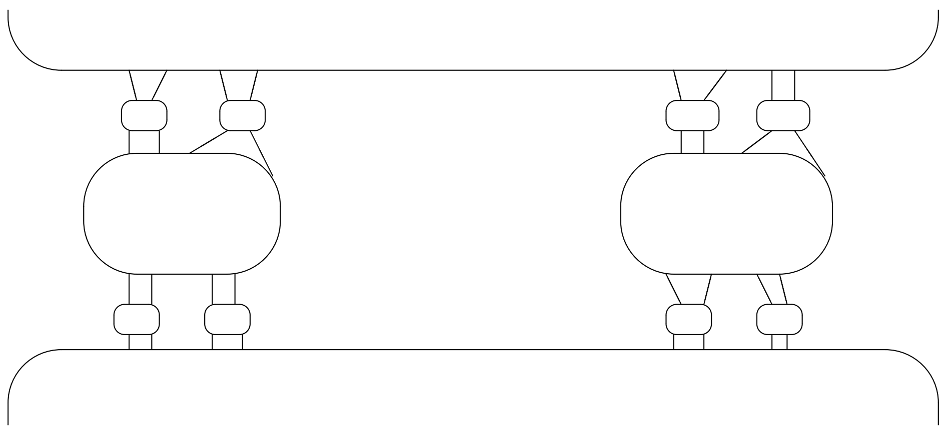

Fig. 3.

The *contiguity arcs* $\mathbf{q}_1$ and $\mathbf{q}_2$ of Γ are the intersections of $\partial(\Gamma)$ with $\partial(\Pi_1)$ and $\partial(\Pi_2)$. The contour of Γ has the form $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$ where $\mathbf{p}_1$ and $\mathbf{p}_2$ are called the *side arcs* of Γ. The ratio $|\mathbf{q}_1|/|\partial(\Pi_1)|$ is called the contiguity degree*contiguity degree* of Π_1 to Π_2 with respect to Γ and is denoted by (Π_1, Γ, Π_2). If $\Pi_1 = \Pi_2 = \Pi$ then (Π, Γ, Π) is a pair of numbers.

Contiguity submaps of a cell to a section of the contour and between sections of the contour are defined in a similar way. (In this paper we do not use notion "degree of contiguity of a section of a contour to" anything.)

We are going to write "bonds" and " contiguity submaps" in place of "k-bonds" and "k-contiguity submaps". Instead of writing "the contiguity submap Γ of a cell Π to ...", we sometimes write "the Γ- contiguity of Π to ...".

The above definition involved the standard decomposition of the contour of a contiguity submap Γ into four sections $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$ where $\mathbf{q}_1 = \Gamma \bigwedge \Pi_1$, $\mathbf{q}_2 = \Gamma \bigwedge \Pi_2$ (or $\mathbf{q}_2 = \Gamma \bigwedge \mathbf{q}$ if $\mathbf{q}$ is a section of $\partial(\Delta)$), we shall write $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2 = \partial(\Pi_1, \Gamma, \Pi_2)$, and so on.

As in [19] (see §15.1) we fix certain real numbers *parametes*

$$\iota << \zeta << \varepsilon << \delta << \gamma << \beta << \alpha \tag{1}$$

between 0 and 1 where "$<<$" means "much smaller". Here "much" means enough to satisfy all the inequalities in Chapters 5, 6 of [19]. We also set

$$\bar{\alpha} = \frac{1}{2} + \alpha, \quad \bar{\beta} = 1 - \beta \ \ , \bar{\gamma} = 1 - \gamma, \quad h = \delta^{-1}, \quad n = \iota^{-1}. \tag{2}$$

2.4 Condition A

The set of positive cells of a map Δ is denoted by $\Delta(2)$ and as before the length of a path in Δ is the number of positive edges in the path. A path $\mathbf{p}$ in Δ is called *geodesic* if $|\mathbf{p}| \leq |\mathbf{p}'|$ for any path $\mathbf{p}'$ combinatorially homotopic to $\mathbf{p}$.

The condition A has three parts:

A1 If Π is a cell of rank $j > 0$, then $|\partial(\Pi)| \geq nj$.

A2 Any subpath of length $\leq \max(j, 2)$ of the contour of an arbitrary cell of rank j in $\Delta(2)$ is geodesic in Δ.

A3 If $\pi, \Pi \in \Delta(2)$ and Γ is a contiguity submap of π to Π with $(\pi, \Gamma, \Pi) \geq \varepsilon$, then $|\Gamma \bigwedge \Pi| < (1 + \gamma)k$, where $k = r(\Pi)$.

A map satisfying conditions A1, A2, A3 will be called an *A-map.* As in [19], Section 15.2, a (cyclic) section $\mathbf{q}$ of a contour of a map Δ is called a *smooth section of rank* $k > 0$ if:

1) every subpath of $\mathbf{q}$ of length $\leq \max(k, 2)$ is geodesic in Δ;

2) for each contiguity submap Γ of a cell π to $\mathbf{q}$ satisfying $(\pi, \Gamma, \mathbf{q}) \geq \varepsilon$, we have $|\Gamma \bigwedge \mathbf{q}| < (1 + \gamma)k$.

Remark 2.1. *This definition implies that every geodesic section* $\mathbf{q}$ *of positive length in* $\partial\Delta$ *is smooth of rank* $k = |\mathbf{q}|$.

It is shown in [19], §§16, 17, that A-maps have several "hyperbolic" properties. (They show that A-maps of rank i are hyperbolic spaces with hyperbolic constant depending on i only.) The next three lemmas are Lemma 15.8, Theorem 16.1 and Theorem 16.2 from [19]. (The proof of Lemma 2.4 is contained in the proof of Theorem 2 [19].)

Lemma 2.2. *In an arbitrary A-map* Δ*, the degree of contiguity of an arbitrary cell* π *to an arbitrary cell* Π *or to an arbitrary smooth section* $\mathbf{q}$ *of the contour* $\partial\Delta$ *via arbitrary contiguity submap is less than* $\bar{\alpha}$. □

Lemma 2.3. *(a) Let* Δ *be a disk A-map whose contour is subdivided into at most* 4 *sections* $\mathbf{q^1}$, $\mathbf{q^2}$, $\mathbf{q^3}$, $\mathbf{q^4}$ *and* $r(\Delta) > 0$. *Then there exists a positive cell* π *and disjoint contiguity submaps* Γ_1, Γ_2, Γ_3, Γ_4 *of* π *to* $\mathbf{q^1}, \ldots, \mathbf{q^4}$ *(some of these submaps may be absent) such that the sum of corresponding contiguity degrees* $\sum_{i=1}^{4}(\pi, \Gamma_i, \mathbf{q}_i)$ *is greater than* $\bar{\gamma} = 1 - \gamma$.

(b)Let Δ *be an annular A-map whose contours are subdivided into at most* 4 *sections* $\mathbf{q^1}$, $\mathbf{q^2}$, $\mathbf{q^3}$, $\mathbf{q^4}$ *regarded as cyclic or ordinary paths, and* $r(\Delta) > 0$. *Then there exists a positive cell* π *and disjoint contiguity submaps* Γ_1, Γ_2, Γ_3, Γ_4 *of* π *to* $\mathbf{q^1}, \ldots, \mathbf{q^4}$ *(up to three of these submaps may be absent) such that the sum of corresponding contiguity degrees* $\sum_{i=1}^{4}(\pi, \Gamma_i, \mathbf{q}_i)$ *is greater than* $\bar{\gamma}$. □

Lemma 2.4. *Let* Δ *be a disk or annular A-map. Assume that there is a cell* Π *in* Δ *and a contiguity submap* Γ *of* Π *to a section* q *of the boundary with* $(\Pi, \Gamma, \mathbf{q}) \geq \varepsilon$. *Then there is a cell* π *of* Δ *and a contiguity submap* Γ' *of* π *to* $\mathbf{q}$ *such that* $r(\Gamma') = 0$ *and* $(\pi, \Gamma', \mathbf{q}) \geq \varepsilon$. □

3 Factor groups of finite exponent

3.1 Axioms

Recall that in Chapter 6 of [19], it is proved that for any sufficiently large odd n, there exists a presentation of the free Burnside group $B(m,n)$ with $m \geq 2$ generators, such that every reduced diagram over this presentation is an A-map. The method from [19] was generalized in [23] to Burnside factors of non-free groups in the following way.

In [19], 0-edges are labeled by 1 only, and (trivial) 0-relators have the form $1 \cdot \ldots \cdot 1$ or $a \cdot 1 \cdot \ldots \cdot 1 \cdot a^{-1} \cdot 1 \cdot \ldots \cdot 1$. In [23], we enlarged the collection of labels of 0-edges and the collection of 0-relations. Namely, we divide the generating set of our group into the set of 0-letters and the set of positive letters. Then in the diagrams, 0-letters label 0-edges. We shall also add defining relators of our group which have the form $uavb^{-1}$ where u, v are words consisting of 0-letters only (we call such words 0-*words*), and a, b are a positive letters, as well as defining relations without positive letters, to the list of zero relations. The corresponding cells in diagrams will be called 0-cells.

Let $\mathbf{G}$ be a group given by a presentation $\langle \mathcal{X} \mid \mathcal{R} \rangle$. Let $\mathcal{X}$ be a symmetric set (closed under taking inverses). Let $\mathcal{Y} \subseteq \mathcal{X}$ be a symmetric subset which we shall call the *set* of *non-zero letters* or Y-*letters*. All other letters in $\mathcal{X}$ will be called 0-*letters*.

By the Y-*length* of a word A, written $|A|_Y$ or just $|A|$, we mean the number of occurrences of Y-letters in A. The Y-length $|g|_Y$ of an element $g \in \mathbf{G}$ is the length of the Y-shortest word w representing element g. A word w in generators is called *minimal* (*cyclically minimal*) in a group if it is not equal (resp., not conjugate) in this group to a shorter word.

Let $\mathbf{G}(\infty)$ be the group given by the presentation $\langle \mathcal{X} \mid \mathcal{R} \rangle$ inside the variety of Burnside groups of exponent n. We will choose a presentation of the group $\mathbf{G}(\infty)$ in the class of all groups, that is we add enough relations of the form $A^n = 1$ to ensure that $\mathbf{G}(\infty)$ has exponent n. This infinite presentation will have the form $\langle \mathcal{X} \mid \mathcal{R}(\infty) = \cup \mathcal{S}_i \rangle$ where sets $\mathcal{S}_i$ will be disjoint, and relations from $\mathcal{S}_i$ will be called relations of *rank* i. Unlike [19, Chapter 6], i runs over $0, 1/2, 1, 2, 3, \ldots$. ($i = 0, 1, 2, \ldots$ in [19, Chapter 6].) The set $\mathcal{R}$ will be equal to $\mathcal{S}_0 \cup \mathcal{S}_{1/2}$. The detailed description of sets $\mathcal{S}_i$ will be given below. We shall call this presentation of $\mathbf{G}(\infty)$ a *graded presentation*.

As in [19], we proved in [23] that (reduced) diagrams over $\mathcal{R}(\infty)$ are A-maps. This goal was achieved if $\mathcal{R}$ satisfies conditions (Z1), (Z2), (Z3) presented below. While listing these conditions, we also fix some notation and definitions.

(Z1) The set $\mathcal{R}$ is the union of two disjoint subsets $\mathcal{S}_0 = \mathcal{R}_0$ and $\mathcal{S}_{1/2}$. The group $\langle \mathcal{X} \mid \mathcal{S}_0 \rangle$ is denoted by $\mathbf{G}(0)$ and is called the group of rank 0. We call relations from $\mathcal{S}_0$ relations of rank 0. The relations from $\mathcal{S}_{1/2}$ have rank 1/2. The group $\langle \mathcal{X} \mid \mathcal{R} \rangle$ is denoted by $\mathbf{G}(1/2)$.

(Z1.1) The set $\mathcal{S}_0$ consists of all relations from $\mathcal{R}$ which have Y-length 0 and all relations of $\mathcal{R}$ which have the form (up to a cyclic shift) $ay_1by_2^{-1} = 1$ where $y_1, y_2 \in \mathcal{Y}^+$, a and b are of Y-length 0 (i.e. 0-words).

(Z1.2) The set $\mathcal{S}_0$ implies all Burnside relations $u^n = 1$ of Y-length 0.

The subgroup of $\mathbf{G}(0)$ generated by all 0-letters is called the 0-*subgroup* of $\mathbf{G}(0)$. Elements from this subgroup are called 0-*elements*. Elements which are not conjugates of elements from the 0-subgroup are called *essential*. An essential element g from $\mathbf{G}(0)$ is called *cyclically Y-reduced* if a shortest word A representing g is cyclically Y-reduced, i.e. no cyclic permutation of A has a subword of lengths 2 equal to a 0-word. It is shown in [23] (Lemma 3.5) that an element g is cyclically Y-reduced if and only if it is cyclically minimal (in rank 0), i.e. it is not a conjugate of a Y-shorter element in $\mathbf{G}(0)$.

Notice that condition (Z1.1) allows us to consider Y-bands in van Kampen diagrams over $\mathcal{S}_0$. Maximal Y-bands do not intersect.

(Z2) The relators of the set $\mathcal{S}_{1/2}$, will be called *hubs*. The corresponding cells in van Kampen diagrams are also called hubs. They satisfy the following properties

(Z2.1) The Y-length of every hub is at least n.

(Z2.2) Every hub is linear in $\mathcal{Y}$, i.e. contains at most one occurrence $y^{\pm 1}$ for every letter $y \in \mathcal{Y}$.

(Z2.3) Assume that each of words v_1w_1 and v_2w_2 is a cyclic permutation of a hub or of its inverse, and $|v_1| \geq \varepsilon |v_1w_1|$. Then an equality $u_1v_1 = v_2u_2$ for some 0-words u_1, u_2, implies in $\mathbf{G}(0)$ equality $u_2w_1 = w_2u_1$ (see Figure 4).

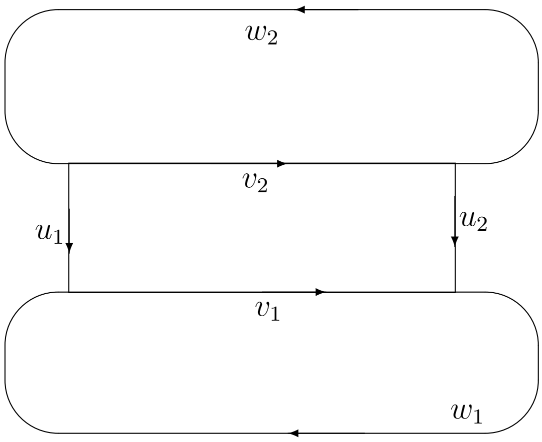

Fig. 4.

A word B over $\mathcal{X}$ is said to be *cyclically minimal in rank* 1/2 if it is not a conjugate in $\mathbf{G}(1/2)$ of a word of smaller Y-length. An element g of $\mathbf{G}(0)$ is called *cyclically minimal in rank* 1/2 if it is represented by a word which is cyclically minimal in rank 1/2.

(Z3) With every essential element $g \in \mathbf{G}(0)$ we associate a subgroup $\mathbf{0}(g) \leq \mathbf{G}(0)$, normalized by g. If g is cyclically Y-reduced, let $\mathbf{0}(g)$ be the maximal subgroup consisting of 0-elements which contains g in its normalizer. By Lemmas 3.3 (2) and 3.4 (3) from [23], arbitrary essential element g is equal to a product vuv^{-1} where u is cyclically Y-reduced. In this case we define $\mathbf{0}(g) = v\mathbf{0}(u)v^{-1}$. It is proved in Lemma 3.7 [23] that $\mathbf{0}(g)$ is well defined.

(Z3.1) Assume that g is an essential cyclically minimal in rank 1/2 element of $\mathbf{G}(0)$ and for some $x \in \mathbf{G}(0)$, both x and $g^{-4}xg^4$ are 0-elements. Then $x \in \mathbf{0}(g)$.

(Z3.2) For every essential cyclically minimal in rank 1/2 element $g \in \mathbf{G}(0)$, there exists a 0-element r such that gr^{-1} commutes with every element of $\mathbf{0}(g)$ [1]

The following statement is given by Lemma 3.12 and Lemma 3.13 in [23].

Lemma 3.1. *Suppose that $\mathcal{R}$ satisfies conditions (Z1.1), (Z1.2), and (Z3.2). Then for every essential element $g \in \mathbf{G}(0)$ which is cyclically minimal in rank 1/2, and every $x \in \mathbf{0}(g)$ we have $(gx)^n = g^n$ and $g^n x = xg^n$.* □

In [23], the approach from [19, Chapter 6] was adapted to diagrams over a presentation satisfying (Z1), (Z2), (Z3). The construction is defined below .

3.2 The construction of a graded presentation

Here we recall several concepts analogous to the concepts in [19], chapter 6. They were introduced in [23] by induction on the rank i.

We say that a certain statement *holds in rank i* if it holds in the group $\mathbf{G}(i)$ given by $\mathcal{R}_i = \cup_{j=0}^{i} \mathcal{S}_j$.

Recall that a word w is A-*periodic* if it is a subword of some power of the word A.

Up to the end of this section, the Y-length of a word W will be called simply length *and will be denoted by $|W|$.*

Let us define simple words of rank $i = 0, 1/2, 1, 2, ...$ and periods of rank $i = 1, 2, 3, \dots$. Let $\mathcal{S}_0 = \mathcal{R}_0$ and $\mathcal{S}_{1/2}$ be the set of hubs. Thus the groups $\mathbf{G}(0)$ and $\mathbf{G}(1/2) = \mathbf{G}$ are defined by relations of the sets $\mathcal{R}_0$ and $\mathcal{R}_{1/2} = \mathcal{S}_0 \cup \mathcal{S}_{1/2}$, respectively. A word A of positive length is said to be *simple* in rank 0 (resp. 1/2) if it is not a conjugate in rank 0 (resp. 1/2) either of a word of length 0 or of a product of the form $B^l P$ for an essential B, where $|B| < |A|$ and $P \in \mathbf{0}(B)$ in rank 0.

Remark 3.2. *It follows from the definition that a simple word in rank 0 is not a conjugate in rank 0 of any shorter word. Also if two words of the same length are conjugate in rank 0 and one of them is simple in rank 0, then the other one is simple in rank 0 as well.*

[1] We noticed in Section 3.1 [23] that (Z3.2) can be replaced with the weaker condition (Z3.2'): For every essential cyclically minimal in rank 1/2 element $g \in \mathbf{G}(0)$, the extension of $\mathbf{0}(g)$ by the automorphism induced by g (acting by conjugation) satisfies the identity $x^n = 1$. Condition (Z2.3') goes back to S.V. Ivanov's paper [8].

Suppose that $i \geq 1$, and we have defined the sets of relators $\mathcal{R}_j$, $j < i$, and the corresponding groups $\mathbf{G}(j)$ of ranks $j < i$. Suppose also that we have defined simple words of ranks j, $j < i$, and periods of rank j, $1 \leq j < i$.

For every $i = 1/2, 1, 2, 3, ..$ let i^- denote $i - 1$ if $i > 1$, $1/2$ if $i = 1$ or 0 if $i = 1/2$.

For every $i \geq 1$ consider the set $\mathcal{X}'_i$ of all words of length i which are simple in rank i^-, and the equivalence $\sim_{i^-}$ given in Lemma 3.3 (we apply it for the smaller rank here). Now choose a set of representatives $\mathcal{X}_i$ of the equivalence classes of $\mathcal{X}'_i$. The words of $\mathcal{X}_i$ are said to be *periods* of rank i. The set of words $\mathcal{R}_i$ (defining the group $\mathbf{G}(i)$ or rank i) is the union of $\mathcal{R}_{i^-}$ and $\mathcal{S}_i = \{A^n,\ A \in \mathcal{X}_i\}$.

Let A be a word of positive length. We say that A is *simple* in rank $i \geq 1$ if it is not a conjugate in rank i either of a word of length 0 or of a word of the form $B^l P$, where B is a period of rank $j \leq i$ or an essential word with $|B| < |A|$, and P represents a word of the subgroup $\mathbf{0}(B)$.

Lemma 3.3. *([23], Lemma 3.24) The following relation $\sim_i$ is an equivalence on the set of all simple in rank i words: $A \sim_i B$ by definition, if there are words X, P, R, where $P \in \mathbf{0}(A)$, $R \in \mathbf{0}(B)$ (in rank 0), such that $AP = XB^{\pm 1}RX^{-1}$ in rank i.* □

Let $\mathcal{R}(\infty) = \cup_{i=0}^{\infty}\mathcal{R}_i$. The set $\mathcal{R}(\infty)$ defines the group $\mathbf{G}(\infty)$, which is the inductive (= direct) limit of the sequence of epimorphisms $\mathbf{G}(0) \to \mathbf{G}(1/2) \to \mathbf{G}(1) \to \dots \to \mathbf{G}(i) \to \dots$

If Π is a cell of a diagram over $\mathcal{R}(\infty)$ with the boundary labeled by a word of the set $\mathcal{S}_j$, then, by definition, $r(\Pi) = j$. The boundary label of every cell of a rank ≥ 1 has a period A defined up to cyclic shifts.

For $j \geq 1$, a pair of distinct cells Π_1 and Π_2 of rank j of a diagram Δ is said to be a *j-pair*, if their counterclockwise contours $\mathbf{p}_1$ and $\mathbf{p}_2$ are labeled by A^n and A^{-n} for a period A of rank j and there is a path $\mathbf{t} = (\mathbf{p}_1)_- - (\mathbf{p}_2)_-$ without self-intersections such that the label $Lab(\mathbf{t})$ is equal, in rank j^-, to an element of $\mathbf{0}(A)$. Then $Lab(\mathbf{t})$ and A^n commute in rank j^- by Lemma 3.1, and so the subdiagram with contour $\mathbf{p}_1\mathbf{t}\mathbf{p}_2\mathbf{t}^{-1}$ can be replaced in Δ by a diagram of rank j^-. As a result, we obtain a diagram with the same boundary label as Δ but of a smaller type.

Similarly, a pair of hubs Π_1, Π_2 with counterclockwise contours $\mathbf{p}_1$, $\mathbf{p}_2$ forms a $\frac{1}{2}$-*pair*, if the vertices $(\mathbf{p}_1)_-$ and $(\mathbf{p}_2)_-$ are connected by a path $\mathbf{t}$ without self-intersections such that the label of the path $\mathbf{p}_1\mathbf{t}\mathbf{p}_2\mathbf{t}^{-1}$ is equal to 1 in $\mathbf{G}(0)$. Consequently, any diagram can be replaced by a reduced diagram having no j-pairs, $j = 1/2, 1, 2, ...$, i.e. by a *g-reduced* diagram with the same boundary label(s)[2].

The following statement is a part of Proposition 3.19 [23].

Lemma 3.4. *If the presentation of a group $\mathbf{G}$ satisfies properties (Z1), (Z2), (Z3) then all g-reduced diagrams over the presentation $\langle \mathcal{X} \mid \mathcal{R}(\infty)\rangle$ are A-maps. Every relator from $\mathcal{S}_i$, $i \geq 1$, is of the form A^n for a cyclically reduced word A of Y-length i. The group $\mathbf{G}(\infty)$ defined by this presentation is the factor group $\mathbf{G}/\langle g^n, g \in \mathbf{G}\rangle$.* □

[2]Recall that in [19], §13.2, such diagrams are called *reduced*. But in this paper we shall consider different kinds of reduced diagrams, so we call diagrams without j-pairs g-reduced; "g" stands for "graded".

Remark 3.5. *If $\mathcal{Y} = \mathcal{X}$, $\mathcal{S}_0 = \mathcal{S}_{1/2} = \emptyset$, then* $\mathbf{G}$ *is the free group with basis* $\mathcal{X}$ *and trivial subgroups* $\mathbf{0}(g)$, *and so Conditions (Z1), (Z2), (Z3) obviously hold. In this case, the factor group* $\mathbf{G}(\infty) = \mathbf{G}/\langle g^n, g \in \mathbf{G}\rangle$ *is the free Burnside group* $B_n(\mathcal{X})$ *of exponent* n *with free set of generators* $\mathcal{X}$. *Therefore g-reduced diagrams over* $\mathbf{G}(\infty) = B_n(\mathcal{X})$ *are A-maps.*

The following two lemmas will be explicitly used in Section 11.

Lemma 3.6. *([23], Lemma 3.25). Every word* X *is conjugate in rank* $i \geq 0$ *either of a* 0*-word or of a word* $A^l P$, *where* $|A| \leq |X|$, A *is either a period of rank* $j \leq i$ *or a simple in rank* i *word and* P *represents in rank* 0 *an element of the subgroup* $\mathbf{0}(A)$.

If X *is not simple in rank* 0 *and is not conjugate to a* 0*-word in rank* 0, *then* $|A| < |X|$. □

Note that the second sentence of Lemma 3.6 is not formulated in Lemma 3.25 [23], but it follows from the first phrase of the proof.

Lemma 3.7. *([23], Lemma 3.32). Let* Δ *be a g-reduced diagram of rank* i *with the contour* $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$, *where* $Lab(\mathbf{q}_1)$ *and* $Lab(\mathbf{q}_2)$ *are periodic words with simple in rank* i *periods* A *and* B, *resp., and* $|A| \geq |B|$. *Assume that* $|\mathbf{p}_1|, \mathbf{p}_2| < \alpha|B|$, $|\mathbf{q}_1| > \frac{3}{4}\delta^{-1}|A|$ *and* $|\mathbf{q}_2| > \delta^{-1}|B|$. *Then the word* A *is a conjugate in rank* i *of a product* $B^{\pm 1}R$, *where* R *represents an element of the subgroup* $\mathbf{0}(B)$, *and* $|A| = |B|$. *Moreover, if the words* $Lab(\mathbf{q}_1)$ *and* $Lab(\mathbf{q}_2)^{-1}$ *start with* A *and* B^{-1}, *resp., then* A *is equal in rank* i *to* $Lab(\mathbf{p}_1)^{-1}B^{\pm 1}R\ Lab(\mathbf{p}_1)$. □

4 Aperiodic endomorphisms

The rules of the S-machines introduced in this paper will use words from an infinite set $\mathbf{W}$ of positive words in the 2-letter alphabet $\{a, b\}$. This set has the following properties. (Below we use symbol $\equiv$ for letter-by-letter equality of words.)

(A) If two words w and v from $\mathbf{W}$ contain equal subwords, i.e. $w \equiv w_1 u w_2$ and $v \equiv v_1 u v_2$, then either $|u| < \frac{1}{5}\min(|w|, |v|)$ or $w_1 \equiv v_1$ and $w_2 \equiv v_2$. In particular, the subgroup generated by $\mathbf{W}$ in the free group $F = F(a, b)$ is a free group with basis $\mathbf{W}$.

(B) Suppose that in some product of words from $\mathbf{W}^{\pm 1}$ after all cancellations, we have a non-empty B-periodic subword V of length $\geq 2|B|$. Then either B is freely conjugate of a some product of words from $\mathbf{W}^{\pm 1}$ or $|V| \leq 11|B|$.

(C) The subgroup $\mathcal{B}$ generated by $\mathbf{W}$ in the free Burnside quotient $B_n(a, b) = F/F^n$ of F is the free Burnside group of exponent n with basis $\mathbf{W}$.

For every large enough odd n, such infinite set $\mathbf{W}$ was constructed by D. Sonkin in [30].

Choose now $(\mathbf{w}_1, \mathbf{w}_1(a), \mathbf{w}_1(b)), (\mathbf{w}_2, \mathbf{w}_2(a), \mathbf{w}_2(b)), \ldots$, disjoint triples of distinct words from $\mathbf{W}$. With every triple $(\mathbf{w}_j, \mathbf{w}_j(a), \mathbf{w}_j(b))$ we associate the subgroup $H(j) = gp\{\mathbf{w}_j(a), \mathbf{w}_j(b)\}$, the left coset $K(j) = \mathbf{w}_j H(j)$ and the endomorphism ψ_j of the free group $F(a, b)$ to $F(a, b)$ (and of the free Burnsine group of exponent n $B_n(a, b)$ to $B_n(a, b)$) such that

$$\psi_j(a) = \mathbf{w}_j(a) \ \ \textit{and} \ \ \psi_j(b) = \mathbf{w}_j(b)$$

Remark 4.1. *It follows from Property (A) of the set* $\mathbf{W}$ *(from (C)) that every* ψ_j *is a monomorphism of* $F(a,b)$ *(of* $B_n(a,b)$*) and* $\psi_j(F(a,b)) \cap \psi_{j'}(F(a,b)) = \{1\}$ *(resp.,* $\psi_j(B_n(a,b)) \cap \psi_{j'}(B_n(a,b)) = \{1\}$*) if* $j \neq j'$*. In other words,* $H(j) \cap H(j') = \{1\}$*, and the images of these two subgroups in* $B_n(a,b)$ *have trivial intersection too.*

Let ϕ_j be the mapping $F(a,b) \to F(a,b)$ and $B_n(a,b) \to B_n(a,b)$ given by the formula $\phi_j(u) = \mathbf{w}_j \psi_j(u)$. Property (A) of the set $\mathbf{W}$ implies $|u| < |\phi_j(u)|$ for every $u \in F(a,b)$.

Remark 4.2. *By Remark 4.1, every mapping* ϕ_j *is injective both on* $F(a,b)$ *and on* $B_n(a,b)$*. Besides we have* $\phi_j(F(a,b)) \cap \phi_{j'}(F(a,b)) = \emptyset$ *(resp.,* $\phi_j(B_n(a,b)) \cap \phi_{j'}(B_n(a,b)) = \emptyset$*) if* $j \neq j'$ *by Property (A) (resp., by (C)) of the set* $\mathbf{W}$*. In other words,* $K(j) \cap K(j') = \emptyset$*, and the images of these two cosets in* $B_n(a,b)$ *have empty intersection too.*

Lemma 4.3. *Assume that* $w_0 \in F(a,b)$ *and for* $i = 1, \dots, t$*, we have* $w_i = \phi_{j_i}^{\epsilon_i}(w_{i-1})$*, for some mappings of the form* ϕ_{j_i}*, where* $\epsilon_i = \pm 1$ *and we write* $w_i = \phi_{j_i}^{-1}(w_{i-1})$ *if* $w_{i-1} = \phi_{j_i}(w_i)$*. If* $w_t = w_0$*, then the reduced form of the word* $\phi_{j_1}^{\epsilon_1} \dots \phi_{j_t}^{\epsilon_t}$ *is empty.*

Proof. Assuming that the word $\phi_{j_1}^{\epsilon_1} \dots \phi_{j_t}^{\epsilon_t}$ is reduced we will induct on t with obvious statement for $t = 0$. By Remark 4.2, we cannot have $\epsilon_{i-1} = 1$ and $\epsilon_i = -1$ for any i. Thus, we may assume that $\epsilon_1 = \dots = \epsilon_l = -1$ for some $l \leq t$ and $\epsilon_{l+1} = \dots = \epsilon_t = 1$. If $1 \leq l < t$, then by Remark 4.2, the equality $w_0 = w_t$ implies $\phi_{i_1} = \phi_{j_t}$. It follows that $u_1 = u_{t-1}$. So the inductive hypothesis can be applied to the series $u_1, \dots, u_{t-1}$.

It remains to assume that $\epsilon_1 = \dots = \epsilon_t = 1$. In this case $t = 0$, since for $t > 0$, we have $|u_0| < |u_1| < \dots < |u_t|$, a contradiction. □

Lemma 4.4. *Let* $w_0, \dots, w_t$ *be words representing elements of the free group* $F(a,b)$*, where for every* $i = 1, \dots, t$*, we have either* $w_i = \phi_j(w_{i-1})$ *or* $w_{i-1} = \phi_j(w_i)$ *for some* $j = j(i)$*. Then there is a sequence of positive words* $u_0, \dots, u_t$ *such that*

(1) $u_i = \phi_j(u_{i-1})$ *or, respectively,* $u_{i-1} = \phi_j(u_i)$ *for the same* ϕ_j *as above,* $i = 1, \dots, t$*;*

(2) $u_l = 1$ *(empty word) for some* $l \leq t$*, and every* u_i *is an image of* 1 *under a product of some* ψ_j*-s;*

(3) if $w_0 = w_t$*, then* $u_0 = u_t$*.*

Proof. We will induct on the length t of the sequence with obvious base $t = 0$.

Assume $t > 0$ and for some i we have $\phi_l(w_{i+1}) = w_i = \phi_j(w_{i-1})$. Then $l = j$ by Remark 4.2, and $w_{i+1} = w_{i-1}$ in $F(a,b)$. Hence one can apply the inductive conjecture to the sequence $w_0, \dots, w_{i-1}, w_{i+2} \dots w_t$ of length $t-2$, obtain a sequence $u_0, \dots, u_{i-1}, u_{i+2} \dots u_t$ with conditions (1)-(3), and define $u_i = \phi_j(u_{i-1})$, $u_{i+1} = u_{i-1}$. obtaining the required sequence of length t.

Therefore one may assume that for some $l \in [0;t]$, the mappings act as follows: $w_l \mapsto \dots \mapsto w_0$ and $w_l \mapsto \dots \mapsto w_t$. Then we define $u_l = 1$, and by induction $u_{i+1} = \phi_{j_{i+1}}(u_i)$ for $i \geq l$ and $u_{i-1} = \phi_{j_i}(u_i)$ for $i \geq l$. Clearly, we obtain Properties (1) and (2). It remains to prove Property (3).

Suppose $1 \leq l < t$. If $w_0 = \phi_j(w_1)$ and $w_t = \phi_{j'}(w_{t-1})$, then $j = j'$ by Remark 4.2. Then the equality $w_0 = w_t$ implies $w_1 = w_{t-1}$ in $F(a,b)$. Since we may assume by induction that $u_1 = u_{t-1}$, we obtain $u_0 = \phi_j(u_1) = \phi_j(u_{t-1}) = u_1$, as required.

It remains to assume that all the mappings are arranged as $w_0 \mapsto \ldots \mapsto w_t$. Then $|w_0| < |w_1| < \ldots < |w_t|$, contrary to the assumption $w_0 = w_t$. Thus, Property (3) is proved. □

Lemma 4.5. *Let $w_0, \ldots, w_t$ be words representing elements of the free group $F(a,b)$, where for every $i = 1, \ldots, t$, we have either $w_i = \phi_j(w_{i-1})$ or $w_{i-1} = \phi_j(w_i)$ for some $j = j(i)$. Then either w_0 and w_t are equal in $F(a,b)$ or these two words represent different elements of $B_n(a,b)$.*

Proof. Assume that for some i, we have $\phi_l(w_{i+1}) = w_i = \phi_j(w_{i-1})$. Then $l = j$ by Remark 4.2 and $w_{i+1} = w_{i-1}$ in $F(a,b)$. Hence inducting on t, one can replace t with $t-2$.

Thus, we may assume that for some $l \in [0;t]$, the mappings are directed as follows: $w_l \mapsto \ldots \mapsto w_0$ and $w_l \mapsto \ldots \mapsto w_t$. Suppose $1 \leq l < t$. Then $w_0 = \phi_j(w_1)$ and $w_t = \phi_j(w_{t-1})$ for the same ϕ_j by Remark 4.2. Then the equality $w_1 = w_{t-1}$ implies $w_0 = w_t$ in $F(a,b)$ and $w_1 \neq w_{t-1}$ implies $w_0 \neq w_t$ in $B_n(a,b)$ since the mapping ϕ_j is injective on $B_n(a,b)$. Hence to complete the proof, one can consider a shorter sequence $w_1, \ldots, w_{t-1}$ again.

It remains to assume that all the mappings are directed as $w_0 \mapsto \ldots \mapsto w_t$. Under this assumption, we will prove that $w_0 \neq w_t$ in $B_n(a,b)$ if $t > 0$. Proving by contradiction, we may assume that we have a counter-example with the shortest first word w_0. It follows that the word w_0 is minimal in $B_n(a,b)$, i.e. it is not equal to a shorter word. Indeed if $w_0 = v_0 v_0'$, where $|v_0| < |w_0|$ and v_0' is trivial in $B_n(a,b)$, then applying the first mapping ϕ_{j_1} we would obtain $v_0 v_0' \to \mathbf{w}_{j_1} \psi(v_0)_{j_1} \psi_{j_1}(v_0')$, which is equal to $\mathbf{w}_{j_1} \psi_{j_1}(v_0)$ in $B_n(a,b)$ since ψ_{j_1} is a homomorphism of $B_n(a,b)$. Therefore $\phi_{j_1}(w_0) = \phi_{j_1}(v_0)$, and by induction, $\phi_{j_t}(\ldots(\phi_{j_1}(w_0))) = \phi_{j_t}(\ldots(\phi_{j_1}(v_0)))$ in $B_n(a,b)$, and v_0 would begin a shorter counter-example.

Let $w_t = \phi_j(w_{t-1})$. If w_0 is also a ϕ_j-image in $F(a,b)$, i.e. $w_0 = \phi_j(w)$ for some word w, then we have a sequence $w \mapsto w_0 \mapsto \ldots \mapsto w_{t-1}$ of the same length but with $|w| < |w_0|$ and $w = w_{t-1}$ in $B_n(a,b)$ by Remark 4.2, a contradiction. Hence we may assume further that w_0 is not a ϕ_j-image in $F(a,b)$.

The equality $w_0 = w_t$ in $B_n(a,b)$ provides us with a g-reduced van Kampen diagram Δ over the presentation of $B_n(a,b)$ defined in Section 3.2 (see Remark 3.5) with boundary path $\mathbf{pq}$, where $Lab(\mathbf{p}) \equiv w_0$, $Lab(\mathbf{q}) \equiv w_t^{-1}$. w_t is a ϕ_j-image, there is a subdiagram Γ of minimal type with contour $\mathbf{pr}$, where $Lab(\mathbf{r})^{-1}$ is a ϕ_j-image in $F(a,b)$. Γ has at least one cell of positive rank, because w_0 is not a ϕ_j-image.

By Lemma 2.3, there is a cell Π in Γ such that the sum of degrees of contiguity of Π to $\mathbf{p}$ and to $\mathbf{r}$ is greater than $\bar{\gamma}$. However the degree of contiguity of Π to the geodesic path $\mathbf{p}$ cannot exceed $\bar{\alpha}$ by Remark 2.1 and Lemma 2.2. Hence the degree of contiguity of Π to $\mathbf{r}$ is $> \bar{\gamma} - \bar{\alpha} > 1/3$. By Lemma 2.4, we have a cell π of positive rank in Γ with immediate contiguity to $\mathbf{r}$ of degree $\geq \varepsilon$ (see Figure 5).

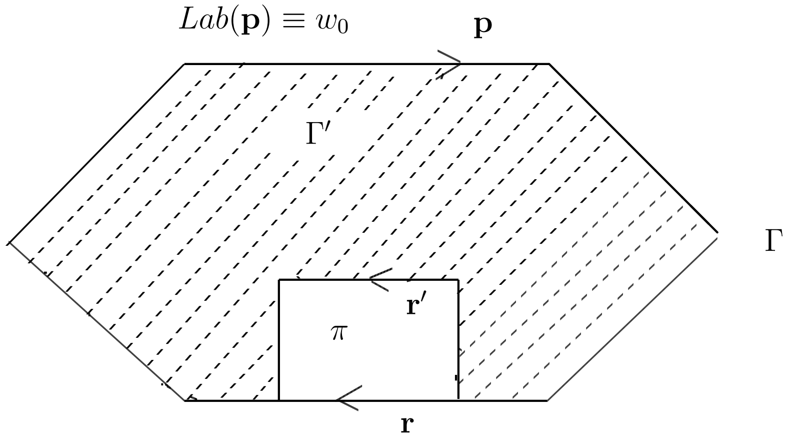

Figure 5

Since (1) the boundary label of π is a power A^n for some cyclically reduced word A, (2) the label of $\mathbf{r}^{-1}$ is a reduced form of a word from the coset $K(j) = \mathbf{w}_j H(j)$ of $H(j)$, and (3) $\varepsilon n > 30$, Properties (A),(B) of the set $\mathbf{W}$ imply that A^{15} is a subword of a word from the subgroup $H(j)$, and A is freely conjugate to a word from the subgroup $H(j)$ of $F(a,b)$, i.e. a cyclic permutation A' of A belongs to $H(j)$. Property (A) of the set $\mathbf{W}$ guarantees that a subpath $\mathbf{x}$ of $\mathbf{r}$ labeled by $(A')^{\pm 14}$ can be replaced with the complement of $\mathbf{x}$ in $\partial\pi$ so that for the obtained path $\mathbf{r}'$, we again have $Lab(\mathbf{r}'^{-1}) \in K(j)$. However the type of the subdiagram Γ' with contour $\mathbf{pr}'$ is strictly less than the type of Γ. Thus, the lemma is proved by contradiction. □

Recall that a subgroup L of a group K is called *malnormal* if $xLx^{-1} \cap L = \{1\}$ for every $x \in K \backslash L$. This property is transitive, i.e., if M is a malnormal subgroup of L and L is malnormal in K, then M is malnormal in K.

Lemma 4.6. *The canonical image $\tilde{H}(j)$ of the subgroup $H(j) \leq F(a,b)$ in the group $B_n(a,b)$ is malnormal in $B_n(a,b)$.*

Proof. Assume we have an equality $z^{-1}uz = v$ in $B_n(a,b)$, where the words u and v represent elements of $H(j)$ with non-trivial images in $B_n(a,b)$. We claim that z represents an element from $\tilde{H}(j)$ too.

According to Remark 3.5, we have a diagram Δ over the presentation of $B_n(a,b)$ with four boundary sections labeled by z, u, z^{-1} and v, resp. Identifying two of them, we obtain an annular diagram Δ_0 with two boundary sections $\mathbf{p}_1$ and $\mathbf{p}_2$ labeled by u and v, where the vertex $o_1 = (\mathbf{p}_1)_- = (\mathbf{p}_1)_+$ is connected with the vertex $o_2 = (\mathbf{p}_2)_- = (\mathbf{p}_2)_+$ by a simple path $\mathbf{q}$ labeled by the word z. If Δ_0 is not a g-reduced diagram, then one can make reductions preserving the labels of $\mathbf{p}_1$ and $\mathbf{p}_2$ and changing $\mathbf{q}$ by $\mathbf{q}'$, where the labels of $\mathbf{q}$ and $\mathbf{q}'$ are equal modulo the relations of $B_n(a,b)$ (see Section 13 in [19]). So we may regard Δ_0 as a g-reduced diagram.

Now we will induct on the number k of positive cells in Δ_0. If $k=0$, then we have the conjugation of u and v in the free group $F(a,b)$. So the statement follows from Property (A) of the set $\mathbf{W}$.

If $k \ge 1$, then by Lemmas 2.3 (b) and 2.4, we get a positive cell π in Δ_0 with immediate conjugacy to $\mathbf{p}_1$ or to $\mathbf{p}_2$ of degree $\ge \varepsilon$, and as in the proof of Lemma 4.5, one can change $\mathbf{p}_1$ (or $\mathbf{p}_2$) by a homotopic path $\mathbf{p}_1'$ such that the word $Lab(\mathbf{p}_1')$ belongs to $H(j)$, it is nontrivial in $B_n(a,b)$, the annular diagram bounded by $\mathbf{p}_1'$ and $\mathbf{p}_2$ has $k-1$ positive cells, and the path $\mathbf{q}'$ connecting $(\mathbf{p}_1')_-$ and $\mathbf{p}_2$ is homotopic to $\mathbf{q}$. So by the inductive conjecture, z belongs to $\tilde{H}(j)$. □

5 Machines

5.1 Turing machines

We will use a model of *recognizing* Turing machine (TM) close to the model from [29]. A TM with k tapes and k heads is a tuple

$$M = \langle A, Y, Q, \Theta, \vec{s}_1, \vec{s}_0 \rangle$$

where A is the input alphabet, $Y = \sqcup_{i=1}^{k} Y_i$ is the tape alphabet, $Y_1 \supset A$, $Q = \sqcup_{i=1}^{k} Q_i$ is the set of states of the machine, Θ is a set of commands, $\vec{s}_1$ is the k-vector of start states, $\vec{s}_0$ is the k-vector of accept states. ($\sqcup$ denotes the disjoint union.) The sets Y, Q, Θ are finite.

We assume that the machine normally starts working with states of the heads forming the vector $\vec{s}_1$, with the head placed at the right end of each tape, and accepts if it reaches the state vector $\vec{s}_0$. In general, the machine can be turned on in any configuration and turned off at any time.

A *configuration* of tape number i of M is a word uqv where $q \in Q_i$ is the current state of the head, u is the word to the left of the head, and v is the word to the right of the head, u, v are words in the alphabet Y_i. A tape is *empty* if u, v are empty words.

A *configuration* U of the machine M is a word

$$\alpha_1 U_1 \omega_1 \alpha_2 U_2 \omega_2 \ldots \alpha_k U_k \omega_k$$

where U_i is the configuration of tape i, and the end-markers α_i, ω_i of the i-th tape are special separating symbols.

An *input configuration* $w(u)$ is a configuration, where all tapes, except for the first one, are empty, the configuration of the first tape (let us call it the *input tape*) is of the form uq, $q \in Q_1$, u is a word in the alphabet A, and the states form the start vector $\vec{s}_1$. The *accept configuration* is the configuration where the state vector is $\vec{s}_0$, the accept vector of the machine, and all tapes are empty.

To every $\theta \in \Theta$, there corresponds a *command* (marked by the same letter θ), i.e., a pair of sequences of words $[V_1, ..., V_k]$ and $[V'_1, ..., V'_k]$ such that for each $j \le k$, either both $V_j = uqv$ and $V'_j = u'q'v'$ are configurations of the tape number j, or $V_j = \alpha_j qv$ and $V'_j = \alpha_j q'v'$, or $V_j = uq\omega_j$ and $V'_j = u'q'\omega_j$, or $V_j = \alpha_j q\omega_j$ and $V'_j = \alpha_j q'\omega_j$ $(q, q' \in Q_j\ )$.

In order to execute this command, the machine checks if every V_i is a subword of the current configuration of the machine, and if this condition holds the machine replaces V_i by V'_i for all $i = 1, \dots, k$. Therefore we also use the notation: $\theta : [V_1 \to V'_1, \dots, V_k \to V'_k]$, where $V_j \to V'_j$ is called the j-th *part* of the command θ.

Suppose we have a sequence of configurations $w_0, ..., w_t$ and a word $H = \theta_1 \dots \theta_t$ in the alphabet Θ, such that for every $i = 1, ..., t$, the machine passes from w_{i-1} to w_i by applying the command θ_i. Then the sequence $w_0 \to w_1 \to \dots \to w_t$ is said to be a *computation* with *history* H. In this case we shall write $w_0 \cdot H = w_t$. The number t will be called the *time* or the *length* of the computation.

A configuration w is called *accepted* by a machine M if there exists at least one computation which starts with w and ends with the accept configuration.

A word u in the input alphabet A is said to be *accepted* by the machine if the corresponding input configuration is accepted. (A configuration with the vector of states $\vec{s}_1$ is never accepted if it is not an input configuration.) The set of all accepted words over the alphabet A is called the *language* $\mathcal{L}_M$ *recognized by the machine* M.

We do not only consider *deterministic* Turing machines, for example, we allow several transitions with the same left side. A Turing machine M is called *symmetric* if with every command $\theta : [V_1 \to V'_1, \dots, V_k \to V'_k]$, it contains the *inverse* command $\theta^{-1} : [V'_1 \to V_1, \dots, V'_k \to V_k]$. Given a deterministic machine M, one can extend the set of commands by adding θ^{-1} for every $\theta \in \Theta$. The obtained (non-deterministic) Turing machine $Sym(M)$ is called the symmetrization of M.

A computation of a machine is called *reduced* if the history H of it admits no cancellations, i.e. H has no 2-letter subwords of the form $\theta\theta^{-1}$. Clearly, every computation can be made reduced (without changing the start or end configurations of the computation) by removing consecutive mutually inverse rules.

We will use the following additional properties of Turing machines; most of them were formulated in [2, 3, 21]

Lemma 5.1. *For every deterministic Turing machine M recognizing a language $\mathcal{L}$ there exists a Turing machine M_0 with the following properties.*

1. *The language recognized by M_0 is $\mathcal{L}$.*
2. *M_0 is symmetric.*
3. *If $\mathcal{C} : w_0 \to \dots \to w_t$ is a reduced computation of M_0 and $w_0 \equiv w_t$, then $t = 0$.*
4. *If a computation $\mathcal{C} : w_0 \to \dots \to w_t$ of M_0 with length $t > 0$ starts and ends with input configurations of M_0 then both w_0 and w_t are accepted by M_0.*
5. *The machine M_0 accepts only when all tapes are empty.*

6. *Every command of M_0 or its inverse has one of the following forms for some i*

$$[q_1\omega_1 \to q_1\omega_1, ..., q_{i-1}\omega_{i-1} \to q_{i-1}\omega_{i-1}, q_i\omega_i \to aq_i\omega_i, q_{i+1}\omega \to q_{i+1}\omega_{i+1}, ...] \quad (3)$$

$$[q_1\omega_1 \to q_1\omega_1, ..., q_{i-1}\omega_{i-1} \to q_{i-1}\omega_{i-1}, q_i\omega_i \to q_i'\omega_i, q_{i+1}\omega \to q_{i+1}\omega_{i+1}, ...] \quad (4)$$

$$[q_1\omega \to q_1\omega_1, ..., q_{i-1}\omega_{i-1} \to q_{i-1}\omega_{i-1}, \alpha_i q_i\omega_i \to \alpha_i q_i'\omega_i, q_{i+1}\omega_{i+1} \to q_{i+1}\omega_{i+1}, ...] \quad (5)$$

where a belongs to the tape alphabet of tape i, q_j, q_j' are state letters of tape j.

7. *The letters used on different tapes are from disjoint alphabets. This includes the state letters.*

Remark 5.2. *(1) Only one part (number i) of the commands (3, 4, 5) is changing.*

(2) If the head observes the right markers at the beginning of a computation, it will observe the right markers during the whole computation. If the head does not observe the right markers at the beginning, then no command is applicable, and so the computation is trivial.

Proof. Properties (1), (2), and (5) are provided by Lemma 2.3 [21] for the machine $M_0 = Sym(M)$.

The same lemma 2.3 [21] (Properties (b),(c)) says that the history of every reduced computation $\mathcal{C}$ of M_0 has the form $H = H_1H_2^{-1}$, where H_1 and H_2 are histories of M, and M can be chosen so that the same configuration cannot occur twice in a computation of M. Hence to prove Property (3) by contradiction, one may assume that both H_1 and H_2 are non-empty. Since $H_1 \neq H_2$, we have $H_1 = H_2H'$ (or $H_2 = H_1H'$) for some non-empty H' since the machine M is deterministic and the computations $\mathcal{C}$ and $\mathcal{C}^{-1}$ start with the same configuration. But then $w_0 \cdot H_2 \equiv (w_0 \cdot H_2)H'$, a contradiction, since H' is a history of M-computation.

To prove Property (4), we first modify the deterministic machine M and obtain a deterministic Turing machine M' working as follows.

After the initial command (or commands) changing the states of the heads only (but preserving tape words), the work of M' is subdivided in three Steps.

Step 1. M' copies the input word from the first tape to two auxiliary tapes with numbers $k+1$ and $k+2$ by deleting letter-by letter from the first tape and inserting the copies of erased letters to the last two tapes. Connecting rule is applied when the first tape is empty. It changes the states of all the heads.

Step 2. M' works as M but uses the $k+1$-st tape as the input one. It keeps the $k+2$-d tape unchanged. The copy of the accept command of M connects this Step with the next one. So the tapes $1, 2, \dots, k+1$ are empty when this command applies.

Step 3. M' letter-by-letter erases the content of the tape number $k+2$ and finally accepts.

Let $M'' = Sym(M')$. By Property (3), we have $w_0 \neq w_t$ and as there, the history is $H = H_1H_2^{-1}$, where H_1 and H_2 are histories of M'-computations. Here both H_1 and H_2 must contain the commands of both Steps 1 and 2 since $w_0 \neq w_t$ and the definition of

Step 1 imply that $w_0 \cdot H' \neq w_t \cdot H''$ for arbitrary histories without commands of Step 2. Furthermore, both H_1 and H_2 have to involve the commands of Step 3, because Step 2 does not change the (different) copies of the input words obtained on the tape number $k+2$. So by the definition of M', both w_0 and w_t are accepted words for M, and therefore for M' and M''.

Property (6) is obtained in Lemma 3.1 [29] by means of division of every tape in two parts, which does not affect (1) - (5), and so we get a modification M'''. However the analogs of (3, 5) are a bit weaker there, e.g. the analog of (3) in [29] allows to change the states of many heads:

$$[q_1\omega \to q_1'\omega, ..., q_{i-1}\omega \to q_{i-1}'\omega, aq_i\omega \to q_i'\omega, q_{i+1}\omega \to q_{i+1}'\omega, ...] \tag{6}$$

But this can be easily improved. For example, if one has $q_1 \to q_1'$ and $q_2 \to q_2'$ for a command θ, then it is possible to introduce a new state letter q and replace θ with three commands θ_1: $q_1 \to q$, $q_2 \to q_2$, θ_2: $q \to q, q_2 \to q_2'$, and θ_3: $q \to q_1', q_2' \to q_2'$ changing each only one state letter. Since the new letter q is involved in these three commands only, it is easy to see that the modified machine keeps the properties (1) - (5). The same trick works if one head inserts/deletes a letter and the same head (or another one) changes the state. Therefore Property (6) follows by induction.

To obtain (7), one can just prescribe the number of the tape to the tape and state letters as the additional index and change the commands of the machine accordingly. The machine M_0 is built. □

Remark 5.3. *(1) The Steps 1 - 3 of the Turing machines $M'' = Sym(M')$ constructed in the proof of Lemma 5.1 (4) (which are also the Steps of M_0) will be applied in the proof of Lemma 6.10.*

(2) The rewriting commands of Step 1 have the form

$$\theta_a : aq_1\omega_1 \to q_1\omega_1, \dots, \alpha_{k+1}q_{k+1} \to \alpha_{k+1}q_{k+1}a, \alpha_{k+2}q_{k+2} \to \alpha_{k+2}q_{k+2}a$$

for every $a \in A$. When we replace such a command with several commands $\theta_{a,1}, \dots, \theta_{a,m}$ using the trick from the proof of item (6) of Lemma 5.1, we introduce a number of auxiliary states of the heads, where every state is used for the transition from some $\theta_{a,i}$ to $\theta_{a,i+1}$ only and does not used in any other command.

Therefore if an M_0-computation $\mathcal{C} : w_0 \to \dots\dots \to w_t$ has no those auxiliary state letters in w_0 and in w_t, then every command $\theta_{a,i}$ can occur in the history H of $\mathcal{C}$ only in the block-subword $b = (\theta_{a,1} \dots \theta_{a,m})^{\pm 1}$. In particular, such a block has a unique command $\theta'(b)^{-1}$ (command $\theta''(b)$) of the form (3) that deletes the letter a from the first sector (resp., inserts the copies of a in sectors $k+1$ and $k+2$.

In cases (3), (4), and (5), we say that the head Q_i is *working* for θ, and so there is exactly one working head for every command of M_0.

5.2 S-machines (modified) as rewriting systems

There are several equivalent definitions of S-machines (see [30]). However we need a more general definition in comparison with that was used in [29], [24] or [25], and below the finite sets of letters $Y(\theta)$ are replaced with finitely generated subgroups of free groups. Besides, some parts of the rules θ having form $aqb \to cq'd$ cannot be written as $q \to (a^{-1}c)q'(db^{-1})$, because now the words a, b, c, d in the tape alphabet are not necessarily belong to $Y(\theta)$. In particular, some rules may now insert/delete words in locked sectors.

A "hardware" of an S-machine $\mathbf{S}$ is a pair (Y, Q), where $Q = \sqcup_{i=1}^{s} Q_i$ and $Y = \sqcup_{i=0}^{s} Y_i$ for some $s \geq 2$. We always set $Y_s = Y_0 = \emptyset$ and if $Q_s = Q_0$ (i.e., the indices of Q_i are counted modulo s), then we say that $\mathbf{S}$ is a *circular* S-machine.

The elements from Q are called *state letters* or *q-letters* of an S-machine, the elements from Y are *tape letters* or a-letters of an S-machine. The sets Q_i (resp. Y_i) are called *parts* of state and tape letters of Q (resp. Y). The number of a-letters (resp. q-letters, θ-letters) in a word W is called the a-length (resp., q-length, θ-length) of W denoted by $|W|_a$ ($|W|_q$, $|W|_\theta$, resp.).

The language of *admissible words* of an S-machine $\mathbf{S}$ (the language of $\mathbf{S}$-admissible words) consists of reduced words W of the form

$$q_1 u_1 q_2 \dots u_{r-1} q_r, \tag{7}$$

where $r \geq 1$, every q_i is a state letter from some part $Q_{j(i)}^{\pm 1}$, u_i are reduced group words in the alphabet of tape letters of the part $Y_{k(i)}$ and for every $i = 1, ..., r-1$, one of the following holds:

- If q_i is from $Q_{j(i)}$ then q_{i+1} is either from $Q_{j(i)+1}$ or is equal to q_i^{-1} and $k(i) = j(i)$.
- If $q_i \in Q_{j(i)}^{-1}$ then q_{i+1} is either from $Q_{j(i)-1}^{-1}$ or is equal to q_i^{-1} and $k(i) = j(i) - 1$.

Every subword $q_i u_i q_{i+1}$ of an admissible word will be called the $Q_{j(i)}^{\pm 1} Q_{j(i+1)}^{\pm 1}$-*sector* of that word. An admissible word may contain many $Q_{j(i)}^{\pm 1} Q_{j(i+1)}^{\pm 1}$-sectors.

Usually parts of the set Q of state letters are denoted by capital letters. For example, a part P would consist of letters p with various indices.

If an admissible word W has the form (7), $W = q_1 u_1 q_2 u_2 ... q_r$, and $q_i \in Q_{j(i)}^{\pm 1}$, $i = 1, ..., r$, u_i are group words in tape letters, then we shall say that the *base* of an admissible word W is the word $Q_{j(1)}^{\pm 1} Q_{j(2)}^{\pm 1} ... Q_{j(r)}^{\pm 1}$. Here Q_i are just symbols which denote the corresponding parts of the set of state letters. Note that, by the definition of admissible words, the base is not necessarily a reduced word.

Instead of saying that the parts of the set of state letters of $\mathbf{S}$ are $Q_1, Q_2, ..., Q_s$ we will write that *the standard base* of the S-machine is $Q_1 ... Q_s$.

The software of an S-machine with the standard base $Q_1 ... Q_s$ is a set of *rules* Θ. Every $\theta \in \Theta$ is given by

(1) a family $\{Y_i(\theta)\}_{i=0}^{s}$ of finitely generated subgroups $Y_i(\theta) \leq F(Y_i)$ of free groups and

(2) a sequence

$$[a_1q_1b_1 \to a'_1q'_1b'_1, ..., a_sq_sb_s \to a'_sq'_sb'_s],$$

where $q_i \in Q_i$, a_i, a'_i are reduced words from the group $F(Y_{i-1})$, b_i, b'_i are reduced words from $F(Y_i)$ (recall that $Y_0 = Y_s = \emptyset$) for every i.

Each component $a_iq_ib_i \to a'_iq'_ib_i$ is called a *part* of the rule. In most cases it will be clear what the sets $Y_i(\theta)$ are. By default we assume that $Y_i(\theta)$ is the whole free group $F(Y_i)$.

Every rule $\theta = [a_1q_1b_1 \to a'_1q'_1b'_1, ..., a_sq_sb_s \to a'_sq'_sb'_s]$ has the inverse one

$$\theta^{-1} = [a'_1q'_1b'_1 \to a_1^{-1}q_1b_1^{-1}, ..., a'_sq'_sb'_s \to a_s^{-1}q_sb_s]$$

It is always the case that $Y_i(\theta^{-1}) = Y_i(\theta)$.

Thus the set of rules Θ of an S-machine is divided into two disjoint parts, Θ^- Θ^+ and Θ^- such that for every $\theta \in \Theta^+$, we have $\theta^{-1} \in \Theta^-$ and for every $\theta \in \Theta^-$, we have $\theta^{-1} \in \Theta^+$ (in particular $\Theta^{-1} = \Theta$, that is any S-machine is symmetric).

The rules from Θ^+ (resp. Θ^-) are called *positive* (resp. *negative*).

To apply a rule $\theta = [a_1q_1b_1 \to a'_1q'_1b'_1, ..., a_sq_sb_s \to a'_sq'_sb'_s]$ to an admissible word $p_1u_1p_2u_2...p_r$ (7) where each $p_i \in Q^{\pm 1}_{j(i)}$, means

- to check for every subword $p_iu_ip_{i+1}$ if the following word belongs to the subgroup $Y_{j(i)}(\theta)$:

  (a) $b^{-1}_{j(i)}u_ia^{-1}_{j(i)+1}$, when $p_i \in Q_{j(i)}$ and $p_{i+1} \in Q_{j(i)+1}$,

  (b) $a_{j(i)+1}u_ia^{-1}_{j(i)+1}$, when $p_{i+1} \in Q_{j(i)+1}$ and $p_i = p^{-1}_{i+1}$,

  (c) $b^{-1}_{j(i)}u_ib_{j(i)}$, when $p_i \in Q_{j(i)}$ and $p_{i+1} = p_i^{-1}$,

  and if this property holds,

- replace each $p_i = q^{\pm 1}_{j(i)}$ by $(a^{-1}_{j(i)}a'_{j(i)}q'_{j(i)}b'_{j(i)}b^{-1}_{j(i)})^{\pm 1}$

- if the resulting word is not reduced or starts (ends) with a-letters, then reduce the word and trim the first and last a-letters to obtain an admissible word again.

It follows from the definitions that if W' is obtained by applying a rule θ (we write $W' = W \cdot \theta$), then the inverse rule is applicable to W' and $W' \cdot \theta^{-1} = W$.

Remark 5.4. *The application of every rule θ of* **S** *is effective since the membership problem is decidable for any finitely generated subgroup of a free group.*

Remark 5.5. *If $a, b, a', b' \in Y_i(\theta)$, then the component $aq_ib \to a'q'_ib'$ of a rule θ can be written in the equivalent form $q_i \to cq'_id$, where c and d are reduced forms of the products $a'a^{-1}$ and $b'b^{-1}$, respectively.*

For example, applying the rule $[q_1 \to a^{-1}q_1'b, q_2 \to cq_2'd]$ to the admissible word $q_1b^{-1}q_2dq_2^{-1}q_1^{-1}$ (where a, b, c, d are just letters) we first obtain the word

$$a^{-1}q_1'bb^{-1}cq_2'ddd^{-1}(q_2')^{-1}c^{-1}b^{-1}(q_1')^{-1}a,$$

then after trimming and reducing we obtain

$$q_1'cq_2'd(q_2')^{-1}c^{-1}b^{-1}(q_1')^{-1}$$

If a rule θ is applicable to an admissible word W, we say that W belongs to the *domain* of θ. In other words, W is *θ-admissible.*

We call an admissible word with the standard base a *configuration* of an S-machine.

We usually (but not always) assume that every part Q_i of the set of state letters contains a *start state letter* and an *end state letter.* Then a configuration is called a *start* (*end*) configuration if all state letters in it are start (end) letters. As Turing machines, some S-machines are *recognizing a language.* In that case we choose an *input* sector, say, the Q_1Q_2-sector, of every configuration. The a-projection of that sector (i.e. the word obtaining by deleting of q-letters) is called the *input* of the configuration. In that case, the end configuration with empty a-projection is called the *accept configuration.*

A *computation* of *length* $t \geq 0$ is a sequence of admissible words $W_0 \to \ldots \to W_t$ such that for every $0 = 1, ..., t-1$ the S-machine passes from W_i to W_{i+1} by applying one of the rules θ_i from Θ. The word $h = \theta_1 \ldots \theta_t$ is called the *history* of the computation. Since W_t is determined by W_0 and the history h, we use notation $W_t = W_0 \cdot h$ and say that W_0 belongs to the domain of h.

If by means a computation, the S-machine can take an input configuration with input u to the accept configuration, we say that the word u is *accepted* by the S-machine. We define *accepted* configurations (not necessarily start configurations) similarly. All accepted words u form the *language recognized by the S-machine.*

A computation is called *reduced* if its history is a reduced word. Clearly, every computation can be made reduced (without changing the start or end configurations of the computation) by removing consecutive mutually inverse rules. The domain of a reduced form of a history h can be larger than the domain of h.

Lemma 5.6. *Suppose the base of an admissible word W is $Q_iQ_i^{-1}$ (resp., $Q_i^{-1}Q_i$). Assume that each rule θ of a computation $W \equiv quq^{-1} \to \ldots \to W' \equiv q'u'(q')^{-1}$ (resp., $W \equiv q^{-1}uq \to \ldots \to W' \equiv (q')^{-1}u'q'$) with history h has a component $a_{\theta,i}q_ib_{\theta,i}^{-1} \to a'_{\theta,i}q_i'(b'_{\theta,i})^{-1}$, Assume that the mapping $\theta \mapsto b_{\theta,i}^{-1}b'_{\theta,i}$ ($\theta \mapsto a_{\theta,i}^{-1}a'_{\theta,i}$) extends to a monomorphism λ of the free Burnside group $\mathcal{B}$ on θ-letters of the history to the free Burnside group $B = B_n(Y_{i+1})$ with basis Y_{i+1} (to $B = B(Y_i)$). Then u' is equal to u modulo the Burnside relations if and only if the λ-image of h belong to the centralizer of u in B.*

Proof. We consider the base $Q_iQ_i^{-1}$ only. By induction on t, we see that u' is equal to the word $v^{-1}uv$ in B, where $v = \lambda(h)$, which proves the lemma. □

If for some rule $\theta = [a_1q_1b_1 \to a_1'q_1'b_1', ..., a_sq_sb_s \to a_s'q_s'b_s']$ of an S-machine **S**, the subgroup $Y_i(\theta)$ is trivial then we say that $\theta^{\pm 1}$ *locks* the Q_iQ_{i+1}-sector. In that case we denote the i-th part of the rule as follows: $a_iq_i \stackrel{\ell}{\to} a_iq_i'$. Since the admissible words are reduced, the definition of rule application implies

Lemma 5.7. *If the i-th component of the rule θ has the form $v_iq_i \stackrel{\ell}{\to} v_i'q_i'$, then the base of any θ-admissible word cannot have subwords $Q_iQ_i^{-1}$ or $Q_{i+1}^{-1}Q_{i+1}$.*

□

Remark 5.8. *The above definition of S-machines can be compared with the definition of multi-tape Turing machines. The main differences are that the heads of Turing machines are near-sighted, but the heads of S-machines are farsighted : they do not "see" the nearest tape letters but know if the whole tape words belong to distinguished cosets associated with the subgroups $Y_i(\theta)$.*

Since S-machines are symmetric, for every computation $\mathcal{C} : W_0 \to W_1 \to \dots \to W_t$, there is an inverse one $\mathcal{C}^{-1} : W_t \to W_{t-1} \to \dots \to W_0$. An S-machine can work with words containing negative letters, and words with "non-standard" order of state letters. In particular, the mirror copy of the computation $\mathcal{C}$ is the computation $W_0^{-1} \to W_1^{-1} \to \dots \to W_t^{-1}$ with the same history. Many properties of computations obviously hold for their mirror copies.

5.3 Primitive machines

For a set of letters A, let A' and A'' be disjoint copies of A, the maps $a \mapsto a'$ and $a \mapsto a''$ identify A with A' and A'', resp. Let $\overrightarrow{Z} = \overrightarrow{Z}(A)$ and $\overleftarrow{Z} = \overleftarrow{Z}(A)$ be the S-machines with tape alphabet $A' \sqcup A''$, state alphabet $\{L\} \cup P \cup \{R\}$, where $P = \{p(1), p(2), p(3)\}$ and the following positive S-rules. For $\overleftarrow{Z}$ we have the rules

$$\xi_1(a) = [L \to L, p(1) \to (a')^{-1}p(1)a'', R \to R], \;\; a \in A$$

Comment: The head can move from right to left, replacing the word in alphabet A' by its copy in the alphabet A''. (However, the inverse rules can move the head from right to left. Also an application of the rule can insert/delete two letters, one letter from the left and one letter from the right of the p-letter.)

$$\xi_2 := [L \stackrel{\ell}{\to} L, p(1) \to p(2), R \to R]$$

Comment: When the head meets L, it turns into $p(2)$.

For $\overrightarrow{Z}$, we define the rules

$$\xi_3(a) = [L \to L, p(2) \to a'p(2)(a'')^{-1}, R \to R]$$

Comment: The head $p(2)$ can move from left to right, replacing the word in A'' by its copy in A'.

$$\xi_4 = [L \to L, p(2) \stackrel{\ell}{\to} p(3), R \to R]$$

Comment: When the head reaches the right end of the tape, it turns into $p(3)$.

Remark 5.9. *For every $a \in A$, $i = 1, 3$, it will be convenient to denote $\xi_i(a)^{-1}$ by $\xi_i(a^{-1})$. It is clear from the definition $\xi_i(a)$ that this does not lead to a confusion.*

Lemma 5.10. *Assume that $\mathcal{C}: W_0 \to ... \to W_t$ is a reduced computation of $\overrightarrow{Z}$ (of $\overleftarrow{Z}$) with history h. Suppose the base $B = B(\mathcal{C})$ of $\mathcal{C}$ contains LP or PR. Then $W_i \neq W_j$ if $0 \le i < j \le t$.*

Furthermore, if W_i contains the subword Lu_ip (subword pu_iR) for some tape word u_i and the corresponding subword of W_j is Lu_jp (resp., pu_jR) for the same p-letter p, then the history of the computation $W_i \to \ldots \to W_j$ is trivial in the free Burnside group of exponent n provided u_j is equal to u_i modulo the Burnside relations.

Proof. It suffices to assume that $B \equiv LP$ or $B \equiv PR$.

Let the machine $\overrightarrow{Z}$ start with a word $W_i \equiv Lu_ip(2)$, ends with $W_j \equiv Lu_jp(2)$, where u_i, u_j are words in A'-letters, and does not apply the rule ζ_4. Then u_j is freely equal to u_iu, where u is the copy $a'_{i+1}\ldots a'_j$ of the reduced history $h_{ij} \equiv \xi_3(a_{i+1})...\xi_3(a_j)$ of this computation, and so $W_j \neq W_i$. If $W_i \equiv Lu_ip(2)$ and $W_j \equiv Lu_jp(3)$, then W_i and W_j have just different state letters.

It is not possible that $W_k \equiv Lu_kp(3)$ for $0 < k < t$ since in this case both transitions $W_{k-1} \to W_k$ and $W_{k+1} \to W_k$ were the applications of the rule ζ_4, and the history h were not reduced. If $t = 2$ and both W_0 and W_2 contain $p(3)$ we have the same contradiction, since the history should be equal to $\zeta_4^{-1}\zeta_4$. If $t > 2$ and $W_0 \equiv Lu'_ip(3)$ and $W_t \equiv Lu_jp(3)$, then $u_j \neq u_i$ as in the previous paragraph since $W_1 \equiv Lu_ip(2)$ and $W_{t-1} \equiv Lu_jp(2)$ in this case.

Thus, the statement of the lemma is proved for the machine $\overrightarrow{Z}$ and the subbase LP. The proof of the Lemma is finished since other properties are similar. □

We also need two simple machines depending on the mappings ψ_j and ϕ_j defined in Section 4. Now $A = \{a, b\}$. The machine $\overleftarrow{Z}(\phi_j)$ replaces a word with the ϕ_j-image of it. As $\overleftarrow{Z} = \overleftarrow{Z}(A)$, the S-machines $\overleftarrow{Z}(\phi_j)$ has tape alphabet $A' \sqcup A''$, state alphabet $\{L\} \cup P \cup \{R\}$, where $P = \{p(1), p(2), p(3)\}$ and the following positive S-rules depending on ψ_j.

$$\xi'_1(c) = [L \to L, p(1) \to (c')^{-1}p(1)\mathbf{w}''_j(c), R \to R], \quad c \in A = \{a, b\},$$

where $\mathbf{w}''_j(c)$ is the copy of $\mathbf{w}_j(c)$ in the alphabet A'', and the finitely generated subgroup $A''(\xi'_1(c)) \le F(A'')$ corresponding to this rule is the the copy $H''(j)$ of the subgroup $H(j) \le F(A)$ in $F(A'')$, i.e. this rule cannot be applied unless the words in the PR-sectors (and in PP^{-1}-, $R^{-1}R$-sectors if any occur) belong to $H''(j)$.

$$\xi'_2 := [L \xrightarrow{\ell} L, p(1) \to \mathbf{w}'_jp(2), R \to R],$$

where $\mathbf{w}'_j$ is the copy of the word $\mathbf{w}_j$ in the alphabet A', $A'(\xi'_2) = \{1\}$ and $A''(\xi'_2) = H''(j)$. So the rule ξ'_2 (the rule $(\xi'_2)^{-1}$) is applicable only if the sector LP is empty (resp., if the tape word of this sector is $\mathbf{w}'_j$).

Comment: The difference in comparison with $\overleftarrow{Z} = \overleftarrow{Z}(A)$ is that the replacements $a' \mapsto \psi_j(a')$ given by the rules $\xi_1'(a)$ $(a \in A)$ use the monomorphism ψ_j, and the rule ξ_2' inserts the word $\mathbf{w}_j'$ into the LP-sector .

The machine $\overrightarrow{Z}(\phi_j)$ has the positive rules

$$\xi_3'(c) = [L\mathbf{w}_j' \to L\mathbf{w}_j', p(2)\mathbf{w}_j''(c) \to \mathbf{w}_j'(c)p(2), R \to R],$$

where $A'(\xi_3'(c)) = H'(j)$ and $A''(\xi_3') = H''(j)$.

The next rule is

$$\xi_4' = [L\mathbf{w}_j' \to L\mathbf{w}_j', p(2) \xrightarrow{\ell} p(3), R \to R]$$

Here $A'(\xi_4') = H'(j)$, and $A''(\xi_4') = \{1\}$.

Comment: The work of the machine $\overleftarrow{Z}(\phi_j)$ followed by the work of $\overrightarrow{Z}(\phi_j)$ replaces the configuration $Lup(1)R$ with $L\phi_j(u)p(3)R$.

We have the following analog of Lemma 5.10.

Lemma 5.11. *Suppose $\mathcal{C}: W_0 \to ... \to W_t$ is a reduced computation of $\overleftarrow{Z}(\phi_j)$ or of $\overrightarrow{Z}(\phi_j)$ for some j. Assume that the base $B = B(\mathcal{C})$ of $\mathcal{C}$ contains LP or PR. Then $W_i \ne W_j$ if $0 \le i < j \le t$.*

Furthermore, if W_i contains the subword Lu_ip (subword pu_iR) for a tape word u_i and the corresponding subword of W_j is Lu_jp (resp., pu_jR) for the same p-letter p, then the history of the computation $W_i \to \ldots \to W_j$ is trivial in the free Burnside group of exponent n provided u_j is equal to u_i modulo the Burnside relations.

Proof. It suffices to repeat the argument of the proof of Lemma 5.10 but instead of a non-trivial copy of the word $a_{i+1}' \ldots a_j'$, we now have a (copy of the) nontrivial ψ_j-image of such word for the machine $\overleftarrow{Z}(\phi_j)$ since the homomorphism ψ_j is injective by Remark 4.1. □

Lemma 5.12. *Let QQ' be a sector of the standard base of one of the machines $\overleftarrow{Z}$, $\overrightarrow{Z}$ $\overleftarrow{Z}(\phi_j)$, $\overrightarrow{Z}(\phi_j)$ and Y' or Y'' be the tape subalphabet of this sector. Assume that a rule η does not lock the sector QQ' and has a word xuy in the domain, where $x \in Q$, $y \in Q'$. Then for every rule ξ of the same machine having at least one word xvy in the domain, we have $Y'(\xi) \le Y'(\eta)$ (resp., $Y''(\xi) \le Y''(\eta)$).*

Proof. This property can be checked by the inspection of every rule. For instance, if $QQ' = LP$ and $\eta = \xi_1'(c)$, then ξ is either $\xi_1'(d)$ for a letter $d \in \{a^{\pm 1}, b^{\pm 1}\}$, or $\xi = \xi_2'$. In both cases the required inequality follows from the definition of these rules. □

Remark 5.13. *Let $\mathcal{C} : W_0 \to W_1 \to \ldots \to W_s$ be a computation of $\overleftarrow{Z}$ or of $\overrightarrow{Z}$, or of $\overrightarrow{Z}(\phi_j)$ with standard base, and u_l', v_l'' be the tape words of the LP- and PR-sectors of the word W_l, $l = 0, \ldots, s$. Let $\pi : F(A') \to F(A'')$ be an isomorphism, such that $\pi(a') = a''$ for every $a \in A$. Then it follows from the definition of Z-machines that the reduced forms of the words $\pi(u_j')v_j''$ are equal for all $l = 0, \ldots, s$.*

We will refer to this property as 'projection argument'. For the computations of the machine $\overleftarrow{Z}(\phi_j)$ having no rules $(\xi_2')^{\pm 1}$ in the history, it states that the reduced form of $\psi_j(u_l)v_l$ is unchanged in a computation with standard base, where u_l and v_l copy the words u_l' and v_l'' in the alphabet $A = \{a, b\}$.

Remark 5.14. *Assume that $W_0 \equiv Lup(1)vR$ and a reduced computation $\mathcal{C} : W_0 \to W_1 \to \dots \to W_s$ of $\overleftarrow{Z}$ starts with $W_0 \to W_1 = W_0 \cdot \xi_1(a)$, $a \in A^{\pm 1}$, where the last letter of u is not a' and the word v does not start with a''. Then we have $W_1 \equiv Lu(a')^{-1}p(1)a''vR$ with reduced subwords $u_1 \equiv u(a')^{-1}$ and $v_1 \equiv a''v$. Since u_1 and v_1 are non-empty, the word W_1 is in the domains of the rules $\xi_1(b)$ ($b \in A^{\pm 1}$) only, and $b \neq a^{-1}$ in the transition $W_1 \to W_2 = W_1 \cdot \xi_1(b)$ since $\mathcal{C}$ is a reduced computation. Hence the products $u_2 \equiv u_1(b')^{-1}$ and $v_2 \equiv b''v_1$ are reduced too, and so on. Therefore all the rules in the history of $\mathcal{C}$ are of the type $\xi_1(c)$ for $c \in A^{\pm 1}$ and both LP- and PR-sectors of W_s are non-empty. Under above assumption, we say that the application of the first rule $\xi_1(a)$ to W_0 is wrong.*

Similarly we define wrong applications of ξ_3-rules of $\overrightarrow{Z}$.

The definition of a wrong application of a rule $\xi_1'(c)$ of the machine $\overleftarrow{Z}(\phi_j)$ is similar, but now we assume that c' is not the last letter of u and the word v is a copy of the ψ_j-image of a reduced word which does not start with c^{-1}. The definition of a wrong application of a rule $\xi_3'(c)$ of the machine $\overrightarrow{Z}(\phi_j)$ is also similar: we assume that u (resp., v) is a a copy of ϕ_j-image (a copy of ψ_j-image) of a reduced word which does not end with c^{-1} (does not start with c).

Lemma 5.15. *Assume that for some rule ξ of $\overleftarrow{Z}$ or of $\overrightarrow{Z}$, or of $\overleftarrow{Z}(\phi_j)$, or of $\overleftarrow{Z}(\phi_j)$, the canonical image $\tilde{A}$ of the subgroup $A'(\xi)$ (or of the subgroup $A''(\xi)$) in $B_n(a', b')$ (resp., in $B_n(a'', b'')$) contains nontrivial in $\tilde{A}$ words u and $z^{-1}uz$ for some $z \in B_n(a', b')$ (resp. $z \in B_n(a'', b'')$). Then the word z belongs in $\tilde{A}$ too.*

Proof. We consider (non-trivial) $A'(\xi)$ only. The statement is obvious if $\tilde{A}$ is the whole $B_n(a', b')$. Otherwise $A'(\xi)$ is equal to a subgroup $H'(j)$, and therefore the statement follows from Lemma 4.6. □

6 Encoding of the rules of Turing machines by primitive machines

Let $M_0 = \langle A, Y(0), Q, \Theta, \vec{s}_1, \vec{s}_0 \rangle$ be a Turing machine satisfying the conditions (1) - (7) of Lemma 5.1 with the tape alphabet $Y(0) = \cup_{i=1}^{k} Y_i(0)$; it recognizes a language $\mathcal{L}$ in the input alphabet A. We will construct an S-machine M_1 that simulates the work of M_0 by applying the mappings ϕ_j instead of the commands of M_0.

Recall that M_0 is symmetric, and every rule θ has at most one active head inserting/deleting a letter or changing the state. Separating the commands of M_0 into two parts, we say that inserting commands are positive. If a command does not change tape words, then we may randomly choose the sign for this command and the opposite sign for the inverse command. We also may assume that there are no duplicate commands, i.e. every command of M_0 is uniquely determined by the set of parts of it, and there is no command $\theta = \theta^{-1}$.

If the tape subalphabet $Y_i(0)$ of M_0 has letters $a_{i1}, ..., a_{i,k_i}$, then we choose disjoint triples of different words $(\mathbf{w}_{i1}, \mathbf{w}_{i1}(a), \mathbf{w}_{i1}(b))..., (\mathbf{w}_{i,k_i}, \mathbf{w}_{i,k_i}(a), \mathbf{w}_{i,k_i}(b))$ in the alphabet $\{a(i), b(i)\}$ in accordance with Section 4, where the words are taken from the (infinite) set $\mathbf{W}$ with Properties (A)-(C), and for every a_{ij}, we define the endomorphisms

$$\psi_{ij} : a(i) \mapsto \mathbf{w}_{ij}(a), \quad b(i) \mapsto \mathbf{w}_{ij}(b)$$

of the groups $F(a(i), b(i))$ and of $B_n(a(i), b(i))$, and define the mappings ϕ_{ij} on these groups by the rule $\phi_{ij}(u) = \mathbf{w}_{ij}\psi_{ij}(u)$.

The tape alphabet of M_1 is $Y = \sqcup_{i=1}^k Y_i$, where $Y_i = Y_i' \sqcup Y_i''$ with $Y_i' = \{a'(i), b'(i)\}$ and $Y_i'' = \{a''(i), b''(i)\}$. The standard base of M_1 is

$$Q_0 P_1 Q_1 P_2 \dots P_k Q_k,$$

where $\sqcup_{i=1}^k Q_i$ is the set of states of M_0, $Q_0 = \{q_0\}$, and

$$P_i = \{p_i, p_{\theta,i}(1), p_{\theta,i}(2), p_{\theta,i}(3) \mid \theta \in \Theta^+, i = 1, \dots, k\}$$

For every positive command θ of M_1 we introduce a set of positive rules $\Xi^+(\theta)$ of M_1. The S-mashine $M(\theta)$ (part of M_1) with the set of rules $\Xi^+(\theta)$ works as follows.

If θ is given by formula (3), i.e. it has a unique working head Q_i with part $q_i\omega \to a_{ij}q_i\omega$, where $a_{ij} \in Y_i(0) = \{a_{i1}, \dots, a_{i,k_i}\}$, then $M(\theta)$ has the following set of rules $\Xi^+(\theta)$.

(1) $p_i \xrightarrow{\ell} p_{\theta,i}(1)$, $p_j \xrightarrow{\ell} p_j$ for $j \neq i$, $q_s \to q_s$ for every q_s from the left-hand side of (3), i.e., the p-letter of the working sector gets a θ-index.

Defining the rules of S-machines below we often omit their idling parts of the form $q \to q$.

(2) This gives start to the work of the copy of $\overleftarrow{Z}(\phi_{ij})$ working with the alphabet Y_i (instead of A) and having the standard base $Q_{i-1}P_iQ_i$ (instead of LPR). When the state letter $p_{\theta,i}(1)$ meets the left state letters q_{i-1}, the copy of the rule $\xi'(2)$ is applicable, i.e.

$$p_{\theta,i}(1) \to \mathbf{w}'_{ij} p_{\theta,i}(2), \;\; \textit{and also} \;\; p_{i-1} \xrightarrow{\ell} p_{\theta,i-1}(1), \tag{8}$$

which activates P_{i-1}-head (the head P_k if $i = 1$, i.e. the i-index is taken modulo k).

(3) The machine $\overleftarrow{Z}(\theta, i-1)$ (a copy of $\overleftarrow{Z}$ working on Y_{i-1}) is switched on now. When $p_{\theta,i-1}(1)$ meets q_{i-2}, a copy of the rule $\xi(2)$ applicable, and by definition, this rule should also switch on the machine $\overleftarrow{Z}(\theta, i-2)$ on the base $Q_{i-3}P_{i-2}Q_{i-2}$, i.e. we have similar work subsequently on the bases between Q_{i-1} and Q_i, Q_{i-2} and Q_{i-1}, … Q_0 and Q_1, Q_{k-1} and Q_k, … Q_i and Q_{i+1}. So machine $\overleftarrow{Z}(\phi_{ij})$ (for some j) works only between Q_{i-1} and Q_i, in other intervals, p-letters just subsequently run from right to left.

(4) The last rule changing $p_{\theta,i+1}(1)$ by $p_{\theta,i+1}(2)$ switches on the machine $\overrightarrow{Z}(\theta, i+1)$, the letter $p_{\theta,i+1}(2)$ runs until it reaches q_{i+1} and becomes $p_{\theta,i+1}(3)$, simultaneously giving start to

$\overrightarrow{Z}(\theta, i+2)$, whose work follows by the subsequent work of $\overrightarrow{Z}(\theta, i+3), \dots, \overrightarrow{Z}(\theta, i-1)$ (i-index is taken modulo k), $\overrightarrow{Z}(\phi_{ij})$, i.e. again, the copy of the machine $\overrightarrow{Z}(\phi_{ij})$ works between Q_{i-1} and Q_i only.

(5) Finally, the rule with parts $p_{\theta,i}(3) \to p_i$ ($i = 1, \dots, k$) erases the θ-index in every p-letter and makes possible the work of a machine $M(\theta')$ (but $M(\theta')$ cannot start working unless θ' can be applied for the states of heads $(q_1, \dots, q_k)$).

Comment. The described work of $M(\theta)$ with the standard base replaces the tape word u in the $Q_{i-1}P_i$-sector with $\phi'_{ij}(u)$ (where ϕ'_{ij} copies ϕ_{ij} in the alphabet Y'_i) if the command θ of M_0 inserts the letter a_{ij} from the left of q_i. Indeed, at Step (2) we can obtain the copy of $\phi'_{ij}(u)$ in the alphabet Y''_i, and the machine $\overrightarrow{Z}(\phi_{ij})$ rewrites it in the alphabet Y'_i at Step (4). The "useless" work of Z-machines in other sectors will be exploited for the control of computations, especially if the base is not standard.

If the command θ of M_0 has type (4), then the work of the machine $M(\theta)$ differs in comparison with the case (4) as follows. At Steps (2) and (4), the work of machines $\overleftarrow{Z}(\phi_{ij})$ and $\overrightarrow{Z}(\phi_{ij})$ is replaced with the work of $\overleftarrow{Z}(\theta, i)$ and $\overrightarrow{Z}(\theta, i)$, respectively; the rule changing $p_{\theta,i+1}(1)$ by $p_{\theta,i+1}(2)$ also changes q_i by q'_i.

Comment The described work of $M(\theta)$ just replaces q_i with q'_i.

If the command θ of M_0 has type (5), then the machine $M(\theta)$ works as follows.

(1') $q_{i-1} \xrightarrow{\ell} q_{i-1}$, $p_i \xrightarrow{\ell} p_{\theta,i}(2)$, $p_{i-1} \xrightarrow{\ell} p_{\theta,i-1}(1)$,

Comment This rule checks that there are no letters between q_{i-1} and p_i and immediately switches on the machine $\overleftarrow{Z}_{\theta,i-1}$.

(2') Then as above the machines $\overleftarrow{Z}_{\theta,i-1},\dots, \overleftarrow{Z}_{\theta,1}, \overleftarrow{Z}_{\theta,k},\dots,\overleftarrow{Z}_{\theta,i+1}$ work subsequently until the rule $q_i \xrightarrow{\ell} q'_i$, $p_{\theta,i+1}(1) \to p_{\theta,i+1}(2)$ completes moving of the p-letters to the left and changes the Q_i-letter.

(3') The obtained set of state letters makes possible the work of the machine $\overrightarrow{Z}_{\theta,i+1}$. In turn, this machine will switches on the machine $\overrightarrow{Z}_{\theta,i+2}$ after the rule $p_{\theta,i}(2) \xrightarrow{\ell} p_{\theta,i}(3)$, and so on. Finally, the rule $p_{\theta,i}(2) \xrightarrow{\ell} p_{\theta,i}(3)$ follows by the rule $p_{\theta,j}(3) \xrightarrow{\ell} p_j$, where the latter one is applicable simultaneously for all $j = 1, \dots, k$.

Definition 6.1. *For every positive word u in the alphabet $Y_i(0)$ of M_0, we define the word $f_i(u)$ in the alphabet $Y'_i = \{a'(i), b'(i)\}$. If $u = 1$ (i.e. u is empty), then we define $f_i(u) = 1$, and by induction, if $u \equiv va_{ij}$ for some $a_{ij} \in Y_i(0)$, then*

$$f_i(u) = \phi'_{ij}(f_i(v)) = \mathbf{w}'_{ij}\psi'_{ij}(f_i(v)),$$

where $\mathbf{w}'_{ij}, \psi'_{ij}, \phi'_{ij}$ copy $\mathbf{w}_{ij}, \psi_{ij}, \phi_{ij}$, resp., in the alphabet Y'_i.

Lemma 6.2. *If $f_i(w) = \phi_{ij}(u)$ for some words w and u, then u is a positive word and the last letter of w is a_{ij}.*

Proof. It follows from Property (A) of the set of positive words $\mathbf{W}$ that ϕ_{ij}-images of words u are positive for positive words u only. The second statement is true by the definition of $f_i(v)$ and Remark 4.2. □

For every configuration $\mathcal{W} = \alpha_1 u_1 q_1 \omega_1 \dots \alpha_k u_k q_k \omega_k$ of the Turing machine M_0, we define the configuration $W = F(\mathcal{W}) \equiv q_0 f_1(u_1) p_1 q_1 \dots q_{k-1} f_k(u_k) p_k q_k$ of the S-machine M_1. The start (stop) configuration of M_1 is the F-image of the start (stop) configuration of M_0.

Lemma 6.3. *If $\mathcal{W}_0 \to \dots \to \mathcal{W}_s$ is a computation of M_0, then there is a computation $F(\mathcal{W}_0) \to \dots \to F(\mathcal{W}_1) \to \dots \to F(\mathcal{W}_{s-1}) \to \dots \to F(\mathcal{W}_s)$ of M_1.*

Proof. Arguing by induction, we may assume that $s = 1$. Let $\mathcal{W}_1 = \mathcal{W}_0 \cdot \theta$. If $\theta = \theta_1$ is a positive command of type (3), and so it changes the i-th sector inserting a letter a_{ij}: $\alpha_i u_i q_i \omega_i \to \alpha_i u_i a_{ij} q_i \omega_i$, then the work of the S-machine $M(\theta)$ defined above can transform the configuration $F(\mathcal{W}_0) = \dots q_{i-1} f(u_i) p_i q_i \dots$ to

$$\dots q_{i-1} \mathbf{w}'_{ij} \psi'_{ij}(f_i(u_i)) p_i q_i \dots = \dots q_{i-1} \phi'_{ij}(f_i(u_i)) p_i q_i \dots = \dots q_{i-1} f_i(u_i a_{ij}) p_i q_i \dots = F(\mathcal{W}_1),$$

as required. If θ^{-1} is a positive command of type (3), then the application of θ is possible only if the last letter of u_i is a_{ij}, that is $u_i \equiv v_i a_{ij}$ for some word v_i. Therefore there is a work of $M(\theta^{-1})$ transforming $F(\mathcal{W}_1) = \dots q_{i-1} f_i(v_i) p_i q_i \dots$ to $F(\mathcal{W}_0)$. The inverse computation takes $F(\mathcal{W}_0)$ to $F(\mathcal{W}_1)$, as required.

If θ has type (4), then $\mathcal{W}_1$ is obtained from $\mathcal{W}_0$ by replacement of q_i with q'_i (or vice versa). Therefore the same is true for the words $F(\mathcal{W}_0)$ and $F(\mathcal{W}_1)$, and the machine $M(\theta)$ can execute the similar replacement as well.

If θ is of type (5), the statement is also clear since the canonical work of $M(\theta)$ is possible because the word u_i is empty and so $f_i(u_i) = 1$. The machine $M(\theta)$ just replaces q_i with q'_i or vice versa. □

Lemma 6.4. *Let $\mathcal{C} : W_0 \to \dots \to W_s$ $(s > 0)$ be a reduced computation of $M(\theta)$ with standard base for some positive command θ of M_0. Suppose the state letters of the configurations W_0 and W_s have no θ-indices while other configurations have θ-indices. Then the computation $\mathcal{C}$ is unique, i.e. there is no other reduced computation of $M(\theta)$ starting with W_0 and ending with W_s. If Q_i is the working head for θ, then the only $Q_{i-1}P_i$- and P_iQ_i-sectors of W_0 and W_s are different, namely,*

(1) for a command θ given by (3), these sectors of W_0 are $q_{i-1} u p_i q_i$ for some reduced word u in the alphabet $\{a'(i), b'(i)\}$ (and some $q_{i-1} \in Q_{i-1}, q_i \in Q_i$), and W_s has the sectors $q_{i-1} \phi'_{ij}(u) p_i q_i$, or vice versa;

(2) for a command θ given by (4), these sectors of W_0 are $q_{i-1} u p_i q_i$ for some reduced word u in the alphabet $\{a'(i), b'(i)\}$, and W_s has the sectors $q_{i-1} u p_i q'_i$, or vice versa;

(3) for a command θ given by (5), the sectors of W_0 are $q_{i-1} p_i q_i$ and W_s has the sectors $q_{i-1} p_i q'_i$, or vice versa.

In particular, if $W_0 = F(\mathcal{W}_0)$ for some $\mathcal{W}_0$, then $W_s = F(\mathcal{W}_0 \cdot \theta^{\pm 1})$.

Proof. Assume that the command θ is given by (3). The history of $\mathcal{C}$ can start from the rule $p_i \xrightarrow{\ell} p_{\theta,i}(1)$ ((1) in the definition of $M(\theta)$) or with the inverse rule with $p_{\theta,i}(3) \xrightarrow{\ell} p_i$ ($i = 1, \dots, k$, (5) in the definition of $M(\theta)$) since the state letters of W_0 have no θ-indices. We consider the first case only.

Since $\mathcal{C}$ is reduced, the next rule has to be the rule of the form $\xi_1'(c)$ of a machine $\overleftarrow{Z}(\phi_{ij})$ for some $j \in \{1, 2\}$. Here $c' = c(ij)$ is the last letter $a(i)^{\pm 1}$ or $b(i)^{\pm 1}$ of u since otherwise the p-letter would never loose its θ-index by Remark 5.14. Hence this rule moves the letter $p_{\theta,i}(1)$ left replacing c' with the copy of the word $\mathbf{w}_{ij}(c)^{\pm 1}$ in the alphabet Y_i''. For the same reason the subsequent rules move it until it meets q_{i-1}. Remark 5.14 now implies that no rules of type $\xi_1'(c)$ are applicable anymore. Hence the next rule has to be of type (8). It changes $p_{\theta,i}(1)$ by $\mathbf{w}_{ij}' p_{\theta,i}(2)$ and switches on the machine $\overleftarrow{Z}(\theta, i-1)$. Then the rules of type $\xi_1(c)$ of $\overleftarrow{Z}(\theta, i-1)$ move left the P_{i-1}-head until it meets q_{i-2} and turns into $p_{\theta,i-1}(2)$ switching on the machine $\overleftarrow{Z}(\theta, i-2)$, and so on, until $\overrightarrow{Z}(\theta, i-1)$ completes work, switches on the machine $\overrightarrow{Z}(\phi_{ij})$ and the final rule executes the transformation $p_{\theta,j}(3) \to p_j$ for all $j = 1, \dots, k$.

Note that the rules of $\overleftarrow{Z}(\phi_{ij})$ letter-by-letter replace the word u with the ψ_{ij}''-image of it and insert $\mathbf{w}_{ij}'$ from the left of the P-head, while the rules of $\overrightarrow{Z}(\phi_{ij})$ just replace subwords v'' from the subgroup $H''(ij)$ by the copies $v' \in H'(ij)$, where the subgroup $H(ij)$ is generated by $\mathbf{w}_{ij}(a)$ and $\mathbf{w}_{ij}(b)$. In other sectors, the words have been replaced with copies twice, and so Property (1), of the lemma follows.

If θ has type (4), then the proof can be obviously modified. If θ has type (5), the above argument works as well, because the canonical work of $M(\theta)$ verifies that the $Q_{i-1}P_i$-sector has no tape letters.

The last statement follows from the previous one and the inductive definition of the functions f_i and F. □

Remark 6.5. *By Lemma 6.4, the computation $\mathcal{C}$ starting with W_0, ending with W_s and satisfying the assumptions of this lemma is unique, but it can start either from the rule of one of the types (1) and (1') or from a rule of one of the types* (5) *and (3') (as in the definition of $M(\theta)$ for a positive rule θ). In the former case, we say that $\mathcal{C}$ is the canonical computation of $M(\theta)$, and in the later case, $\mathcal{C}$ is the canonical computation of $M(\theta^{-1})$. Thus, for any θ (positive or negative), the canonical computation of $M(\theta)$ is determined by the command θ and the first admissible word W_0.*

Lemma 6.6. *A computation $\mathcal{C} : W_0 \to \dots \to W_s$ of the machine $M(\theta)$ with a base $Q_{i-1}P_iQ_i$ has the following property. If the state letters of the words W_0 and W_s have no θ-indices but other words W_k have θ-indices, then the $Q_{i-1}P_i$-sector words w_0 and w_s of W_0 and W_s satisfy one of the conditions: either $w_s \equiv w_0$ or $w_s = \phi_{ij}'(w_0)$, or $w_0 = \phi_{ij}'(w_s)$ for some mapping ϕ_{ij}'.*

Proof. If the working head of the command θ of M_0 is $Q_{i'}$ with $i' \neq i$, then only rules of $\overleftarrow{Z}(\theta, i)$ and $\overrightarrow{Z}(\theta, i)$ can change the $Q_{i-1}P_i$- and P_iQ_i-sectors, and $w_s \equiv w_0$ by the projection argument (Remark 5.13) since the sector P_iQ_i is locked by the rules having no θ-indices.

Assume now that Q_i is the working command of M_0 and, as in the proof of Lemma 6.4, assume that we have the part $p_i \stackrel{\ell}{\to} p_{\theta,i}$ in the first rule of $\mathcal{C}$. Then as there, we see that the machine $\overleftarrow{Z}(\phi_{ij})$ has to start and complete its work switching on the machine $\overleftarrow{Z}(\theta, i-1)$, and so on. However we have no P_{i-1} (and many other heads) in the base now. So after an application of a number of rules of $\overleftarrow{Z}(\theta, i-1)$, which change nothing since the base is $Q_{i-1}P_iQ_i$, the machine $\overleftarrow{Z}(\phi_{ij})$ can be switched on again. (Also such a return work can happen after an idling work of $Z(\theta, i-2)$ and return work of $Z(\theta, i-1)$, and so on.)

So the computation $\mathcal{C}$ can finish either with the work of $\overrightarrow{Z}(\phi_{ij})$ as in Lemma 6.4, and we get $w_s = \phi'_{ij}(w_0)$, or the last admissible word w_s looses θ-indices after a return work of $\overleftarrow{Z}(\phi_{ij})$ starting with the rule inverse to (8). But in the latter case all the rules of this second switching on of $\overleftarrow{Z}(\phi_{ij})$ has to be uniquely determined (for the similar reason as for the first switching on), and the second work of $\overleftarrow{Z}(\phi_{ij})$ is just inverse to the first one. Hence we obtain $w_s = w_0$, and the lemma is proved. □

A computation $\mathcal{C} : W_0 \to \ldots \to W_s \to \ldots \to W_t$ is called the *product* $\mathcal{C}'\mathcal{C}''$ of two subcomputations $\mathcal{C}' : W_0 \to \ldots \to W_s$ and $\mathcal{C}'' : W_s \to \ldots \to W_t$.

Lemma 6.7. *Let $\mathcal{C} : W_0 \to \ldots \to W_s$ be a reduced computation of the S-machine M_1, where $W_0 = F(\mathcal{W}_0)$ for some configurations $\mathcal{W}_0$ of M_0 and the state letters of W_s have no θ-indices. Then there is a reduced computation $\mathcal{W}_0 \to \ldots \to \mathcal{W}_t$ of the Turing machine M_0 with a history $H = \theta_1 \ldots \theta_t$ such that $\mathcal{C} = \mathcal{C}_1 \ldots \mathcal{C}_t$, where $\mathcal{C}_i$ is the canonical computation of $M(\theta_i)$ starting and ending with configurations $F(\mathcal{W}_{i-1})$ and $F(\mathcal{W}_i)$, and $W_s = F(\mathcal{W}_t)$.*

Proof. We can decompose $\mathcal{C} = \mathcal{C}_1 \ldots \mathcal{C}_t$, where $\mathcal{C}_l : W_{i_{l-1}} \to \ldots \to W_{i_l}$ is a computation, without θ-indices in the state letters of the first configuration $W_{i_{l-1}}$ and the last one W_{i_l} only. Then by Lemma 6.4, $\mathcal{C}_i$ is a canonical computation of a single S-machine $M(\theta_i)$ since the θ-index can be changed only after a rule erasing θ-index. The product $\theta_1 \ldots \theta_t$ is reduced, because by Remark 6.5 the equality $\theta_i = \theta_{i-1}^{-1}$ would imply that the histories of the computations $\mathcal{C}_{i-1}^{-1}$ and $\mathcal{C}_i$ start with the same rule, contrary to the assumption that $\mathcal{C}$ is a reduced computation.

By the assumption, $W_0 = F(\mathcal{W}_0)$. Assume by induction that for $l > 0$ we have a computation $\mathcal{W}_0 \to \ldots \to \mathcal{W}_{l-1}$ of M_0 starting with $\mathcal{W}_0$ and ending with $\mathcal{W}_{l-1}$ such that $W_{i_{l-1}} = F(\mathcal{W}_{l-1})$.

Assume first that θ_i has type (3). If θ_i is a positive rule, we can define $\mathcal{W}_l = \mathcal{W}_{l-1} \cdot \theta_i$ and $W_{i_l} = F(\mathcal{W}_l)$ by Lemma 6.4 and the inductive definition of the functions f_i and F.

If θ_i is negative, the word $W_{i_{l-1}}$ is obtained from W_{i_l} after a canonical computation of $M(\theta_i^{-1})$ by Lemma 6.4. If W_{i_l} has a subword $q_{i-1}up_iq_i$, then the configuration $W_{i_{l-1}} = F(\mathcal{W}_{l-1})$ contains the subword $q_{i-1}\phi_{ij}(u)p_iq_i$ by Lemma 6.4 (1). By the inductive conjecture, the subword $\phi_{ij}(u)$ of $W_{i_{l-1}}$ is equal to $f_i(v)$ for some v. Hence by Lemma 6.2, the word v ends with the letter a_{ij} i.e. $v \equiv v'a_{ij}$, and so $\phi_{ij}(u) = f_i(v) = \phi_{ij}(f_i(v'))$, which implies $u \equiv f_i(v')$ since the mapping ϕ_{ij} is injective by Remark 4.2. It follows that $\mathcal{W}_{l-1}$ belongs to the domain of the rule θ_i, and we define $\mathcal{W}_l = \mathcal{W}_{l-1} \cdot \theta_i$. Then by the definition of θ_i, the word $\mathcal{W}_l$ has the subword $p_i v' q_i \omega_i$, and the equality $u \equiv f_i(v')$ implies that $W_{i_l} = F(\mathcal{W}_l)$, as desired.

If θ_i is a command of type (4), then both θ and $M(\theta)$ just replace q_i with q_i', and the statement becomes clear. If θ_i is a command of type (5), then the word $\mathcal{W}_l$ has the required form as well by Lemma 6.4 (2), since 1 is the f_i-image of the empty word. $\square$

Lemma 6.8. *If $\mathcal{L}$ is the language recognized by M_0, then the language recognized by M_1 is $f_1(\mathcal{L}) = \{f_1(u) \mid u \in \mathcal{L}\}$.*

Proof. Given an accepting computation

$$\alpha_1 u q_1 \omega_1 \alpha_2 q_2 \omega_2 \dots \alpha_k q_k \omega_k \to \dots \to \alpha_1 q_1' \omega_1 \alpha_2 q_2' \omega_2 \dots \alpha_k q_k' \omega_k$$

of M_0, Lemma 6.3 provides us an M_1-accepting computation

$$q_0 f_1(u) p_1 q_1 p_2 q_2 \dots p_k q_k \to \dots \to q_0' p_1 q_1' p_2 q_2' \dots p_k q_k'$$

since f_i-image of the empty word is 1. Therefore the word $f_1(u)$ is accepted by M_1.

Let now a word U in the alphabet $\{a'(1), b'(1)\}$ be accepted by a computation $\mathcal{C} : W_0 \to \dots \to W_t$ of M_1, where $W_0 \equiv q_0 U p_1 q_1 \dots$. Note that the accept word W_t is of the form $F(\mathcal{W})$ since 1 is the f_i-image of the empty word. Applying Lemma 6.7 to the computation $\mathcal{C}^{-1}$, we see that that $U = f_1(u)$ for some positive word u. Now the same lemma applied to $\mathcal{C}$, provides us with an accepting computation $\mathcal{W}_0 \to \dots \to \mathcal{W}_t$ of M_0, where $\mathcal{W}_0$ is of the form $\alpha_1 u q_1 \omega_1 \alpha_2 q_2 \omega_2 \dots \alpha_k q_k \omega_k$, i.e. the word u is accepted by M_0. $\square$

Lemma 6.9. *There are no reduced computations $\mathcal{C} : W_0 \to \dots \to W_s$ of the S-machine M_1 with standard base such that $s > 0$, $W_0 \equiv W_s$, and the state letters of W_0 have no θ-indices.*

Proof. Proving by contradiction, we consider the factorization $\mathcal{C} = \mathcal{C}_1 \dots \mathcal{C}_m$, where $\mathcal{C}_l : W_{i_{l-1}} \to \dots \to W_{i_l}$ and each $\mathcal{C}_l$ is a canonical computation of some $M(\theta_l)$. We have $m \ge 1$, and it follows from Remark 6.5 that the product $\theta_1 \dots \theta_m$ is reduced.

If a head Q_i is the working head for a rule θ_l of the form (3), then by Lemma 6.4, the tape word $w_{i_{l-1}}$ in the sector $Q_{i-1}P_i$ of $W_{i_{l-1}}$ is replaced with a word $w_{i_l} = \phi'_{ij}(w_{i_{l-1}})$ in W_{i_l} (or vice versa) by $M(\theta_l)$. If the head Q_i is not working for the command θ_l or θ_l has type (5), then the restriction of the canonical computation $\mathcal{C}_l$ to sectors $Q_{i-1}P_iQ_i$ does not change the tape words of these sectors by the projection argument (i.e. $W_{i_{l-1}}$ has the same sectors $Q_{i-1}P_iQ_i$ as W_{i_l}).

Thus by Lemma 6.6, we have a sequence of words $w_0, w_{i_1}, \dots, w_{i_m}$ in the sectors $Q_{i-1}P_i$, where for every pair $w_{i_{k-1}}, w_{i_k}$ we have either $w_{i_{k-1}} = w_{i_k}$ or w_{i_k} is a ϕ'_{ij}-image of $w_{i_{k-1}}$ for some mapping ϕ'_{ij} of $F(a'(i), b'(i))$, or vice versa. Then we can obtain a sequence of positive words $u_0, u_{i_1}, \dots, u_{i_m}$ with properties (1) - (3) of Lemma 4.4. In particular, $u_0 = u_{i_m}$ if $w_0 = w_{i_m}$.

It follows that all the computations $\mathcal{C}_1, \dots, \mathcal{C}_m$ can be modified and replaced with canonical computations $\mathcal{C}_1', \dots, \mathcal{C}_m'$, where the restriction of $\mathcal{C}'_l$ to the sector $Q_{i-1}P_i$ transforms the word $U_{i_{l-1}}$ to U_{i_l}, where every U_{i_k} is obtained from W_{i_k} by replacement of the subword w_{i_k} with u_{i_k}.

The modified computation $\mathcal{C}'_l$ differs from $\mathcal{C}_l$ not only in the sectors $Q_{i-1}P_iQ_i$; but in the other sectors we just vary the number of the rules changing nothing there, since the modifying rules work in the sectors $Q_{i-1}P_iQ_i$ only. So one can subsequently make such changes for every i and obtain an M_1-computation $\mathcal{C}' : V_0 \to \dots \to V_t$ with $V_0 \equiv V_t$. Here $V_0 = F(\mathcal{W})$ for some M_0-admissible word $\mathcal{W}$ by Lemma 4.4, because every tape subword v_i of V_0 is the image of 1 under a product of mappings of the form ϕ'_{ij}.

Note that after the modification of $\mathcal{C}$, the product $\mathcal{C}'_1 \dots \mathcal{C}'_m$ is reduced since $\mathcal{C}'_l$ and $\mathcal{C}_l$ start/end with the same rule inserting/erasing the θ-index. Since $V_0 = F(\mathcal{W})$, we obtain by Lemma 6.7, a reduced computation of M_0 with positive length and equal the first and the last configurations. But this contradicts Property (3) of Lemma 5.1. □

Lemma 6.10. *Let $\mathcal{C} : W_0 \to \dots \to W_s$ be a reduced M_1-computation with $s > 0$, standard base and history h. Assume that both words W_0 and W_s are input configurations. Then both of them are M_1-accepted and uniquely determined by h.*

Proof. By Lemma 6.9, we have $W_s \neq W_0$, and so the corresponding input words u_0 and u_s are different.

Let $\mathcal{C} = \mathcal{C}_1 \dots \mathcal{C}_t$ be the factorization, where every factor $\mathcal{C}_l$ is the canonical computation of some machine $M(\theta)$. The product $H = \theta_1 \dots \theta_t$ is reduced by Remark 6.5.

Since $u_0 \neq u_s$, there are commands in H sending a letter from the input tape to the tapes $k-1$ and k (Step 1 according to Remark 5.3 (1), but we now subtract 2 from the numbers of auxiliary tapes). Let $\mathcal{D} = \mathcal{C}'_1 \dots \mathcal{C}'_l$ be the (first) maximal subcomputation of $\mathcal{C}$ corresponding to the commands of this kind, where every $\mathcal{C}'_r$ $(1 \le r \le l)$ is a product of the factors $\mathcal{C}_m$-s corresponding to a block of θ-commands of M_0, according to the definition of block given in Remark 5.3.

Since the sector $Q_{k-1}P_k$ is empty in the beginning of $\mathcal{D}$, the computation $\mathcal{C}'_1$ replaces it with $\phi'_{kj}(1) = \mathbf{w}'_{ij}$ for some j by Lemma 6.4 (1). Similarly $\mathcal{C}'_2$ can replace the word $\mathbf{w}'_{ij}$ with a longer $\phi'_{kj'}$ image of it (not by $\phi^{-1}_{kj'}$-image by Remark 4.1, since $j' \neq j$, because $\mathcal{C}'_1$ and $\mathcal{C}'_2$ do not correspond to mutually inverse subwords of H). It follows that one never obtains 1 again in the sector $Q_{k-1}P_k$ during $\mathcal{D}$, and so $\mathcal{D}$ is not a suffix of $\mathcal{C}$, because W_s has this sector empty.

Since the transition of M_0 to Step 2 is possible with empty first tape only (after a command of type (5)), we have nothing in the sector Q_0P_1 at the end of the computation $\mathcal{D}$, according to the definitions of the machines $M(\theta)$-s (see the rules (1') - (3')). Hence the input word u_0, and so the input configuration W_0 is completely determined by the histories of $\mathcal{D}$ and $\mathcal{C}$. Furthermore by Lemma 6.7, the input configuration W_0 has the form $F(\mathcal{W}_0)$ for an input word of M_0, because $\mathcal{D}^{-1}$ starts with the empty word $1 = f_1(1)$ in Q_0P_1-sector. We have a similar form for W_s, which is also determined by the history of $\mathcal{C}$.

Therefore we can apply Lemma 6.7 to the entire computation $\mathcal{C}$ now. The obtained reduced computation $\mathcal{W}_0 \to \dots \to \mathcal{W}_t$ of M_0 has accepted configurations $\mathcal{W}_0$ and $\mathcal{W}_t$ by

Lemma 5.1 (4). Therefore by Lemma 6.8, the word W_0 (and similarly, the word W_s) is accepted by M_1, as required. $\square$

7 Main machine

7.1 Translator machine

An auxiliary input-output S-machine T is designed for translation of arbitrary language $\mathcal{L}$ of positive words in a finite alphabet $A = \{a_1, \dots, a_k\}$ into the language $f(\mathcal{L})$.

The tape alphabet of T is $A \sqcup Y' \sqcup Y''$, where $Y' = \{a', b'\}$ and $Y'' = \{a'', b''\}$. The set of state letters of T is $\{K\} \sqcup \{L\} \sqcup P \sqcup \{R\}$, where $P = \{p, p_a(1), p_a(2), p_a(3) \mid a \in A\}$.

The input configuration of T is $KuLpR$, where the input word u is a reduced word in the alphabet $A^{\pm 1}$, and the output configuration is $KLvpR$, where the output word v is a word in Y'.

The set of positive rules of T consists of the following rules of 'submachines' $T(a_j)$-s, corresponding to letters $a_j \in A$, $j = 1, \dots, k$.

(1) The rule

$$K \to K,\ L \to L,\ p \xrightarrow{\ell} p_{a_j}(1),\ R \to R$$

prescribes an a-index a_j to p and gives start to

(2) the work of the machine $\overleftarrow{Z}(a_j)$, which is a copy of $\overleftarrow{Z}(\phi_{1j})$ working with the alphabet $Y = Y' \sqcup Y''$, i.e. the P-head starts moving left replacing every letter in the LP-sector with the ψ''_{1j}-image of the copy of it in the PR-sector. (So the subgroup $Y''(\xi'_{a_j,1})$ associated with every rule $\xi'_{a_j,1}$ is $H''(1j)$, the image of $F(a'', b'')$ under the homomorphism ψ''_{1j}.) When the state letter $p_{a_j}(1)$ meets L,

(3) the next rule $\xi'(2) = \xi'_2(a_j)$ is applicable:

$$K \to K,\ a_j L \xrightarrow{\ell} L,\ p_{a_j}(1) \to \mathbf{w}'_{1j} p_{a_j}(2),\ R \to R$$

(4) Then the machine $\overrightarrow{Z}(a_j)$ (a copy of $\overrightarrow{Z}(\phi_{1j})$ working on Y) is switched on.

When $p_{a_j}(2)$ meets R, a 'copy' of the rule $\xi'(4)$ is applicable: $p_{a_j}(2) \xrightarrow{\ell} p_{a_j}(3)$, and the rule $p_{a_j}(3) \xrightarrow{\ell} p$ removes the a-index..

Comment. The work of the machine $T(a_j)$ in sectors LPR is similar to the work of M_1 (moreover, it is a copy of the work of $M(\theta)$ restricted to the base $Q_{i-1}P_iQ_i$ if the command θ of M_0 adds a_j to the a-word on tape number i) , i.e. it can replace the tape word u of the LP sector with the copy of $\phi'_{1j}(u)$ in the alphabet Y'. Such a computation also deletes the last tape letter in the sector KL if it ends with a_j.

Remark 7.1. *The definition of canonical computations C of the machine $T(a_j)$ is the same as the definition of the canonical work of $M(\theta)$ in Remark 6.5. (The index θ is replaced with index a_j of a p-letter). We say that the canonical computation of $T(a_j)$ starts with the* *application of the rule* $p \xrightarrow{\ell} p_{a_j}(1)$ *and the canonical computation of* $T(a_j^{-1})$ *starts with the application of the rule* $p \xrightarrow{\ell} p_{a_j}(3)$.

Lemma 7.2. *There exists a computation* $W_0 \to \ldots \to W_s$ *of* T *starting with an input configuration* $W_0 \equiv KuLpR$ *and ending with an output configuration* $KLvpR$ *if and only if* u *is a positive word in the alphabet* A *and* $v = f_1(u^*)$, *where* $u^* \equiv a_{i_t} \ldots a_{i_1}$ *is the mirror copy of* $u \equiv a_{i_1} \ldots a_{i_t}$.

Proof. If $u \equiv a_{i_1} \ldots a_{i_t}$ a positive word, then starting with the input configuration, after the canonical work of the machine $T(a_{i_t})$, we obtain the configuration $Ka_{i_1}a_{i_2} \ldots a_{i_{t-1}}L\phi_{1,i_t}(1)pR$. Then one can subsequently switch on the machines $T(i_{t-1}), \ldots, T(i_1)$ and obtain

$$KL\phi_{1,i_1}(\phi_{1,i_2}(\ldots \phi_{1,i_t}(1)\ldots))pR = KLf_1(u^*)pR$$

by the inductive definition of f_1.

Conversely, given a reduced T-computation $\mathcal{C} : W_0 \equiv KuLpR \to \ldots \to KLvpR$, we consider the sectors LPR, where the machine T works as the machine $M(\theta)$ for a command θ of the form (5).

Therefore the argument of Lemma 6.7 gives $v = f_1(w)$ for some (positive) w and the factorization $\mathcal{C} = \mathcal{C}_1 \ldots \mathcal{C}_t$, where $\mathcal{C}_l$ is the canonical work of some $T(a_{j_l}^{\pm 1})$. Hence this factorization corresponds to a reduced word u_0 in the alphabet A, and since every $T(a_j^{\pm 1})$ multiplies the sector KL by $(a_j)^{\mp 1}$ from the right, we have $u_0 \equiv u^*$, because the word u is erased in the final configuration.

Let us prove now that there are no negative letters in $u \equiv a_{i_1} \ldots a_{i_t}$. Assume $a_{i_t} = a_j^{-1}$ for some j. Then the computation should start with the canonical work of $T(a_j^{-1})$ inserting a_j in the sector KL, since otherwise it will increase the length of the sector KL forever by Remark 5.14. So $T(a_j^{-1})$ replaces 1 in the sector LP with the inverse image under $\phi'_{j,1}$, which does not exists since arbitrary $\phi'_{j,1}$-image has positive length. Thus, we have $a_{i_t} = a_j$. If $a_{i_{t-1}} = a_{j'}^{-1}$, then $j' \neq j$, and similarly we obtain $\phi_{j,1}(1) = \phi_{j',1}(*)$, contrary to Remark 4.2. Proceeding by induction, we see that the whole word u is positive and $w \equiv u_0 \equiv u^*$. $\square$

7.2 The machine M

Let L be a large integer. Consider now $L-1$ copies of the machine M_1, denote them by $M_1^{(i)}$, $i = 1, \ldots, L-1$. We denote the state and tape letters of $M_1^{(i)}$ accordingly, by adding superscript (i) to all tape and state letters and to all rules. Let $\Xi(M_1)$ be the set of positive rules of M_1 and B be the standard base of M_1, $B^{(i)}$ be the copy of the word B with new superscript (i) added to all letters. We now consider the S-machine M_2 with the rules

$$\xi(M_2) = [\xi^{(1)}, .., \xi^{(L-1)}], \;\; \xi \in \Xi(M_1)$$

(we shall denote $\xi(M_2)$ by ξ too) and the standard base

$$\{t^{(1)}\}B^{(1)}\{t^{(2)}\}B^{(2)}...\{t^{(L-1)}\}B^{(L-1)}\{t^{(L)}\}, \tag{9}$$

where $t(i)$-s are just separating letters. The start/stop words are defined accordingly (every letter in the standard base is replaced by the corresponding letter in the stop word of $M_1^{(i)}$).

For every admissible word W of M_1 with the standard base we denote by $W(M_2)$ the corresponding admissible word $t^{(1)}W^{(1)}t^{(2)}W^{(2)}\dots$ of M_2 with the standard base (of M_2), where $W^{(i)}$ is the i-th copy of W. By definition, $W(M_2)$ is an input of M_2 if W is an input word of M_1, and so M_2 has $L-1$ input sectors.

The letters in the copy $W^{(i)}$ of the word W are equipped with the extra superscript (i). Thus every a-letters and every q-letter of M_2 has this extra index. We call it the *superscript* of a letter.

By definition the part of a rule ξ of M_2 coincides with corresponding part of some rule $\xi^{(i)}$, except for t-letters $t^{(1)},\dots t^{(L)}$, which are separating but never working: $t^{(i)}\to t^{(i)}$ for every rule ξ.

Notice that for every rule ξ of M_2 and every admissible word W of M_1 with the standard base of M_1, we have $W\cdot\xi=W'$ if and only if $W(M_2)\cdot\xi=W'(M_2)$. So there is no sense of the obtained extension of M_1 if we are interested in the properties of computations, but the inequality $L>>1$ will be helpful for the study of groups and diagrams associated with machines.

Next, we want to connect M_2 to the translator machine, and at first we extend the base of M_2. By definition, the standard base of the S-machine M_3 is obtained from the standard base of M_2 by adding one t-letter $t^{(0)}$ from the left: $\{t^{(0)}\}\{t^{(1)}\}B^{(1)}\dots$ But M_3 works exactly as M_2 since every rule ξ of M_2 is extended by trivial part $t^{(0)}\stackrel{\ell}{\to}t^{(0)}$. i.e. the new sector $t^{(0)}t^{(1)}$ is always locked by M_3. However now we are able to connect M_3 with the translator T.

Let M_4 be the machine with the standard base of the same form as M_3, that is

$$\{t^{(0)}\}\{t^{(1)}\}B^{(1)}\{t^{(2)}\}B^{(2)}\dots\{t^{(L)}\}$$

The analogy $\bar\xi$ of every rule ξ of T works in the sector $t^{(0)}t^{(1)}$ as the rule ξ in the sector KL, and $\bar\xi$ copies in every two sectors $Q_0^{(i)}P_1^{(i)}Q_1^{(i)}$ $(i=1,\dots,L-1)$, the work of ξ in the sectors LPR. Thus M_4 is built to *simultaneously* translate the content of the sector $t^{(0)}t^{(1)}$ to each of the input sectors $Q_0^{(i)}P_1^{(i)}$ of the machines M_2.

To connect the S-machine M_4 and M_3 we introduce the connection rule $\xi(43)$ which locks all the sectors of M_3 except for the input sectors $Q_0^{(i)}P_1^{(i)}$ and replaces all state letters of M_4 with the corresponding letters of M_3. (However we will not introduce more indices, since it will be clear from the context if a rule ξ belongs to M_3 or to M_4.)

Denote by M_5 the obtained union of the machines M_4 and M_3. The input configuration of M_5 is the input configuration of M_4 (and so the input sector of M_5 is $t^{(0)}t^{(1)}$) and the accept configuration of M_5 is the accept configuration of M_3. The main machine M is the circular form of M_5: we identify $t^{(L)}=t^{(0)}$, and the superscripts are taken modulo L for M. So the standard base $B(M)$ of M is the standard base $B(M_3)$ of M_3 However computations of M can have arbitrary long finite base, for example, $B^{(L-1)}t^{(L-1)}B^{(L)}t^{(0)}t^{(1)}B^{(1)}\dots$.

Lemma 7.3. *Let $Q'Q''$ be a sector with a tape alphabet $\bar{Y}$ of the standard base of M. Assume that there exist two words $q'uq''$ and $q'vq''$ with base $Q'Q''$ belonging to the domains of two rules η and ξ, resp., and η does not lock the sector $Q'Q''$. Then*

(1) we have $\bar{Y}(\xi) \le \bar{Y}(\eta)$;

(2) if ξ locks the sector $Q'Q''$, then the application of ξ changes either state q' or q''.

Proof. The inclusion is obvious if $\bar{Y}(\xi)$ is the trivial subgroup or $\bar{Y}(\eta) = \bar{Y}$. Otherwise both rules ξ and η belong to a copy of one of the machines $\overleftarrow{Z}$, $\overrightarrow{Z}(\phi_j)$ $\overleftarrow{Z}(\phi_j)$, $\overrightarrow{Z}(\phi_j), T(a_j)$. In the later case $Q', Q'' \ne K$ since $\bar{Y}(\xi)$ is nontrivial subgroup and $\bar{Y}(\eta) \ne \bar{Y}$, and so the machine $T(a_j)$ copies $\overrightarrow{Z}(\phi_j)$ and $\overleftarrow{Z}(\phi_j)$ in the sector $Q'Q''$. (See Comment after the definition of $T(a_j)$.) Therefore Statement (1) follows from Lemma 5.12. Statement (2) can be checked by inspection of the list of M-rules. (For example, for the pair $\xi_1(c), \xi_2$ and $QQ' = Q^{(j)}_{i-1}P^{(j)}_i$, the rule ξ_2 changes the p-letter.) □

Let $\mathcal{L}$ be a recursively enumerable set of positive words in some alphabet $A = \{a_1, .., a_k\}$. Then the set $\mathcal{L}^* = \{w^* \mid w \in \mathcal{L}\}$ (where $w^* \equiv a_{i_t} \dots a_{i_1}$ if $w \equiv a_{i_1} \dots a_{i_t}$) is recursively enumerable as well. Therefore there is a Turing machine M_0 recognizing the language $\mathcal{L}^*$, and one may assume that M_0 satisfies the conditions (1) - (7) from Lemma 5.1. We now assume that the construction of M_1, of other auxiliary S-machines, and of the main machine M are based of the Turing machine M_0 recognizing the language $\mathcal{L}^*$.

So by Lemma 6.8, the machine M_1 recognizes the language $f_1(\mathcal{L}^*)$. It follows from the definition of the machines M_2 and M_3 that the language of their accepted words is also $f_1(\mathcal{L}^*)$. The difference in comparison with M_1 is that the standard bases of M_2 and M_3 have $L-1$ input sectors, and an input configuration is accepted iff all the input words are copies (with $L-1$ different superscripts) of the same word from $f_1(\mathcal{L}^*)$. This observation leads to

Lemma 7.4. *The language recognized by the machine M is $\mathcal{L}$.*

Proof. If $u \in \mathcal{L}$, then by Lemma 7.2, there is a computation of the translator machine T, and so of the machine M_4, transforming the input configuration $t_0ut_1 \dots$ of M with input u in the input configuration of M_3 with input words $f_1(\mathcal{L}^*)$ (with different superscripts), which is accepted by M_3 as we noticed above. Therefore the word u is accepted by M.

Conversely, assume that an input configuration with input u is accepted by a reduced computation $\mathcal{C} : W_0 \to \dots \to W_s$ of $\mathbf{M}$. The history h of $\mathcal{C}$ has the form $h_1\xi(43)h_2$, where h_1 is the history of the translator machine T. Assume that h_2 contains the rule $\xi(43)^{-1}$, and so $h_2 = h_3\xi(43)^{-1}h_4$, where h_3 is the history of the machine M_3. Then by Lemma 6.10, the subcomputation with history h_3 starts with an accepted input configuration of the machine M_3. Thus the last configuration of the subcomputation with history h_1 has the form $t_0t_1vp_1q_1 \dots$, where $v = f_1(w^*)$ for some $w \in \mathcal{L}$ by Lemma 6.8. Moreover by Lemma 7.2, we have $w \equiv u$, since the mapping f_1 is injective. So $u \in \mathcal{L}$, as desired.

Now assume that h_2 does not contain rule $\xi(43)^{-1}$, i.e. it is the history of an accepting computation of M_3. Then we obtain the subword $t_0t_1vp_1q_1$ in the beginning of the subcomputation with the history h_2 as in the previous paragraph, and so $u \in \mathcal{L}$ again. □

Lemma 7.5. *Let $\mathcal{C} : W_0 \to W_1 \to \dots \to W_s$ be a reduced computation of M with a base $Q_{i-1}^{(j)}P_i^{(j)}Q_i^{(j)}$ and a history h. Assume that the state p-letters of W_0 and W_s have neither θ-indices no a-indices, and the tape words u_0 and u_s of the $Q_{i-1}^{(j)}P_i^{(j)}$-sectors of the words W_0 and W_s are equal modulo the Burnside relations. Then $u_s \equiv u_0$.*

Proof. The history h of the computation $\mathcal{C}$ can be factorized as $(\chi_0)h_1\chi_1h_2\chi_2\dots h_t(\chi_t)$, where $\chi_j = \xi(43)^{\pm 1}$ or χ_j is empty and each h_l corresponds to the canonical work of $M(\theta^{\pm 1})$ for some command θ of M_0 or to the canonical work of $T(a_j^{\pm 1})$. Let $W_{i_l} = W_0(\chi_0)h_1\chi_1h_2\chi_2\dots h_l\chi_l$. If $u_0, \dots, u_t$ are tape words of the sector $Q_{i-1}^{(j)}P_i^{(j)}$ of the words $W_0 = W_{i_0}, \dots, W_{i_t}$, then by Lemma 6.6, for every pair (u_{l-1}, u_l), we have either $u_l = \phi'_{ij}(u_{l-1})$ for some mapping ϕ_{ij} or $u_{l-1} = \phi'_{ij}(u_l)$, or $u_l \equiv u_{l-1}$. Since u_s and u_0 are equal modulo the Burnside relations, we have $u_s \equiv u_0$ by Lemma 4.5, and the statement of the lemma follows. □

Lemma 7.6. *Assume that for some rule ξ of of M and the tape alphabet $\bar{Y}$ of some sector, the canonical image $\tilde{Y}$ of the subgroup $\bar{Y}(\xi)$ in the free Burnside group $B_n(\bar{Y})$ contains both nontrivial in $\tilde{Y}$ words u and $z^{-1}uz$ for some $z \in B_n(\bar{Y})$. Then the word z belongs to $\tilde{Y}$ too.*

Proof. The statement is obvious if $Y(\xi) = \bar{Y}$. Otherwise ξ is a copy of a rule of one of the machines $\overleftarrow{Z}$, $\overrightarrow{Z}$, $\overleftarrow{Z}(\phi_j)$, $\overleftarrow{Z}(\phi_j)$. Therefore the statement follows from Lemma 5.15. □

8 Interference of Burnside relations

In this section, we want to show that changing tape words by equal words in the free Burnside group, one cannot essentially spoil computations of M. From now, the language $\mathcal{L}$ is the language of positive words in the generators of the group G, which represent the identity of G. Since the group G from the formulation of Theorem 1.1 is recursively presented group of finite exponent, this language is recursively enumerable and arbitrary relation $w = 1$ of G is a consequence of the relations with left-hand sides from $\mathcal{L}$, i.e. $G = \langle A \mid \mathcal{L} \rangle$.

We say that two words in the tape subalphabet $\bar{Y}$ of some sector are *congruent* if they represent the same element of the group G if $\bar{Y}$ is the input alphabet of M, and they represent the same element of the free Burnside group of exponent n with basis $\bar{Y}$ if $\bar{Y}$ is a tape alphabet of a non-input sector.

An admissible word is said to be *regular* if it has no subwords qwq^{-1}, where q is a q-letter and w is an a-word congruent to 1. An M-computation $W_0 \to \dots \to W_s$ is called *regular* if W_0 (or, equivalently some W_j for $0 \le j \le s$) is regular.

Definition 8.1. *We say that two regular M-admissible words W and W' are congruent ($W' \cong W$) if they have the same vector of state letters $q(1)\dots q(l)$ and their tape words w_j, w'_j between $q(j)$ and $q(j+1)$ are congruent for every $j = 1, \dots, l-1$.*

Remark 8.2. *Note that if two words W and W' are congruent and η-admissible for some rule η, then the words $W \cdot \eta$ and $W' \cdot \eta$ are also congruent, because multiplication of the congruent tape words by the same words from the left/right preserves the congruence.*

Definition 8.3. *We say that* $\mathcal{C} : W_0 \to \ldots \to W_s$ *is a quasi-computation with history* $h = \eta_1 \ldots \eta_s$ *if every word* W_r *is regular and* η_{r+1}*-admissible* ($r = 0, \ldots, s-1$)*, and for every* $r = 1, \ldots, s-1$*, we have* $W_r \cong W_{r-1} \cdot \eta_l$ *and* $W_s = W_{s-1}\eta_s$*. If the history* h *is reduced, then the quasi-computation* $\mathcal{C}$ *is called reduced.*

It follows from Remark 8.2 that for every quasi-computation $W_0 \to \ldots \to W_s$ with history h, there is a reduced quasi-computation $W_0 \to \ldots \to W'_{s'}$ whose history is the reduced form of h and $W'_{s'} \cong W_s$. The *inverse quasi-computation* $\mathcal{C}^{-1}$ with history h^{-1} is $V_0 \to \ldots \to V_s$, where $V_0 \equiv W_s$, $V_s \equiv V_0$, and $V_i \equiv W_{s-i-1} \cdot \eta_{s-i}$ for $i = 1, \ldots s-1$.

Lemma 8.4. *Let* $\mathcal{C} : W_0 \to W_1 \to W_2$ *be a reduced quasi-computation with history* $h = \eta_1\eta_2$ *and base* $Q'Q''$ *for some base letters* Q' *and* Q''*. If the rule* η_2 *(the rule* η_1*) does not lock this sector, then there is a regular* M*-computation* $\mathcal{C}' : W'_0 \to W'_1 \to W'_2$ *(resp.,* $W'_2 \to W'_1 \to W'_0$*) with history* h *(with history* h^{-1}*), where* $W'_0 \equiv W_0$*,* $W'_1 \equiv W'_0 \cdot \eta_1$*, and* $W'_2 \cong W_2$ *(resp.,* $W'_2 \equiv W_2$*,* $W'_1 \equiv W'_2 \cdot \eta_2^{-1}$*, and* $W'_0 \cong W_0$*).*

Proof. Let $\bar{Y}$ be the tape alphabet of the sector $Q'Q''$. Since the word W_1 is η_2-admissible and has the same state letter as the word $W_0 \cdot \eta_1$ (which is η_1^{-1}-admissible), we have $\bar{Y}(\eta_1^{-1}) \leq \bar{Y}(\eta_2)$ by Lemma 7.3. Therefore the word $W'_1 = W_0 \cdot \eta_1$ is η_2-admissible. Since it is congruent to W_1, the word $W'_2 = W'_1 \cdot \eta_2$ is congruent to $W_1 \cdot \eta_2 = W_2$ by Remark 8.2.

The second statement of the lemma follows from the first one after one replaces $\mathcal{C}$ with $\mathcal{C}^{-1}$. □

Lemma 8.5. *Let* $\mathcal{C} : W_0 \to \ldots \to W_s$ *be a reduced computation of* M *with base* $t^{(i)}B^{(i)}t^{(i+1)}$*, (which is the subbase of the standard base of* M*), where* $t_i \neq t_0$*. Assume that* $|W_0|_a = |W_s|_a = 0$ *and each of* W_0*,* W_s *has the vector of state letters, which is a part of either start or end configuration of* M*. Then* $\mathcal{C}$ *is a restriction of a quasi-computation* $\mathcal{D} : V_0 \to \ldots \to V_s$ *with standard base, where each of* V_0*,* V_s *is either an accepted input configuration of* M *or the accept configuration, and the restriction of* $\mathcal{D}$ *to the subbase* $t_1 \ldots t_L$ *is a computation.*

Proof. We obviously can extend $\mathcal{C}$ to the base $t^{(1)}B^{(1)}t^{(2)} \ldots t^{(L-1)}B^{(L-1)}t^{(L)}$ since the rules of M act uniformly on each subbase $t^{(k)}B^{(k)}t^{(k+1)}$, $k \neq 0$. We will use the same notation $\mathcal{C}$ for such an extension. So it remains to extend $\mathcal{C}$ to the sector t_0t_1.

The history h can be written as $h = h_1\chi_1h_2\chi_2 \ldots h_l$, where $\chi_k = \chi(43)^{\pm 1}$ ($k = 1, \ldots, l-1$), and $h_1, \ldots, h_l$ are the histories of either M_3 or M_4. We will induct on the number l, and so we make the assumption of the lemma weaker allowing the last admissible word be in the domain of rule $\chi(43)^{\pm 1}$ too (which unlock sectors $Q_0^{(j)}P_1^{(j)}$). So by the inductive conjecture, we can extend the computation with history $h_1\chi_1h_2\chi_2 \ldots h_{l-1}$ and obtain either the empty sector t_0t_1 at the end on the quasi-computation or a word u from the language $\mathcal{L}$ in this sector if h_{l-1} is a history of M_4.

If h_l is a history of M_3, then before the aplication of χ_l , one can replace u with the congruent empty word since $u = 1$ in G the further extension is obvious since every rule of M_3 locks the sector t_0t_1.

Let now h_l be a history of M_4. We first assume that $\chi_{l-1} = \chi(43)^{-1}$. Then either h_{l-1}^{-1} is a history of an accepting computation $\mathcal{E} : U \to \ldots \to U'$ of M_3 (if $l = 2$) or it is a computation of M_3 starting and ending with input configurations. In both cases U is an accepted input configuration of M_3 (see Lemma 6.10 for the second case). Thus, the input tape word v of U (in the sector $Q_0^{(1)}P_1^{(1)}$) belongs to the language $f_1(\mathcal{L}^*)$ (see Section 7.2). So there is $u \in \mathcal{L}$ such that $v = f_1(u^*)$, and by Lemma 7.2, one can extend the subcomputation of $\mathcal{C}$ with history h_l to the sector t_0t_1, and obtain at the end the accepted input configuration $t_0ut_1 \dots t_L$.

It remains to assume that $l = 1$. Then h_l is a product $g_1 \dots g_k$, where every g_r is the history of the canonical work of a submachine $T(a_{i_r}^{\pm 1})$ for some letter a_{i_r}. Note that g_{s-1} and g_s cannot be associated with a and a^{-1} for some $a \in Y_0^{\pm 1}$ since the computation $\mathcal{C}$ is reduced. Hence the word $a_{i_1}^{\pm 1} \dots a_{i_k}^{\pm 1}$ is reduced too. The canonical work of every machine $T(a_{i_r})$ corresponds to the application of the mapping ϕ_{1,i_r} to the sector $Q_{i-1}^{(j)}P_i^{(j)}$. Hence if $k \geq 1$, then we have a contradiction with Lemma 4.3. If $k = 0$, then the sector t_0t_l can be left empty under the computation with history h_l, as required. □

We say that a history h is *stable* if all the rules of h do not change the states of the heads of the S-machine M. According to the definition of M this means that all rules from h are the copies of the rules $\xi_1(*)$ or all of them are copies of the rules $\xi_3(*)$ of Z-machines, or all the rules are copies of the rules $\xi_1'(*)$ of the machine $\overleftarrow{Z}(\phi_{ij})$, or all the rules are copies of the rules $\xi_3'(*)$ of the machine $\overrightarrow{Z}(\phi_{ij})$. (Only one working p-head from the standard base of M_1 or of T is moving, although the standard base of M has $L-1$ copies of such a moving head).

A reduced history is called *simple* if it has no non-empty maximal stable subhistories h, such that the word h is trivial in the free Burnside group of exponent n.

Lemma 8.6. *For every quasi-computation $W_0 \to \dots \to W_s$ with a history h, there is a quasi-computation $W_0' \to \dots \to W_{s'}'$ with a simple history h' such that $W_0' \cong W_0$, $W_{s'}' \cong W_s$, and h' is equal to h modulo the Burnside relations.*

Proof. Assume that h is not simple, i.e. a nontrivial maximal stable sub-quasi-computation $\mathcal{D} : W_j \to \dots \to W_l$ has history h_0 trivial in the free Burnside group. Since $\mathcal{D}$ is stable, the words W_j and W_l have the same vector of state letters. They also have congruent tape subword in each sector. Indeed, if the working head P_i of $\mathcal{D}$ works as the running head in one of the machines $\overleftarrow{Z}(\theta, i)$, $\overrightarrow{Z}(\theta, i)$, then it consequently inserts/deletes the copies of letters from h_0 in the tape alphabets from the left and from the right, and so $\mathcal{D}$ does not change sector words modulo the Burnside relations. If P_i works as in $\overrightarrow{Z}(\phi_{ij})$, then every rule from h_0 multiplies the sector tape words by the (copy of the) corresponding generator of the group $H(j)$. So $\mathcal{D}$ changes tape words by equal words modulo the Burnside relations too. If P_i works as in $\overleftarrow{Z}(\phi_{ij})$, then one may assume that $\mathcal{D}$ corresponds to the positive command θ and use both above arguments.

Since we have $W_l \cong W_j$, the sub-quasi-computation $\mathcal{D}$ can be removed from $\mathcal{C}$. The obtained quasi-computation can be made reduced. This procedure proves the lemma. □

Two reduced histories $h \equiv h_1 g_1 h_2 g_2 \dots g_l h_l$ and $h' \equiv h'_1 g_1 h'_2 g_2 \dots g_l h'_l$, where h_i, h'_i are all maximal stable subhistories (here h_1 or/and h_l can be empty), are called *congruent* if $h'_i = h_i$ $(i = 1, \dots, l)$ in the free Burnside group of exponent n.

Lemma 8.7. *Let $\mathcal{C} : W_0 \to W_1 \to W_2$ be a reduced quasi-computation with a history $h = \eta_1\eta_2$ and the base $(Q_{i-1}^{(j)} P_i^{(j)} Q_i^{(j)})^{\pm 1}$ for some $i \in \{1, \dots, k\}$, $j \in \{1, \dots, L-1\}$. Suppose that the rule η_2 locks one of the sectors $Q_{i-1}^{(j)} P_i^{(j)}$, $P_i^{(j)} Q_i^{(j)}$, but not both.*

Then there is a reduced regular M-computation $\mathcal{C}'$ starting with W_0 whose simple history h' is congruent to h and the last word of $\mathcal{C}'$ is congruent to W_2.

Proof. (a) We will omit superscripts and assume that η_2 belongs to $\overleftarrow{Z}(\phi_{ij})$ and locks the sector $Q_{i-1}P_i$. Then we assume first that $W_0 \cdot \eta_1 = q_{i-1} u p_{\theta,i}(1) v q_i$ for some a-words u and v. The congruent word is $W_1 \equiv q_{i-1} p_{\theta,i}(1) v' q_i$ (and so it is η_2-admissible). We consider an extension $\mathcal{D}$ of the computation $W_0 \to W_0 \cdot \eta_1$ by a stable computation of $\overleftarrow{Z}(\phi_{ij})$ with the moving state letter $p_{\theta,i}(1)$, which erases the word u from the left. Since u is equal to the empty word in the free Burnside group, the same extension replaces v by a word $v''v$, where v'' is trivial modulo the Burnside relations since v'' is (a copy of) the image of u under a homomorphism ψ'_{ij}.

If $\mathcal{E}$ is the reduced form of $\mathcal{D}$, we have that the history of $\mathcal{E}$ is congruent to h and the last admissible word is equal to (the reduced form of) $q_{i-1} p_{\theta,i}(1) v'' v q_i$ The only difference in comparison with W_1 is the occurrence $v''v$ instead of v'. Since $v''v$ and v' are equal modulo the Burnside relations, it remains to obtain $\mathcal{C}'$ applying Lemma 8.4 to the quasi-computation in sector $P_i Q_i$ whose history is product of the last rule of $\mathcal{E}$ and η_2.

If $W_0 \cdot \eta_1 = q_{i-1} u p_{\theta,i}(2) v q_i$ and therefore we have $W_1 = q_{i-1} \mathbf{w}'_{ij} p_{\theta,i}(2) v' q_i$, where $u = \mathbf{w}'_{ij} u'$, $u' \in H'(ij)$, and u is equal to $\mathbf{w}'_{ij}$ modulo the Burnside relations. Then instead of $\overleftarrow{Z}(\phi_{ij})$ in the previous paragraph, now we exploit the machine $\overrightarrow{Z}(\phi_{ij})$, which however moves left erasing the word u' and copying this word from the right of the P_i-head in the $P_i Q_i$-sector.

If the rule η_2 locks the sector $P_i Q_i$, one can also argue as above. For example, if $W_0 \cdot \eta_1 = q_{i-1} u p_{\theta,i}(1) v q_i$ and v is trivial in the free Burnside group, then by Property (C), it is a trivial word in the free Burnside subgroup $H''(ij)$ freely generated by the words $\mathbf{w}''_i(a)$ and $\mathbf{w}''_{ij}(b)$. Hence there is an extension $\mathcal{D}$ of the computation $W_0 \to W_0 \cdot \eta_1$ by a stable computation with the moving state letter $p_{\theta,i}(1)$ of $\overleftarrow{Z}(\phi_{ij})$, which erases the word v. The history of this moving is trivial modulo the Burnside relation, and so the word u will be replaced by a congruent word, and one can finish as above but using that the sector $Q_{i-1}P_i$ is unlocked by η_2 now.

If η_2 belongs to a copy of $\overleftarrow{Z}$ and $\overrightarrow{Z}$, then in the sectors $Q_{i-1}P_iQ_i$, we have the work of 'more primitive' machines copying, and the proof is easier. $\square$

Remark 8.8. *In the first (the second) variant of Lemma 8.4 and in Lemma 8.7, the constructed computation $\mathcal{C}'$ depends on the first word W_0 (resp., the last word W_s) and the history of $\mathcal{C}$ only. In particular, the replacement of W_1 with a congruent word can change the quasi-computation $\mathcal{C}$, but cannot change the computation $\mathcal{C}'$.*

Lemma 8.9. *Let $\mathcal{C} : W_0 \to \dots \to W_s$ be a quasi-computation with simple history $h = \eta_1 \dots \eta_s$ and the base $Q^{(j)}_{i-1}P^{(j)}_i Q^{(j)}_i$ for some $i \in \{1, \dots, k\}$, $j \in \{1, \dots, L-1\}$. Suppose that*

(1) exactly one rule of the history locks both $Q^{(j)}_{i-1}P^{(j)}_i$- and $P^{(j)}_i Q^{(j)}_i$-sectors, namely, the first rule η_1 or

(2) exactly two rules of the history lock both $Q^{(j)}_{i-1}P^{(j)}_i$- and $P^{(j)}_i Q^{(j)}_i$-sectors, namely, the first rule η_1 and the last rule η_s.

Then there is a reduced regular M-computation $\mathcal{C}'$ starting with W_0 whose (simple) history h' is congruent to h such that the last admissible word of $\mathcal{C}'$,

(a) under the assumption (1), is congruent to W_s and

(b) the under assumption (2), is W_s.

Proof. We will omit the superscripts (j) in the proof. Inducting on t, we want for every prefix $h(t)$ of h of length t, $1 \le t \le s$, to construct a regular reduced computation $\mathcal{C}'(t)$ of M such that the history $h'(t)$ of $\mathcal{C}'(t)$ is congruent to $h(t)$ and the last admissible word of $\mathcal{C}'(t)$ is congruent to W_t. Since under assumption (2), the last rule of of $h = h(s)$ locks all the sectors, the last word of the computation with history $h'(s)$ must be equal to W_s. If $t = 1$ this property is trivial with $h'(1) = h(1)$.

Assume that $t \ge 2$ and the required $\mathcal{C}'(t-1) : V_0 \to \dots \to V_{i_{t-1}}$ is constructed, where $V_{i_{t-1}} \cong W_{t-1}$. If the rule η_t unlocks both sectors $Q_{i-1}P_i$ and P_iQ_i, then we can apply Lemma 8.4 to the quasi-computation $V_{i_{t-1}-1} \to W_{t-1} \to W_t$ and obtain the required computation $\mathcal{C}'(t) : V_0 \to \dots \to V_{i_{t-1}} \to W'_t$, where $W'_t \cong W_t$ with history $h'(t)$ congruent to $h(t)$. Similarly, if η_t locks only one of the sectors, then we can apply Lemma 8.7 to the quasi-computation $V_{i_{t-1}-1} \to W_{t-1} \to W_t$ and obtain $\mathcal{C}'(t)$.

Assume now that η_t locks both sectors. It follows from the assumption (2) that $V_{i_{s-1}} \equiv q_{i-1}upvq_i \cong W_{s-1}$, where both u and v are congruent to the empty word since the rule η_s locks both sectors $Q_{i-1}P_i$ and P_iQ_i and the word W_{s-1} is η_s-admissible. Therefore the words u and v are freely trivial by Lemma 6.6 and Lemma 4.5. We obtain the required computation $\mathcal{C}'$ extending $\mathcal{C}'(s-1)$ with $V_{i_{s-1}} \to V_{i_{s-1}} \cdot \eta_s \equiv W_s$. □

Definition 8.10. *We call the base $Q(1)Q(2)\dots Q(k)$ of a computation or a quasi-computation $\mathcal{C} : W_0 \to W_1 \to \dots \to W_t$ revolving if $k > 1$ and $Q(1) = Q(k)$.*

If the computation $\mathcal{C}$ with revolving base has history h, then for every $j = 1, \dots k$ there is the computation (quasi-computation) $\mathcal{C}' : W'_0 \to W'_1 \to \dots \to W'_t$ with history h and the revolving base $Q(j)Q(j+1)\dots Q(k)Q(2)\dots Q(j-1)Q(j)$, where every word W'_i is obtained from W_i by the corresponding cyclic permutation of the sectors; $\mathcal{C}'$ is a *cyclic permutation* of $\mathcal{C}$. So $Q(1)Q(2)\dots Q(k)$ can be regarded as the circular base of $\mathcal{C}$.

Lemma 8.11. *Let $\mathcal{C} : W_0 \to \dots \to W_s$ be a quasi-computation with a revolving base, with simple history $h = \eta_1 \dots \eta_s$, and let the copies of the word $(Q^{(j)}_{i-1}P^{(j)}_i Q^{(j)}_i)^{\pm 1}$ (with equal or different superscripts but with the same $i \ge 1$) occur in the cyclic base of $\mathcal{C}$ $m > 0$ times. Suppose that at least one rule η_r locks both $Q^{(j)}_{i-1}P^{(j)}_i$- and $P^{(j)}_i Q^{(j)}_i$-sectors. Then there is a quasi-computation $\mathcal{C}' : W'_0 \to \dots \to W'_{s'}$ such that*

(a) the simple history h' of $\mathcal{C}'$ is congruent to h;

(b) the restriction of $\mathcal{C}'$ to every subbase of the form $(Q_{i-1}^{(j)}P_i^{(j)}Q_i^{(j)})^{\pm 1}$ is a reduced regular M-computation;

(c) if the restriction $\mathcal{D}$ of $\mathcal{C}$ to a subbase $Q'Q''$, where neither Q' nor Q'' is equal to $P_i^{(*)}$, is a computation of M, then the restriction $\mathcal{D}'$ of $\mathcal{C}'$ to this subbase is also a regular computation of M;

(d) the first and the last admissible words of $\mathcal{C}'$ are congruent to W_0 and W_s, respectively;

(e) if the state letters of an admissible word W_l' have neither θ-indices nor a-indices, then either the $(Q_{i-1}^{(*)}P_i^{(*)})^{\pm 1}$-sectors of W_l' are all empty or they are all non-trivial in the free Burnside group of exponent n.

Proof. Since the rule η_r locks the sectors $(Q_{i-1}^{(j)}P_i^{(j)})^{\pm 1}$ and $(P_i^{(j)}Q_i^{(j)})^{\pm 1}$, the letters $P_i^{(j)}$ (with arbitrary superscript (j)) can occur in the revolving base (taken up to cyclic permutations) solely between $(Q_{i-1}^{(j)})^{\pm 1}$ and $(Q_i^{(j)})^{\pm 1}$ by Lemma 5.7.

Assume first that

(1) exactly one rule of the history locks both $Q_{i-1}^{(j)}P_i^{(j)}$- and $P_i^{(j)}Q_i^{(j)}$-sectors, namely, the first rule η_1, or

(2) exactly two rules of the history lock both $Q_{i-1}^{(j)}P_i^{(j)}$- and $P_i^{(j)}Q_i^{(j)}$-sectors, namely, the first rule η_1 and the last rule η_s.

Since in the beginning all m pairs of the sectors are empty, the computation $\mathcal{C}(t)$, we defined by induction in the proof of Lemma 8.9, is the same (up to superscripts) for all pairs of such sectors. (See Remark 8.8.) Therefore the statements (a), (b), and (d) follow from Lemma 8.9.

Consider now a sector $Q'Q''$, where neither Q' nor Q'' is equal to $P_i^{(*)}$ and the restriction of $\mathcal{C}$ to this sector is an M-computation. Then the modification of the work of the $P_i^{(*)}$-heads needed for the inductive construction of $\mathcal{C}'(t)$, does not change the configurations in the $Q'Q''$-sector at all by the definition of M; it can just extend the computation in this sector by a longer trivial computation. This proves the statement (c) under the above assumption.

Assume now that the history h of $\mathcal{C}$ has $l > 0$ rules $\chi_1, \dots, \chi_l$ simultaneously locking all $Q_{i-1}^{(*)}P_i^{(*)}$- and $P_i^{(*)}Q_i^{(*)}$-sectors. So, $h \equiv h_0\chi_1h_1\dots\chi_lh_l$, where the subhistories $h_0, \dots, h_l$ contain no such rules for given i. Then we consider the quasi-computations with subhistories $h_0\chi_1$, $\chi_1h_1\chi_2$,... and apply the statements of Lemmas 8.9 and 8.11 under assumptions (1) or (2) formulated above. (More precisely, the assumption (1) works for the inverse subcomputation with history $\chi_1^{-1}h_0^{-1}$.)

To obtain Property (e), we apply Lemma 6.6 to the subcomputation of $\mathcal{C}'^{\pm 1}$ (in the sectors $(Q_{i-1}^{(j)}P_i^{(j)}Q_i^{(j)})^{\pm 1}$) starting with the rule locking both sectors and ending with W_l'. Then the statement follows from Lemma 4.5. Thus, the lemma is proved. □

Lemma 8.12. *Let $\mathcal{C} : W_0 \to \dots \to W_s$ be a quasi-computation with a revolving base, a simple history $h = \eta_1 \dots \eta_s$, and let the copies of the word (with equal or different superscripts but with the same subscript i) occur in the base of $\mathcal{C}$ $m > 0$ times. Suppose the base of $\mathcal{C}$ has no subwords of the form $(P_i^{(j)})^{\pm 1}(P_i^{(j)})^{\mp 1}$ and at least one rule of h locks one of the sectors $Q_{i-1}^{(j)}P_i^{(j)}$ and $P_i^{(j)}Q_i^{(j)}$ but no rule locks both these sectors.*

Then there is a reduced M-computation $\mathcal{C}'$ with Properties (a) - (e) of Lemma 8.11.

Proof. It follows from the assumption of the lemma that every occurrence of $(P_i^{(j)})^{\pm 1}$ in the revolving base appears in a subword $(Q_{i-1}P_iQ_i)^{\pm 1}$. We first assume that there is a rule locking $Q_{i-1}P_i$-sectors in h and there is a rule locking P_iQ_i-sectors. The word W_0 has subwords of the form $(q_{i-1}upvq_i)^{\pm 1}$ (where we omit the superscripts and the indices at p). For different subbases $Q_{i-1}P_iQ_i$ of W_0, the pairs (u, v) can be different but the first components of such pairs are congruent. Indeed, all of them are empty in the word W_l obtained after the application of the rule locking the $Q_{i-1}P_i$-sectors, and the application of arbitrary rule preserves congruence of words by Remark 8.2. The second components are also congruent by the same reason.

For the beginning, we can therefore replace all such pairs in different subwords with the pair (u, v), changing the admissible word W_0 by a congruent word, and so the word W_1 of the quasi-computation can be left unchanged. Indeed, this can be done by Lemma 8.4 if the rule η_1 does not lock $Q_{i-1}P_i$-sector (P_iQ_i-sector); otherwise the words in all $Q_{i-1}P_i$-sectors (resp., all P_iQ_i-sectors) of W_0, are copies of the same word.

Now one can repeat the inductive construction of $\mathcal{C}'(t)$ and $\mathcal{C}'$ given in the proof of Lemma 8.9. Since the initial subwords of W_0 with the base of the form $Q_{i-1}P_iQ_i$ are just copies of each other, the computations $\mathcal{C}'(t)$ can be built simultaneously for all subbases $(Q_{i-1}P_iQ_i)^{\pm 1}$ by Remark 8.8. So we obtain Properties (a) - (d) of Lemma 8.11 arguing as in Lemma 8.9.

If one of the two sectors $Q_{i-1}P_i$ and P_iQ_i is unlocked by all the rules from h, then there is no need to modify W_0 making the a-words in such sector equal.

To obtain Property (e) from Lemma 8.11, we consider a word W_l' without θ-and a-indices and two computations $\mathcal{C}_1 : W_l' \to \ldots \to W_s'$ and $\mathcal{C}_2 : W_l' \to \ldots \to W_0'$. If a $Q_{i-1}^{(*)}P_i^{(*)}$-sector of W_l' contains a non-empty word w trivial in the free Burnside group of exponent n, then we can replace W_l' by a word W_l'' with empty $Q_{i-1}^{(*)}P_i^{(*)}$-sectors and replace $W_l' \cdot \eta_{l+1}'$ with a congruent word $W_l'' \cdot \eta_{l+1}'$. So, we obtain a quasi-computation starting with W_l''. Then again, this quasi-computation gives a computation $\mathcal{C}'_1$ starting with W_l''. Similarly we obtain a computation $\mathcal{C}'_2$. The computation $(\mathcal{C}'_2)^{-1}\mathcal{C}'_1$ has Property (e), because by Lemmas 6.6 and 4.5, any word W_k without θ-indices in this computation has empty $Q_{i-1}^{(*)}P_i^{(*)}$-sector word w provided w is trivial modulo the Burnside relations. $\square$

Lemma 8.13. *Let $\mathcal{C} : W_0 \to \ldots \to W_s$ be a quasi-computation with a revolving base having no subwords of the form $(P_i^{(j)})^{\pm 1}(P_i^{(j)})^{\mp 1}$ and with a simple history $h = \eta_1 \ldots \eta_s$. Then there is a quasi-computation $\mathcal{C}'$ starting with a word congruent to W_0, ending with a word congruent to W_s, whose simple history h' is congruent to h, and the restriction of $\mathcal{C}'$ to every sector $Q'Q''$ is a regular computation provided $Q' \notin \{t^{(0)}, (t^{(1)})^{-1}\}$.*

If the base (or a cyclic permutation of it) has a subword $Q_{i-1}^{(j)}P_i^{(j)}Q_i^{(j)}$ and $\mathcal{C}$ has an admissible word W_r without θ-indices and a-indices, then in every such admissible word W_r' of $\mathcal{C}'$, every $Q_{i-1}^{(*)}P_i^{(*)}$-sector is either empty or non-trivial in the free Burnside group of exponent n.

Proof. If we have a 2-letter subbase of the form $Q_iQ_i^{-1}$ or $Q_i^{-1}Q_i$, then such sectors cannot be locked by Lemma 5.7, and the restrictions of $\mathcal{C}$ to such base can be replaced with computations having the same history, due to Lemma 8.4.

By the definition of admissible words and the assumption of the lemma, a base letter $(P_i^{(j)})^{\pm1}$ can occur only in the subbases $(Q_{i-1}^{(j)}P_i^{(j)}Q_i^{(j)})^{\pm1}$ of the revolving base.

If the history h of $\mathcal{C}$ has rules locking the sectors $Q_{i-1}^{(j)}P_i^{(j)}$ and rules locking the sectors $P_i^{(j)}Q_i^{(j)}$, then the restrictions of $\mathcal{C}$ to such subbases can be replaced with computations by Lemmas 8.11 and 8.12.

In the remaining cases, there are no rules in h locking the sectors $P_i^{(j)}Q_i^{(j)}$ or there are no rules locking the sectors $Q_{i-1}^{(j)}P_i^{(j)}$. If nothing is locked, then we refer to Lemma 8.4 again. Otherwise one can apply Lemma 8.12. To complete the proof, we recall that by Lemmas 8.11 (c) and 8.12, the restrictions of the quasi-computation $\mathcal{C}$ to the subbases $(Q_{i-1}^{(j)}P_i^{(j)}Q_i^{(j)})^{\pm1}$ with different subscripts i can be modified one-by-one. □

9 Revolving quasi-computations

We say that a reduced computation (a quasi-computation) $W_0 \to \dots \to W_s$ with a history $\eta_1 \dots \eta_s$ is *idling* if $W_{i-1} \cdot \eta_i \equiv W_{i-1}$ for every $i = 1, \dots, s$. An admissible word W of M is *passive* if there exists a rule of M such that the computation $W \to W \cdot \eta$ is idling.

Lemma 9.1. *Let $\mathcal{C} : W_0 \to \dots \to W_t$ be a quasi-computation of M with history h nontrivial modulo the Burnside relations, with all rules from M_3 or all rules from M_4, and with revolving base $B(\mathcal{C})$. Assume that W_0 and W_t are passive words with equal vectors of state letters. Then either*

(1) $W_0 \cong W_t$ or

(2) all the state letters of all words W_l $(0 \le l \le t)$ are equal to the same $p^{\pm1}$, where $p \in P_i^{(j)}$ for some i, j. Furthermore, for some command θ of M_0, all the transitions $W_j \to W_{j+1}$ are obtained by application of the rules $\xi_1(\theta, i)(*)^{\pm1}$ of a machine $\overrightarrow{Z}(\theta, i)$ or all of them are obtained by application of the rules $\xi_3(\theta, i)(*)^{\pm1}$ of a machine $\overleftarrow{Z}(\theta, i)$, or all of them are obtained by application of the rules $\xi_1'(\theta, i)(*)^{\pm1}$ of a machine $\overleftarrow{Z}(\phi_{ij})$, or all of them are obtained by application of the rules $\xi_3'(\theta, i)(*)^{\pm1}$ of a machine $\overrightarrow{Z}(\phi_{ij})$.

Proof. If the base $B(\mathcal{C})$ has P-letters only, then we have Property (2) since a p-letter can be changed only by a rule locking a P_iQ_i-sector or locking a $Q_{i-1}P_i$-sector. (We often omit superscripts.) So one may assume that there is Q-letter in the base. Since the words W_0 and W_t are passive, every p-letter from $P_i^{\pm1}$ stays in these words either near a letter $Q_i^{\pm1}$ or near a letter from $Q_{i-1}^{\pm1}$.

Assume first that the history h is simple. Let we have a subword $(Q_{i-1}P_iP_i^{-1}Q_{i-1}^{-1})^{-1}$ (or $Q_i^{-1}P_i^{-1}P_iQ_i$) in the revolving base. Then the P_i-letters cannot start running since they cannot stop running returning to the Q-head by Lemmas 5.10, because the maximal stable subhistory of a simple history cannot be trivial in the free Burnside group. So the three sectors $Q_i^{-1}P_i^{-1}P_iQ_i$ of W_0 do not change, up to congruence, in $\mathcal{C}$.

By the projection argument (see Remark 5.13), we have for W_0 and W_t, the same a-projections of the subwords in subsectors $Q_{i-1}P_iQ_i$ unless the machine $\overleftarrow{Z}(\phi_{ij})$ works. Note also, that these subwords cannot be of the forms $q_{i-1}u'p_iq_i$ and $q_{i-1}p_iu''q_i$ with non-empty u', u'' since W_0 and W_t are passive. So we obtain Condition (1) if no $\overleftarrow{Z}(\phi_{ij})$ works, because the sector t_0t_1 is locked by M_3 and sectors t_0t_1 and $t_1^{-1}t_1$ are impossible for a computation of M_4 having passive words.

It remains to consider the work of $\overleftarrow{Z}(\phi_{ij})$ under the assumption that Property (1) is false. The P_i-head of one of such submachine has to run through the whole sector from Q_i to Q_{i-1} since W_0 and W_t are passive. (It cannot start running from a Q-head and return to the same head by Lemma 5.11.) This submachine starts working of a machine $M(\theta)$ for a command θ of type (3). After the P_i-head reaches Q_{i-1}, it locks the sectors $Q_{i-1}P_i$ and $P_{i-1}Q_{i-1}$ and gives start to $\overleftarrow{Z}(\theta, i-1)$, which also completes its canonical work since W_0 and W_t are passive. Similarly, we obtain the work of each Z-machine from the definition of $M(\theta)$, and every sector of the base will be locked by at least one rule. Hence the revolving base has to contain the standard base of M_1, which contradicts to the assumption that the words W_0 and W_t are passive.

If h is not simple, then we can replace it with a simple history h' by Lemma 8.6. If we have got Condition (2) for the quasi-computation with history h', the same property holds for $\mathcal{C}$ since only a stable quasi-computation can have a base consisting of P-letters. □

Definition 9.2. *We say that a regular computation of M or a quasi-computation $\mathcal{C} : W_1 \to \dots \to W_t$ is revolving if*

(a) the base $B(\mathcal{C})$ is revolving,

(b) $W_0 \cong W_t$,

(c) the history of h is simple and non-trivial modulo the Burnside relations, and

(d) no subword of the (cyclic) word $W_0^{\pm 1}$ is congruent to an M-accepted configuration.

Lemma 9.3. *Let $\mathcal{C} : W_0 \to \dots \to W_s$ be an idling revolving quasi-computation with cyclically reduced history h, and a quasi-computation $\mathcal{C}' : W_0' \to \dots \to W_{s'}'$ have history h' freely equal to ghg^{-1} for some word g. If $W_0' \equiv W_0$, then there is a computation $W_0 \to \dots \to W_t$ with history g satisfying either Condition (1) or Condition (2) of Lemma 9.1. In the latter case both $\mathcal{C}$ and $\mathcal{C}'$ satisfy Condition (2) as well.*

Proof. Note that either all the rules of h belong to M_3 or all of them belong to M_4 since the connecting command $\xi(43)^{\pm 1}$ changes all state letters an so it cannot occur in idling computations.

We have $h' \equiv fh_0f^{-1}$, where the right-hand side is reduced and $h_0 \equiv h_1h_2$ is a cyclic permutation of $h \equiv h_2h_1$ with non-empty h_1. So in the free group, we have $ghg^{-1} = fh_1hh_1^{-1}f^{-1}$. Therefore $g^{-1}fh_1$ belongs to the centralizer of h, and if h is a power of r, where r is not a proper power, then we obtain $g = fh_1r^l$ for some integer l.

Since $h' \equiv fh_0f^{-1}$, we have a quasi-computation $W_0' \to \dots \to V$ with history f extended to a quasi-computation $W_0' \to \dots \to V \to \dots V'$ with history fh_0, which starts with a quasi-computation $W_0' \to \dots \to V \to \dots V''$ with history fh_1.

Recall that by Lemma 7.3, every sector of the base is either locked by every rule of h or is unlocked by every rule of h, because h is idling. Since the rules of non-empty h_1 and r are contained in h, one can extend the above quasi-computation and obtain a quasi-computation $\mathcal{C}' : W_0' \to \dots \to U$ with history $fh_1r^l = g$.

Note that the first word $W_0' \equiv W_0$ is passive since the nontrivial idling computation $\mathcal{C}$ starts with W_0. The word U is also passive since the last rule of r^{-1} (i.e. the last rule of h) or the last rule of h_1 (if $l = 0$) is applicable to U, and it changes nothing since the word U has the same base as the words of $\mathcal{C}$. Now the reference to Lemma 9.1 completes the proof of the lemma, because Condition (2) for a non-trivial g implies the same condition for $\mathcal{C}$ and $\mathcal{C}'$ since all these quasi-computations have the same base. $\square$

Lemma 9.4. *Let $\mathcal{C} : W_0 \to W_1 \to \dots \to W_t$ be a non-idling revolving quasi-computation of M with a base $B(\mathcal{C})$ and cyclically minimal in $B(\Xi^+)$ history h corresponding to the rules of the machine M_3. Then $\mathcal{C}$ satisfies Condition (2) of Lemma 9.1.*

Proof. Assume first that the history h is stable. Then every rule of h belongs to some $\overleftarrow{Z}(\theta, i)$ (or to $\overrightarrow{Z}(\theta, i)$, or to $\overleftarrow{Z}(\phi_{ij})$, or to $\overrightarrow{Z}(\phi_{ij})$). Then the letters staying from the right and from the left of $P_i = P_i^{(j)}$ in the revolving base $B(\mathcal{C})$ cannot be either Q_{i-1}- or Q_i-letters by the assumption $W_t \cong W_0$ and Lemmas 5.10 and 5.11, because the stable history h is not congruent to the empty word. Hence by the definition of admissible word, we should have the subword $P_i^{-1}P_iP_i^{-1}$ in $B(\mathcal{C})$. (The base $B(\mathcal{C})$, being revolving, can be regarded as cyclic word). Using the same argument, we obtain $P_iP_i^{-1}P_iP_i^{-1}P_i$, and the entire base $B(\mathcal{C})$ is $(P_iP_i^{-1}P_iP_i^{-1}\dots P_i)^{\pm 1}$. Therefore p will never be replaced with other state letter during the quasi-computation since only a rule locking the sector $Q_{i-1}P_i$ or the sector P_iQ_i can change the states of the head P_i. Thus, under the above assumption, the lemma is proved.

We may now assume that h is not stable. Since $\mathcal{C}$ is revolving, one can construct a quasi-computation $\mathcal{C}^m$ with history h^m for arbitrary $m \ge 1$. Every stable subhistory of h^m is a subhistory of a cyclic permutation of h, because h is not stable. Therefore h^m is simple history since h is a cyclically minimal word. Moreover, we can construct an infinite in both direction "quasi-computation" $\mathcal{D}$ whose history is periodic with period h.

Since the history h is not stable, the p-letter of a machine $\overleftarrow{Z}(\theta, i)$ (or $\overrightarrow{Z}(\theta, i)$, or $\overleftarrow{Z}(\theta, i)$, or $\overleftarrow{Z}(\phi_{ij})$), starts working with some W_l. The P_i-head eventually stops working since the histories of $\mathcal{C}$ and $\mathcal{D}$ are not stable. It stops running when the next rule locks the sector $Q_{i-1}P_i$ of the revolving base (or the sector P_iQ_i) taken up to cyclic permutations.

Then by the definition of M_1 and M, the sector $P_{i-1}Q_{i-1}$ is also locked by the last rule, we have P_{i-1} in the revolving base, and the machine $\overleftarrow{Z}(\theta, i-1)$ has to start working. (There is no difference if we obtain P_{i+1} and $\overleftarrow{Z}(\theta, i+1)$ starts working.) In turn, the machine $\overleftarrow{Z}(\theta, i-1)$ will soon or later switch on the next one, and so on. (Here we keep in mind that the P-head cannot start and return to the same Q-head by Lemmas 5.10, because the maximal stable subhistory of a simple history cannot be trivial in the free Burnside group.) Reversing the history, we see that the machine $\overleftarrow{Z}(\theta, i)$ was switched on right after the machine $\overleftarrow{Z}(\theta, i+1)$ stopped working. It follows that the quasi-computation with history h^m includes for some m, the canonical work of every $\overleftarrow{Z}(\theta, i)$, $i = 1, \dots, k$.

Therefore every sector of the base $B(\mathcal{C})$ is locked by at least one rule from h. We obtain similar conclusion when starting with $\overrightarrow{Z}(\theta, i)$, or $\overleftarrow{Z}(\theta, i)$, or $\overleftarrow{Z}(\phi_{ij})$. Thus, by Lemma 5.7, the revolving base $B(\mathcal{C})^{\pm 1}$ must be (a cyclic permutation) of a power of the standard base. Besides, it follows that the same machine $M(\theta)$ cannot canonically work permanently, and there are rules in h erasing the θ-index of the state letters. So the quasi-computation $\mathcal{C}$ has a word W_l without θ-indices.

Consider the quasi-computation $\mathcal{E} : W_l \to \dots \to W_0 \to \dots \to W_l$ of positive length. By Lemma 8.13, there is a computation $\mathcal{F}$ of positive length starting and ending with words W and W' congruent to W_l and having history $h(\mathcal{F})$ congruent to $h(\mathcal{E})$. Moreover, by Lemma 8.13, we may assume that all $Q_{i-1}P_i$-sectors of W and W' are either empty or non-trivial in the free Burnside group. Now by Lemma 7.5 for $\mathcal{F}$, we have $W \equiv W'$, and Lemma 6.9 for $\mathcal{F}$ gives a contradiction. So h is stable, and the lemma is proved. $\square$

Lemma 9.5. *Let $\mathcal{C} : W_0 \to \dots \to W_t$ be revolving quasi-computation whose history h is cyclically minimal modulo the Burnside relations and contains a rule of the machine M_4. Then h has the rules of M_4 only and either*

(1) $\mathcal{C}$ is an idling quasi-computation, or

(2) all the state letters of all words W_l ($0 \le l \le t$) are equal to the same $p^{\pm 1}$, where $p \in P_1^{(j)}$ for some j, all rules of h are rules $\xi_1'(a_k)^{\pm 1}$ of a machine $\overleftarrow{Z}(\phi_{k1})$, or the rules $\xi_3'(a_k)^{\pm 1}$ of a machine $\overrightarrow{Z}(\phi_{k1})$.

Proof. If the maximal M_4-quasi-computation $\mathcal{C}'$ of $\mathcal{C}$ is idling, then all the rules of h have to belong to M_4 since the states of the heads are not changed by $\mathcal{C}'$. Then we have option (1). So we may assume further that $\mathcal{C}'$ is not idling.

Assume first that every rule of h belongs to the copy of a machine $T(a_l^{\pm 1})$. Since $\mathcal{C}$ is not idling, some P-letter is working in this quasi-computation as it follows from the definition of T and M_4. (When the rule $\zeta_2'(a_j)$ inserts/deletes a letter from the left of L it locks the sector LP, and so there is a working P-head if L is working head.)

Then (as in Lemma 9.4) the letters staying from the right and from the left of $P = P_1^{(j)}$ in the revolving base $B(\mathcal{C})$ cannot be $Q_0^{(j)}$ or $Q_1^{(j)}$ by the assumption $W_t \cong W_0$ and Lemmas 5.10 and 5.11, because the history is simple. Hence we obtain the word $P^{-1}PP^{-1}$ in $B(\mathcal{C})$,..., and the entire base $B(\mathcal{C})$ is $(PP^{-1}PP^{-1}\dots P)^{\pm 1}$. Therefore the letter p will never be replaced with other state letter during the quasi-computation since a rule changing the states of the head $P_1^{(j)}$ locks the sector $Q_0^{(j)}P_1^{(j)}$ or the sector $P_1^{(j)}Q_1^{(j)}$. Thus, the statement of the lemma follows.

Second, proving by contradiction, we assume now that all the rules of h belong to M_4, but do not belong to a copy of the same $T(a_l^{\pm 1})$. It follows that the $P_1^{(j)}$-head eventually stops running in $\mathcal{C}$ and meets a neighbor $Q_1^{(j)}$-head (or $Q_0^{(j)}$-head). Since the $P_1^{(j)}$-head changes the state when it ends running, and $W_t \cong W_0$, $P_1^{(j)}$-head has to meet another neighbor $Q_0^{(j)}$ (resp. $Q_1^{(j)}$). (The P-head cannot start and return to the same Q-head by Lemma 5.11.)

Now we consider the restriction $\mathcal{D} : V_0 \to \ldots \to V_t$ of $\mathcal{C}$ to the subbase $Q_0^{(j)} P_1^{(j)} Q_1^{(j)}$. We have a configuration V_m without a-indices in the computation $\mathcal{D}$. (Otherwise $\mathcal{D}$ has to correspond to the work of a single machine $T(a_k^{\pm 1})$.)

There exists a quasi-computation $\mathcal{E} : V_m \to \ldots \to V_0 \to \ldots \to V_m$. By Lemma 8.13, there is a computation $\mathcal{E}'$ of positive length starting and ending with words V and V' congruent to V_m and having history $h(\mathcal{E}')$ congruent to $h(\mathcal{E})$. Moreover, by Lemma 8.13, we may assume that all $Q_0^{(j)} P_1^{(j)}$-sectors of W and W' are either empty or non-trivial in the free Burnside group. By Lemma 7.5, for $\mathcal{E}'$, we have $V \equiv V'$.

The history $h(\mathcal{E}')$ is a product $h_1 \ldots h_l$, where every h_r is the history of the canonical work of a submachine $T(a_{i_r}^{\pm 1})$ for some letter a_{i_r}. Note that h_{s-1} and h_s cannot be associated with a and a^{-1} for some $a \in Y_0^{\pm 1}$ since the computation $\mathcal{C}$ is reduced. Hence the word $a_{i_1}^{\pm 1} \ldots a_{i_l}^{\pm 1}$ is reduced to. Recall that the canonical work of every machine $T(a_{i_r})$ corresponds to the application of the mapping ϕ_{1,i_r} to the sector $Q_0^{(j)} P_1^{(j)}$. Hence we have a contradiction with Lemma 4.3.

Assume finally that there are rules of the machine M_3 in h, and so there is the rule $\xi^{\pm 1}(43)$ connecting the machines M_4 and M_3. This rule locks all sectors except for the input sectors of the machine M_3, and therefore there is a subbase $(P_1^{(i)} Q_1^{(i)})^{\pm 1}$ in the revolving base.

A rule of M_3 next to $\xi(43)^{\pm 1}$ starts moving the $P_1^{(i)}$-head, and this head has to meet $Q_0^{(i)}$ and so on, which again means that the base of $\mathcal{C}$ (without the first or the last letter) is a cyclic permutation of a power of the standard subbase.

Since the rule $\xi(43)$ changes the state letters, we have both $\xi(43)$ and $\xi(43)^{-1}$ in the history. So there is a subhistory $\xi(43) h' \xi(43)^{-1}$ in h or in a cyclic permutation of h, where h' is a history of M_3-quasi-computation.

Since the sector $t_0 t_1$ is locked by M_3, by Lemma 8.13, we have an M-computation $V \to \ldots \to V'$ of M_3 with reduced history between the applications of $\xi(43)$ and $\xi(43)^{-1}$. By the definition of $\xi(43)$, only input sectors of these words can be non-empty. Thus the restrictions U and U' to the standard base of M_2 are input words. By Lemma 6.10, both of them are accepted by M_2, and so V and V' are (without the first or the last state letter) accepted by M_3, since the sectors with different superscripts have to be copies of each other being determined by the history. Hence the word W_0 (without the first or the last state letter) or a cyclic permutation of $W_0^{\pm 1}$ is congruent to a word accepted by M contrary to the assumption of the lemma saying that the quasi-computation $\mathcal{C}$ is revolving. □

Lemma 9.6. *Let $\mathcal{C}_1 : W_0 \to W_1 \to \ldots \to W_s$ be a revolving quasi-computation with a history h cyclically minimal in the free Burnside group $B_n(\Xi^+)$. Let $\mathcal{C}_2 : W_0 \equiv W_0' \to W_1' \to \ldots \to W_{s'}'$ be another revolving quasi-computation whose history h' is freely equal to ghg^{-1} for a reduced word g. Then there is computation with a history g': $W_0 \to \ldots \to W_0 \cdot g' \cong W_0$, where g' is equal to g modulo the Burnside relations.*

Proof. If $\mathcal{C}_1$ is an idling quasi-computation, then by Lemma 9.3, either we obtain the required property of g or $\mathcal{C}_1$ satisfies Condition (2) of Lemma 9.1. If $\mathcal{C}_1$ has a rule of M_4, then we come to the same conclusion by Lemma 9.5. If $\mathcal{C}_1$ is not idling computation of M_3, then again we obtain Condition (2) by Lemma 9.4. It remains to assume that $\mathcal{C}_1$ satisfies Condition (2). Since $\mathcal{C}_2$ has the same base, it satisfies Condition (2) as well.

Let $\mathcal{D} : p^{-1}u_0p \to \ldots \to p^{-1}u_tp$ (the sector index is omitted) be the restriction of $\mathcal{C}_1$ to a $P^{-1}P$-sector. (The case PP^{-1} is similar.) Note that this computation satisfies the assumption of Lemma 5.6, and using notation of that lemma we obtain that $\lambda(h)$ belongs to the centralizer $\mathcal{Z}$ of u_0 in B. Similarly, $\mathcal{Z}$ contains $\lambda(ghg^{-1})$. The centralizer of nontrivial element u_0 is the unique cyclic subgroup of order n containing u_0 (see [19], Theorem 19.5), and so $\lambda(g) \in \mathcal{Z}$ by [19], Theorem 19.6. Observe that $\mathcal{Z}$ is also the centralizer of the non-trivial $\lambda(h)$. It follows that the images $\lambda(g)$ and $\lambda(h)$ commute, and so do g and h since λ is a monomorphism.

The projection of the equality $ghg^{-1} = h$ to the free Burnside subgroup $\mathcal{B}_0$ generated by the rules of the stable computation with history h gives us $g'h(g')^{-1} = h$, where the rules of g' are among the generators of $\mathcal{B}_0$, and therefore there is a stable computation $W_0 \to \ldots \to W_0 \cdot g'$ with history g'. Since $\lambda(g')$ commutes with $\lambda(h)$, we have $\lambda(g') \in \mathcal{Z}$ too, and therefore $W_0 \cdot g' \cong W_0$ by Lemma 5.6. Besides, $\lambda(g')$ belongs to the centralizer $\mathcal{Z}' \leq \mathcal{Z}$ of $\lambda(h)$ in $\lambda(\mathcal{B}_0)$, and since both subgroups $\mathcal{Z}'$ and $\mathcal{Z}$ have order n, they are equal, and therefore $\lambda(g) \in \lambda(\mathcal{B}_0)$. Since λ is a monomorphism, we have $g \in \mathcal{B}_0$, and so g' is equal to g modulo the Burnside relations. □

10 Groups and diagrams related to the machine M

10.1 Construction of the embedding

We are going to introduce a group M associated with the machine M. Since we will use the main properties of M (given by Lemmas 8.5, 8.13 and 9.6) only and shall not use many details, we re-denote the machine M accepting the language $\mathcal{L}$ of positive relators of the group G by $\mathbf{M}$ and simplify notation as follows.

(1) For the set of (state) q-letters $\{t^{(0)}\} \sqcup \{t^{(1)}\} \sqcup Q_0^{(1)} \sqcup P_1^1 \sqcup Q_1^{(1)} \sqcup \ldots \sqcup \{t^{(L)}\}$ introduced in Subsection 7.2, we also will use the uniform notation $\mathbf{Q} = \sqcup_{j=0}^N \mathbf{Q}_j$, where $\mathbf{Q}_0 = \{t^{(0)}\}$, $\mathbf{Q}_1 = \{t^{(1)}\}$, $\mathbf{Q}_2 = Q_0^{(1)}$,...,$\mathbf{Q}_N = \{t^{(L)}\} = \{t^{(0)}\}$. The number of t-letters L is chosen large enough, so that $N \geq n$.

(2) The set of a-letters of $\mathbf{M}$ is$\mathbf{Y} = \sqcup_{j=1}^N \mathbf{Y}_j$ including the input alphabet $A \subset \mathbf{Y}_1$.

(3) The set of rules of $\mathbf{M}$ is now denoted by $\mathbf{\Theta}$.

(4) If u is an input word in the alphabet $A^{\pm 1}$, then the corresponding input configuration of $\mathbf{M}$ is denoted by $\Sigma(u)$.

The finite set of generators of the group M consists of *q-letters* from $\mathbf{Q}$, *a-letters* from $\mathbf{Y}$, and θ-letters from N copies $\mathbf{\Theta}_\mathbf{j}^+$ of $\mathbf{\Theta}^+$, i.e., for every $\theta \in \mathbf{\Theta}^+$, we have N generators $\theta_0, \theta_1, \ldots, \theta_N$, where $\theta_N = \theta_0$.

The relations of the group M correspond to the rules of the machine M; for every $\theta = [U_0 \to V_0, \ldots U_N \to V_N] \in \mathbf{\Theta}^+$, we have

$$U_i\theta_{i+1} = \theta_i V_i, \qquad \theta_i a = a\theta_i, \quad i = 1, ..., N \tag{10}$$

for every *word* a from the set of free generators of the subgroup $\mathbf{Y}_i(\theta) \le F(\mathbf{Y}_i)$. The first type of relations will be called (θ, q)-*relations*, the second type of relations are (θ, a)-*relations*.

Finally, the group $\tilde{G}$ is given by the generators and relations of the group M and two more relations, namely the *hub*-relations

$$\Sigma_0 = 1, \text{ and } \Sigma_1 = 1, \tag{11}$$

where Σ_0 is the accept word of the machine $\mathbf{M}$ and Σ_1 is the start word $\Sigma(1)$ (where the input subword u is empty).

Suppose a configuration W' of $\mathbf{M}$ is obtained from a configuration W by an application of a rule $\theta : [U_0 \to V_0, \dots, U_N \to V_N]$. This implies that $W = U_0 w_1 U_1 \dots w_N U_N$, $W' = V_0 w_1 V_1 \dots w_N V_N$, where every w_i is a word in the alphabet $\mathbf{Y_i}(\theta)$, and therefore $W' = \theta_0^{-1} W \theta_0$ in M by Relations (10), since $\theta_0 \equiv \theta_N$.

Now suppose $u \in \mathcal{L}$. Then u is recognized by the machine $\mathbf{M}$ by Lemma 7.4, and so the word $\Sigma(u)$ is accepted by the machine $\mathbf{M}$ and therefore it is conjugate to Σ_0 in the group M. Consequently, we have that $\Sigma(u) = 1$ in $\tilde{G}$ by (10, 11).

Note that the word $\Sigma(u)$ is obtained from Σ_1 by inserting the subword u. It follows from (11) that $u = 1$ in $\tilde{G}$ too. Since we identify the alphabet A of the input sector $t_0 t_1$ with the alphabet of the generators of $\tilde{G}$, Dyck's lemma implies the following.

Lemma 10.1. *The mapping $a \mapsto a$ $(a \in A)$ extends to a homomorphism of the group G to $\tilde{G}$.*

10.2 Bands and trapezia

To study (van Kampen) diagrams over the group M we shall use their simpler subdiagrams such as bands and trapezia, as in [20], [29], [3], etc.

We will consider *q-bands*, where the set $\mathcal{Z}$ from the definition in Subsection 2.2 is one of the sets $\mathbf{Q}_i$ of state letters for the S-machine $\mathbf{M}$, and *θ-bands* for every $\theta \in \mathbf{\Theta}$, where $\mathcal{Z} = \{\theta_0 . \theta_1, \dots, \theta_N\}$.

The following is a consequence of the definition of the relations of M.

Lemma 10.2. *([25], Lemma 5.4). Let $\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f}\mathbf{q}_2^{-1}$ be the boundary of a θ-band $\mathcal{T}$ with bottom $\mathbf{q}_1$ and top $\mathbf{q}_2$ in a reduced diagram.*

(1) If the start and the end edges $\mathbf{e}$ and $\mathbf{f}$ have different labels, then $\mathcal{T}$ has (θ, q)-cells.

(2) For every (θ, q)-cell π_i of $\mathcal{T}$, one of its boundary q-edges belongs to $\mathbf{q}_1$ and another one belongs to $\mathbf{q}_2$. □

To construct the top (or bottom) path of a $\mathcal{Z}$-band $\mathcal{T}$, at the beginning one can just form a product $\mathbf{x}_1 \dots \mathbf{x}_n$ of the top paths $\mathbf{x}_i$-s of the cells $\pi_1, \dots, \pi_n$ (where each π_i is a $\mathcal{Z}$-bands of length 1). If this product is not freely reduced, then one obtains the freely reduced form after a number of folding of the subsequent edges with mutually inverse labels.

Remark 10.3. *(1) No θ-letter is being canceled in the word $W \equiv Lab(\mathbf{x}_1) \dots Lab(\mathbf{x}_n)$ if $\mathcal{T}$ is a q-band since otherwise two neighbor cells of the band would make the diagram non-reduced.*

So for every subword $\theta(1)u\theta(2)$ of the reduced top/bottom label V of a q-band, the subword u in a-letters is uniquely defined by the pair $(\theta(1),\theta(2))$ of θ-letters.

We will call an arbitrary reduced word V in θ- and a-letters θ-defined if for every subwords $\theta(1)u\theta(2)$ and $\theta(3)u'\theta(4)$, the equalities $\theta(1)=\theta(3)$ and $\theta(2)=\theta(4)$ imply the equality of the middle a-words: $u'\equiv u$.

(2) By Lemma 10.2 (2), there are no cancellations of q-letters of W if $\mathcal{T}$ is a θ-band.

If $\mathcal{T}$ is a θ-band then a few cancellations of a-letters (but not q-letters) are possible in W. We will always assume that the top/bottom label of a θ-band is a reduced form of the word W.

In the next lemma, by trimmed word w we mean the maximal subword of w starting and ending with q-letters.

Lemma 10.4. *(1) The trimmed bottom and top labels W_1 and W_2 of any reduced θ-band $\mathcal{T}$ containing at least one (θ,q)-cell are $\mathbf{M}$-admissible and $W_2\equiv W_1\cdot\theta$.*

(2) If W_1 is a θ-admissible word and $W_2=W_1\cdot\theta$, then there is a reduced θ-band with trimmed bottom label W_1 and trimmed top label W_2.

Proof. (1) By Lemma 10.2 (2), we have $W_1\equiv q_1^{\pm1}u_1q_2^{\pm1}\dots u_kq_{k+1}^{\pm1}$, where $q_j^{\pm1}$ and $q_{j+1}^{\pm1}$ are the labels of q-edges of some cells $\pi(j)$ and $\pi(j+1)$ such that the subband connecting these cells has no (θ,q)-cells. Therefore by Lemma 10.2 (1), all the θ-edges of $\mathcal{T}$ between $\pi(j)$ and $\pi(j+1)$ have the same label. It follows from the list of (θ,a)-relations that the subwords $q_iu_iq_{i+1}$ of the word W_1 satisfy the definition of admissible word with $u_i\in F(\mathbf{Y}_{j(i)})$ if $q_i\in\mathbf{Q}_{j(i)}$ or $q_i^{-1}\in\mathbf{Q}_{j(i)+1}$. Hence W_1 is admissible, and so is W_2. The word u_i is a reduced form of a product $bc_1\dots c_sa$, where the words $c_1,\dots,c_s$ are the bottom labels of (θ,a)-cells of the band $\mathcal{T}$, and so $c_1\dots c_s\in\mathbf{Y}_{j(i)}(\theta)$, b is the subword in the part $*q_ib\to *$ (or in $b^{-1}q_i^{-1}*\to *$) of the rule θ and a is the subword of the part $aq_{i+1}*\to *$ (or in the part $*q_{i+1}^{-1}a^{-1}\to *$) of θ. In all these cases, the word u_i satisfies the requirement for applying of the rule θ. (For example, if $q_i\in\mathbf{Q}_{j(i)}$ and $q_{i+1}\in\mathbf{Q}_{j(i)+1}$, then $b^{-1}u_ia^{-1}\in\mathbf{Y}_{j(i)}(\theta)$, as required; the other cases are similar.) Thus, the word W_1 belongs to the domain of the rule θ. Likewise W_2 belongs to the domain of θ^{-1} and finally, $W_2=W_1\cdot\theta$ since the bottom and top labels of the (θ,q)-cells of $\mathcal{T}$ are, resp., the left-hand sides and the right-hand sides of the parts of θ.

(2) If we have $q_i\in\mathbf{Q}_{j(i)}$ and $q_{i+1}\in\mathbf{Q}_{j(i)+1}$, then one can construct a part of $\mathcal{T}$ taking two cells Π_i and Π_{i+1} corresponding to the (θ,q)-relations involving the letters q_i and q_{i+1}, respectively, and connecting Π_1 and Π_2 by a band of (θ,a)-cells, such that the bottom labels $c_1,\dots,c_s$ of these (θ,a)-cells form a product freely equal to the word $b^{-1}u_ic^{-1}$ (we use the notation of item (1) here). This is possible since by the definition of θ-application, we have $b^{-1}u_ic^{-1}\in\mathbf{Y}_{j(i)}(\theta)$. The trimmed bottom label of the obtained θ-band (including Π_1 and Π_1) is equal to $q_iu_iq_{i+1}$.

The reader can easily verify that the same conclusion is true if one or both letters q_i and q_{i+1} are negative. Since such a construction works for every $i=1,\dots,k$, one obtains a longer θ-band with trimmed bottom and top labels W_1 and W_2, resp. □

We say that a reduced θ-band is *regular* if the trimmed top and bottom labels of it are regular admissible words.

We shall consider the projections of words in the generators of M onto $\mathbf{\Theta}$ (all θ-letters map to the corresponding element of $\mathbf{\Theta}$, all other letters map to 1), and the projection onto the alphabet $\{\mathbf{Q}_0 \sqcup \ldots \sqcup \mathbf{Q_{N-1}}\}$ (every q-letter maps to the corresponding $\mathbf{Q}_i$, all other letters map to 1).

Definition 10.5. The projection of the label of a side of a q-band onto the alphabet $\mathbf{\Theta}$ is called the *history* of the band. The projection of the label of a side of a θ-band onto the alphabet $\{\mathbf{Q}_0, ..., \mathbf{Q}_{N-1}\}$ is called the *base* of the band, i.e., the base of a θ-band is equal to the base of the label of its top or bottom.

We call a $\mathcal{Z}$-band *maximal* if it is not contained in any other $\mathcal{Z}$-band. Counting the number of maximal $\mathcal{Z}$-bands in a diagram we will not distinguish the bands with boundaries $\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f}\mathbf{q}_2^{-1}$ and $\mathbf{f}\mathbf{q}_2^{-1}\mathbf{e}^{-1}\mathbf{q}_1$, and so every $\mathcal{Z}$-edge belongs to a unique maximal $\mathcal{Z}$-band.

We say that a $\mathcal{Z}_1$-band and a $\mathcal{Z}_2$-band *cross* if their connecting lines cross (i.e. these bands have a common cell) and $\mathcal{Z}_1 \cap \mathcal{Z}_2 = \emptyset$. We say that a band is an *annulus* if its connecting line is a closed curve. In this case the start and the end edges of the band coincide (see Figure 6a).

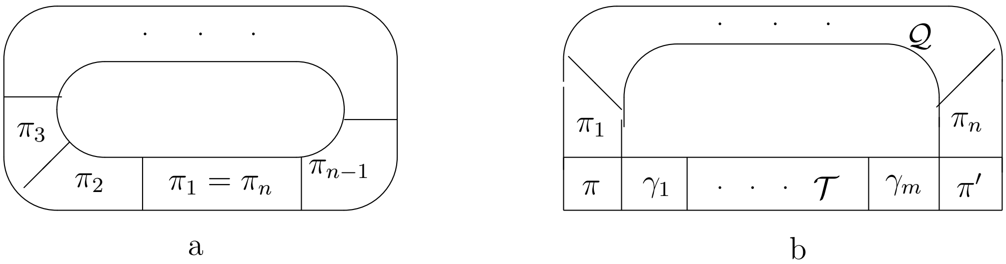

Fig. 6.

Sometimes we specify the types of bands as follows. A q-band corresponding to one of the letter $\mathbf{Q}_i$ of the base is called a $\mathbf{Q}_i$-band. For example, we will consider t_i-*band* (or just t-band) corresponding to the part $\{t_i\}$.

By definition, the group $H(\mathbf{Y})$ is the factor group of M obtaining from the presentation of M by adding of a-relations, i.e. all relations $w = 1$, where the word w in the alphabet $\mathbf{Y}^{\pm 1}$ is trivial in the free Burnside group $B_n(\mathbf{Y})$, and all relations $w = 1$, where w is a word in the subalphabet $A = \mathbf{Y}_1$ trivial in the group G. So there is a homomorphism $G \to H(\mathbf{Y})$. The $H(\mathbf{Y})$-diagrams are considered over this presentation. Let us call such a disk diagram Δ *minimal* $H(\mathbf{Y})$-diagram if it is reduced, every θ-band in it is regular and Δ contains no θ-annuli. A reduced diagram over the presentation of M (i.e. an $H(\mathbf{Y})$-diagram without a-cells) is said to be minimal if it has no θ-annuli.

Lemma 10.6. *For every disk diagram Δ over $H(\mathbf{Y})$ (over M), there is a minimal diagram Δ' over $H(\mathbf{Y})$ (resp., over M) with the same boundary label and the following properties. Every θ-band of Δ shares at most one cell with any q-band; Δ has no q-annuli.*

Proof. Proving by contradiction, we consider a counter-example Δ with minimal number of cells.

Assume that a θ-band $\mathcal{T}$ and a q-band $\mathcal{Q}$ cross each other two times. By the minimality of the counter-example, these bands have exactly two common cells π and π', and Δ has no cells outside the region bounded by $\mathcal{T}$ and $\mathcal{Q}$ (see Figure 7b).Then $\mathcal{Q}$ has exactly two cells since otherwise a maximal θ-band starting with a cell π'' of $\mathcal{Q}$, where $\pi'' \notin \{\pi, \pi'\}$, has to end on $\mathcal{Q}$, bounding with a part of $\mathcal{Q}$ a smaller counter-example. (We use that a θ-band cannot end on an a-cell.) For the similar reason, $\mathcal{T}$ has no (θ, q)-cells except for π and π'. Therefore by Lemma 10.2 (2), these two cells have the same labels of θ-edges, so these two neighbor in $\mathcal{Q}$ cells are just mirror copies of each other, since the label of a q-edge together with the label of a θ-edge completely determine the boundary label of a (θ, q)-cell. Thus, the diagram is not reduced, a contradiction.

If Δ has a q-annulus $\mathcal{Q}$, then the boundary $\partial\Delta$ is the outer boundary component of $\mathcal{Q}$, and there must be a θ-band starting and ending on $\partial\Delta$. It crosses $\mathcal{Q}$ two times, a contradiction.

Assume that Δ has a θ-annulus $\mathcal{T}$. Then $\mathcal{T}$ has no (θ, q)-cells since otherwise some q-band crosses the annulus $\mathcal{T}$ twice. Notice that every (θ, a)-cell corresponds to a commutator relation (10), and so the inner label and the outer label of $\mathcal{T}$ are equal. Hence one can identify the two boundary components of $\mathcal{T}$ and remove all the (θ, a)-cells of $\mathcal{T}$. Such a surgery preserves the boundary label of Δ but decreases the number of (θ, a)-cells. So after several surgeries of this type, we obtain the required minimal diagram Δ'.

Assume now that a diagram Δ over $H(\mathbf{Y})$ has a non-regular θ-band $\mathcal{T}$. Then by Lemma 10.2, there is a subband $\mathcal{T}'$, whose the first and the last cells π and π' have mutual inverse boundary labels, they are connected with a subband $\mathcal{T}''$ having no (θ, q)-cells, and the top/bottom labels of $\mathcal{T}''$ are congruent to 1. Then $\mathcal{T}''$ can be replaced with a subdiagram with a-cells only, so that π and π' get adjacent θ-edges. Hence this pair of cells becomes cancellable. Decreasing the number of (θ, q)-cells, we obtain the desired diagram Δ'. □

Corollary 10.7. *The homomorphism $G \to H(\mathbf{Y})$ given by the mapping $a \mapsto a$ is injective.*

Proof. One should prove that an equality $w = 1$ in $H(\mathbf{Y})$ for a word w in the alphabet $A^{\pm 1}$ implies the same equality in G.

By Lemma 10.6, there exists a minimal $H(\mathbf{Y})$-diagram Δ with boundary label w. Then neither θ-band nor q-band can start on $\partial\Delta$ since w has neither θ- nor q-edges. It follows from Lemma 10.6 that Δ has neither (θ, q)- nor (θ, a)-cells, and so the equality $w = 1$ follows from a-relations only.

Note that the replacement of letters from $\mathbf{Y} \backslash A$ with 1 in an a-relation provides us with an a-relation again. So the equality $w = 1$ is a consequence of a-relations depending on letters from A only. These a-relations hold in G since G is a group of exponent n. Hence we get $w = 1$ in G. □

Definition 10.8. *Let Δ be a reduced diagram over M (resp., a minimal diagram over $H(\mathbf{Y})$), which has boundary path of the form $\mathbf{p}_1^{-1}\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2^{-1}$, where $\mathbf{p}_1$ and $\mathbf{p}_2$ are sides of q-bands, and*

$\mathbf{q}_1$, $\mathbf{q}_2$ *are maximal parts of the sides of θ-bands such that $Lab(\mathbf{q}_1)$, $Lab(\mathbf{q}_2)$ start and end with q-letters.*

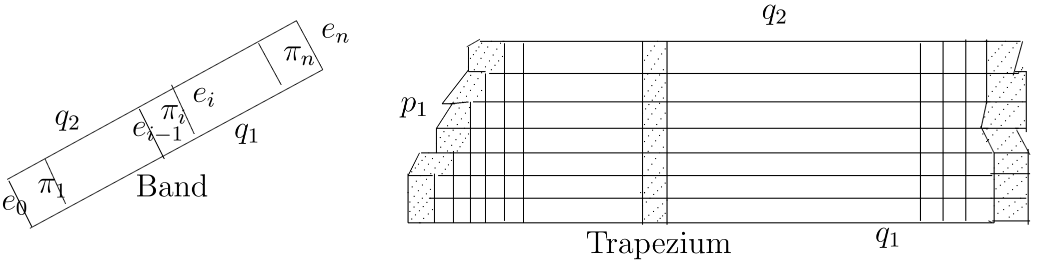

Figure 7

Then Δ is called a trapezium *(respectively, quasi-trapezium). The path $\mathbf{q}_1$ is called the* bottom*, the path $\mathbf{q}_2$ is called the* top *of the trapezium, the paths $\mathbf{p}_1$ and $\mathbf{p}_2$ are called the* left and right sides *of the (quasi-)trapezium. The history of the q-band whose side is $\mathbf{p}_2$ is called the* history *of the (quasi-)trapezium; The base of $Lab(\mathbf{q}_1)$ is called the* base *of the (quasi-)trapezium.*

Remark 10.9. *Notice that the top (bottom) side of a θ-band $\mathcal{T}$ does not necessarily coincide with the top (bottom) side $\mathbf{q}_2$ (side $\mathbf{q}_1$) of the corresponding trapezium of height 1, and $\mathbf{q}_2$ ($\mathbf{q}_1$) is obtained from $\mathbf{top}(\mathcal{T})$ (resp. $\mathbf{bot}(\mathcal{T})$) by trimming the first and the last a-edges if these paths start and/or end with a-edges.*

By Lemma 10.6, any maximal θ-band of a (quasi-)trapezium Δ connects the left side and the right side of Δ. So one can enumerate them from the bottom to the top of Δ: $\mathcal{T}_1, ..., \mathcal{T}_h$. The following lemma claims that every trapezium (quasi-trapezium) simulates a computation of $\mathbf{M}$ (resp., quasi-computation). (Similar statements can be found in [23]. For the formulations (1) and (3) below, it is important that $\mathbf{M}$ is an S-machine. The analog of this statement is false for Turing machines. - See [22] for a discussion.)

Lemma 10.10. *(1) Let Δ be a trapezium with history $\theta_1 \dots \theta_d$ ($d \geq 1$). Assume that Δ has consecutive maximal θ-bands $\mathcal{T}_1, \dots \mathcal{T}_d$, and the words U_j and V_j are the trimmed bottom and*

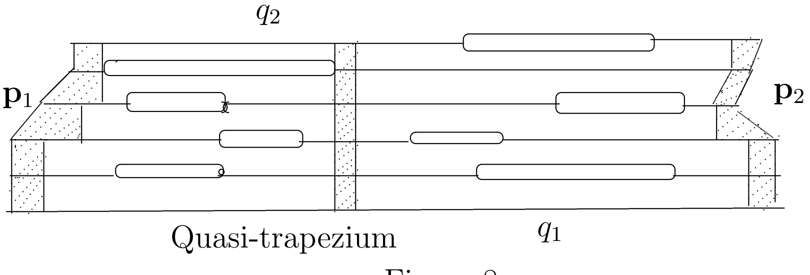

Figure 8

the trimmed top labels of $\mathcal{T}_j$, $(j = 1, \ldots, d)$. *Then the history of* Δ *is a reduced word,* U_j, V_j *are admissible words for* M, *and*

$$V_1 = U_1 \cdot \theta_1,\ U_2 \equiv V_1, \ldots, U_d \equiv V_{d-1},\ V_d = U_d \cdot \theta_d.$$

(2) For every reduced computation $U \to \ldots \to U \cdot h \equiv V$ *of* $\mathbf{M}$ *with* $|h| \geq 1$ *there exists a trapezium* Δ *with bottom label* U, *top label* V, *and with history* h.

(3) Let Δ *be a quasi-trapezium with history* $\theta_1 \ldots \theta_d$ $(d \geq 1)$ *and consecutive maximal* θ*-bands* $\mathcal{T}_1, \ldots \mathcal{T}_d$. *Let the words* U_j *and* V_j *are the trimmed bottom and the trimmed top labels of regular* θ*-bands* $\mathcal{T}_j$, $(j = 1, \ldots, d)$. *Then the history of* Δ *is a reduced word,* U_j, V_j *are regular admissible words for* M, *and*

$$V_1 = U_1 \cdot \theta_1,\ U_2 \cong V_1, \ldots, U_d \cong V_{d-1},\ V_d = U_d \cdot \theta_d.$$

(4) For every reduced quasi-computation $U_0 \to U_1 \to \ldots \to U_d$ *with history* $h = \theta(1) \ldots \theta(d)$ *where* $d \geq 1$ *there exists a quasi-trapezium* Δ *with bottom label* U_0, *top label* U_d, *and with history* h.

Proof. (1) The equalities $V_i = U_i \cdot \theta_i$ follow from Lemma 10.4 (1). The equalities $V_{i-1} \equiv U_i$ are true because every cell of Δ is a θ-cell, and so there are no cells between the top of the band $\mathcal{T}_{i-1}$ and the bottom of $\mathcal{T}_{i-1}$.

(2) Given a reduced computation $\mathcal{C} : U \equiv W_0 \to \ldots \to W_s \equiv V$, by Lemma 10.4 (2), one can construct a reduced band $\mathcal{T}_i$ for every subcomputation $W_{i-1} \to W_i$. Since the trimmed top label W_i of $\mathcal{T}_i$ coincides with the trimmed bottom label of $\mathcal{T}_{i+1}$, the obtained bands can be glue up. The obtained diagram Δ is reduced since the history of $\mathcal{C}$ is reduced, and so the cells from different θ-band cannot form reducible pairs. Clearly, the history of $\mathcal{C}$ is equal to the history of the constructed trapezium.

(3) We have regular θ-bands $\mathcal{T}_i$-s corresponding to the letters θ_i of the history and the equalities $V_i = U_i \cdot \theta_i$ as in the proof of Statement (1), but now there are *a-cells* corresponding to a-relations between the neighbor bands $\mathcal{T}_i$ and $\mathcal{T}_{i+1}$. Since the q-edges of the (θ, q)-cells from the neighbor θ-bands belong to the same q-bands, we have the equalities of the form $v_{ij} = u_{i+1,j}$ modulo the a-relations for the corresponding sector words of the words V_i and U_{i+1}. However v_{ij} and $u_{i+1,j}$ are the words in the same subalphabet $\mathbf{Y}_k \subset \mathbf{Y}$ since the words V_i and U_{i+1} are admissible by Lemma 10.4 (1). They are regular since every θ-band of a quasi-trapesium is regular. Hence the homomorphism mapping all the letters from $\mathbf{Y} \backslash \mathbf{Y_k}$ to 1 shows that the words v_{ij} and $u_{i+1,j}$ are equal modulo the Burnside relations of the group $B_n(\mathbf{Y}_k)$ (and modulo the relations of G if the sector is the input one for $\mathbf{M}$). Hence $U_{i+1} \cong V_i$, as required.

(4) As in the proof of Statement (2), one can start constructing Δ with the subsequent regular θ-bands $\mathcal{T}_1, \ldots$ identifying the q-edges of the neighbor bands. The congruence $W_i \cong W_{i-1} \cdot \theta_i$ implies that for the corresponding sector subwords $v_{i-1,j}$ and u_{ij} of these two words, we have equalities modulo the Burnside relations of $B_n(\mathbf{Y}_k)$, $k = k(j)$ (or modulo G-relations). It remains to glue up each hole having boundary label $v_{i-1,j} u_{ij}^{-1}$ with the corresponding a-cell. □

11 More Burnside relations

11.1 Burnside relations without q-letters

In this subsection, we study presentations of some auxiliary groups. Let us start with the group $H(a)$ generated by the set $\cup_{j=1}^N \mathbf{\Theta}_\mathbf{j}^+ \cup \mathbf{Y}$ and defined by all (θ, a)-relations and all a-relations. It follows from these relations that there is a retraction of $H(a)$ onto the free group $F(\cup_j \mathbf{\Theta}_\mathbf{j}^+)$. Furthermore, $H(a)$ is the multiple HNN-extension of the subgroup $K(a)$ generated by $\mathbf{Y}$ with stable letters from $\cup_j \mathbf{\Theta}_\mathbf{j}^+$. If θ_i belongs to some $\mathbf{\Theta}_\mathbf{i}^+$, then the subgroup $\tilde{\mathbf{Y}}_i(\theta)$ of $K(a)$ associated with θ_i is the image of the subgroup $\mathbf{Y}_i(\theta)$ in the free Burnside group $B_n(\mathbf{Y}_i)$ (in G if $i=1$) and the conjugation by θ_i is identical on $\tilde{\mathbf{Y}}_i(\theta)$.

Lemma 11.1. *Every subgroup $\tilde{\mathbf{Y}}_i(\theta)$ is malnormal in the group $K(a)$.*

Proof. The group $K(a)$ is generated by the set $\mathbf{Y}$ of all a-letters. The a-relations can be imposed as follows. At first we impose the relations on the words in the subalphabet $\mathbf{Y}_1 = A$ and impose the Burnside relations on the words in the subalphabet $\mathbf{Y}_j$ for every $j = 2, \dots, N$. We obtain the free product P of G and free Burnside groups $B_n(\mathbf{Y}_j)$, $j = 2, \dots, N$. Then imposing all the Burnside relations on P we obtain the group $K(a) \cong P/P^n$, which is called the free product of the groups G and $B_n(\mathbf{Y}_j)$ ($j = 2, \dots, N$) in the Burnside variety $\mathcal{B}_n$. A few properties of free multiplication in $\mathcal{B}_n$ can be found in Subsection 36.1 of [19]. The operation $o_{i\in J}^n G_j$ studied there coincides with the free product operation in $\mathcal{B}_n$ if the factors G_j-s belong to $\mathcal{B}_n$. In particular, we have $G_j \cap xG_jx^{-1} = \{1\}$ if $x \notin G_j$, i.e. the factors G_j of a free product in $\mathcal{B}_n$ are malnormal subgroups. In particular we have that the subgroups G and $B_n(\mathbf{Y}_j)$ ($j = 2, \dots, N$) are malnormal in $K(a)$.

It follows from the definition of the machine $\mathbf{M}$, that every subgroup $\mathbf{Y}_j(\theta)$ is either equal to the free group $F(\mathbf{Y}_j)$ or trivial, or is the copy of some $H(ij)$ (we have no latter variant if $j = 1$, i.e. if $\mathbf{Y}_j(\theta) = G$). Therefore its image $\tilde{\mathbf{Y}}_i(\theta)$ is either obviously malnormal in G or in $B_n(\mathbf{Y}_j)$ or it is malnormal in one of these groups by Lemma 4.6. Now the statement of the lemma follows from the previous paragraph since the malnormality is a transitive property. □

We will verify the conditions (Z1), (Z2), and (Z3) from Section 3.1 for the group $H(a)$ under the assumption that a-letters are zero-letters and θ-letters are non-zero letters. So one can talk of reduced, cyclically reduced and essential words as they were defined in Section 3.1. Now $\mathcal{R}_0 = \mathcal{S}_0$ is the union of the set (θ, a)-relations and the set of a-relations; the set $\mathcal{S}_{1/2}$ is empty.

Lemma 11.2. *The presentation of $H(a)$ satisfies the conditions (Z1), (Z2) and (Z3).*

Proof. Since the set $\mathcal{S}_{1/2}$ is empty, Condition (Z2) holds. Condition (Z1) follows from the definition of $H(a)$. It remains to verify Condition (Z3).

Assume that g is an essential cyclically reduced word and for some $x \in K(a)\backslash\{1\}$, $g^{-4}xg^4$ is also a 0-element, that is $g^{-4}xg^4 \in K(a)$. Then there are no pinches in g^4 regarded as a product in the HNN-extension $H(a)$, and every time one obtains only one pinch in the middle when reducing the product $g^{-4}xg^4$ to the normal form in the HNN-extension. Hence the product $g^{-1}xg = x'$ is a 0-element too.

Let $g = y_0\theta(1)y_1\dots\theta(k)y_k$ be the normal form of g in the HNN extension $H(a)$, where $\theta(1),\dots,\theta(k)$ are stable letters or their inverses ($k \geq 1$). Then the element $y_0^{-1}xy_0$ must belong to the subgroup $\tilde{\mathbf{Y}}_i(\theta)$ for some i and θ, and so $\theta(1)$ commutes with $y_0xy_0^{-1}$ in $H(a)$. Similarly we obtain that $\theta(2)$ commutes with $y_1^{-1}y_0^{-1}xy_0y_1$, and so on, and $x' = y^{-1}xy$, where $y = y_0y_1\dots y_k$.

We also have the next pinch $\theta(1)^{-1}y_0^{-1}x'y_0\theta(1)$, and therefore the product

$$y_0^{-1}x'y_0 = y_0^{-1}y^{-1}xyy_0 = (y_0^{-1}y^{-1}y_0)(y_0^{-1}xy_0)(y_0^{-1}yy_0)$$

is also in $\tilde{\mathbf{Y}}_i(\theta)$. By Lemma 11.1, we have $y_0^{-1}yy_0 = y_1\dots y_ky_0 \in \tilde{\mathbf{Y}}_i(\theta)$ as well. Hence $\theta(1)$ commutes with $y(1)^{-s}(y_0^{-1}xy_0)y(1)^s$ for every s, where $y(1) = y_1\dots y_ky_0$.

Changing g by a cyclic permutations, we similarly obtain that arbitrary $\theta(m)$ commutes with $y(m)^{-s}(y_0y_1\dots y_{m-1})^{-1}x(y_0y_1\dots y_{m-1})y(m)^s$ for every s, where $y(m) = y_m\dots y_ky_0\dots y_{m-1}$. Using these commutativity relations k times and taking into account that every $y(m)$ is a cyclic permutation of the product y, we have

$$g^{-1}y^{1-s}xy^{s-1}g = (y_0\theta(1)y_1\dots\theta(k)y_k)^{-1}y^{1-s}xy^{s-1}(y_0\theta(1)y_1\dots\theta(k)y_k) =$$

$$(y_1\theta(2)\dots\theta(k)y_k)^{-1}y(1)^{1-s}y_0^{-1}xy_0y(1)^{s-1}(y_1\theta(2)\dots\theta(k)y_k) = \dots = y^{-s}xy^s,$$

because $y_0\dots y_{k-1}y(k)^{s-1}y_k = y^s$. Hence by induction on s, we obtain equalities $g^{-s}xg^s = y^{-s}xy^s$ for every $s \geq 0$. Since $y \in K(0)$, Conditions $(Z3.1)$ holds. Condition $(Z3.2)$ follows as well since the definition of y does not depend on the choice of x in $\mathbf{0}(g)$. □

Lemma 11.2 and Lemma 3.4 provide us with

Corollary 11.3. *There exists a graded presentation of the factor group $H(a,\infty)$ of $H(a)$ over the subgroup generated by all n-th powers of elements of $H(a)$, such that every g-reduced diagram over this presentation satisfies Property A from Section 2.3.*

□

We denote by $H(\theta,a)$ the group $H(a,\infty)$ from Corollary 11.3, that is the group generated by $\cup_j\mathbf{\Theta}_\mathbf{j}^+ \cup \mathbf{Y}$, which is subject to all Burnside relations $w^n = 1$ in this alphabet (and so $H(\theta,a) \in \mathcal{B}_n$), all relations of G in the subalphabet A, and all (θ,a)-relations.

Lemma 11.4. *The mapping identical on the set of generators A of the group G and trivial on other generators from $\cup\mathbf{\Theta}_\mathbf{j}^+ \cup \mathbf{Y}$ is a retraction of the group $H(\theta,a)$ onto the subgroup G.*

Proof. Indeed, this mapping preserves all defining relations of $H(\theta,a)$ since $G^n = \{1\}$. □

Lemma 11.5. *Let $\theta(1),\dots,\theta(m)$ be different rules from Θ^+ and $\theta_j(1),\dots,\theta_j(m)$ be their copies from some set $\mathbf{\Theta}_\mathbf{j}^+$. Assume that $x(i) = c(i)\theta_j(i)d(i)$ for some group words c_i, d_i over $\mathbf{Y}$ $(i = 1,\dots m)$. Then*

(a) the words $x(i)$ freely generate a free Burnside subgroup, which is a retract of $H(\theta,a)$,

(b) The mapping $x(i) \mapsto \theta(i)$ $(i = 1,\dots m)$ defines an isomorphism between the free Burnside groups $B_n(x(1),\dots,x(m))$ and $B_n(\theta(1),\dots,\theta(m))$.

Proof. We preserve the set of defining relations of $H(\theta, a)$ if we replace all the generators, except for $\theta_j(1), \ldots, \theta_j(m)$, with 1. Hence there is a homomorphism of $H(\theta, a)$ onto the free Burnside group $B_n(\theta_j(1), \ldots, \theta_j(m))$ such that $x(i) \mapsto \theta_j(i)$. The mapping $\theta_j(i) \mapsto x(i)$ $(i = 1, \ldots, m)$ also defines a homomorphism since all the relations $w^n = 1$ hold in $H(\theta, a)$, which proves the lemma. $\square$

Corollary 11.6. *(a) A product of several letters $x(i)^{\pm 1}$ $(i = 1, \ldots, m)$ is minimal (cyclically minimal) word w in $H(\theta, a)$ if and only if the image of w under the mapping $x(i) \to \theta(i)$ $(i = 1, \ldots m)$ is minimal (cyclically minimal) in $B_n(\mathbf{\Theta})$.*

(b) The top label of a q-band is trivial in the group $H(\theta, a)$ if and only if the bottom label is trivial in $H(\theta, a)$.

Proof. (a) This follows from Lemma 11.5, because a-letters have length 0.

(b) The top label is freely equal to a product of the words having the form $x(i)^{\pm 1} = (c(i)\theta(i)d(i))^{\pm 1}$ for $\theta \in \mathbf{\Theta}^+_{\mathbf{j}}$, where $c(i), d(i)$ are words over $\mathbf{Y}$. The bottom label is a product of the words $y(i)$ of the similar form with middle terms from $\mathbf{\Theta}^+_{\mathbf{j+1}}$, or from $\mathbf{\Theta}^+_{\mathbf{j-1}}$ and the corresponding factors of these products have the same θ in the middle. It follows from Lemma 11.5 (a) that the top label is trivial in $H(\theta, a)$ if and only if the bottom one is trivial. $\square$

We say that a reduced (disk or annular) diagram over $H(a)$ is *minimal* if it contains no θ-annuli. Similarly, a reduced diagram Δ over $H(\theta, a)$ is *minimal* if it is g-reduced and has no θ-annuli.

Lemma 11.7. *(1) For every (disk or annular) diagram Δ over $H(a)$ (over $H(\theta, a)$), there is a minimal diagram Δ' over $H(a)$ (over $H(\theta, a)$) with the same boundary label(s).*

(2) If Δ is an annular diagram with boundary components $\mathbf{p}$ and $\mathbf{q}$, a path $\mathbf{s}$ connects the vertices $\mathbf{p}_-$ and $\mathbf{q}_-$ in Δ, and the word $Lab(\mathbf{p})$ is not a conjugate of a word of θ-length 0 in $H(a)$ (resp., in $H(\theta, a)$), then there is a path $\mathbf{s}'$ in the diagram Δ', such that $Lab(\mathbf{s}) = Lab(\mathbf{s}')$ in $H(\theta, a)$.

Proof. In a disk diagram Δ, θ-bands can be eliminated exactly as in the proof of Lemma 10.6 since their top and bottom have the same label. The same trick works for annular diagrams, since if a θ-annulus surrounds the hole of Δ, then we have an annular subdiagram with boundary components $\mathbf{p}$ and the top (or the bottom) of this annulus. It follows that $Lab(\mathbf{p})$ is conjugate to a word of θ-length 0, a contradiction.

Note that the modifications of Δ (i.e. eliminating of θ-band and g-reductions) are 'local', i.e. the path $\mathbf{s}$ can be replaced with a homotopic path which does not cross the modified subdiagram. Thus, Statement (2) follows. $\square$

Let now $H(\mathbf{\Theta})$ be the factor group of $H(\mathbf{Y})$ over the all relations $w^n = 1$, where w is a word in generators $\mathbf{\Theta} \cup \mathbf{Y}$. In other words, $H(\mathbf{\Theta})$ is the factor group of the free product $H(\theta, a) \star F(\mathbf{Q})$ over all (θ, q)-relations.

We will call a reduced disk diagram Δ over $H(\mathbf{\Theta})$ *minimal* if it has no q-annuli and every subdiagram of Δ over $H(\theta, a)$ or over $H(\mathbf{Y})$ is minimal.

Lemma 11.8. *For every disk diagram Δ over $H(\mathbf{\Theta})$, there exists a minimal diagram Δ' with the same boundary label.*

Proof. Assume Δ contains a q-annulus $\mathcal{Q}$. Proving by contradiction we may assume that the inner subdiagram Δ_1 of $\mathcal{Q}$ has no q-annuli. Since the top/bottom of $\mathcal{Q}$ has no q-edges, Δ_1 has no (θ, q)-cells, i.e., it is a diagram over $H(\theta, a)$. So the label of the inner boundary component of $\mathcal{Q}$ has trivial in $H(\theta, a)$ label. By Corollary 11.6 (b), the label of the outer boundary component of the annulus $\mathcal{Q}$ is also trivial in $H(\theta, a)$. So the annulus $\mathcal{Q}$ together with Δ_1 can be replaced with a diagram without (θ, q)-cells. One can continue decreasing the number of (θ, q)-cells and obtain the required minimal diagram Δ' using Lemmas 11.8 and 10.6. □

Lemma 11.9. *The canonical mapping $a \mapsto a$ defines an embedding of the group G in $H(\mathbf{\Theta})$.*

Proof. This mapping defines a homomorphism $G \to H(\mathbf{\Theta})$ by Corollary 10.7. Assume that a word v in the generators A of G is trivial in $H(\mathbf{\Theta})$ and consider a minimal diagram Δ for this equality. By Lemma 11.8, it has no q-annuli, and therefore Δ has no (θ, q)-cells at all since the boundary $\partial\Delta$ has no q-edges. Hence $v = 1$ in the group $H(\theta, a)$. By Lemma 11.4, $v = 1$ in G, which completes the proof. □

Lemma 11.10. *The mapping sending every θ-generator $\theta_j \in \mathbf{\Theta}_{\mathbf{j}}^+$ of $H(\mathbf{\Theta})$ (of $H(\mathbf{Y})$) to the corresponding history letter $\theta \in \mathbf{\Theta}^+$ and every q- and a-generator of $H(\mathbf{\Theta})$ (of $H(\mathbf{Y})$) to* 1*, extends to a homomorphism of $H(\mathbf{\Theta})$ (resp., of $H(\mathbf{Y})$) to the free Burnside group $B_n(\mathbf{\Theta}^+)$ (to the free group $F(\mathbf{\Theta}^+)$, resp.).*

For every j, the restriction of this mapping to the subgroup generated by $\mathbf{\Theta}_{\mathbf{j}}^+$ is an isomorphism.

Proof. The image of every relator of $H(\mathbf{\Theta})$ under this mapping is either 1 or a word of the form u^n. Therefore the first statement follows from Dyck's lemma. For the restriction to $H(\mathbf{\Theta}_{\mathbf{j}}^+)$, there is an obvious inverse homomorphism given by $\theta \mapsto \theta^{(j)}$, which proves the second statement of the lemma for $H(\mathbf{\Theta})$. The proof for the group $H(\mathbf{Y})$ is similar. □

11.2 Diagrams with one maximal q-band

Lemma 11.11. *(1) Let Δ be a g-reduced annular diagram over $H(\mathbf{\Theta})$ with boundary components $\mathbf{s}_1$ and $\mathbf{t}_1\mathbf{t}_2\mathbf{s}_2$, containing a unique maximal q-band $\mathcal{Q}$ with history h, start/end edges $\mathbf{t}_1$ and $\mathbf{t}_2$ and boundary $\mathbf{t}_1\mathbf{y}_1\mathbf{t}_2\mathbf{y}_2$, where $\mathbf{s}_1$ is contained in the region bounded by the closed path $\mathbf{y}_1$ (left part of Figure 9). Assume that $h \neq 1$ in the free Burnside group $B_n(\Theta^+)$ and the annular subdiagram Δ_1 with boundary components $\mathbf{s}_1$ and $\mathbf{y}_1$ is minimal over $H(\theta, a)$ and contains a positive cell π with contiguity degree $\geq \varepsilon$ to $\mathbf{y}_1$ via a contiguity subdiagram Γ of rank 0, where $\mathbf{y}_1$ is regarded as the path starting and ending at the vertex $(\mathbf{t}_1)_+$.*

Then there exists a minimal annular diagram Δ' with boundary components $\mathbf{s}'_1$ and $\mathbf{t}'_1\mathbf{t}'_2\mathbf{s}'_2$, where $Lab(\mathbf{s}'_1) \equiv Lab(\mathbf{s}_1)$, $Lab(\mathbf{s}'_2) \equiv Lab(\mathbf{s}_2)$, $Lab(\mathbf{t}'_1) \equiv Lab(\mathbf{t}_1)$, $Lab(\mathbf{t}'_2) \equiv Lab(\mathbf{t}_2)$ (right part of Figure 9), such that *(a) Δ' has a unique maximal q-band $\mathcal{Q}'$ with boundary $\mathbf{t}'_1\mathbf{y}'_1\mathbf{t}'_2\mathbf{y}'_2$,*

(b) the history h' of $\mathcal{Q}'$ is equal to h modulo the Burnside relations,

(c) the type of the annular subdiagram Δ'_1 over $H(\theta, a)$ with boundary components $\mathbf{s}'_1$ and $\mathbf{y}'_1$ is less than $\tau(\Delta_1)$.

(d) If a simple path $\mathbf{z}$ connects the vertices $(\mathbf{s}_1)_-$ and $(\mathbf{s}_2)_-$ in Δ_1, then there exists a simple path $\mathbf{z}'$ connecting the vertices $(\mathbf{s}'_1)_-$ and $(\mathbf{s}'_2)_-$ in Δ'_1 such that the words $Lab(\mathbf{z}')$ and $Lab(\mathbf{z})$ are equal in the group $H(\theta, a)$.

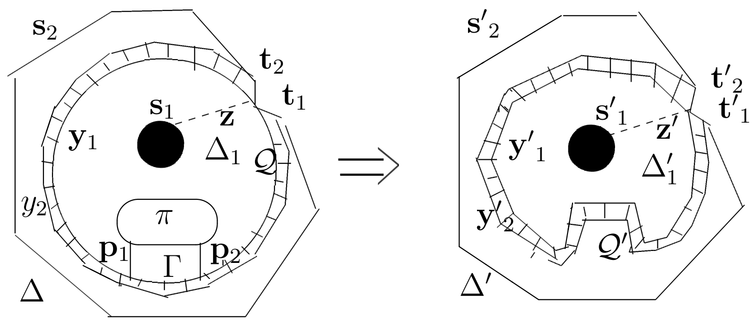

Figure 9

(2) Similar statement holds for disk diagrams Δ, Δ_1 and Δ' and a single maximal q-band $\mathcal{Q}$, where Δ has the boundary $\mathbf{t}_1\mathbf{s}_1\mathbf{t}_2\mathbf{s}_2$ and $\mathcal{Q}$ has boundary $\mathbf{t}_1\mathbf{y}_1\mathbf{t}_2\mathbf{y}_2$ (Figure 10).

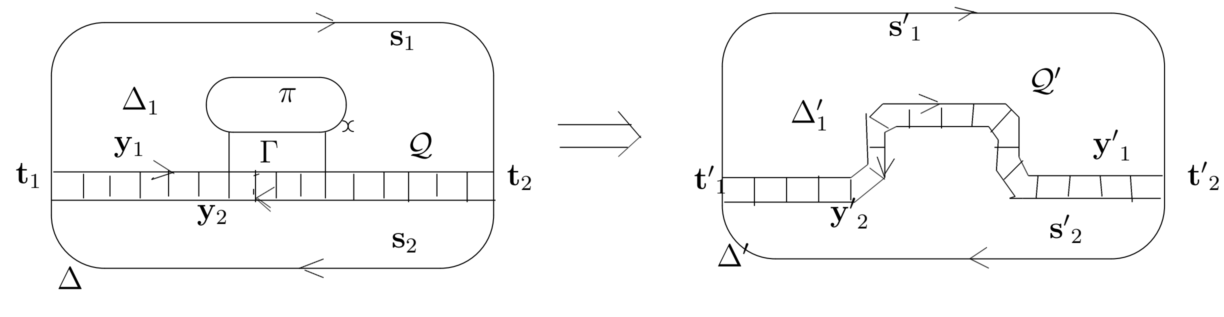

Figure 10

Proof. (1) Let $\partial(\pi, \Gamma, \mathbf{y}_1) = \mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$; then $Lab(\mathbf{q}_1)$ is a $u^{\pm 1}$-periodic word, where u^n is the boundary label of $\partial\pi$. We will assume that it is u-periodic. So $|\mathbf{q}_1|_\theta > \varepsilon n|u|_\theta$ and $|\mathbf{p}_1|_\theta = |\mathbf{p}_2|_\theta = 0$ since $r(\Gamma) = 0$ (when regarding a-edges as 0-edges).

Since u is a simple period of some rank $k > 0$, no θ-band can start and also end on $\mathbf{q}_1$. The same property holds for the subpath $\mathbf{q}_2$ of $\mathbf{y}_1$ by Lemma 10.6 (1). Therefore every maximal θ-band of Γ starts on $\mathbf{q}_1$ and ends on $\mathbf{q}_2$. It follows that the words $Lab(\mathbf{q}_1)$ and

$Lab(\mathbf{q}_2)$ have equal θ-projections; in particular, $|\mathbf{q}_2|_\theta = |\mathbf{q}_1|_\theta$ and $Lab(\mathbf{q}_2^{-1})$ is a v-periodic word with $|v|_\theta = |u|_\theta$, because the side label of a reduced q-band is θ-defined (see Remark 10.3 (1)). Thus, the word v can be chosen so that the θ-projection of v is the θ-projection of u.

Replacing now Γ with a smaller contiguity diagram Γ' with $\partial\Gamma' = \mathbf{p}'_1\mathbf{q}'_1\mathbf{p}'_2\mathbf{q}'_2$, we may assume that $Lab(\mathbf{q}'_1) \equiv u^l$, $Lab(\mathbf{q}'_2)^{-1} \equiv v^l$, $|\mathbf{q}'_1|_\theta = |\mathbf{q}'_2|_\theta$ and $l > \frac{\varepsilon n}{2}$. Therefore

$$x_1 u^l = v^l x_2 \tag{12}$$

in the group $H(a)$, where $|x_1|_\theta = |x_2|_\theta = 0$. The word v is cyclically reduced since the history h is reduced.

Remark 11.12. *If $\mathcal{Q}_0$ is a subband of $\mathcal{Q}$ with history h_0^2 corresponding to v^2 and $\mathcal{Q}_0 = \mathcal{Q}_1\mathcal{Q}_2$, where $\mathcal{Q}_1$ has history h_0, then the top and the bottom of $\mathcal{Q}_1$ have equal labels since both occurrences of h_0 start with the same rule. This makes possible the following surgery: one can construct a q-band by inserting, between $\mathcal{Q}_1$ and $\mathcal{Q}_2$, a q-band with history of the form h_0^d.*

Assume that the word v is not simple in rank 0. Then by Lemma 3.6, there is a simple in rank 0 word w such that for some word x, we have

$$v = xw^b g x^{-1} \tag{13}$$

in rank 0, where $g \in \mathbf{0}(w)$ and $|w|_\theta \le |v|_\theta$. Since the word v is cyclically reduced in rank 0, one can choose x with θ-length 0. Indeed, otherwise we have that either the product xw or the product wgx^{-1} (and therefore wx^{-1} since $g \in \mathbf{0}(w)$) has a 2-letter subword equal to a 0-word in rank 0. It follows that xwx^{-1} is equal to $yw'y^{-1}$, where $|w'|_\theta = |w|_\theta$ and $|y|_\theta < |x|_\theta$. By Remark 3.2, the word w' is also simple. We obtain $v = yw'^b g' y^{-1}$, with $g' = (y^{-1}x)g(y^{-1}x)^{-1}$ and $g' \in \mathbf{0}(w')$ (see Condition (Z3)). So the element x in (13) can be shorten.

Now it follows from (13) that $v^l = xw^{bl}x'x^{-1}$ for some $x' \in \mathbf{0}(w)$ since w normalizes the subgroup $\mathbf{0}(w)$. Together with (12), this implies the equality

$$v^l x_2 = xw^{bl}x'x^{-1}x_2 = x_1 u^l \tag{14}$$

in rank 0, where both u and w are simple in rank 0 and x_2, x, x', x_1 have θ-length 0. Since $l > \frac{\varepsilon n}{2} > \delta^{-1}$, we can apply Lemma 3.7 to equality (14) and conclude that $u = x''w^{\pm 1}x_3(x'')^{-1}$ in rank 0 for some word x'' and some $x_3 \in \mathbf{0}(w)$. We obtain a contradiction because u is simple in rank 0 but $|w|_\theta < |v|_\theta = |u|_\theta$.

Thus, v is a simple word of rank 0, and we may apply Lemma 3.7 to the diagram Γ' with periodic labels u^l and v^{-l} of the contiguity arcs. It follows that $v = Lab(\mathbf{p}'_1)u^{\pm 1}zLab(\mathbf{p}'_1)^{-1}$ in rank 0 for some $z \in \mathbf{0}(u)$. By Lemma 3.1 for $H(a)$, we have $(u^{\pm 1}z)^n = u^{\pm n}$. Hence

$$v^n = Lab(\mathbf{p}'_1)u^{\pm n}Lab(\mathbf{p}'_1)^{-1} \tag{15}$$

in $H(a)$.

The right-hand side of (15) is the label of the closed path starting at the vertex $(\mathbf{p}'_1)_-$ and bounding $\partial\pi$. So one can insert it in $\mathbf{y}_1$ to bypath π and obtain an annular diagram E of smaller type than Δ_1. To obtain Δ'_1 we glue up the diagram of rank 0 corresponding to the equality (15) to E and obtain the desired Δ'_1.

However passing from Δ_1 to Δ'_1, we insert the subpath labeled by v^n in $\mathbf{y}_1$. So we have to change the q-band $\mathcal{Q}$ accordingly, i.e. using Remark 11.12, we insert in $\mathcal{Q}$ a q-band with history $h_0^{\pm n}$ with possible reduction of the history of the obtained q-band $\mathcal{Q}'$. Now we see that the path $\mathbf{y}_2$ is modified too. But the label of $\mathbf{y}'_2$ is equal to the label of $\mathbf{y}_2$ modulo the relations of $H(\theta, a)$, and so one can glue up a number of cells corresponding to these relations to $\mathbf{y}'_2$ so that these additional cells and the diagram Δ_2 with boundary $\mathbf{s}_2\mathbf{y}_2$ form a diagram Δ'_2.

The diagram Δ' is built from Δ'_1, $\mathcal{Q}'$ and Δ'_2 (with subsequent reductions if needed). Property (d) follows from Lemma 11.7 (2). Indeed, if $Lab(\mathbf{s}_1)$ is conjugate of a word of length 0 in $H(\theta, a)$, then so is $Lab(\mathbf{q}_1)$. Then Lemmas 11.5 (b) and 11.10 imply that h is a conjugate of 1 in $B_n(\Theta^+)$, despite of the assumption of the lemma.

(2) The proof is similar. □

Lemma 11.13. *Let $\mathcal{D}$ be a reduced q-band with boundary $\mathbf{x}_1\mathbf{y}_1\mathbf{x}_2\mathbf{y}_2$, where $\mathbf{y}_1$ and $\mathbf{y}_2$ are the top and the bottom of $\mathcal{D}$. Then*

(a) there exists a q-band $\mathcal{D}'$ with boundary $\mathbf{x}'_1\mathbf{y}'_1\mathbf{x}'_2\mathbf{y}'_2$, where $Lab(\mathbf{x}'_i) = Lab(\mathbf{x}_i)$ $(i = 1, 2)$, the top $\mathbf{y}'_1$ of $\mathcal{D}'$ has label equal to $Lab(\mathbf{y}_1)$ in $H(\theta, a)$ and $Lab(\mathbf{y}'_1)$ is minimal in $H(\theta, a)$ word.

(b) there exists a reduced q-band $\mathcal{D}^0$ with boundary $\mathbf{x}^0_1\mathbf{y}^0_1\mathbf{x}^0_2\mathbf{y}^0_2$, where $Lab(\mathbf{x}^0_j) \equiv Lab(\mathbf{x}_j)$ $(j = 1, 2)$, such that the top $\mathbf{y}^0_1$ having label equal to $Lab(\mathbf{y}_1)$ in $H(\theta, a)$ and $Lab(\mathbf{y}^0_1) \equiv uvu^{-1}$, where v is cyclically minimal in $H(\theta, a)$ word.

Proof. (a) Consider the first version. One may assume that the history h of $\mathcal{D}$ is nontrivial in $B_n(\Theta)$ since otherwise by Lemmas 11.5 (b) and 11.10, $Lab(\mathbf{y}_1) = 1$ in $H(\theta, a)$, and the statement of the lemma is obviously true.

Let S_1 be a minimal word equal to $Lab(\mathbf{y}_1)$ in $H(\theta, a)$ and Δ_1 a minimal diagram over $H(\theta, a)$ with boundary $\mathbf{s}_1\mathbf{y}_1$, where $Lab(\mathbf{s}_1) \equiv S_1^{-1}$. We attach Δ_1 to $\mathcal{D}$ along the path $\mathbf{y}_1$, and denote the obtained diagram by Δ.

If Δ_1 has a positive cell, then by Corollary 11.3 and Lemma 2.3, Δ_1 has a cell Π with the sum of contiguity degrees to $\mathbf{s}_1$ and to $\mathbf{y}_1$ at least $\bar{\gamma}$. However the degree of contiguity to $\mathbf{s}_1$ is less than $\bar{\alpha}$ by Lemma 2.2 and Remark 2.1, since $\mathbf{s}_1$ is a geodesic path in Δ_1. Hence the contiguity degree to $\mathbf{y}_1$ is at least $\bar{\gamma} - \bar{\alpha} > 1/3$. Therefore by Lemma 2.4, Δ_1 has a positive cell π and a diagram of contiguity Γ of rank 0 of π to $\mathbf{y}_1$ with $(\pi, \Gamma, \mathbf{y}_1) > \varepsilon$. Using now Lemma 11.11 (2), we can replace Δ_1 with a diagram Δ'_1 of smaller type with boundary $\mathbf{s}_1\mathbf{y}'_1$ and replace $\mathcal{D}$ with a reduced q-band $\mathcal{D}'$ having the same labels of the start and end edges and having top label $Lab(\mathbf{y}'_1)$ equal to $Lab(\mathbf{y}_1)$ in $H(\theta, a)$. We can continue decreasing the types of $\Delta'_1, \Delta''_1, \dots$ until we get a diagram $\Delta^{(k)}$ of rank 0 and a reduced q-band $\mathcal{D}_k$ whose top label $Lab(\mathbf{y}_1^{(k)})$ is equal to $Lab(\mathbf{y}_1)$ in $H(\theta, a)$ and equal to S_1 in rank 0. Hence $|\mathbf{y}_1^{(k)}|_\theta = |\mathbf{s}_1|_\theta$ and the top label of $\mathcal{D}' = \mathcal{D}_k$ is a minimal in $H(\theta, a)$ word, as required.

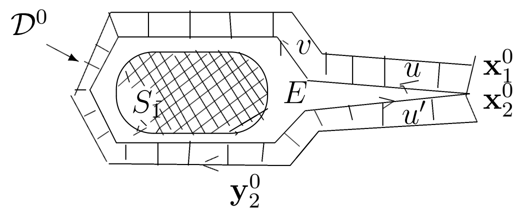

Figure 11

(b) The word $Lab(\mathbf{y}_1)$ is conjugate in $H(\theta, a)$ to a cyclically minimal word S_1, and so there exists a minimal annular diagram Δ_1 over $H(\theta, a)$ whose inner boundary component is labeled by S_1^{-1} and the outer component of the boundary has label $Lab(\mathbf{y}_1)$. Therefore one can obtain an annular diagram Δ identifying the boundary segments of Δ_1 and $\mathcal{D}$ labeled by $Lab(\mathbf{y}_1)$. Then we can reduce the type of Δ_1 as in item (a) (but using the statement (1) of Lemma 11.11), and the modifications of $\mathbf{y}_1$ do not change its label modulo the relations of $H(\theta, a)$. Finally we will replace Δ_1 with an annular diagram E without positive cells, where one of the boundary label is the cyclically minimal word S_1^{-1} and therefore every θ-edge of these boundary is adjacent to an edge of $\mathbf{y}_1^0$. (See Figure 11.)

Hence $Lab(\mathbf{y}_1^0)$ is of the form uvu', where v is equal to a cyclic permutation of S_1 u and u' are the labels of side arcs of some subbands of the band $\mathcal{D}^0$ obtained after the modifications of $\mathcal{D}$, and the θ-edges of these arcs have to be mutually adjacent in the diagram E of rank 0. It follows that $u' \equiv u^{-1}$ since side labels of q-bands are θ-defined. Statement (b) is proved. □

11.3 Axioms for $H(\mathbf{\Theta})$

In this section we start with the group $\mathbf{G}(0) = H(\mathbf{\Theta})$. The q-letters are now non-zero letters, while θ- and a-letters are regarded as 0-letters. The group $H(\mathbf{\Theta})$ satisfies all Burnside relations $u^n = 1$ of q-length zero, so we have Property (Z1.2) defined in Section 3.1. Property (Z1.1) follows from the definition of (θ, q)-relations.

By definition, the group $\mathbf{G}(1/2)$ is the factor group of $\mathbf{G}(0)$ modulo the hub relations.

Lemma 11.14. *Property (Z2) holds in the group* $\mathbf{G}(1/2)$.

Proof. Since we have chosen $N \geq n$ for the number of q-letters in the hubs, Property (Z2.1) holds. Since all q-letters of the hubs are pairwise distinct, we also obtain Property (Z2.2).

Then we will use the notation and the diagram Δ from the definition of (Z2.3). Let us denote by Γ the subdiagram of Δ with boundary $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$, where $Lab(\mathbf{p}_1) \equiv u_1$, $Lab(\mathbf{q_1}) \equiv v_1$, $Lab(\mathbf{p}_2) \equiv u_2^{-1}$, $Lab(\mathbf{q}_2) \equiv v_2^{-1}$ (see Fig. 5 in Subsection 3.1). We may assume that Γ is a minimal diagram over the group $\mathbf{G}(0) = H(\mathbf{\Theta})$.

Since we have $|v_1|_q \geq \varepsilon|v_1w_1|_q$ and one can choose $L > 3\varepsilon n$, at least three maximal t-bands start in Γ on $\mathbf{q}_1$. By Property (Z2.2) they cannot end on $\mathbf{q}_1$. They cannot end on $\mathbf{p}_1$ or $\mathbf{p}_2$ either since $|\mathbf{p}_1|_q = |\mathbf{p}_2|_q = 0$. So they end on $\mathbf{q}_2$, and therefore there are two maximal t-bands $\mathcal{Q}_i$ and $\mathcal{Q}_{i+1}$ connecting $\mathbf{q}_1$ and $\mathbf{q}_2$ and corresponding to t-letters $t^{(i)}$ and $t^{(i+1)}$ with $\{i, i+1\} \neq \{0, 1\}$.

To prove that the boundary label of Δ is trivial in $\mathbf{G}(0) = H(\mathbf{\Theta})$, one may cut off subdiagrams over $H(\mathbf{\Theta})$ from Δ or glue up such diagrams to Δ. We will define such surgeries below. At the end, we will obtain a spherical diagram, whose boundary label is the empty word obviously equal to 1 in $H(\mathbf{\Theta})$, as desired.

First of all we replace the subdiagram Γ with the smaller subdiagram Γ_i bounded by $\mathcal{Q}_i$, $\mathcal{Q}_{i+1}$, and the subpaths $\bar{\mathbf{q}}_1$ and $\bar{\mathbf{q}}_2$ of $\mathbf{q}_1$ and $\mathbf{q}_2$, resp. By Lemma 11.8, Γ_i contains no q-annuli. So every maximal q-band of Γ_i connects $\bar{\mathbf{q}}_1$ and $\bar{\mathbf{q}}_2$. Let $\mathcal{D}_0 = \mathcal{Q}_i, \dots, \mathcal{D}_m = \mathcal{Q}_{i+1}$ be all these consecutive q-bands starting on $\bar{\mathbf{q}}_1$.

By Lemma 11.13, it is possible to replace the band $\mathcal{D}_0$ with a band $\mathcal{D}'_0$ having minimal in $H(\theta, a)$ label of the top. (This operation can also change the boundary label of the whole diagram modulo the relations of $H(\theta, a)$.) By Lemma 11.5 (b), the bottom label of $\mathcal{D}'_0$ is also a minimal word. We may keep notation $\mathcal{D}_0$ for the modified q-band.

Consider now the minimal subdiagram E over $H(\theta, a)$ between $\mathcal{D}_0$ and $\mathcal{D}_1$ bounded by these two q-bands and by subpaths $\mathbf{x}_1$, $\mathbf{x}_2$ of $\bar{\mathbf{q}}_1$ and $\bar{\mathbf{q}}_2$. If E has a positive cell, then by Lemma 2.3, it has a cell Π with the sum of contiguity degrees to $\mathbf{x}_1$, $\mathbf{x}_2$ and to the sides (i.e. top/bottoms) of $\mathcal{D}_0$ and $\mathcal{D}_1$ at least $\bar{\gamma}$. However the first two contiguity degrees are 0 since $\mathbf{x}_1$ and $\mathbf{x}_2$ have no θ-edges, and the degree of contiguity to the side of $\mathcal{D}_0$ is $< \bar{\alpha}$ by Lemma 2.2 and Remark 2.1 since the label of this side is minimal in $H(\theta, a)$. Hence the degree of contiguity of Π to the side of $\mathcal{D}_1$ is $> \bar{\gamma} - \bar{\alpha} > 1/3$, and therefore by Lemma 2.4, there exists a positive cell π in E with contiguity diagram of rank 0 to the side of $\mathcal{D}_1$ and with degree of contiguity at least ε. Now Lemma 11.11 helps us to reduce the type of E (but the type of diagram over $H(\theta, a)$ between $\mathcal{D}_1$ and $\mathcal{D}_2$ may increase). Repeating such operation we may obtain a modified subdiagram E without cells of positive ranks. Thus, we may assume that E is a diagram over the group $H(a)$.

It follows that the projections of side labels of the bands $\mathcal{D}_0$ and modified $\mathcal{D}_1$ (for which we keep the same notation) on the alphabet $\mathbf{\Theta}$ are equal, and therefore the side labels of $\mathcal{D}_1$ are also minimal over $H(\theta, a)$ by Lemma 11.5 (b). Therefore one can eliminate now all positive cells from the subdiagram settled between $\mathcal{D}_1$ and $\mathcal{D}_2$ at the expense of the modifications between $\mathcal{D}_2$ and $\mathcal{D}_3$, and so on. Thus one may assume now that the whole Γ_i is a diagram over the group $H(\mathbf{Y})$.

By Lemma 10.6, every maximal θ-band of Γ_i crosses every q-band of Γ_i exactly once, and so we obtain a quasi-trapezium Δ_i bounded by $\mathcal{Q}_i$, $\mathcal{Q}_{i+1}$ and containing all θ-bands of Γ_i. By Lemma 10.10, Δ_0 corresponds to some quasi-computation $\mathcal{C}$, and $\mathcal{C}$ should start/end with the start/end rules of $\mathbf{M}$ since the labels of the paths $\bar{\mathbf{q}}_1$ and $\bar{\mathbf{q}}_2$ are the subwords of hub relation. Since these rules lock all sectors between $t^{(i)}$ and $t^{(i+1)}$, there are no a-cells between $\bar{\mathbf{q}}_1$ ($\bar{\mathbf{q}}_2$) and the nearest θ-band, i.e. $\Delta_i = \Gamma_i$.

Since the side label of $\mathcal{D}_0 = \mathcal{Q}_i$ is minimal in $H(\theta, a)$, the history of Γ_i is simple by

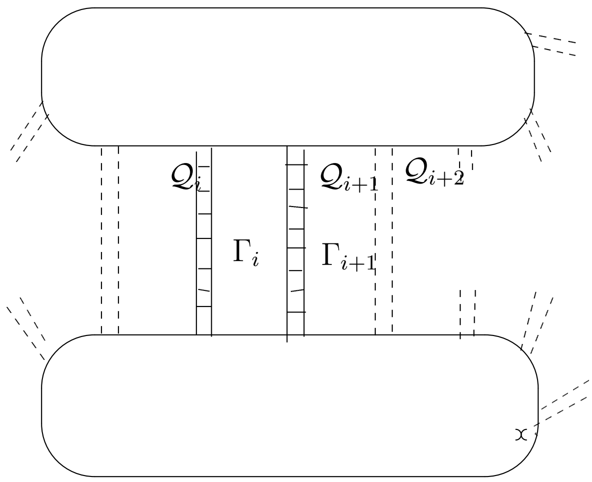

Figure 12

Corollary 11.6 (b), and by Lemmas 8.13 and 10.6, there is a trapezium Γ_i' with the same top and bottom labels whose history is congruent to the history of Γ_i, and so the side labels are equal in $H(\theta, a)$ to the side labels of Γ_i by Lemma 11.5 (a). Hence the replacement Γ_i by Γ_i' does not change the boundary label of the whole diagram modulo the relations of $H(\mathbf{\Theta})$. So we further assume that Γ_i is a trapezium and $\mathcal{C}$ is a computation. (See Figure 12.)

We can now construct a copy Γ_{i+1} of Γ_i by adding 1 to all superscripts of edges if $i+2 \neq 0$ (*mod L*). This diagram over $H(\mathbf{Y})$ can be inserted between two the hubs (with identification the t-band $\mathcal{Q}_{i+1}$ in Γ_i and Γ_{i+1}). Similarly we can connect the hubs by t-bands $\mathcal{Q}_0$,..., $\mathcal{Q}_L$, and the holes between them will be filled in with the diagrams Γ_l-s, but with one exception: to obtain a spherical diagram one should fill in the hole between the t-bands $\mathcal{Q}_0$ and $\mathcal{Q}_1$ with bases $t^{(0)}$ and $t^{(1)}$. Indeed this is possible by Lemma 10.10 since the computation $\mathcal{C}$ can be extended by Lemma 8.5 so that the hole between $\mathcal{Q}_0$ and $\mathcal{Q}_1$ is glued up with a quasi-trapezium too. This completes the construction of the spherical diagram and the proof of Lemma 11.14. □

Lemma 11.15. *Let w be an essential element represented by a cyclically minimal in rank $1/2$ word W and x be a 0-element such that $w^{-4}xw^4$ is a 0-element y in $H(\mathbf{\Theta})$. Then $x \in \mathbf{0}(w)$ and there is a 0-element c in $H(\mathbf{\Theta})$ such that the product wc^{-1} commutes with every element from $\mathbf{0}(w)$.*

Proof. We subdivide the proof in several steps.

0. Note that if $w' = ywy^{-1}$ and $x' = yxy^{-1}$ for some 0-element y, then the statement of the lemma holds for the pair (w, x) iff it holds for (w', x') since $\mathbf{0}(ywy^{-1}) = y\mathbf{0}(w)y^{-1}$.

1. It follows that we may assume that the word W starts with some q-letter q_0. Let Δ be a minimal diagram over $H(\mathbf{\Theta})$ with boundary $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$, where the section $\mathbf{p}_1$ has labels representing x, $Lab(\mathbf{q}_1) \equiv Lab(\mathbf{q}_2)^{-1} \equiv W^4$, and $|\mathbf{p}_1|_q = |\mathbf{p}_2|_q = 0$.

Since the word W is cyclically minimal in $H(\mathbf{\Theta})$, there is no maximal q-band in Δ, which starts and ends on $\mathbf{q}_1$ (on $\mathbf{q}_2$). (Otherwise W were conjugate of a word of length $|W|_q - 2$.) Hence each of them starts on $\mathbf{q}_1$, ends on $\mathbf{q}_2$, and one can enumerate them counting from $\mathbf{p}_1$: $\mathcal{D}_1, \dots, \mathcal{D}_m$ $(m \ge 4)$. So we have a diagram over $H(\theta, a)$ between $\mathbf{p}_1$ and $\mathcal{D}_1$, it can be assumed empty, and so x is represented by the side label of $\mathcal{D}_1$. We may assume that the history of $\mathcal{D}_1$ is non-trivial in $B_n(\mathbf{\Theta})$, since otherwise $x = 1$ in $H(\mathbf{\Theta})$ by Lemmas 11.5 (a) and 11.10, and so $x \in \mathbf{0}(w)$.

2. Let $\mathbf{y}_1\mathbf{x}_1\mathbf{y}_2\mathbf{x}_2$ be the boundary of $\mathcal{D}_1$ with sides $\mathbf{y}_1 = \mathbf{p}_1$ and $\mathbf{y}_2$. Since the labels of $\mathbf{q}_1$ and $\mathbf{q}_2^{-1}$ start with the same letter, we have $Lab(\mathbf{x}_2) \equiv Lab(\mathbf{x}_1)^{-1}$. So by Lemma 11.13 (b), one can construct a q-band $\mathcal{D}_1^0$ with boundary $\mathbf{y}_1^0\mathbf{x}_1^0\mathbf{y}_2^0\mathbf{x}_2^0$, where $Lab(\mathbf{x}_2^0) \equiv Lab(\mathbf{x}_1^0)^{-1}$, $Lab(\mathbf{y}_2^0)$ is equal to $Lab(\mathbf{y}_2)$ in $H(\theta, a)$, $Lab(\mathbf{y}_2^0)$ is vsv^{-1} for some word v, with cyclically minimal s in the group $H(\theta, a)$. Similarly, we have $Lab(\mathbf{y}_1^0) \equiv utu^{-1}$ with cyclically minimal t by Lemma 11.5 (b). Therefore we preserve the boundary label of Δ, when replacing the band $\mathcal{D}_1$ with the band $\mathcal{D}_1^0$ (and inserting an auxiliary diagrams over $H(\theta, a)$ to the top/bottom of $\mathcal{D}_1^0$). Let E be the modified Δ. We may assume that all q-bands of E are reduced, all the subdiagrams over $H(\theta, a)$ are minimal, and E has no q-annuli by Lemma 11.8.

Note that $|u|_q = 0$, and now we can replace the partition $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2$ of $\partial\Delta$ with the partition $\mathbf{y}(\mathbf{z}_1\mathbf{q}_1)\mathbf{p}_2(\mathbf{q}_2\mathbf{z}_2)$, where $Lab(\mathbf{y}) \equiv t$ and $Lab(\mathbf{z}_1^{-1}) \equiv Lab(\mathbf{z}_2) \equiv u$. Furthermore, adding edges to the boundary, we have a diagram E' with boundary $\mathbf{y}(\mathbf{z}_1\mathbf{q}_1\mathbf{z}_3)(\mathbf{z}_3^{-1}\mathbf{p}_2\mathbf{z}_4)(\mathbf{z}_4^{-1}\mathbf{q}_2\mathbf{z}_2)$, where $Lab(\mathbf{z}_3) \equiv Lab(\mathbf{z}_4) \equiv u$. This partition corresponds to simultaneous conjugation of the element x and w by the word u of q-length 0. Therefore we may replace (x, w) with the conjugate pair (x', w'), where $x' = t$ (see item **0**). Finaly, the beginning $\mathbf{z}_1\mathbf{x}_1$ of the path $\mathbf{z}_1\mathbf{q}_1\mathbf{z}_3$ is homotopic in E' to the path $\mathbf{lr}$, where $\mathbf{l}$ is a q-edge cutting the band $\mathcal{D}_1^0$ and $\mathbf{r}$ is a subpath of $\mathbf{y}_2^{-1}$. Let E'' be the diagram, where the boundary subpath $\mathbf{z}_1\mathbf{x}_2$ is replaced by $\mathbf{lr}$, and similarly we replace the subpath $\mathbf{x}_2\mathbf{z}_2$. (See Figure 13.) The replacement of E' with E'' does not change the pair (x', w'), but in E'' the element x' is represented by a cyclically minimal word and the word W' representing w' starts with the q-letter $Lab(\mathbf{l})$.

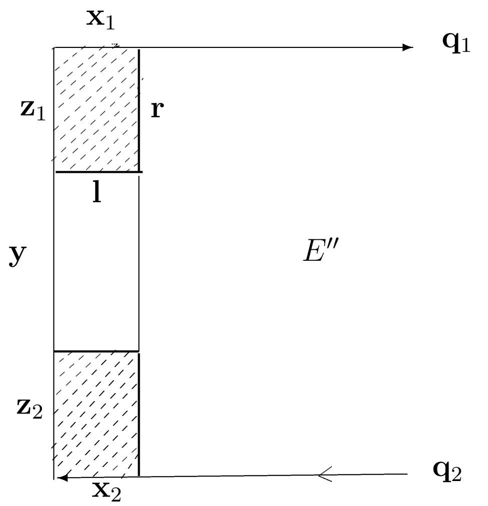

Figure 13

Remark 11.16. *The replacement of the pair (x,w) with the pair (x',w') (and the diagram Δ with E'') is possible if the history of $\mathcal{D}_1$ is reduced but not necessarily cyclically reduced. Let us call such a modification of the pair (x,w) and the diagram Δ the cyclic reduction of the band $\mathcal{D}_1$ in Δ.*

The advantage of the obtained diagram is that the band $\mathcal{D}_1$ is replaced with the band having cyclically minimal in $H(\theta,a)$ labels of the top and the bottom and the modified paths $\mathbf{q}_1$ and $\mathbf{q}_2^{-1}$ start with q-edges. Now we may re-establish the original notation $x, w, \mathbf{p}_1, \dots$ but assume that the first maximal q-band $\mathcal{D}_1$ of Δ has labels of the top and bottom **cyclically minimal** in the group $H(\theta,a)$.

If the θ-letters of $Lab(\mathbf{p}_1)$ belong to a subalphabet $\mathbf{Q}_j$, then the label of the opposite side of the band $\mathcal{D}_1$ has θ-letters from $\mathbf{Q}_{j'}$, where $j' = j \pm 1$ $(modN)$ by the definition of (θ,q)-relations.

3. Let $\bar{\mathbf{q}}_1$ (resp., $\bar{\mathbf{q}}_2^{-1}$) be the subpath of $\mathbf{q}_1$ (of $\mathbf{q}_2^{-1}$) starting with the vertex $(\mathbf{q}_1)_-$ (with $(\mathbf{q}_2)_+$) and labeled by Wq_0, where q_0 is the first letter of W. All q-edges of $\bar{q}_1$ are connected with q-edges of the subpath $\bar{\mathbf{q}}_2$ by the q-bands $\mathcal{D}_1,\dots,\mathcal{D}_s$, where $s-1=|W|_q$. Let Γ be the subdiagram of Δ bounded by $\bar{\mathbf{q}}_1$, $\bar{\mathbf{q}}_2$, $\mathcal{D}_1$ and $\mathcal{D}_s$ ($\mathcal{D}_1$ and $\mathcal{D}_s$ are included in Γ).

Since $Lab(\bar{\mathbf{q}}_1) \equiv Lab(\bar{\mathbf{q}}_2)^{-1}$, identifying $\bar{\mathbf{q}}_1$ and $\bar{\mathbf{q}}_2^{-1}$ in $\partial\Gamma$, except for the last edges of these paths labeled by q_0, one obtains an annular diagram $\Gamma(0)$. The q-bands $\mathcal{D}_1,\dots,\mathcal{D}_{s-1}$ becomes q-annuli in Γ_0. We denote by $\mathcal{Q}_1,\dots,\mathcal{Q}_{s-1}$ the reduced forms of these annuli, where $\mathcal{Q}_1=\mathcal{D}_1$ by item **3**. We do not change $\mathcal{D}_s$, and so there is a cutting path $\mathbf{p}$ connecting two vertices on $\mathcal{Q}_1$ and $\mathcal{D}_s$, whose label is equal to Wq_0 in $H(\boldsymbol{\Theta})$.

The *chamber* Γ_1 bounded by $\mathcal{Q}_1$ and $\mathcal{Q}_2$ can be made minimal annular diagram over $H(\theta,a)$ according to Lemma 11.7 (2). Since the side labels of $\mathcal{Q}_1$ are cyclically minimal in $H(\theta,a)$, there are no positive (over the presentation of $H(\theta,a)$) cells in this chamber with contiguity degree $\ge \bar{\alpha}$ to the side of $\mathcal{Q}_1$ by Lemma 2.2. Using Lemma 11.11 (as in Lemma 11.13), one can reduce the type of Γ_1 and finally transform it in an annular diagram over $H(a)$ (at the expense of the modification of the next chamber Γ_2). This transformation does not change the band $\mathcal{D}_1$, and according to Lemma 11.11 (d), it preserves the label of a simple path homotopic to $\mathbf{p}$ modulo the relations of $H(\Theta)$.

Recall that by item **1**, the history of $\mathcal{Q}_1$ is nontrivial in $B_n(\Theta)$. Therefore the retraction from Lemma 11.5 (a) shows that the side labels of $\mathcal{Q}_1$ are not conjugate in $H(\theta,a)$ to a word of θ-length 0. Hence applying Lemma 11.7 (2), we may assume that Γ_1 is a minimal diagram over $H(a)$.

Since both $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are reduced annuli, every maximal θ-band of Γ_1 must start on the boundary of $\mathcal{Q}_1$ and end on the boundary of $\mathcal{Q}_2$. Since the boundary labels of $\mathcal{Q}_1$ are cyclically minimal in $H(\theta,a)$ the same is true for the boundary labels of $\mathcal{Q}_2$ by Lemma 11.5(b). Therefore the next chamber Γ_2 between $\mathcal{Q}_2$ and $\mathcal{Q}_3$ can be transformed in a minimal annular diagram over $H(a)$ too. The iteration of this procedure makes the boundary labels of all q-annuli of $\Gamma(0)$ cyclically minimal over $H(\theta,a)$ and all the chambers Γ_i minimal annular diagrams over the group $H(a)$. According to Lemma 11.11 (d), the procedure preserves, modulo the relations of $H(\mathbf{\Theta})$, the label of a cutting simple path $\mathbf{p}'$ connecting $\mathbf{p}_-$ and $\mathbf{p}_+$ in the modified annular diagram $\Gamma(0)$, i.e. the word $Lab(\mathbf{p}')$ represents the element wq_0, and $|\mathbf{p}'|_q = |\mathbf{p}|_q$.

4. By Lemma 10.6, no θ-band crosses a q-annulus $\mathcal{Q}_i$ twice in $\Gamma(0)$ ($i = 1, \dots, s-1$), and so the maximal θ-bands starting on $\mathcal{Q}_1$ end on $\mathcal{Q}_s$. Therefore one can connect the vertex $\mathbf{p}'_- = (\mathbf{p}_1)_-$ with the vertex on the band $\mathcal{Q}_s$ along a θ-band of $\Gamma(0)$. So $\mathbf{p}'$ is homotopic in $\Gamma(0)$ to a path $\mathbf{p}''$, which is a product $\mathbf{p}(1)\mathbf{p}(2)$ of a side $\mathbf{p}(1)$ of a maximal θ-band and a path $\mathbf{p}(2)$ whose label S is freely equal to a subword of a power of the (outer) side label of the q-band $\mathcal{Q}_s$. (We use that the two boundary q-edges of $\mathcal{Q}_s$ have the same label q_0.)

Cutting the diagram $\Gamma(0)$ along the path $\mathbf{p}''$, we obtain a modification Γ' of Γ with the additional properties:

(1) Γ' has no positive cells over $H(\theta, a)$, i.e. Γ' is a minimal diagram over $H(\mathbf{Y})$,

(2) Every maximal θ-band of Γ' starting on $\mathbf{p}_1$ crosses each of the q-bands $\mathcal{D}_1, \dots, \mathcal{D}_s$ and

(3) $Wq_0 = W(1)W(2)$, where, $|W(1)|_q = s$, $W(1)$ has no θ-letters, and $W(2)$ is a product of the words $x(i)$-s defined in Lemma 11.5 for the set $\mathbf{\Theta}^+_{\mathbf{j}'}$ from the end of item **2**.

Note that the factorization of Wq_0 with these properties is unique, that is the factor $W(2)$ is uniquely defined modulo the relations of $H(\mathbf{Y})$ by Lemma 11.10, since the words Wq_0 and $W(2)$ have the same image under the homomorpism to $F(\mathbf{\Theta})$ defined in Lemma 11.10.

Simplifying the notation, we will assume that the diagram Γ enjoys the properties (1) - (3) itself.

5. Since all the θ-bands starting on $\mathbf{p}_1$ cross $\mathcal{D}_s$ in Γ and the band $\mathcal{D}_s$ is reduced we conclude that $\mathcal{D}_s$ is a reduced form of a product $\mathcal{F} = \mathcal{F}_1\mathcal{F}_2\mathcal{F}_3$, where the band $\mathcal{F}_2$ is a copy of $\mathcal{D}_1$ and the history of the band $\mathcal{F}_3$ is inverse to the history of $\mathcal{F}_1$, because the side labels of these two bands are $W(2)^{\pm 1}$. Futhermore, one can construct a new diagram E as follows.

Take the subdiagram E_1 formed by all maximal θ-bands of Γ starting with $\mathbf{p}_1$ and by a-cells between them (if any). The top/bottom labels of E_1 are equal to $W(1)$, i.e. they are minimal words in rank 1/2. Hence they cannot contain a subword of q-length 2 equal in $H(\mathbf{Y})$ to a word of q-length 0. It follows by induction that all the θ-bands of E_1 are regular, and therefore E_1 is a quasi-trapezium. (See Figure 14.)

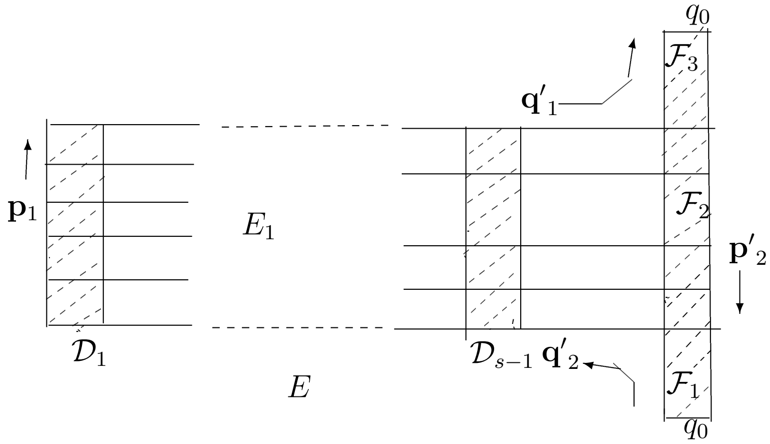

Figure 14

The maximal q-bands of E_1 are $\mathcal{D}_1, \dots, \mathcal{D}_{s-1}, \mathcal{F}_2$. Then we attach to $\mathcal{F}_2$ two q-bands $\mathcal{F}_1$ and $\mathcal{F}_3$. The boundary path of E is $\mathbf{p}_1\mathbf{q}'_1\mathbf{p}'_2\mathbf{q}'_2$, where $\mathbf{p}'_2$ is the side label of $\mathcal{F}$ (equal to the outer side label of $\mathcal{D}_s$). The labels of $\mathbf{q}'_1$ and $(\mathbf{q}'_2)^{-1}$ represent the same element wq_0 in $H(\mathbf{Y})$. So (changing notation) we may assume that these paths are labeled by Wq_0 and $\Gamma = E$.

The clock-wise boundary label of $\mathcal{F}_1$ is $q_0^{-1}Sq_0S'$, where S and S' are side labels having θ-letters from the alphabets $\mathbf{Q}_j$ and $\mathbf{Q}_{j'}$, respectively.

By Lemma 10.10, there is a revolving quasi-computation $\mathcal{C}$ corresponding to the quasi-trapezium E_1 with top/bottom labels $W(1)$ and history h equals the history of $\mathcal{D}_1$. Removing the q-band $\mathcal{F}_2$ from E_1, we see that the word $W(1)q_0^{-1}$ commutes with x in $H(\mathbf{Y})$.

6. Consider now the subpath $\tilde{\mathbf{q}}_1$ of $\mathbf{q}_1$ starting with the last edge of $\bar{\mathbf{q}}_1$ and having label W^3. The maximal q-bands $\mathcal{Q}_s, \dots, \mathcal{Q}_m$ connect it with the subpath $\tilde{\mathbf{q}}_2^{-1}$ of $\mathbf{q}_2^{-1}$. The diagram $\tilde{\Gamma}$ is bounded by $\mathcal{Q}_s$, $\mathcal{Q}_m$, $\tilde{\mathbf{q}}_1$ and $\tilde{\mathbf{q}}_2$.

It corresponds to the pair (y, w), where y is the label of the outer (in $\tilde{\Gamma}$) side of the band $\mathcal{D}_s$ with θ-edges labeled in $\mathbf{Q}_j$. We first replace $\mathcal{D}_s$ with $\mathcal{F}$, and then make the reduction of $\mathcal{F}$, removing $\mathcal{F}_1$ and $\mathcal{F}_3$ with mutually inverse histories h_1 and h_3 according to Remark 11.16. So we replace the pair (y, w) with (x, u), where $u = S^{-1}wS$ in $H(\mathbf{Y})$ and x is represented by a cyclically minimal label of the side of $\mathcal{F}_2$ (which is a copy of $\mathcal{D}_1$).

As in item **4**, for the word Uq_0 representing uq_0 in $H(\mathbf{\Theta})$, we have a decomposition $Uq_0 = U(1)U(2)$, where U and $U(2)$ have equal images in $F(\mathbf{\Theta}^+)$. Since $U = S^{-1}WS$, the image of $U(2)$ is equal to the image of $S^{-1}W(2)S$ in $F(\mathbf{\Theta}^+)$. Therefore $U(2) = (S')^{-1}W(2)S'$ in $H(\mathbf{Y})$ by Lemmas 11.5 and 11.10. Taking into account that $Sq_0(S')^{-1} = q_0$ in $H(\mathbf{Y})$ (since the boundary of $\mathcal{F}_1$ is trivial in $H(\mathbf{Y})$), we have

$$U(1) = Uq_0U(2)^{-1} = S^{-1}WSq_0(S')^{-1}W(2)^{-1}S' = S^{-1}(Wq_0W(2)^{-1})S' = S^{-1}W(1)S' \quad (16)$$

in $H(\mathbf{Y})$.

As in **5**, there exists a quasi-trapezium with top/bottom label equal to $U(1)$ and history h. Removing the last q-band (which is a copy of the first q-band of this quasi-trapezium), we obtain that $U(1)q_0^{-1}$ represents an element commuting with x. It follows from (16) that $W(1)q_0^{-1} = SU(1)(S')^{-1}q_0^{-1} = S(U(1)q_0^{-1})S^{-1}$ commutes with SxS^{-1} in $H(\mathbf{Y})$.

7. Thus, there is a minimal diagram over $H(\mathbf{Y})$ with a boundary $\mathbf{x}_1\mathbf{y}_1\mathbf{x}_2\mathbf{y}_2$, where the labels of $\mathbf{x}_1$ and $\mathbf{x}_2^{-1}$ are the reduced forms of SxS^{-1} and the labels of $\mathbf{y}_1$ and $\mathbf{y}_2^{-1}$ are equal to $W(1)q_0^{-1}$. The q-edges of $\mathbf{y}_1$ are connected with corresponding edges of $\mathbf{y}_2$ by q-bands, and one can add one more q-band (which is a copy of the first q-band) obtaining, by Lemma 10.6, a quasi-trapezium with top/bottom label $W(1)$. Since a side label of this quasi-trapezium is equal to SxS^{-1}, the history g of it is the reduced form of the product $h_1h_2h_1^{-1}$.

So, by Lemma 10.10 we have two revolving quasi-computations starting with the same $W(1)$ and histories h_2 (the history of the quasi-trapezium E_1 from item **5**) and g. Note that that the word $W(1)$ contains no cyclic subwords V congruent to $\mathbf{M}$-accepting configurations. Indeed, otherwise Lemma 10.10 allows us to construct a quasi-trapezium with top label V and bottom label equal to $\Sigma(0)$. Gluing up the hub to the bottom, we see that V is equal to 1 in rank 1/2. Hence the word W is not cyclically minimal in rank 1/2, a contradiction.

It follows from Lemma 9.6 that there is a revolving computation $\mathcal{C}$ with base $W(1)$ and history h' congruent to h_1.

The side label of the corresponding trapezium is equal in $H(\mathbf{Y})$ to S, and so $W' = W(1)q_0^{-1}$ commutes with S in $H(\mathbf{Y})$. Therefore $W^n = (W'S^{-1})^n = (W')^nS^n = (W')^n$ in $H(\Theta)$, because S has q-length 0. Since W' commutes with x, this implies $W^nx = xW^n$ in $H(\boldsymbol{\Theta})$.

To prove that w^kxw^{-k} is a 0-element, it suffices now to consider the maximal q-bands in a diagram of the equality $w^{ns}x = xw^{sn}$ for $sn > k$ (as in item **1**.)

8. Recall that the value of $W(1)$, and therefore the value of W' in $H(\boldsymbol{\Theta})$ does not depend on an element x from $\mathbf{0}(w)$ (see the end of item **4**). We have that the 0-word S is equal to $(W')^{-1}W$ in $H(\boldsymbol{\Theta})$, does not depend on x too. Hence the word $WS^{-1} = W'$ commutes with *every* element from $\mathbf{0}(w)$, and choosing an element $c \in H(\Theta)$ represented by the word S, we complete the proof. □

12 End of proofs

12.1 Proofs of Theorem 1.1 and Proposition 1.5

Proof of Theorem 1.1. Let a finitely generated group G be a subgroup of a group H finitely presented in $\mathcal{B}_n$. Note that the set R of all words in the generators $a_1, ..., a_k$ of H, which are equal to 1 in H, is recursively enumerable since in class of all groups the set of defining relations of H consists of a finite set and the recursively enumerable set of powers $w^n = 1$ for group words w in the alphabet $\{a_1, ..., a_k\}$. Each generator b_i of G from the finite set of generators $\{b_1, ..., b_l\}$ is equal in H to a word $w_i = w_i(a_1, \dots a_k)$, and to obtain the list of words in b_i-s equal to 1, it suffices to obtain the list of the corresponding products of w_i-s equal to 1 in G. Since the procedure of extracting such products from the list R is effective, the group G is recursively presented, i.e. it is given by a recursively enumerable set of relations in the generators $b_1, ..., b_l$.

For the nontrivial part of Theorem 1.1, we assume that there is a recursively enumerable set of defining relations for a finitely generated group G of exponent n, and we have a presentation $G = \langle A \mid \mathcal{L} \rangle$ as in the beginning of Section 8. The axioms (Z1), (Z2), (Z3) were checked for the group $\mathbf{G}(0) = H(\mathbf{\Theta})$ and for the factor group $\mathbf{G}(1/2) = \mathbf{G}$ of $\mathbf{G}(0)$ in Subsection 11.3. Therefore the group $\mathbf{G}(\infty)$ defined in Subsection 3.2 satisfies the identity $x^n = 1$ by Lemma 3.4. It is finitely presented in the variety $\mathcal{B}_n$ since the initial group $\tilde{G}$ given by relations (10), (11) is finitely presented, and the additional relation used in the constructions of groups $H(a)$, $H(\theta, a)$, $H(\mathbf{\Theta})$ and $H = \mathbf{G}(\infty)$ are either a-relations (which follow from the relations of $\tilde{G}$ by Lemma 10.1) or relations of the form $u^n = 1$.

It remains to prove that the canonical homomorphism $G \to H$ is injective. For this goal, we assume that a word w in generators from A is equal to 1 in H and prove that $w = 1$ in G. So there is a g-reduced diagram Δ over H with boundary $\mathbf{p}$ labeled by w. If Δ has a positive cell, then by Lemmas 2.3 and 2.4 (and Lemma 3.4), there is a positive cell π and a contiguity subdiagram Γ of rank 0 such that $(\pi, \Gamma, \mathbf{p}) \geq \varepsilon$. Let $\mathbf{q}_1 = \Gamma \bigwedge \pi$ and $\mathbf{q}_2 = \Gamma \bigwedge \mathbf{p}$. By Condition (A2) (see the definition of A-map), the path $\mathbf{q}_1$ is reduced in Γ, and so every q-edge $\mathbf{e}$ of it is adjacent with a q-edge of q_2^{-1}. Such a q-edge does exists in $\mathbf{q}_1$ since $\varepsilon n > 1$, but does not exist in $\mathbf{q}_2$, a contradiction. It follows that Δ has no positive cells, and so it is a diagram over $H(\mathbf{\Theta})$. So w is trivial in the group G by Lemma 11.9, as required. $\Box$

Proof of Proposition 1.5 (1) Let G' be a homomorphic image of the group G. We can modify the definition of congruent admissible words, replacing equality of the words in the input sector $\{t_0\}\{t_1\}$ (and also in $\{t_0\}\{t_0\}^{-1}$ and $\{t_1\}^{-1}\{t_1\}$) modulo the relations of G with equality modulo the bigger set of relations of G'. The wording of other definitions do not change (although the content of other concepts depending on the definition of congruence changes, we obtain different set of regular admissible word, and so on).

Respectively, we enlarge the set of a-relations now. Let M' be the factor group of M by all relations of G'. (So M' is not necessarily finitely presented.) Respectively we obtain the modified groups $\hat{M}'$, $H'(a)$, $H'(\mathbf{Y})$, and so on.

In the modified Statement of Lemma 10.1 we have a homomorphism of G' to $\hat{M}'$. This just follows from the new definition of a-relations. Respectively, we obtain injective homomorphism $G' \to H'(\mathbf{Y})$, $G' \to H'(\mathbf{\Theta})$, and $G' \to H'$ in Corollary 10.7, Lemma 11.9, and the above proof of Theorem 1.1. There are no changes in the formulations and proofs of other statements.

Thus, an arbitrary homomorphism $G \to G'$ extends to a homomorphism $H \to H'$. This means that the embedding $G \hookrightarrow H$ is a CEP-embedding, and Proposition 1.5(1) is proved for the embedding given by Theorem 1.1. The proofs of Corollaries 1.2 and 1.3 given in Introduction also provide us with the CEP for the embeddings in $G \hookrightarrow E$ and $G_i \hookrightarrow E$, because the CEP is transitive and the retract of a group is a CEP subgroup.

(2)Assume that two words v and w in the alphabet A are conjugate in the group $H =$

$\mathbf{G}(\infty)$, and $v \neq 1$. So we have a g-reduced annular diagram Δ over H. Then the assumption that Δ has a positive cell can be disproved exactly as in the proof of Theorem 1.1. Thus, Δ is a diagram over $H(\mathbf{\Theta})$. By Lemma 11.8, a q-annulus (if any exists in Δ) surrounds the hole of the annulus. So we have an annular diagram over $H(\theta, a)$ for the conjugacy of v and the boundary label V of such an annulus. However the retraction from Lemma 11.5 maps the word v of θ-length zero to 1. It follows that $V = 1$ in $H(\theta, a)$, and therefore $v = 1$ too, contrary the above assumption.

Hence Δ has no q-annuli and (θ, q)-cells since q-bands cannot start/end on the boundary on Δ. Therefore Δ is a diagram over $H(\theta, a)$. By Lemma 11.7, one may assume that Δ is minimal, and so it has no θ-annuli. If Δ has a positive cell Π, then we obtain a contradiction using Lemmas 3.4, 2.3. 2.4, exactly as we used these lemmas in the proof of Theorem 1.1 (but applying them to the graded diagram over $H(\theta, a)$ now), because there are no θ-edges on the boundary of Δ. Hence Δ is a diagram over $H(a)$ without (θ, a)-cells, since the boundary has no θ-edges. Thus, every nontrivial cell of Δ is an a-cell. Therefore the words v and w are conjugate in the group G. The statement (2) of Proposition 1.5 follows for the embedding $G \hookrightarrow H$. We also have the Frattini property for the embeddings in Corollaries 1.2 and 1.3 since following their proofs in Introduction, one should just keep in mind that the Frattini property of subgroups is transitive and every retract has Frattini property. $\square$

12.2 Embedding in a simple group

Let G be a group with a finite set of generators $X = \{x_1, \dots, x_m\}$ and with decidable word problem. Assume that its exponent divides a large odd integer n_0. We want to embed G in a finitely generated recursively presented, simple group of exponent n_0. At first, we recall the embedding in a simple group from $\mathcal{B}_0$ defined in [19]. One may assume that G is non-trivial since the trivial group is a subgroup of cyclic groups of prime orders, which are simple.

We start with a free product $G(0) = G \star G'$, where $G' = \langle Y \rangle = \langle y_1, \dots, y_l \rangle$ is arbitrary non-trivial finitely generated group of exponent n_0 with decidable word problem, e.g. a cyclic group of order n_0. The length $|W|$ is the length of the normal form W in the free product $G(0)$. Below we define by induction the groups $G(i)$, $i = 1, 2, \dots$; each of them is a factor group of $G(0)$ over a set of relations $\mathcal{R}_i$, where $\mathcal{R}_0 = \mathcal{R}_1 = \emptyset$. Also we will define the sets $\mathcal{X}_i$ of periods of rank i, simple in rank i words and some auxiliary sets. The set of simple words of ranks 0 and 1 are empty, Also we define $\mathcal{X}_0 = \mathcal{X}_1 = \emptyset$.

Let us fix non-trivial elements $a_1 \in G$ and $a_2 \in G'$. Since both G and G' have decidable word problem, there is an effective enumeration $a_1, a_2, a_3, \dots$ of all non-trivial elements from $G \cup G'$. For every a_i, we have a word in X or a word in Y representing this element.

For $i \geq 2$, let $\mathcal{X}'_i$ be the set of simple in rank $i-1$ (i.e. in the group $G(i-1)$) words of length i. Let $\mathcal{X}_i$ be a subset of $\mathcal{X}'_i$ maximal with respect to the following property. If A, B are different words from $\mathcal{X}_i$, then A is not a conjugate of $B^{\pm 1}$ in rank $i-1$. The words from $\mathcal{X}_i$ are called periods of rank i.

The set of relators $\mathcal{S}_i$ of rank i contains relators of the first type and relators of the second type. The set of relators of the first type is $\{A^{n_0} \mid A \in \mathcal{X}_i\}$. (In addition to (1), (2), we use two more parameters from [19] in this section: d satisfying $\delta^{-1} << d << \iota^{-1}$ and $n_0 >> \iota^{-1}$.)

For every $A \in \mathcal{X}_i$, we fix a maximal set $\mathcal{Y}_A$ such that (1) every word $T \in \mathcal{Y}_A$ has length satisfying $1 \leq |T| < d|A|$ and (2) each double coset of the pair of cyclic subgroups $\langle A \rangle, \langle A \rangle$ contains at most one word T from $\mathcal{Y}_A$, and this word has minimal length among the words representing elements of $\langle A \rangle T \langle A \rangle$.

The relators of the second type of rank i are defined for every period $A \in \mathcal{X}_i$ as follows. If a_1 is not contained in the cyclic subgroup $\langle A \rangle \leq G(i-1)$, then for every $T \in \mathcal{Y}_A$ outside of $\langle A \rangle a_1 \langle A \rangle$, we introduce the relation

$$a_1 A^n T A^{n+3} \dots T A^{n+3h-3} = 1$$

If a_2 is not contained in $\langle A \rangle \leq G(i-1)$ nor in $\langle A \rangle a_1 \langle A \rangle$, then for every $T \in \mathcal{Y}_A$ outside $\langle A \rangle a_2 \langle A \rangle$ we add another relation

$$a_2 A^{n+1} T A^{n+4} \dots T A^{n+3h-2} = 1$$

If s is the maximal number of the letters a_i occurring in the normal form of A, and $a_{s+1} \notin \langle A \rangle \langle a_1, a_2 \rangle \langle A \rangle$, then for every $T \in \mathcal{Y}_A$, we introduce the relation

$$a_{s+1} A^{n+2} T A^{n+5} \dots T A^{n+3h-1} = 1$$

We regard all cyclic permutations of the normal forms of these words and their inverses as relators of the second type, provided their first and last syllables belong to different free factors of $G(0)$.

Then we define $\mathcal{R}_i = \mathcal{R}_{i-1} \cup \mathcal{S}_i$, and $G(i)$ is the factor-group of $G(0)$ modulo the relations from $\mathcal{R}_i$. Finally, a word A is called simple in rank $i \geq 1$ if it is not conjugate in rank i of a word from G or G', not conjugate to a power of a shorter word, and not conjugate to a power of some period of rank $\leq i$.

The group $G(\infty)$ is the factor group of $G(0)$ over the normal closure of $\mathcal{R}(\infty) = \cup_{i=0}^{\infty} \mathcal{R}_i$.

Since the group presentations are now defined over free products, so are diagrams, i.e. every edge is labeled now with an element of one of free factors. More detailed description of diagrams over free product can be found in Section V.9 of [13] and also in Subsection 33.2 of [19]. The definition of g-reduced (or just reduced) diagram of rank i (i.e., over $G(i)$) sounds as the definition from Subsection 3.2. In the next lemmas we implicitly use Lemma 34.2 [19], which says that the the diagrams for presentations (with Condition R) over free products enjoy the properties of diagrams for the presentations (with Condition R) over free groups.

Lemma 12.1. *([19], Theorem 35.1) If the odd integer n_0 is large enough, then the group $G(\infty)$ is simple and the canonical epimorphism $G(0) \to G(\infty)$ is injective on the free factors G and G' of $G(0)$. Every proper subgroup of $G(\infty)$ is either cyclic of order dividing n_0 or it is contained in a subgroup conjugate either to G or to G'. (In particular, we have $G(\infty) \in \mathcal{B}_{n_0}$.)*

To prove Theorem 1.6, we should make recursive the choice of the relations of $G(\infty)$. Denote by $\mathcal{R}_{ij}$ the subset of $\mathcal{R}_i$ with the following restriction on the syllables $u_l \in G$, $v_l \in G'$ of the normal form of a relator $u_1v_1 \dots u_rv_r$ from $\mathcal{R}_i$. The length of u_l (of v_l) with respect to the set of generators X (set Y) of G (of G', resp.), is at most j. It follows from the definition of $\mathcal{R}_i$, that the cardinality of every subset $\mathcal{R}_{ij}$ is explicitly bounded in terms of i, j. The same is true for the subsets W_{kj} of all normal forms of length at most k with syllables of length at most j with respect to X and Y.

Lemma 12.2. *If Π is a positive cell in a reduced diagram Δ, then $|\partial\Delta| > (1-\alpha)|\partial\Pi|$.*

Proof. The claim is contained in Lemmas 34.14, 26.5 and 23.16 of [19]. (Lemma 34.14 says that the presentation of $G(i)$ satisfies Condition R, Lemma 26.5 says that a reduced diagram over a presentation with Condition R is a B-map, and Lemma 26.5 provides us with the required inequality for B-maps. We do not define R-condition and B-maps here since we do not use them explicitly. The reader can find everything in [19].) □

Lemma 12.3. *If $w \in W_{kj}$ and $w = 1$ in rank i, then there exists a chain $w \equiv w_0 \to w_1 \to \dots \to w_t \equiv 1$, where every transition $w_l \to w_{l+1}$ is either (a) a cancellation of several syllables in w_l or (b) cancellations followed by a merging of two neighbor syllables in a nontrivial one, or (c) replacement of subword of w_l equal to a prefix $a_1b_1...a_kb_k$ of a defining word $a_1b_1 \dots a_sb_s$ with the word $b_s^{-1}a_s^{-1} \dots b_{k+1}^{-1}a_{k+1}^{-1}$ followed by merging/cancellations of syllables.*

The relators used for transformations (c) all belong to $\mathcal{R}_{IJ}$, where $I < 2k$ and J is also effectively bounded in terms of k and j.

Proof. The statement is obvious if $i = 0$ since the word problem is solvable in both free factors G and G'. If a diagram Δ for the equality $w = 1$ has a positive cell Π, then there is such a cell having common boundary edges with $\partial\Delta$, and removing such a cell one applies the transformation of type (c). The condition $I < 2k$ follows from Lemma 12.2. The cubic upper bound k^3 for the number of positive cells in Δ is given in Subsection 28.2 of [19]. (The version for diagrams labeled over free products needs no changes in comparison with diagrams labeled over free groups.) Therefore we have an effective upper bound for the number of transformations of types (a), (b), (c). Note the merging can increase the length of the obtained syllable, but it at most doubles the upper bound for the length of syllables in the free factors. Thus, the effective bound for J is obtained too. □

The following Lemma is similar.

Lemma 12.4. *Let $w \in W_{kj}$ and w is a conjugate of v in rank i, where v is a simple word of rank i (or a power of a period A of rank $i' \le i$), then there exists a chain $w \equiv w_0 \to w_1 \to \dots \to w_t \equiv u$, where u is a cyclic permutation of v and every transition $w_l \to w_{l+1}$ is either (a) a cancellation of several syllables in w_l or (b) cancellations followed by a merging of two syllables in a nontrivial one, or (c) replacement of a subword equal to a prefix $a_1b_1 \dots a_kb_k$ of a defining word $a_1b_1 \dots a_sb_s$ with the word $b_s^{-1}a_s^{-1} \dots b_{k+1}^{-1}a_{k+1}^{-1}$ followed by merging/cancellations of syllables.*

The relators used for transformations (c) all belong to $\mathcal{R}_{IJ}$, where $I < 3k$ and J is also effectively bounded in terms of k and j. Besides we have $u \in W_{IJ}$.

Proof. Again, the statement is obvious if $i = 0$. For $i > 0$ there is an annular diagram Δ for the conjugacy of w and v with a cutting path of length $\leq \gamma(|w| + |v|)$ (Lemmas 26.5 and 22.1 in [19]). The inequality $|v| \leq \bar{\beta}^{-1}|w|$ follows from Lemmas 13.3, 26.5 and Theorem 22.4 from [19]. (Lemmas 13.3 and 26.5 say that the boundary component labeled by v can be made smooth, and Theorem 22.4 provides the required inequality for a smooth boundary component.)

Therefore by Lemma 12.2, the perimeter of every cell in Δ is less than $(|u| + |v|)\bar{\beta}^{-1}(1-\alpha)^{-1} < 3k$ and the number of cells is bounded by k^3, which allows us to finish the proof as in Lemma 12.3. □

Lemma 12.5. *Let A be a simple in rank $i \geq 0$ word.*

(a) If a word V belongs to the cyclic subgroup $\langle A \rangle$ of $G(i)$, then we have $V = A^k$ in $G(i)$ for $|k| < \bar{\beta}^{-1}|V||A|^{-1}$

(b) If $\max\{|U|, |V|\} \leq d|A|$ *and the word U belongs to the double coset $\langle A \rangle V \langle A \rangle$ but does not belong to the subgroup $\langle A \rangle$ of $G(i)$, then $U = A^k V A^l$ in $G(i)$, where* $\max\{|k|, |l|\} < 2\gamma^{-1}\delta^{-1}d$.

Proof. Property (a) is contained in the statements of Lemma 26.5 and Theorem 22.4 [19]. To prove (b) we consider an equality $U = A^k V A^l$ in $G(i)$. If $\min\{|k|, |l|\} \geq \gamma^{-1}\delta^{-1}d$, then $|U| + |V| \leq 2d|A| < \frac{\gamma^{-1}\delta^{-1}}{2} \min\{|A^k|, |A^l|\}$, and so U belongs to the cyclic subgroup $\langle A \rangle$ by Lemma 25.6 [19], a contradiction. Therefore $\min\{|k|, |l|\} < \gamma^{-1}\delta^{-1}d$, whence by Theorem 22.4 [19], we have $\max\{|k|, |l|\} \leq \bar{\beta}^{-1}(\gamma^{-1}\delta^{-1}d + 2d) < 2\gamma^{-1}\delta^{-1}d$. □

Lemma 12.6. *There is an effective construction of the group $G(\infty)$ so that the word and the conjugacy problems are decidable in this group.*

Proof. Although the set $\mathcal{X}'_i$ of simple in rank $i-1$ words of length i can be infinite, the subset $\mathcal{X}'_{ij}$ containing the words with length of every syllable at most j in the generators X or Y of groups G and G' is finite for every j. However if a word W from the finite set W_{ij} (of bounded size in terms of i and j) is conjugate of a power B^l of a period B of rank $\leq i-1$, then we may assume that $|B^l| \leq \bar{\beta}^{-1}i$ by Theorem 22.4 and Lemma 26.5 of [19]. By Lemma 12.4, this property can be effectively verified, provided the relations of the set $\mathcal{R}_{i-1,J}$ are defined for the number J effectively bounded in terms of i and j. Under the same assumption one can effectively define the subset $\mathcal{X}_{ij}$ of periods in $\mathcal{X}'_{ij}$ and obtain the relations of the first type belonging to $W_{s,j}$ for $s = in_0$. To define the relations of the second type from $\mathcal{R}_{ij}$, one should define the sets $\mathcal{Y}_A$ for every period A from $\mathcal{X}_{ij}$. The effective way to do this is given by Lemma 12.5, provided the relations of the set $\mathcal{R}_{i-1,J}$ are defined for the number J effectively bounded in terms of i and j. The same lemma helps us to construct all relations of the second type from $\mathcal{R}_{ij}$.

Now to enumerate the relations of the whole set $\mathcal{R}(\infty)$, we have the following algorithm using a constructive enumeration of the pairs (i, j) by natural numbers. We can go along the enumeration of pairs with the following deviations. Assume that at a point (i, j) we have sufficient data of the form $\mathcal{R}_{i-1,J}$ $(J = J(i, j))$ to construct the set $\mathcal{R}_{ij}$. (This can be effectively verified since J is explicitly bounded in terms of i and j.) Then we construct the set $\mathcal{R}_{ij}$ and move to the next pair. Otherwise we choose the first pair (i', j'), for which we have sufficient data but the set $\mathcal{R}_{i'j'}$ has not yet been constructed, we construct $\mathcal{R}_{i'j'}$ and return to (i, j). Of course, no data is needed for $i \leq 2$ since there are no relations of rank ≤ 1. Besides, there is no pair (i, j) such that we will never get the data to define $\mathcal{R}_{ij}$. Indeed, if we have i minimal with this property, then every $\mathcal{R}_{i-1,J}$ will be defined soon or later, which will provide us with the necessary data at point (i, j), a contradiction.

Thus, the set of relations $\mathcal{R}(\infty)$ is recursively enumerable. If a word W is trivial in $G(\infty)$, then $W = 1$ in rank $i \leq |W|$ by Lemma 12.2. Hence the decidability of the word problem in $G(\infty)$ follows from Lemma 12.3. Similarly, the rank of a conjugacy diagram for two words U and V can be bounded by $2(|U| + |V|)$ by Lemma 25.4[19], and decidability of the conjugacy problem follows from Lemma 12.4. □

Proof of Theorem 1.6. Let G be a finitely generated group from the variety $\mathcal{B}_{n_0}$, where n_0 is a large enough odd integer. By Lemma 12.1, G is a subgroup of the finitely generated simple group $G(\infty)$ of exponent n_0. By Lemma 12.6, the group $G(\infty)$ can be defined by a recursive set of defining relations. Therefore $G(\infty)$ is embeddable in a group finitely presented in the variety $\mathcal{B}_{n_0}$ by Theorem 1.1, and the part "only if" of Theorem 1.6 is proved.

The part "if" is easier and it exploits Kuznetsov's observation [12]. Let G be a finitely generated group, and we have the embeddings $G \leq S \leq H$, where S is simple and H is a finitely presented in $\mathcal{B}_{n_0}$ group. Since G is a finitely generated subgroup of the recursively presented group H, the set of words in generators of G equal to 1 is recursively enumerable. To prove that the word problem is decidable in G, it suffices to show that the set of non-trivial in G words is also recursively enumerable.

Given arbitrary word w in the generators of G, one can rewrite it and obtain a word w' in the generators of H. In the factor group $H(w)$ of H over the additional relator w', the image of S is trivial if $w \neq 1$ in G (and in S) since S is a simple group. If $w = 1$ in G (and in H), then obviously $H(w) = H$, and every nontrivial element of S remains non-trivial in $H(w)$.

Therefore choosing one element $s \in S\backslash\{1\}$ (given as a word in the generators of H) one has $s = 1$ in $H(w)$ iff $w \neq 1$ in G. Since every $H(w)$ is a finitely generated recursively presented group, the word s is contained in the recursively enumerable list $R(w)$ of trivial in $H(w)$ words iff $w \neq 1$ in G. Clearly there is an algorithm scanning all the lists $R(w)$ and marking those words w, for which s gets into $R(w)$. This algorithm enumerates all non-trivial in G words. □

Acknowledgement. The author is grateful to the referee and to Nathaniel Tenpas for useful remarks.

References

[1] S.I. Adian, The Burnside problem and identities in groups, Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], 95, 311 pp., Berlin - New York, Springer, 1979.

[2] J.-C. Birget, Time-complexity of the word problem for semigroups and the Higman embedding theorem, Intern. J. Algebra Comput 8 (1998), 235-294.

[3] J.-C. Birget, A. Yu. Olshanskii, E.Rips, M.Sapir, Isoperimetric functions of groups and combinatorial complexity of the word problem , Annals of Mathematics , 156 (2002), no. 2, 467–518.

[4] W.W. Boone, Certain simple unsolvable problems in group theory, Proc. Kon. Ned. Acad. Wetensch, A, (I) 57 (1954), 231-237, (II) 57 (154), 492-497, (III) 58 (1955), 252-256, (IV) 58 (1955), 571-577, (V) 60 (1957), 22-27, (VI) 60 (1957), 222-232.

[5] W.W. Boone, G. Higman, An algebraic characterization of the solvability of the word problem, J. Austral. Math. Soc. 18 (1974), 41-53.

[6] N.J. Fine and H.S. Wilf, Uniqueness theorems for periodic functions, Proc. Amer. Math. Soc., 16 (1965), 109114.

[7] G. Higman, Subgroups of finitely presented groups. Proc. Roy. Soc. Ser. A, 292 (1961), 455-475.

[8] S.V. Ivanov, Burnside problem on periodic groups, Bull. AMS 27 (1992), no.2, 257-260.

[9] S.V. Ivanov, The free Burnside groups of sufficiently large exponent, Intern. J. Algebra Comput., 4 (1994), no. 1-2, 1 - 308.

[10] O. Kharlampovich, The word problem for the Burnside varieties, Journal of Algebra, 173 (1995), no. 3, 613-621.

[11] Kourovka Notebook (Unsolved Problems in Group Theory), no. 12, Novosibirsk, 1992.

[12] Kuznetsov A.V., Algorithms as operations in algebraic systems, Uspekhi Matematicheskikh Nauk, 13 (1958), 240-241 (in Russian)

[13] R. C. Lyndon and P. E. Schupp. *Combinatorial group theory.* Springer-Verlag, 1977.

[14] M. Morse, Recurent geodesies on a surface of negative curvature, Trans. Amer. Math. Soc. 22 (1921), 84-100.

[15] P.S. Novikov, On the algorithmic insolvability of the word problem in group theory. American Mathematical Society Translations, Ser. 2, 9 (1958), 1-122 or Trudy Mat. Inst. im. Steklov, 44 (155), 143 pp.

[16] P.S. Novikov, On periodic groups [Russian], Dokl. Akad. Nauk SSSR, 127 (1959), 749-752.

[17] P.S. Novikov, S.I. Adian, Infinite periodic groups, I, II, III (Russian), Izv. Akad. Nauk, Ser. Mat., 32 (1968), 212-244, 251-524, 709-731.

[18] A. Yu. Olshanskii, On Novikov - Adian's theorem, Matem. Sbornik, 118 (1982),no. 2,203–235 (in Russian).

[19] A. Yu. Olshanskii, Geometry of defining relations in groups, Moscow, Nauka, 1989, 448 p.p. (in Russian, English translation by Kluwer Academic Publ., 1991)

[20] A. Yu. Olshanskii, On the subgroup distortion in finitely presented groups, Matem. Sbornik, 188 (1997), N 11, 73–120

[21] A. Yu. Olshanskii, Groups with undecidable word problem and almost quadratic Dehn function, Journal of Topology, 5, 4(2012), 785–886 (with Appendix written by M.V.Sapir).

[22] A.Yu. Olshanskii, M.V. Sapir, Length and area functions in groups and quasi–isometric Higman embeddings, Intern. J. Algebra and Comput., 11 (2001), no. 2, 137–170.

[23] A.Yu. Olshanskii, M.V. Sapir, Non-amenable finitely presented torsion-by-cyclic groups, Publ. Math. IHES, 96 (2003), no. 6, p.p. 43–169.

[24] A.Yu. Olshanskii, M.V. Sapir, Groups with small Dehn functions and bipartite chord diagrams, Geometric and Functional Analysis, 16 (2006), 1324–1376.

[25] A.Yu. Olshanskii, M.V. Sapir, Conjugacy problem in groups with quadratic Dehn function, Bulletin of Mathematical Sciences, 10(2020), 103 pp.

[26] E.Prouchet, Memoir sur quelques relations entre les puissonies des nombres, C.R.Acad. Sci. Paris Ser 1, 33(1851), 225.

[27] J.Rotman, An introduction to the theory of groups, 3d edition, Allyn and Bacon Inc., Boston, Mass, 1984.

[28] M.V. Sapir, On the word problem in periodic groups varieties, Intern. Journal of Algebra and Comput., 1 (1991), no.1, 115-126.

[29] M. V. Sapir, J. C. Birget, E. Rips, Isoperimetric and isodiametric functions of groups, Annals of Mathematics, 157, 2(2002), 345-466.

[30] D. Sonkin, CEP-subgroups of free Burnside groups of large odd exponents, Comm. Algebra, 31 (2003), no. 10, 4687-4695.

[31] A.Thue, Uber unendliche Zeichenreihen, Norske Vid. Selsk. Skr. Math. Not. Kl. 7 (1906), 1-22.

[32] R.J. Thompson, Embeddings into finitely generated simple groups which preserve the word problem, Word problems, II (Conf. on Decision Problems in Algebra, Oxford, 1976), 401-441, Stud. Logic Foundations Math., 95, North-Holland, Amsterdam - New York, 1980.

Subject index

Alexander Yu. Ol'shanskii
Department of Mathematics
Vanderbilt University and
Department of Higher Algebra,
MEHMAT, Moscow State University
alexander.olshanskiy@vanderbilt.edu