Conjugacy problem in groups with quadratic Dehn function

A.Yu. Olshanskii, M.V. Sapir*

Abstract

We construct a finitely presented group with quadratic Dehn function and undecidable conjugacy problem. This solves E. Rips' problem formulated in 1992.

Key words: generators and relations in groups, finitely presented group, Dehn function of group, S-machine, conjugacy problem, van Kampen diagram.

AMS Mathematical Subject Classification: 20F05, 20F06, 20F65, 20F69, 03D10.

Contents

*Both authors were supported in part by the NSF grant DMS-1500180. The first author was also supported by RFFI grant 15-01-05823

1 Introduction

We assume that the reader is familiar with basic facts and definitions about van Kampen (disk and annular (Schupp) diagrams over group presentations. We remind some of it in Section 5.2 below, see also books [12, 14, 27]).

The Dehn function of a finitely presented group $G = \langle X \mid R \rangle$ is the smallest function $f(n)$ such that for every word w of length at most n in the alphabet $X \cup X^{-1}$, which is equal to 1 in G there exists a van Kampen diagram over the presentation of G with boundaary label w and area at most $f(n)$. It is well known [6, 7] that the Dehn functions of different finite presentations of the same group are equivalent, where we call two functions $f(n), g(n)$ equivalent if for some constants $A, B, C, D \geq 1$ we have

$$\frac{1}{A} f\left(\frac{n}{B}\right) - Cn - D < g(n) < Af(Bn) + Cn + D.$$

As usual, we do not distinguish equivalent functions.

The Dehn function of a group is an important asymptotic invariant. From the algorithmic point of view, smaller Dehn function means more tractable word problem (see, for example, the Introduction of [28] for details). Moreover as was shown in [1] a not necessarily finitely presented finitely generated group has word problem in NP if and only if it is a subgroup of a finitely presented group with polynomial Dehn function (a similar result holds for other computational complexity classes [1]). From the geometric point of view the Dehn function measures the "curvature" of the group: linear Dehn functions correspond to negative curvature, quadratic Dehn function correspond to zero curvature, etc.

More precisely, a finitely presented group is hyperbolic if and only if it has a subquadratic (hence linear) Dehn function [6, 2, 15]. In particular, the conjugacy problem in such groups is decidable [6].

It is also known that groups with quadratic Dehn functions exhibit certain "non-generic" non-positive curvature behavior as far as geometric and algorithmic properties are concerned. For example their asymptotic cones are simply connected [24]. For large classes of groups with quadratic Dehn functions, the conjugacy problem is decidable. In fact it is true for all known examples of groups with quadratic Dehn functions such as bi-automatic groups [5], $SL_n(\mathbb{Z}), n \geq 5$ [29, 8], groups acting geometrically on CAT(0) spaces [4], the R. Thompson group F [9, 10], free-by-cyclic groups [3, 21], etc. The decidability of conjugacy problem was proved in a completely different way in each of these cases and it is natural to ask if every group with quadratic Dehn function has decidable conjugacy problem and there is a uniform proof of that fact. That question was first formulated by Rips in the early 90s (some of the important results mentioned above had not appeared yet at that time).

Problem 1.1 (Rips)**.** Does every finitely presented group with quadratic Dehn function have decidable conjugacy problem?

In fact Rips had a "quasi-proof" showing that the answer should be positive. That "quasi-proof" first appeared in [21]. Basically the idea is the following (see details in [21]). If the conjugacy problem in a group $G = \langle X \mid R\rangle$ is undecidable, then for arbitrary $n \in \mathbb{N}$ for some pairs of words (u, v) in the alphabet X of length $\leq n$, there exists a minimal area annular diagram Δ with boundaries labeled by u, v and no path connecting the two boundaries of length smaller than any given recursive function $f(n)$. Let q be a simple path connecting the boundaries of Δ, $t = |q|$. Then there are simple closed paths $p_1, ..., p_m$ of Δ surrounding the hole such that $p_i, ..., p_j$ do not intersect if $i \neq j$ and $m > c_1 t$ for some constant c_1. The area of Δ is at least a constant times $\sum |p_i|$. If "many" lengths $|p_i|$ are less than $c \log t$ where $c = \frac{1}{2|X|}$, then two of the paths p_i, p_j $(i \neq j)$ have the same labels. That allows us to identify p_i, p_j and remove the annular subdiagram of Δ bounded by p_i, p_j, decreasing the area of Δ, a contradiction. Therefore "many" lengths $|p_i|$ are at least $c_2 \log t$ for some constant c_2. Hence the area of Δ is at least $c_3 t \log t$ for some constant c_3. If we cut Δ along the path q, we obtain a disk van Kampen diagram Δ' with boundary path subdivided into four parts $q_1 p_1 q_2^{-1} p_2^{-1}$ where $|p_1|, |p_2| \leq n$ and the labels of q_1 and q_2 coincide with the label of q. The area of Δ' is at least $c_3 t \log t$. Since the labels of q_1, q_2 are the same, we can glue t/n copies of Δ' together to obtain a van Kampen diagram Δ'' with perimeter bounded from above by a linear function in t and area bounded below by $c_3 t^2 \log t/n$ since t is bounded below by any given recursive function in n, n is insignificant compared to t. The diagram Δ'' can be assumed reduced. So we found a reduced van Kampen diagram of perimeter $\sim t$ and area $\sim t^2 \log t$. Hence the Dehn function cannot be smaller than $n^2 \log n$.

The incorrectness of this "quasi-proof" is in the last phrase. Indeed, there may be a smaller area van Kampen diagram with the same boundary label as Δ''. Still there is a lot of flexibility in choosing Δ and the path q in it. It looks like it would require infinite number of relations to ensure that all the boundary paths of various diagrams Δ'' have feelings with much fewer cells than Δ''. In particular, if G satisfies some mild form of asphericity, the proof should work. Rips conjectured that this should be true for all finitely presented groups. In [21] we confirmed this conjecture for a wide class of multiple HNN extensions of free groups. We also constructed in [21] a multiple HNN extension of a free group with undecidable conjugacy problem and the minimal possible Dehn function $n^2 \log n$.

Nevertheless, in this paper, we give a negative answer to Rips' question (and hence disprove Rips's conjecture as well):

Theorem 1.2. *There exists a finitely presented group with undecidable conjugacy problem and quadratic Dehn function.*

As in several of our previous papers ([28, 1, 21, 18] the construction is based on an S-machine (we call it $\mathbf{M}_5$) which can be viewed as a computing device with undecidable halting problem or as a group which is a multiple HNN extension of a free group. S-machines were first introduced by Sapir in [28] (see Section 2.1 below for the definition used here and [26] for various other definitions).

In order to describe some ideas of our proof in more details, let us start with a simple example of an S-machine $\mathbf{S}$ (That S-machine first appeared in [22]. The corresponding group was the first example of a group with polynomial Dehn function, linear isodiametric function and non-simply connected asymptotic cones answering a question of C. Druţu.) It is a rewriting system [27] with alphabet $\{a, q, a^{-1}, q^{-1}\}$ and two "same" rules $\theta_i \colon q \to aq$ and their inverses $\theta_i^{-1} \colon q \to a^{-1}q$, $i = 1, 2$. The rewriting system works with group words in $\{a, q\}$. And applying a rule $\theta_i^{\pm 1}$ means replacing every letter q^ϵ (where $\epsilon = \pm 1$ by $(a^{\pm 1})q)^\epsilon$ and then reducing the word. The S-machine $\mathbf{S}$ can also be viewed as a multiple HNN extension of the free group $\langle a, q\rangle$:

$$\langle a, q, \theta_1, \theta_2 \mid q^{\theta_i} = aq, a^{\theta_i} = a, i = 1, 2\rangle.$$

(Note that this is far from the only way to interpret an S-machines as groups. We are using a different interpretation in this paper, and the most complicated one so far was used in [20]. But the main principle is still the same.)

As the name S-machine suggests, we can also consider $\mathbf{S}$ as a kind of Turing machine with tape letter a, state letter q and commands θ_1, θ_2 (and their inverses). Then we can consider *computations*. Say,

$$q^{-1}aqaq \stackrel{\theta_1}{\to} q^{-1}aqaaq \stackrel{\theta_2}{\to} q^{-1}aqaaaq \stackrel{\theta_1^{-1}}{\to} q^{-1}aqaaq \stackrel{\theta_2^{-1}}{\to} q^{-1}aqaq \qquad (1.1)$$

is a *reduced* computation of $\mathbf{S}$. At the same time if we consider $\mathbf{S}$ as an HNN-extension of the free group, then this computation corresponds the van Kampen diagram on Figure 1.

This diagram is called the *trapezium* corresponding to the computation (1.1). Three things need to be noticed from this diagram.

1. The trapezium looks like a rectangle with the first word and the last word of the computation on the bottom and top side. All other words of the computation are on the horizontal paths of the trapezium, and θ's conjugate each of these words to the next one.

2. The vertical sides of the trapezium are labeled by the same words: the *history* of the computation (in the case of (1.1) it is $\theta_1\theta_2\theta_1^{-1}\theta_2^{-1}$).

3. The trapezia has three types of *bands* (also called in the literature *corridors*), i.e. sequences of cells where each two consecutive cells share an edge with a prescribed letters: horizontal $\theta_i^{\pm 1}$-bands, vertical q-bands and a-bands. The median lines of these bands serve as "walls" in van Kampen diagrams over S-machines, provide necessary rigidity and are crucial for all applications of S-machines.

Now let us continue our description of the construction and proof of Theorem 1.2.

As any S-machine viewed as a group, $\mathbf{M}_5$ contains a-letters, q-letters and θ-letters. Some of the words containing only q- and a-letters are called *input words*. Among them, there is one word W_0 which does not contain the a-letters. All other input words are obtained from W_0 by inserting a power of a single letter a into the *input sector* of W_0.

If we view $\mathbf{M}_5$ as multiple HNN-extension M_5 of a free group, it easily follows from undecidability of the halting problem by $\mathbf{M}_5$ that the group M_5 has undecidable conjugacy problem (the existence of such S-machines was proved in [28], a similar construction

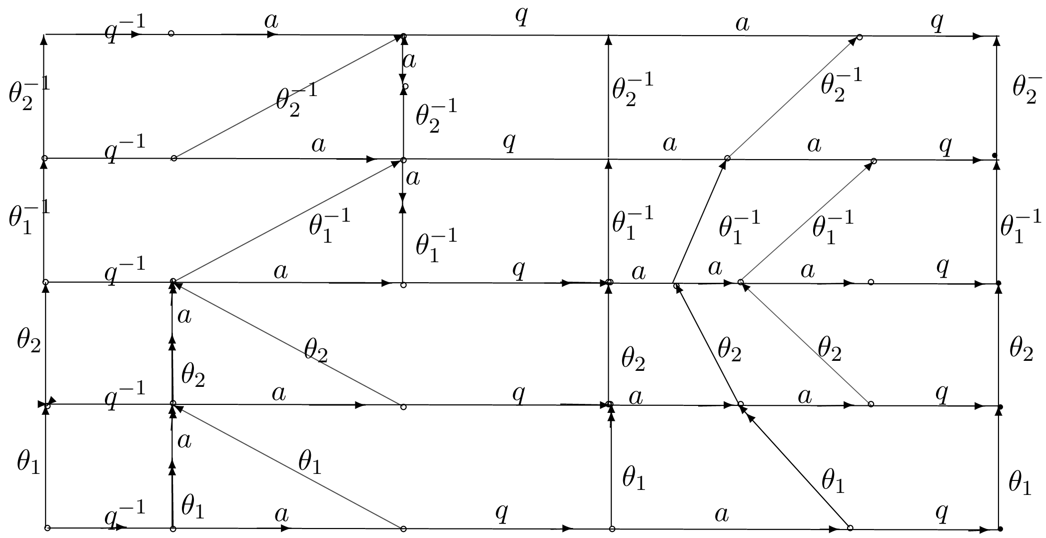


Figure 1: The trapezium corresponding to a computation of $\mathbf{S}$.

of groups with undecidable conjugacy problem was suggested by C. Miller [13], but C. Miller did not use S-machines). Therefore by [21], the Dehn function of M_5 is at least $n^2 \log n$.

But we prove here that most of the area in van Kampen diagrams of large area over M_5 is concentrated in a few standard trapezia which we call *big trapezia*.

The phenomenon that large area of a van Kampen diagram is concentrated in a few large standard subdiagrams is interesting. For example, we proved similar facts for van Kampen diagrams over presentations satisfying the small cancelation condition $C(p) - T(q)$ in the the CAT(0) case $\frac{1}{p} + \frac{1}{q} = \frac{1}{2}$ in [23]. In that case the geometric meaning of existence of large standard subdiagram is very close to a popular topic in CAT(0) geometry: "every quasi-flat in the universal cover of the presentation complex is close to a flat" (see a discussion in [21]). In the case of S-machines, we proved similar facts in [21] and [18], in both cases, as in the present paper, these were crucial steps in the proofs.

The big trapezia over M_5 must correspond to "very long" computations of $\mathbf{M}_5$. The S-machine $\mathbf{M}_5$ is constructed in such a way that long computations $\mathcal{C}$ are "normal", that is the start word w of $\mathcal{C}$ can be reached by the S-machine $\mathbf{M}_5$ from some (determined uniquely by the computation) input word u of $\mathbf{M}_5$. That is another crucial property of $\mathbf{M}_5$.

Then we lower the Dehn function from $\geq n^2 \log n$ to $O(n^2)$. For this goal we embed M_5 in a larger group G. We extend the machine $\mathbf{M}_5$ to $\mathbf{M} = \mathbf{M}_6$ and add hub relations to the obtained group $M = M_6$. The hub is the product of $L \gg 1$ copies of the end word; for convenience, we use two hubs in this paper, but the second hub relation follows from the other relations of G. (Note that hub relations are usually used in constructions of groups with undecidable word problem, but the word problem in G is decidable.) The hubs and the disks (that are hubs surrounded by θ-annuli) make the areas of trivial over M words quadratic with respect of the presentation of G (another important idea). Therefore the presentation of G is highly non-aspherical: the boundaries of the large normal trapezia can be filled both by diagrams with $\sim n^2\log n$ cells are by diagrams with at most $\sim n^2$ cells.

The new S-machine $\mathbf{M}$ is obtained by augmenting $\mathbf{M}_5$ with a simple S-machine $\mathbf{M}_{12}$ (the union of Steps 1 and 2 in the definition of $\mathbf{M}$ given in Subsection 4.1) which starts with a specific input word W_0 with no tape letters, and produces (nondeterministically) an arbitrary input word u of $\mathbf{M}_5$ by by inserting the a power of a into the input sector (by a command similar to θ_1 of $\mathbf{S}$). This augmentation provides us with the property that arbitrary configuration of a "long computation" of $\mathbf{M}$ can be reached with linear time and space either from W_0 or from the stop configuration of $\mathbf{M}$. Afterwards this linearity guaranties quadratic estimates of the areas of both disks and big trapezia over the presentation of G. The linearity is achieved by, in particular, adding many so-called history sectors where the history of a computation is non-deterministically written before the actual computation executing that history starts.

In order to connect $\mathbf{M}_{12}$ with the S-machine $\mathbf{M}_5$ and obtain the main S-machine M, we need one rule, called $\theta(23)$ which changes the state letters to the start state letters of $\mathbf{M}_5$. However the standard interpretation of $\mathbf{M}$ as a group would make the conjugacy problem decidable in the group M. So the rule $\theta(23)$ is interpreted in G as turning all copies of the input word into identical words (by erasing extra indices). This new "irregular" interpretation induces the study of some non-reduced computations, i.e. the history of an "elligible" computation may contain (many) subwords $\theta(23)\theta(23)^{-1}$.

The proof that G has quadratic Dehn function is much harder than the proof of undecidability of the conjugacy problem. We use several tools developed in [28, 21, 17, 18] and more. As in all our papers where estimates of the Dehn function are produced, we need to consider diagrams with and without hubs separately. This is done in Sections 6 and 7 respectively. In both cases, one of the main ideas is to assign to the boundary of every van Kampen diagram Δ over the presentation of G a certain numeric invariant $\mu(\Delta)$ (the *mixture* from [18]) which is bounded from above by a quadratic function in terms of the perimeter. We had a somewhat similar numeric invariant called *dispersion* in [21] but that invariant does not work well for diagrams with hubs.

To obtain a quadratic estimate for diagrams Δ over M, we have to consider an artificial G-areas instead of areas, and just at the end of this paper we replace the diagrams of quadratic G-area over M with diagrams with hubs, having quadratic (usual) areas over G. The quadratic upper bound for G-area is obtained by induction over the (modified) perimeter n of Δ. We perform surgeries on the diagram, so that each surgery makes the diagram look more "standard" and smaller. Our inductive argument estimates the G-area in terms of some linear combination of n^2 and the mixture $\mu(\Delta)$. Although we are not able to choose just one of these two summands for induction, the final upper bound of the G-area is $O(n^2)$, because of the aforementioned quadratic estimate of the mixture in terms of n.

In the case of diagrams with hubs, we estimate a similar linear combination, but the inductive parameter is not the (modified) perimeter n but the sum $\Sigma = n + \sigma(\Delta)$. The invariant $\sigma(\Delta) = \sigma_\lambda(\Delta$ was invented in [18]. It is defined by the design formed by maximal bands of two types in Δ. The important and non-trivial feature of the σ-invariant is the linear inequality $\sigma_\lambda(\Delta) = O(n)$, and so the quadratic upper bound of the form $O(\Sigma^2)$ is also quadratic in terms of the perimeter n.

In fact in both cases (over M or over G), the proof proceeds by taking a minimal counterexample diagram Δ and then by performing surgeries trying to find a smaller counterexample. This provides more and more useful information about Δ, until finally one of the surgeries succeeds and we show that Δ could not have been a minimal counterexample.

For instance, in Section 7 where diagrams with hubs are considered, we need to remove one of the disks from the diagram. As in our previous papers (starting with [28] and [16]), we use hyperbolicity of certain graph associated with hubs (hubs are vertices, q-bands connecting hubs are edges), and find a hub connected to the boundary of the whole diagram by almost all bands starting on the hub. This gives a subdiagram of Δ consisting of a subdiagram called a *clove* and a disk. We would like to remove that subdiagram from Δ producing a smaller counterexample.

A similar task was solved in [28]. It is one of the most non-trivial parts of [28]. Using it, we decomposed a diagram in [28] into a few disks of small total perimeter, and a diagram without hubs, it was called the *snowman decomposition*. But that task is now much harder than in [28]. The reason is that in [28], after removing the clove and the disk, we needed to show that the perimeter of the diagram decreases and the perimeter of the removed disk (only the disk) is linearly bounded by the difference of the perimeters of the old and new diagrams. For the quadratic upper bound this is not enough. We need to get a linear lower bound of the difference in terms of the whole piece that we cut off (the clove and the disk). That can be achieved not always. If not, we get a new information about the disk and the clove and remove the disk together with a certain sub-clove. The mixture and σ_λ invariant help achieve it at the end.

Some estimates used in this paper are very similar to the estimates in [18] More precisely for every function $f(n)$ satisfying certain conditions, a finitely presented group G_f with Dehn function $n^s f(n)^3$ (where $s \geq 2$) is constructed in [18]. In particular, if $s = 2$ and $f(n)$ is a constant, then G_f has quadratic Dehn function. Although the group G_f in [18] is very different from the group G in this paper, the underlying S-machines have similar enough properties, so that we could use identical and almost identical proofs of several lemmas (which indicates that there is a general theory of S-machines for which this paper and [18] are applications). For the sake of completeness, we include these lemmas here.

2 S-machines

2.1 S-machines as rewriting systems

There are several equivalent definitions of S-machines (see [26]) We are going to use the following definition which is easily seen to be equivalent to the original definition from [28] (essentially the same definition was used in [21]):

A "hardware" of an S-machine $\mathbf{S}$ is a pair (Y, Q), where $Q = \sqcup_{i=0}^{n} Q_i$ and $Y = \sqcup_{i=1}^{n} Y_i$ for some $n \geq 1$. Here and below $\sqcup$ denotes the disjoint union of sets.

We always set $Y_n = Y_0 = \emptyset$ and if $Q_n = Q_0$ (i.e., the indices of Q_i are counted mod n, then we say that $\mathbf{S}$ is a *circular* S-machine.

The elements from Q are called *state letters*, the elements from Y are *tape letters*. The sets Q_i (resp. Y_i) are called *parts* of Q (resp. Y).

The *language of admissible words* consists of reduced words W of the form

$$q_1 u_1 q_2 \dots u_s q_{s+1}, \tag{2.2}$$

where 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 $i = 1, ..., s$ 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 $k(i) = j(i) + 1$.
- 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 $k(i) = j(i)$.

Every subword $q_i u_i q_{i+1}$ of an admissible word (2.2) 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.

For every word W, if we delete all non-$Y^{\pm 1}$ letters from W we get the Y-projection of the word W. The length of the Y-projection of W is called the Y*-length* and is denoted by $|W|_Y$. 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 (2.2), $W = q_1 u_1 q_2 u_2 ... q_s$, and $q_i \in Q_{j(i)}^{\pm 1}$, $i = 1, ..., s$, u_i are group words in tape letters, then we shall say that the *base* of W is the word $Q_{j(1)}^{\pm 1} Q_{j(2)}^{\pm 1} ... Q_{j(s)}^{\pm 1}$. 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 **S** are $Q_0, Q_1, ..., Q_n$ we will write that the *the standard base* of the S-machine is $Q_0 ... Q_n$.

The *software* of an S-machine with the standard base $Q_0 ... Q_n$ is a set of *rules* Θ. Every $\theta \in \Theta$ is a sequence $[q_0 \to a_0 q_0' b_0, ..., q_n \to a_n q_n' b_n]$ and a subset $Y(\theta) = \sqcup Y_j(\theta)$, where $q_i \in Q_i$, a_i is a reduced word in the alphabet $Y_{i-1}(\theta)$, b_i is a reduced word in $Y_i(\theta)$, $Y_i(\theta) \subseteq Y_i$, $i = 0, ..., n$ (recall that $Y_0 = Y_n = \emptyset$).

Each component $q_i \to a_i q_i' b_i$ is called a *part* of the rule. In most cases the sets $Y_j(\theta)$ will be equal to either Y_j or $\emptyset$. By default $Y_j(\theta) = Y_j$.

Every rule

$$\theta = [q_0 \to a_0 q_0' b_0, ..., q_n \to a_n q_n' b_n]$$

has an inverse

$$\theta^{-1} = [q_0' \to a_0^{-1} q_0 b_0^{-1}, ..., q_n' \to a_n^{-1} q_n b_n]$$

which is also a rule of **S**. It is always the case that $Y_i(\theta^{-1}) = Y_i(\theta)$ for every i. Thus the set of rules Θ of an S-machine is divided into two disjoint parts, Θ^+ and Θ^- such that for every $\theta \in \Theta^+$, $\theta^{-1} \in \Theta^-$ and for every $\theta \in \Theta^-$, $\theta^{-1} \in \Theta^+$ (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 $\theta = [q_0 \to a_0 q_0' b_0, ..., q_n \to a_n q_n' b_n]$ as above to an admissible word $p_1 u_1 p_2 u_2 ... p_s$ (2.2) where each $p_i \in Q_{j(i)}^{\pm 1}$ means

- check if u_i is a word in the alphabet $Y_{j(i)+1}(\theta)$ when $p_i \in Q_{j(i)}$ or if it is a word in $Y_{j(i)}(\theta)$ when $p_i \in Q_{j(i)}^{-1}$ $(i = 1, \ldots, s-1)$; and if this property holds,
- replace each $p_i = q_{j(i)}^{\pm 1}$ by $(a_{j(i)} q_{j(i)}' b_{j(i)})^{\pm 1}$,
- if the resulting word is not reduced or starts (ends) with Y-letters, then reduce the word and trim the first and last Y-letters to obtain an admissible word again.

For example, applying the rule $[q_1 \to a^{-1}q_1'b, q_2 \to cq_2'd]$ to the admissible word $q_1b^{-1}q_2dq_2^{-1}q_1^{-1}$ 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 (i.e., W belongs to the *domain* of θ) then we denote the result of application of θ to W by $W \cdot \theta$. Hence each rule defines an invertible partial map from the set of configurations to itself, and one can consider an S-machine as an inverse semigroup of partial bijections of the set of admissible words.

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

We usually 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, usually the Q_0Q_1-sector, of every configuration. The Y-projection of that sector is called the *input* of the configuration. In that case, the end configuration with empty Y-projection is called the *accept* configuration. If the S-machine (viewed as a semigroup of transformations as above) can take an input configuration with input u to the accept configuration, we say that u is *accepted* by the S-machine. We define *accepted* configurations (not necessarily start configurations) similarly.

A *computation* of *length* $t \ge 0$ is a sequence of admissible words $W_0 \to \cdots \to W_t$ 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 $H = \theta_1 \dots \theta_t$ is called the *history* of the computation. Since W_t is determined by W_0 and the history H, we use notation $W_t = W_0 \cdot H$.

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.

Note, though, that in this paper, unlike the previous ones, we consider non-reduced computations too because these may correspond to reduced van Kampen diagrams under our present interpretation of S-machines in groups.

The *space* of a computation $W_0 \to \cdots \to W_t$ is $\max_{i=0}^{t} ||W_i||$, where $||W_i||$ is the maximal length of W_i.

If for some rule $\theta = [q_0 \to a_0q_0'b_0, ..., q_n \to a_nq_n'b_n] \in \Theta$ of an S-machine $\mathbf{S}$ the set $Y_{i+1}(\theta)$ is empty (hence in every admissible word in the domain of θ every Q_iQ_{i+1}-sector has no Y-letters) then we say that θ locks the Q_iQ_{i+1}-sector. In that case we always assume that b_i, a_{i+1} are empty and we denote the i-th part of the rule $q_i \xrightarrow{\ell} a_iq_i'$. If the Q_iQ_{i+1}-sector is locked by θ then we also assume that a_{i+1} is empty too.

Remark 2.1. It is easy to see that the substitution $[q_i \xrightarrow{\ell} aq_i', q_{i+1} \to q_{i+1}'b]$ is equivalent to the substitution $[q_iq_{i+1} \to aq_i'q_{i+1}'b]$. Thus for the sake of brevity we will allow parts of rules of the form $q_i...q_j \to aq_i'...q_j'b$. If the rule locks the Q_sQ_{s+1}-sector where Q_s is the part of state letters containing q_j, q_j', then we write $q_i...q_j \xrightarrow{\ell} aq_i'...q_j'b$ (in that case b is empty).

The above definition of S-machines resembles the definition of multi-tape Turing machines (see [28]). The main differences are that every state letter of an S-machines is blind: it does not "see" tape letters next to it (two state letters can see each other if they stay next to each other). Also S-machines are symmetric (every rule has an inverse), can work with words containing negative letters, and words with "non-standard" order of state letters.

It is important that S-machines can simulate the work of Turing machines. This non-trivial fact, especially if one tries to get a polynomial time simulation, was first proved in [28]. but we do not need a restriction on time, and it would be more convenient for us to use an easier S-machine from [21].

Let $\mathbf{M}_0$ be a deterministic Turing machine accepting a non-recursive language $\mathcal{L}$ of words in the one-letter alphabet $\{\alpha\}$.

Lemma 2.2. *([21]) There is a recognizing S-machine $\mathbf{M}_1$ whose language of accepted input words is $\mathcal{L}$. In every input configuration of $\mathbf{M}_1$ there is exactly one input sector, the first sector of the word, and all other sectors are empty of Y-letters.*

We say that two recognizing S-machines are *equivalent* if they have the same language of accepted configurations.

We can simplify rules of any S-machine in the obvious way.

Lemma 2.3. *Every S-machine $\mathbf{S}$ is equivalent to an S-machine $\mathbf{S}'$, where*

(*) every part $q_i \to aq_ib$ of an S-rule of $\mathbf{S}'$ has $||a|| \le 1$, $||b|| \le 1$, i.e., both words a and b are just letters from $Y^{\pm 1}$ or empty words;

(**) moreover $\mathbf{S}'$ can be constructed so that for every rule $\theta = [q_0 \to a_0q_0'b_0, ..., q_n \to a_nq_n'b_n]$ of $\mathbf{S}'$, we have $\sum_i(||a_i|| + ||b_i||) \le 1$.

For example, a rule $[q \to aq'b]$ is equivalent to the set of two rules $[q \to aq'']$, $[q'' \to q'b]$ where q'' is a new state letter added to the part containing q and q'.

Thus, applying Lemma 2.2 we will assume that the machine $\mathbf{M}_1$ satisfies Property (**).

2.2 Some elementary properties of S-machines

The base of an admissible word is not always a reduced word. However the following is an immediate corollary of the definition of admissible word.

Lemma 2.4. *If the i-th component of the rule θ has the form $q_i \overset{\ell}{\to} a_iq_i'$, then the base of any admissible word in the domain of θ cannot have subwords $Q_iQ_i^{-1}$ or $Q_{i+1}^{-1}Q_{i+1}$.*

In this paper we are often using copies of words. If A is an alphabet and W is a word involving no letters from $A^{\pm 1}$, then to obtain a *copy* of W in the alphabet A we substitute letters from A for letters in W so that different letters from A substitute for different letters. Note that if U' and V' are copies of U and V respectively corresponding to the same substitution, and $U' \equiv V'$, then $U \equiv V$, where '≡' means leter-by-letter equality of words. We also use copies of S-machines (defined in the same way).

The following two lemmas also immediately follow from definitions (see details in [18, Lemmas 2.6,2.7]).

Lemma 2.5. *Suppose that the base of an admissible word W is Q_iQ_{i+1}. Suppose that each rule of a reduced computation starting with $W \equiv q_iuq_{i+1}$ and ending with $W' \equiv q_i'u'q_{i+1}'$ multiplies the Q_iQ_{i+1}-sector by a letter on the left (resp. right). And suppose that different rules multiply that sector by different letters. Then*

(a) the history of computation is a copy of the reduced form of the word $u'u^{-1}$ read from right to left (resp. of the word $u^{-1}u'$ read from left to right). In particular, if $u \equiv u'$, then the computation is empty;

(b) the length of the history H of the computation does not exceed $||u|| + ||u'||$;

(c) for every configuration $q_i''u''q_{i+1}''$ of the computation, we have

$$||u''|| \le \max(||u||, ||u'||).$$

Lemma 2.6. *Suppose the base of an admissible word W is Q_iQ_{i+1}. Assume that each rule of a reduced computation starting with $W \equiv q_iuq_{i+1}$ and ending with $W' \equiv q_i'u'q_{i+1}'$ multiplies the Q_iQ_{i+1}-sector by a letter on the left and by a letter from the right. Suppose different rules multiply that sector by different letters and the left and right letters are taken from disjoint alphabets. Then*

(a) for every intermediate configuration W_j of the computation, we have $||W_j|| \le \max(||W||, ||W'||)$

(b) the length of the history H of the computation does not exceed $\frac{1}{2}(||u|| + ||u'||)$.

The next statement is Lemma 3.7 from [17].

Lemma 2.7. *Suppose the base of an admissible word W of an S-machine $\mathbf{S}$ is $Q_iQ_i^{-1}$ (resp., $Q_i^{-1}Q_i$). Suppose that each rule θ of a reduced computation starting with $W \equiv q_iuq_i^{-1}$ (resp., $q_i^{-1}uq_i$), where $u \ne 1$, and ending with $W' \equiv q_i'u'(q_i')^{-1}$ (resp., $W' \equiv (q_i')^{-1}u'q_i'$) has a part $q_i \to a_\theta q_i' b_\theta$, where b_θ (resp., a_θ) is a letter, and for different θ-s the b_θ-s (resp., a_θ-s) are different. Then the history of the computation has the form $H_1H_2^kH_3$, where $k \ge 0$, $||H_2|| \le \min(||u||, ||u'||)$, $||H_1|| \le ||u||/2$, and $||H_3|| \le ||u'||/2$.*

Lemma 2.8. *Suppose that a reduced computation $W_0 \to W_1 \to \dots \to W_t$ of an S-machine $\mathbf{S}$ satisfying (*) has a 2-letter base and the history of the form $H \equiv H_1H_2^kH_3$ ($k \ge 0$). Then for the Y-projection w_i of W_i ($i = 0, 1, \dots, t$) , we have the inequality*

$$||w_i|| \le ||w_0| + ||w_t|| + 2||H_1|| + 3||H_2|| + 2||H_3||$$

.

Proof. By (*) we have that the absolute value of $||w_i|| - ||w_{i-1}||$ is at most 2 for every $i = 1, \dots t$. Therefore for $i \le ||H_1||$, we have $||w_i|| \le ||w_0|| + 2||H_1||$. Similarly, $||w_i|| \le ||w_t|| + 2||H_3||$ for $i \ge t - ||H_3||$. It remains to assume that $||H_1|| < i < t - ||H_3||$.

Denote the words w_i with $i = ||H_1|| + j||H_2||$, by u_j, $j = 0, 1, \dots, k$ and the corresponding words W_i by U_j. Then there exist two words v_l, v_r such that for every s from 1 to k, $u_s = v_lu_{s-1}v_r$ in a free group for some Y-words v_l and v_r depending on H_2. Hence $u_j = v_l^ju_0v_r^j$, where both v_l and v_r have length at most $||H_2||$ by (*).

By [20, Lemma 8.1], the length of an arbitrary word U_j then is not greater than $||v_l|| + ||v_r|| + ||U_0|| + ||U_k||$ provided $0 \le j \le k$.

Now we need to estimate the lengths of W_i ($i = |||H_1||, ..., t - ||H_3||$, such that w_i which are not equal to any u_j. Choose j such that the absolute value of $i - j||H_2||$ does not $||H_2||/2$. Then the absolute value of $||w_i|| - ||u_j||$ does not exceed $||H_2||$ by (*), and therefore $||W_i|| \le ||v_l|| + ||v_r|| + ||U_0|| + ||U_k|| + ||H_2||$. Since $||U_0|| \le ||w_0|| + 2||H_1||$ and $||U_k|| \le ||w_t|| + 2||H_3||$, we obtain

$$\begin{aligned} ||w_i|| &\le ||v_l|| + ||v_r|| + ||w_0|| + ||w_t|| + 2||H_1|| + 2||H_3|| + ||H_2|| \\ &\le ||w_0| + ||w_t|| + 2||H_1|| + 2||H_3|| + 3||H_2|| \end{aligned}$$

for every i, as required. □

2.3 The highest parameter principle

In this paper, we estimate length and space of computations of S-machines, and also areas and other numerical invariants of van Kampen diagrams. The following constants will be used in the estimates throughout this paper.

$$\begin{aligned} &\lambda^{-1} \ll m \ll N \ll c_0 \ll c_1 \ll c_2 \ll c_3 \ll c_4 \ll c_5 \ll L_0 \ll L \ll K \ll \\ &J \ll \delta^{-1} \ll c_6 \ll c_7 \ll N_1 \ll N_2 \ll N_3 \ll N_4 \end{aligned} \tag{2.3}$$

where $\ll$ means "much smaller".

For each inequality in this paper involving several of these constants, let D be the biggest constant appearing there. The inequality always can then be rewritten in the form

$$D \ge \text{some expression involving smaller constants.}$$

This *highest parameter principle* [14] makes the system of inequalities used in this paper consistent.

3 Auxiliary S-machines and constructions

3.1 Running state letters

For every alphabet Y we define a "running state letters" S-machine $\mathbf{LR}(Y)$. We will omit Y if it is obvious or irrelevant. The standard base of $\mathbf{LR}(Y)$ is $Q^{(1)}PQ^{(2)}$ where $Q^{(1)} = \{q^{(1)}\}$, $P = \{p^{(i)}, i = 1, 2\}$, $Q^{(2)} = \{q^{(2)}\}$. The state letter p with indices runs from the the state letter $q^{(2)}$ to the state letter $q^{(1)}$ and back. The S-machine $\mathbf{LR}$ will be used to check the "structure" of a configuration (whether the state letters of a configuration are in the appropriate order), and to recognize a computation by its history.

The alphabet of tape letters Y of $\mathbf{LR}(Y)$ is $Y^{(1)} \sqcup Y^{(2)}$, where $Y^{(2)}$ is a (disjoint) copy of $Y^{(1)}$. The positive rules of $\mathbf{LR}$ are defined as follows.

- $\zeta^{(1)}(a) = [q^{(1)} \to q^{(1)}, p^{(1)} \to a^{-1}p^{(1)}a', q^{(2)} \to q^{(2)}]$, where a is any positive letter from $Y = Y^{(1)}$ and a' is the corresponding letter in the copy $Y^{(2)}$ of $Y^{(1)}$.

  Comment. The state letter $p^{(1)}$ moves left replacing letters a from $Y^{(1)}$ by their copies a' from $Y^{(2)}$.

- $\zeta^{(12)} = [q^{(1)}p^{(1)} \to q^{(1)}p^{(2)}, q^{(2)} \to q^{(2)}]$.

  Comment. When $p^{(1)}$ meets $q^{(1)}$, $p^{(1)}$ turns into $p^{(2)}$.

- $\zeta^{(2)}(a) = [q^{(1)} \to q^{(1)}, p^{(2)} \to ap^{(2)}(a')^{-1}, q^{(2)} \to q^{(2)}]$

  Comment. The state letter $p^{(2)}$ moves right towards $q^{(2)}$ replacing letters a' from $Y^{(2)}$ by their copies a from $Y^{(1)}$.

The start (resp. end) state letters of **LR** are $\{q^{(1)}, p^{(1)}, q^{(2)}\}$ (resp. $\{q^{(1)}, p^{(2)}, q^{(2)}\}$).

Remark 3.1. Note that each of the rules $(\zeta^j)^{\pm 1}(a)$, $(j = 1, 2)$ either moves the state letter p left or moves it right, or deletes one letter from left and one letter from right, or insert letters from both sides of itself. In the later case, the next rule of a computation must be again $\zeta(j)^{\pm 1}(b)$ for some b, and if the computation is reduced, it again must increase the length of the configuration by two. This observation implies

Remark 3.2. Note that no rule of **LR** changes the projection of a configuration onto the free group with basis $Y^{(1)}$ if the state letters are mapped to 1 and the letters from $Y^{(2)}$ are mapped to their copies from $Y^{(1)}$. This will be later referred to as the *projection argument.*

Lemma 3.3. *Let $\mathcal{C}\colon W_0 \to \cdots \to W_t$ be a reduced computation of the S-machine* **LR** *with the standard base. Then*

(1) if $|W_i|_Y > |W_{i-1}|_Y$ for some $i = 1, \dots, t-1$, then $|W_{i+1}|_Y > |W_i|_Y$;

(2) $|W_i|_Y \le \max(|W_0|_Y, |W_t|_Y)$ for every $i = 0, 1, \dots t$;

(3) if $W_0 \equiv q^{(1)}up^{(1)}q^{(2)}$ and $W_t \equiv q^{(1)}vp^{(2)}q^{(2)}$ for some words u, v, then $u \equiv v$, $|W_i|_Y = |W_0|_Y$ for every $i = 0, \dots, t$, $t = 2k+1$, where $k = |W_0|_Y$, and the sector $Q^{(1)}P$ is locked in the transition $W_k \to W_{k+1}$. Moreover if W_0 and W_t have the form $q^{(1)}up^{(1)}q^{(2)}$ and $q^{(1)}vp^{(2)}q^{(2)}$, then the history H of $\mathcal{C}$ is a copy of the word $\bar{u}\zeta(12)(\bar{u}')^{-1}$ where $\bar{u}$ is the mirror image of u and $\bar{u}'$ is a copy of $\bar{u}$. Thus W_0, W_t, H uniquely determine each other in that case.

(4) if $W_0 \equiv q^{(1)}up^{(1)}q^{(2)}$ and $W_t \equiv q^{(1)}vp^{(1)}q^{(2)}$ for some u, v or $W_0 \equiv q^{(1)}up^{(2)}q^{(2)}$ and $W_t \equiv q^{(1)}vp^{(2)}q^{(2)}$ then $u \equiv v$ and the computation is empty ($t = 0$);

(5) if $W_0 \equiv q^{(1)}up^{(1)}q^{(2)}$ or $W_0 \equiv q^{(1)}p^{(1)}uq^{(2)}$, or $W_0 \equiv q^{(1)}up^{(2)}q^{(2)}$, or $W_0 \equiv q^{(1)}p^{(2)}uq^{(2)}$ for some word u, then $|W_i|_Y \ge |W_0|_Y$ for every $i = 0, \dots, t$.

Proof. For every $i = 0, ..., t$ let $W_i = q^{(1)}u_ip^{(l_i)}v_iq^{(2)}$ where u_i is a word in Y, v_i is a word in Y' (it is easy to check by induction on i that this is true for every i).

Suppose that $|W_{i-1}|_Y < |W_i|_Y$ for some i. That means that the i-th rule in the computation is of the form $(\zeta^{(k)}(a))^{\pm 1}$. This rule multiplies u_{i-1} by a letter $a^{\pm 1}$ on the right, and multiplies v_{i-1} by a copy of the inverse of that letter on the left, and these letters do not cancel in u_i, v_i. In particular both u_i and v_i are not empty. Hence $\zeta^{(12)}$ does not apply to W_i. Thus the rule in $W_i \to W_{i+1}$ is $(\zeta^{(j)}(b))^{\pm 1}$ (with the same j) and it multiplies $u_i = u_{i-1}a$ by $b^{\pm 1}$ on the right and multiplies v_i by a copy of the inverse of that letter on the left. Since the computation is reduced, $b \ne a^{-1}$. Therefore $|W_{i+1}|_Y > |W_i|_Y$. Continuing in this manner, we establish (1).

To establish (2), we can choose the shortest word W_j in the computation and apply (1) to the computation $W_j \to \cdots \to W_t$ and the inverse computation $W_j \to \cdots \to W_0$.

Suppose that the assumptions of (3) hold. Then $u \equiv v$ by the projection argument. Since $\zeta^{(12)}$ locks Q^1P-sector, the p-letter must reach $q^{(1)}$ moving always left to change $p^{(1)}$ by $p^{(2)}$, and so $W_k \equiv q^{(1)}p^{(1)}\dots$. If the next rule of the form $\zeta^{(1)}(a)^{\pm 1}$ could increase the length of the configuration, we would obtain a contradiction with Property (1). Since the computation is reduced, the next rule is $\zeta^{(12)}$, and arguing in this way, one uniquely reconstructs the whole computation in case (3) for given W_0 or W_t, and vice versa, the history H determines both u and v. Propery (4) holds for same reasons.

By the projection argument, we have $|q^{(1)}up^{(1)}q^{(2)}|_Y = ||u|| \le |W_i|_Y$ if the first assumptions of (5) holds. The other cases of (5) are similar. □

The projection argument also immediately gives:

Lemma 3.4. *If $W_0 \to \dots \to W_t$ is a reduced computation of* **LR** *with base*

$$Q^{(1)}PP^{-1}(Q^{(1)})^{-1} \text{ or } (Q^{(2)})^{-1}P^{-1}PQ^{(2)}$$

and

$$W_0 \equiv q^{(1)}p^{(i)}u(p^{(i)})^{-1}(q^{(1)})^{-1}(i=1,2)$$

or

$$W_0 \equiv (q^{(2)})^{-1}(p^{(i)})^{-1}v(p^{(i)})q^{(2)}(i=1,2)$$

for some words u, v, then $|W_j|_Y \ge |W_0|_Y$ for every $j = 0, \dots, t$.

Remark 3.5. We will also use the right analog **RL** of **LR**. The base of **RL** is Q_1RQ_2. The state letter r first moves right from $q^{(1)}$ to $q^{(2)}$ and then left. Lemmas "left-right dual" to Lemmas 3.3 and 3.4 as well as Remark 3.2 are true for **RL** as well.

Remark 3.6. For every $m \ge 1$, we will also need the S-machine $\mathbf{LR}_m$, that repeats the work of **LR** m times. That is the S-machine $\mathbf{LR}_m$ runs the state letter p back and forth between $q^{(2)}$ and $q^{(1)}$ m times. Every time p meets $q^{(1)}$ or $q^{(2)}$, the upper index of p increases by 1 after the application of the rule $\zeta^{(i,i+1)}$ $(i = 1, \dots, 2m-1)$, so the highest upper index of p is $(2m)$. A precise definition of $\mathbf{LR}_m$ is obvious and is left to the reader. (Recall that m is one of the system of parameters used in this paper (see Section 2.3).)

Remark 3.7. The analog of Lemma 3.3 holds for $\mathbf{LR}_m$. In particular, if

$$W_t \equiv q^{(1)}vp^{(2m)}q^{(2)}$$

in the formulaion of (3), then $t = 2mk + 2m - 1$ (the proof is essentially the same and is left to the reader).

3.2 Adding history sectors

We will add new (history) sectors to our S-machine $\mathbf{M}_1$. If we ignore the new sectors, we get the hardware and the software of the S-machine $\mathbf{M}_1$. The new S-machine $\mathbf{M}_2$ will start with a configuration where in every history sector a copy of the history H of a computation of $\mathbf{M}_1$ is written. Then it will execute H on the other (working) sectors simulating the work of $\mathbf{M}_1$, while in the history sector, a state letter moves scans the history, one symbol at a time. Thus if a computation with the standard base starts with a configuration W and ends with configuration W', then the length of the computation does not exceed $||W|| + ||W'||$.

Here is a precise definition of $\mathbf{M}_2$. Recall that the S-machine $\mathbf{M}_1$ satisfies the condition (**) of Lemma 2.3 and has hardware (Q, Y), where $Q = \sqcup_{i=0}^{n} Q_i$, and the set of rules Θ. The new S-machine $\mathbf{M}_2$ has hardware

$$Q_{0,r} \sqcup Q_{1,\ell} \sqcup Q_{1,r} \sqcup Q_{2,\ell} \sqcup Q_{2,r} \sqcup \dots \sqcup Q_{n,\ell}, \quad Y_h = Y_1 \sqcup X_1 \sqcup Y_2 \sqcup \dots \sqcup X_{n-1} \sqcup Y_n$$

where $Q_{i,\ell}$ and $Q_{i,r}$ (left and right) copies of Q_i X_i is a disjoint union of two copies of Θ^+, namely $X_{i,\ell}$ and $X_{i,r}$. (The sets $Q_{0,\ell}$, $Q_{n,r}$ are empty.) Every letter q from Q_i has two copies $q^{(\ell)} \in Q_{i,\ell}$ and $q^{(r)} \in Q_{i,r}$. By definition, the start (resp. end) state letters of $\mathbf{M}_2$ are copies of the corresponding start (end) state letters of $\mathbf{M}_1$. The $Q_{0,r}Q_{1,\ell}$-sectors are the *input sectors* of configurations of $\mathbf{M}_2$.

The positive rules θ_h of $\mathbf{M}_2$ are in one-to-one correspondence with the positive rules θ of $\mathbf{M}_1$. If $\theta = [q_0 \to a_0q_0'b_0, ..., q_n \to a_nq_n'b_n]$ is a positive rule of $\mathbf{M}_1$, then each part $q_i \to a_iq_i'b_i$ is replaced in θ_h by two parts

$$q_{i,\ell} \to a_iq_{i,\ell}'h_{\theta,i}^{-1}$$

and

$$q_{i,r} \to \overline{h}_{\theta,i}q_{i,r}'b_i,$$

where $h_{\theta,i}$ (resp., $\overline{h}_{\theta,i}$) is a copy of θ in the alphabet $X_{i,\ell}$ (in $X_{i,r}$, respectively).

If θ is the start (resp. end) rule of $\mathbf{M}_1$, then for any word in the domain of θ_h (resp. θ_h^{-1} all Y-letters in history sectors are from $\sqcup_i X_{i,\ell}$ (resp. $\sqcup X_{i,r}$).

Thus for every rule θ of $\mathbf{M}_1$, the rule θ_h of $\mathbf{M}_2$ acts in the $Q_{i,r}Q_{i+1,\ell}$-sector in the same way as θ acts in the Q_iQ_{i+1}-sector. In particular, Y-letters which can appear in the $Q_{i,r}Q_{i+1,\ell}$-sector of an admissible word in the domain of θ_h are the same as the Y-letters that can appear in the Q_iQ_{i+1}-sector of an admissible word in the domain of θ. Hence if θ locks Q_iQ_{i+1}-sectors, then θ_h locks $Q_{i,r}Q_{i+1,\ell}$-sectors.

Comment. Every computation of S-machine $\mathbf{M}_2$ with history H and the standard base coincides with the a computation of $\mathbf{M}_1$ whose history is a copy of H if one observes it only in *working* sectors $Q_{i,r}Q_{i+1,l}$. In the standard base of $\mathbf{M}_2$ the *working sectors* $Q_{i,r}Q_{i+1,\ell}$ alternate with *history* sectors $Q_{i,\ell}Q_{i,r}$. Every positive rule θ_h multiples the content of the history $Q_{i,\ell}Q_{i,r}$-sector by the corresponding letter $\overline{h}_{\theta,i}$ from the right and by letter $h_{\theta,i}^{-1}$ from the left. Thus if the S-machine $\mathbf{M}_2$ executes the history written in the history sectors, then the history word H in letters from $X_{i,\ell}$ gets rewritten into the copy of H in letters from $X_{i,r}$. Say, if the copy of the history H was written in a history sector as $h_1h_2h_3$, then during the computation with history H it will transform as follows:

$$h_1h_2h_3 \to h_2h_3\overline{h}_1 \to h_3\overline{h}_1\overline{h}_2 \to \overline{h}_1\overline{h}_2\overline{h}_3.$$

Let $I_1(\alpha^k)$ be a start configuration of $\mathbf{M}_1$ (i.e.,a configuration in the domain of the start rule of $\mathbf{M}_1$) with α^k written in the input sector (all other sectors do not contain Y-letters). Then the corresponding start configuration $I_2(\alpha^k, H)$ of $\mathbf{M}_2$ is obtained by first replacing each state letter q by the product of two corresponding letters $q^{(\ell)}q^{(r)}$, and then inserting a copy of H in the *left alphabet* $X_{i,\ell}$ in every history $Q_{i,\ell}Q_{i,r}$-sector. End configurations $A_2(H)$ of $\mathbf{M}_2$ are defined similarly, only the Y-letters in the history sectors must be from the *right alphabet* $X_{i,r}$.

Lemma 3.8. *(1) If a word α^k is accepted by the Turing machine $\mathbf{M}_0$, then for some word H, there is a reduced computation $I_2(\alpha^k, H) \to \cdots \to A_2(H)$ of the S-machine $\mathbf{M}_2$.*

(2) If there is a computation $I_2(\alpha^k, H) \to \cdots \to A_2(H')$ of $\mathbf{M}_2$, then the word α^k is accepted by $\mathbf{M}_0$ and $H' \equiv H$.

Proof. (1) The word α^k is accepted by the S-machine $\mathbf{M}_1$ by Lemma 2.2. If H is the history of the accepting computation of $\mathbf{M}_1$, then the computation of $\mathbf{M}_2$ with history H starting with $I_2(\alpha^k, H)$ ends with $A_2(H)$ since $\mathbf{M}_2$ works as $\mathbf{M}_1$ in the working sectors and replaces the letters from the left alphabets by the corresponding letters from the right alphabets in the history sectors.

(2) If $I_2(\alpha^k, H)\cdot H'' = A_2(H')$ for some history H'' of $\mathbf{M}_2$ then the word α^k is accepted by $\mathbf{M}_0$ by Lemma 2.2 and the fact that $\mathbf{M}_2$ works as $\mathbf{M}_1$ in the working sectors. Note that both H and H' must be the copies of H'', because the word $I_2(\alpha^k, H)$ has no letters from right alphabets, $A_2(H')$ has no letters from left alphabets, and every rule multiplies the Y-projection of every history sector by a letter from $X_{i,\ell}^{-1}$ (from $X_{i,r}$) on the left (resp., on the right). □

The sectors of the form $Q_{i,\ell}Q_{i,\ell}^{-1}$ and $Q_{i,r}^{-1}Q_{i,r}$ (in a non-standard base) are also called *history* sectors. History sectors help obtaining a linear estimate of the space of every computation $W_0 \to \cdots \to W_t$ in terms of $||W_0|| + ||W_t||$.

Lemma 3.9. *Let $W_0 \to \cdots \to W_t$ be a reduced computation of $\mathbf{M}_2$ with base $Q_{i,\ell}Q_{i,r}$ and history H. Assume that all the Y-letters of W_0 belong to only one of the alphabets $X_{i,\ell}$ or $X_{i,r}$. Then $||H|| \le |W_t|_Y$ and $|W_0|_Y \le |W_t|_Y$*

Proof. Let $W_i = q_i v_i q_i'$, $i = 0, ..., t$, and assume that v_0 has no letters from $X_{i,r}$. Then $v_t = uv_0u'$, where u is a copy of H^{-1} in the alphabet $X_{i,\ell}$ and u' is a copy of H in $X_{i,r}$. So no letter of u' is cancelled in the product uv_0u', Therefore $|W_t|_Y \ge ||u'|| = ||H||$ and $|W_t|_Y \ge |W_0|_Y$. □

Lemma 3.10. *For any reduced computation $W_0 \to \cdots \to W_t$ of S-machine $\mathbf{M}_2$ with base of length at least 3, we have $|W_i|_Y \le 9(|W_0|_Y + |W_t|_Y)$ $(0 \le i \le t)$.*

Proof. Let $Q_{i_1}^{\pm 1} \dots Q_{i_k}^{\pm 1}$ be the base of the computation. We can divide the base into several subwords of length 3 or 4, each containing one history sector. Thus we can assume that k is equal to 3 or 4 and that the base contains one history sector. Without loss of generality, that history sector is either a $Q_{i,\ell}Q_{i,r}$-sector or a $Q_{i,\ell}Q_{i,\ell}^{-1}$-sector or a $Q_{i,r}^{-1}Q_{i,r}$-sector.

Consider two cases.

1. The history sector has the form $Q_{i,\ell}Q_{i,r}$. By Lemma 2.6, we have $||H|| \le \frac{1}{2}(|W_0|_Y + |W_t|_Y)$. It follows from property (*) of Lemma 2.3 that $|\,|W_{i+1}|_Y - |W_i|_Y\,| \le 6$ for every i. Therefore

$$|W_i|_Y \le \max(|W_0|_Y, |W_t|_Y) + 3||H|| \le$$

$$\max(|W_0|_Y, |W_t|_Y) + \frac{3}{2}(|W_0|_Y + |W_t|_Y) \le \frac{5}{2}(|W_0|_Y + |W_t|_Y)$$

2. The history sector is either a $Q_{i,\ell}Q_{i,\ell}^{-1}$-sector or a $Q_{i,r}^{-1}Q_{i,r}$-sector. Then one can apply Lemma 2.7 to the history sector and obtain the factorization $H \equiv H_1H_2^cH_3$, with $c \ge 0$, $||H_2|| \le \min(||u_0||, ||u_t||)$, $||H_1|| \le ||u_0||/2$, and $||H_3|| \le ||u_t||/2$, where u_0 and u_t are the Y-projections of the history sectors of W_0 and W_t, respectively. Since every W_i has at most three sectors, applying Lemma 2.8 to each of them, we obtain:

$$|W_i|_Y \le |W_0|_Y + |W_t|_Y + 3(2||H_1|| + 3||H_2|| + 2||H_3||) \le$$

$$|W_0|_Y + |W_t|_Y + 3|W_0|_Y + 9\min(|W_0|_Y, |W_t|_Y) + 3|W_t|_Y \le 9(|W_0|_Y + |W_t|_Y).$$

□

Lemma 3.11. *Suppose that a reduced computation $W_0 \to \dots \to W_t$ of the S-machine $\mathbf{M}_2$ starts with an admissible word W_0 having no letters from the alphabets $X_{i,l}$ (resp., from the alphabets $X_{i,r}$) . Assume that the length of its base B is bounded from above by a constant N_0, and B has a history subword $Q_{i,\ell}Q_{i,r}$. Then there is a constant $c = c(N_0)$ such that $|W_0|_Y \le c|W_t|_Y$.*

Proof. Let $V_0 \to \dots \to V_t$ be the restriction of the computation to the $Q_{i,\ell}Q_{i,r}$-sector. By Lemma 3.9, we have $t \le |V_t|_Y$ and $|V_0|_Y \le |V_t|_Y$.

It follows from (*) that

$$|W_0|_Y \le |W_t|_Y + 2N_0 t \le |W_t|_Y + 2N_0|V_t|_Y \le (2N_0+1)|W_t|_Y$$

It suffices to choose $c = 2N_0 + 1$. □

3.3 Adding running state letters

Our next S-machine will be a composition of $\mathbf{M}_2$ with **LR** and **RL**. The running state letters will control the work of $\mathbf{M}_3$.

First we replace every part Q_i of the state letters in the standard base of $\mathbf{M}_2$ by three parts $P_iQ_iR_i$ where P_i, R_i contain the running state letters. Thus if $Q_0...Q_s$ is the standard base of $\mathbf{M}_2$ then the standard base of $\overline{\mathbf{M}}_2$ is

$$P_0Q_0R_0P_1Q_1R_1 \dots P_sQ_sR_s, \tag{3.4}$$

where P_i (resp., R_i) contains copies of running P-letters (resp. R-letters) of **LR** (resp. **RL**), $i = 0, \dots, s$.

For every rule θ of $\mathbf{M}_2$, its i-th part $[q_i \to a_iq_i'b_i]$ is replaced in $\overline{\mathbf{M}}_2$ with

$$[p^{(i)}q_ir^{(i)} \to a_ip^{(i)}q_i'r^{(i)}b_i], (i = 0, \dots, s), \tag{3.5}$$

where $p^{(i)} \in P_i, r^{(i)} \in R_i$) do not depend on θ.

Comment. Thus, the sectors P_iQ_i and Q_iR_i are always locked. Of course, such a modification is useless for solo work of $\mathbf{M}_2$. But it will be helpful when one constructs a composition of $\overline{\mathbf{M}}_2$ with **LR** and **RL** which will be turned on after certain rules of $\overline{\mathbf{M}}_2$ are applied.

If Q_iQ_{i+1}-sector is a history sector of $\mathbf{M}_2$, then Q_iR_i-, R_iP_i-, P_iQ_i-sectors are history sectors of $\overline{\mathbf{M}}_2$. Accordingly the $Q_iQ_i^{-1}$-sectors ($R_iR_i^{-1}$-sectors, etc.) of admissible words with nonstandard bases will be called history sectors of $\overline{\mathbf{M}}_2$ too. (Alternatively, history sectors of admissible words of $\overline{\mathbf{M}}_2$ are those sectors which can contain letters from left or right alphabets.) The R_0P_1-sectors of admissible words are the input sectors. The $R_0R_0^{-1}-$ and $P_1^{-1}P_1$-sectors are also input sectors of admissible words of $\overline{\mathbf{M}}_2$.

If B is the base of some computation $\mathcal{C}$ of $\overline{\mathbf{M}}_2$, and UV is a 2-letter subword of B such that UV-sectors of admissible words in $\mathcal{C}$ are history (resp. working, input) sectors, then we will call UV a history (resp. working, input) subword of B.

3.4 $\mathbf{M}_3$

The next S-machine $\mathbf{M}_3$ is the composition of the S-machine $\overline{\mathbf{M}}_2$ with **LR** and **RL**. The S-machine $\mathbf{M}_3$ has the input, working and history sectors, i.e. the same base as $\overline{\mathbf{M}}_2$, although the parts of this base have more state letters than the corresponding parts of

$\overline{\mathbf{M}}_2$. It works as follows. Suppose that $\mathbf{M}_3$ starts with a start configuration of $\overline{\mathbf{M}}_2$, a word α^k in the input R_0P_1-sector, copies of a history word H in the alphabets $X_{i,\ell}$ in the history sectors, all other sectors empty of Y-letters. Then $\mathbf{M}_3$ first executes $\mathbf{RL}$ in all history sectors (moves the running state letter from R_i in the history sectors right and left), then it executes the history H of $\overline{\mathbf{M}}_2$. After that the Y-letters in the history sectors are in $X_{i,r}$ and $\mathbf{M}_3$ executes copies of $\mathbf{LR}$ in the history sectors (moves the running state letters left then right). After that $\mathbf{M}_3$ executes a copy of H backwards, getting to a copy of the same start configuration of $\overline{\mathbf{M}}_2$, runs $\mathbf{RL}$, executes a copy of the history H of $\overline{\mathbf{M}}_2$, runs a copy of $\mathbf{LR}$, etc. It stops after m times running $\mathbf{RL}, \overline{\mathbf{M}}_2, \mathbf{LR}, \overline{\mathbf{M}}_2^{-1}$ and running $\mathbf{RL}$ one more time.

Thus the S-machine $\mathbf{M}_3$ is a concatenation of $4m+1$ S-machines $\mathbf{M}_{3,1} - \mathbf{M}_{3,4m+1}$. After one of these S-machines terminates, a transition rule changes its end state letters to the start state letters of the next S-machine. All these S-machines have the same standard bases as $\overline{\mathbf{M}}_2$.

The configuration $I_3(\alpha^k, H)$ of $\mathbf{M}_3$ is obtained from $I_2(\alpha^k, H)$ by adding the control state letters $r_i^{(1)}$ and $p_i^{(1)}$ according to 3.4.

Set $\mathbf{M}_{3,1}$ is a copy of the set of rules of the S-machine $\mathbf{RL}$, with *parallel work* in all history sectors, i.e., every subword $Q_{i-1}R_{i-1}P_i$ of the standard base, where $Q_{i-1}Q_i$ is a history sector of $\mathbf{M}_2$, is treated as the base of a copy of $\mathbf{RL}$, that is R_{i-1} contain the running state letters which run between state letters from Q_{i-1} and P_i. Each rule of Set $\mathbf{M}_{3,1}$ executes the corresponding rule of $\mathbf{RL}$ simultaneously in each history sector of $\mathbf{M}_2$. The partition of the set of state letters of these copies of $\mathbf{RL}$ in each history sector is $X_{i,\ell} \sqcup X_{i,r}$ for some i (that is state letters from R_{i-1} first run right replacing letters from $X_{i,\ell}$ by the corresponding letters of $X_{i,r}$ and then run left replacing letters from $X_{i,r}$ by the corresponding letters of $X_{i,\ell}$.

The transition rule $\chi(1,2)$ changes the state letters by the state letters of start configurations of $\overline{\mathbf{M}}_2$. The admissible words in the domain of $\chi(1,2)^{\pm 1}$ have all Y-letters from the left alphabets $X_{i,\ell}$. The rule $\chi(1,2)$ locks all sectors except the history sectors $R_{i-1}P_i$ and the input sector. It does not apply to admissible words containing Y-letters from right alphabets.

Set $\mathbf{M}_{3,2}$ is a copy of the set of rules of the S-machine $\overline{\mathbf{M}}_2$.

The transition rule $\chi(2,3)$ changes the state letters of the stop configuration of $\overline{\mathbf{M}}_2$ by their copies in a different alphabet. The admissible words in the domain of $\chi(2,3)^{\pm 1}$ have no Y-letters from the left alphabets $X_{i,\ell}$. The rule $\chi(2,3)$ locks all sectors except for the history sectors $R_{i-1}P_i$. It does not apply to admissible words containing Y-letters from right alphabets.

Set $\mathbf{M}_{3,3}$ is a copy of the set of rules of the S-machine $\mathbf{LR}$, with parallel work in the same sectors as Set $\mathbf{M}_{3,1}$ (and the partition of Y-letters in each history sector $X_{i,r} \sqcup X_{i,\ell}$).

The transition rule $\chi(3,4)$ changes the state letters of the stop configuration of $\overline{\mathbf{M}}_2$ by their copies in a different alphabet. The admissible words in the domain of $\chi(3,4)^{\pm 1}$ have no Y-letters from the left alphabets $X_{i,l}$. The rule $\chi(3,4)$ locks all non-history sectors.

Set $\mathbf{M}_{3,4}$**.** The positive rules of Set $\mathbf{M}_{3,4}$ are the copies of the negative rules of the S-machine $\overline{\mathbf{M}}_2$.

The transition rule $\chi(4,5)$ changes the state letters of the start configuration of $\overline{\mathbf{M}}_2$ by their copies in a different alphabet. The admissible words in the domain of $\chi(4,5)^{\pm 1}$ have no Y-letters from the right alphabets $X_{i,r}$. The rule $\chi(4,5)$ locks all non-history and non-input sectors.

Sets $\mathbf{M}_{3,5},\dots,\mathbf{M}_{3,8}$ consist of rules that are copies of the rules of the Sets $\mathbf{M}_{3,1},\dots,$ $\mathbf{M}_{3,4}$, respectively.

...

Sets $\mathbf{M}_{3,4m-3},\dots,\mathbf{M}_{3,4m}$ consist of copies of the steps $\mathbf{M}_{3,1},\dots,\mathbf{M}_{3,4}$, respectively.

Set $\mathbf{M}_{3,4m+1}$ is a copy of Set $\mathbf{M}_{3,1}$. The end configuration for Set $\mathbf{M}_{3,4m+1}$, $A_3(H)$, is obtained from a copy of $A_2(H)$ by inserting the control letters according to (3.4).

The transition rules $\chi(i,i+1)$ are called χ-rules.

We say that a configuration W of the S-machine $\mathbf{M}_3$ is *tame* if every P- or R-letter is next to some Q-letter in W.

Lemma 3.12. *Let $\mathcal{C}\colon W_0\to\cdots\to W_t$ be a reduced computation of $\mathbf{M}_3$ consisting of rules of one of the copies of* **LR** *or* **RL** *with standard base. Then*

(a) $|W_j|_Y\le\max(|W_0|_Y,|W_t|_Y)$ for every configuration W_j of $\mathcal{C}$; moreover, $|W_0|_Y\le\cdots\le|W_t|_Y$ if W_0 is tame;

(b) $t\le||W_0||+||W_t||-2$, moreover, $t\le 2||W_t||-2$ if W_0 is tame.

Proof. (a) Let W_r be a shortest word of the computation $\mathcal{C}$. Then either $|W_r|_Y=|W_{r+1}|_Y=\cdots=|W_t|_Y$, or $|W_r|_Y=|W_{r+1}|_Y=\cdots=|W_s|_Y<|W_{s+1}|_Y$ for some s. It follows that the number of sectors increasing their lengths by two at the transition $W_s\to W_{s+1}$ is greater than the number of the sectors decreasing the lengths by 2. Now it follows from Lemma 3.3 (1) that the lengths of the Y-projections will keep increasing: $|W_{s+1}|_Y<|W_{s+2}|_Y<\dots$. So for every $j\ge r$, we have $|W_j|_Y\le|W_t|_Y$. Similarly, we have $|W_r|_Y\le|W_0|_Y$ for $j\le r$. If the word W_0 is tame, then it is the shortest configuration by the projection argument.

(b) If the rules do not change the lengths of configurations, then every control letter runs right and left only one time by Lemma 3.3 (4), and the inequality follows. If $||W_r||<||W_{r+1}||$ for some r, then every next transition keeps increasing the length by Lemma 3.3 (1), and so the inequality holds as well.

□

Lemma 3.13. *Let $\mathcal{C}\colon W_0\to\cdots\to W_t$ be a reduced computation of $\mathbf{M}_3$. Then for every i, there is at most one occurrence of the rules $\chi(i,i+1)^{\pm1}$ in the history H of $\mathcal{C}$ provided the base of $\mathcal{C}$ has a history $(R_{j-1}P_j)^{\pm1}$-sector.*

Proof. Arguing by contradiction, we can assume that $H=\chi(i,i+1)^{\pm1}H'\chi(i,i+1)^{\mp1}$, where H' is a copy of the history of a computation of either **LR** or **RL** or $\overline{\mathbf{M}}_2$. The first two case contradict Lemma 3.3 (4). The later case is also impossible. Indeed, consider any history subword $(R_{j-1}P_j)^{\pm1}$ of the base of the computation. Then the Y-projection of the $(R_{j-1}P_j)^{\pm1}$ -sector of W_1 must be a word either in the $X_{j,\ell}$ or in $X_{j,r}$ (depending on the parity of i). Without loss of generality assume that it is $X_{j,\ell}$. Then the computation $W_1\to,\dots,\to W_{t-1}$ multiplies the Y-projection of the $(R_{j-1}P_j)^{\pm1}$ -sector of W_1 by a word in $X_{j,\ell}$ and a reduced word in $X_{j,r}$. Hence the $(R_{j-1}P_j)^{\pm1}$ -sector of W_{t-1} contains letters from a right alphabet, hence W_{t-1} cannot be in the domain of $\chi(i,i+1)^{\pm1}$, a contradiction. □

Lemma 3.14. *Let $\mathcal{C}\colon W_0\to\cdots\to W_t$ be a reduced computation of $\mathbf{M}_3$. Suppose also that the base of $\mathcal{C}$ is standard, then*

(a) if the history of $\mathcal{C}$ has the form $\chi(i,i+1)H'\chi(i+4,i+5)$, then the word W_0 is a copy of W_t;

(b) two subcomputations $\mathcal{C}_1$ and $\mathcal{C}_2$ of $\mathcal{C}$ with histories $\chi(i,i+1)H'\chi(i+4,i+5)$ and $\chi(j,j+1)H''\chi(j+4,j+5)$ have equal lengths; moreover a cyclic permutation of $\mathcal{C}_2$ is a copy of $\mathcal{C}_1$;

(c) there is a constant $c_1 = c_1(\mathbf{M}_3)$ such that $|W_j|_Y \le c_1 \max(|W_0|_Y, |W_t|_Y)$ for $j = 0,1,\dots,t$; moreover, $|W_j|_Y \le c_1|W_t|_Y$ if W_0 is a tame configuration. (Recall that c_1 is one of the parameters from Section 2.3.).

Proof. (a) Without loss of generality we assume that $i = 1$. Consider the projection H_χ of the history H of $\mathcal{C}$ onto the alphabet of χ-rules of $\mathbf{M}_3$. By the definition of $\mathbf{M}_3$, if $\chi = \chi(j,j+1)^{\pm 1}$ is a letter in H_χ, then the next letter in H_χ is either χ^{-1} or $\chi(j-1,j)^{\pm 1}$ or $\chi(j+1,j+2)$. By Lemma 3.13, for the every letter χ, the word H_χ contains at most one occurrence of $\chi^{\pm 1}$. This implies that $H_\chi \equiv \chi(1,2)\chi(2,3)\chi(3,4)\chi(4,5)\chi(5,6)$.

Therefore the history of $\mathcal{C}$ has the form

$$\chi(1,2)H_1\chi(2,3)H_2\chi(3,4)H_3\chi(4,5)H_4\chi(5,6),$$

for some subhistories H_1, H_2, H_3, H_4 which do not contain χ-rules. By the definition of $\mathbf{M}_3$, each H_i is the history of a computation of a copy of one of the S-machines: $\overline{\mathbf{M}}_2, \mathbf{LR}, \mathbf{RL}$ (because rules of any two of these mahines have disjoint domains). This implies that H_1, H_2, H_3, H_4 are histories of computations of copies of $\overline{\mathbf{M}}_2, \mathbf{LR}, \overline{\mathbf{M}}_2, \mathbf{RL}$, respectively.

Let UV be a history 2-letter subword in the base B of the computation $\mathcal{C}$. The Y-projection u of the UV-sector of W_1 is a word in a left alphabet, while the Y-projection of the UV-sector of $W_1 \cdot H_1$ is a word in the corresponding right alphabet. Each rule θ of H_1 multiples the Y-projection of the UV-sector by a letter from the left alphabet on the left and by a letter from the right alphabet on the right. The two letters correspond to the rule θ. Therefore u must be a copy of H_1. In particular, this implies that the Y-projections of all history sectors of W_1 and $W_1 \cdot H_1$ are copies of H_1.

Applying Lemma 3.3 (3) to the subcomputation $W_1\cdot H_1\chi(2,3) \to \dots, W_1\cdot H_1\chi(2,3)H_2$ and considering the history UV-sector again, we deduce that H_2 is a copy of

$$\bar{H}_1\zeta^{(12)}(\bar{H}_1')^{-1}$$

where $\bar{H}_1$ is the mirror image of H_1 and $\bar{H}_1'$ is a copy of H_1. Moreover H_2 is uniquely determined by $W_1 \cdot H_1$, hence by W_1.

Similar arguments work for the rest of the computation $\mathcal{C}$: H_3 is a copy of H_1^{-1} and H_4 is a copy of $H_1\zeta^{(12)}H_1'$. This implies (a).

(b) follows from the same argument as (a).

(c) If the history H of $\mathcal{C}$ does not have χ-rules, then $\mathcal{C}$ is a computation of a copy of one of the S-machines $\overline{\mathbf{M}}_2, \mathbf{LR}, \mathbf{RL}$ and we can apply Lemmas 3.12 (b) and 3.11.

Suppose that H contains a χ-rule. Then $H = H_1H_2H_3$ where H_1, H_3 do not contain χ-rules, but H_2 starts and ends with χ-rules (it is possible that $||H_2|| = 1$). Let $W_k = W_0 \cdot H_1$, $W_s = W_0 \cdot H_1H_2 = W_t \cdot H_3^{-1}$. Then W_k is tame being in the domain of a χ-rule. Hence by Lemmas 3.12 (b) and 3.11 for every i between 0 and k $|W_i|_Y$ does not exceed $c|W_0|_Y$ for some constant c. The same argument shows that for i between s and t $|W_i|_Y$ does not exceed $c|W_t|_Y$. The proof of part (a) describes the subcomputation $W_k \to \dots \to W_s$ in detail. This description and Lemma 3.11 imply that for i between k and s, $|W_i|_Y$ does not exceed a constant times the maximum of $|W_k|_Y$ and $|W_s|_Y$. This implies (c).

□

Lemma 3.15. *(1) If a word α^k is accepted by the Turing machine $\mathbf{M}_0$, then for some word H, there is a reduced computation $I_3(\alpha^k, H) \to \cdots \to A_3(H)$ of the S-machine $\mathbf{M}_3$.*

(2) If there is a computation $\mathcal{C}\colon I_3(\alpha^k, H) \to \cdots \to A_3(H')$ of $\mathbf{M}_3$, then the word α^k is accepted by $\mathbf{M}_0$ and $H' \equiv H$.

Proof. (1) is obvious from the definition of $\mathbf{M}_3$ (see the informal definition of $\mathbf{M}_3$ at the beginning of Section 3.4): H is a copy of the history of a computation of $\overline{\mathbf{M}}_2$ accepting $I_2(\alpha^k)$ (which exists by Lemma 3.8 (1)).

(2) The word $I_3(\alpha^k, H)$ is in the domain of a rule from $\mathbf{M}_{3,1}$ while $I_3(H')$ is in the domain of a rule from $\mathbf{M}_{3,4m+1}$. For different i, j domains of rules from $\mathbf{M}_{3,i}$ and $\mathbf{M}_{3,j}$ are disjoint and if rules of sets $\mathbf{M}_{3,i}$ and $\mathbf{M}_{3,i+1}$ appear in a computation, the computation must also contain the χ-rule $\chi(i, i+1)$. Therefore the projection of the history of $\mathcal{C}$ onto the alphabet of χ-rules must contain a subword $\chi(1,2)\chi(2,3)$ Hence $\mathcal{C}$ must contain a subcomputation $\mathcal{D}$ with history of the form $\chi(1,2)H_1\chi(2,3)$, where H_1 is the history of a computation of a copy of $\overline{\mathbf{M}}_2$ of the form $I_2(\alpha^\ell, H) \to \cdots \to A_2(H'')$ for some ℓ, H'' and the rules in $\mathcal{C}$ applied before this $\chi(1,2)$ are from $\mathbf{M}_{3,1}$. Since rules of $\mathbf{M}_{3,1}$ do not modify the input sector, $k = \ell$. Therefore α^k is accepted by $\overline{\mathbf{M}}_2$. By Lemma 3.8 then α^k is accepted by $\mathbf{M}_0$ and $H'' \equiv H$. The fact that $H' \equiv H$ is proved in the same way as in Lemma 3.8 (2). □

3.5 $\mathbf{M}_4$ and $\mathbf{M}_5$

Let B_3 be the standard base of $\mathbf{M}_3$ and B_3' be its disjoint copy. By $\mathbf{M}_4$ we denote the S-machine with standard base $B_3(B_3')^{-1}$ and rules $\theta(\mathbf{M}_4) = [\theta, \theta]$, where $\theta \in \Theta$ and Θ is the set of rules of $\mathbf{M}_3$. So the rules of $\Theta(\mathbf{M}_4)$ are the same for $\mathbf{M}_3$-part of $\mathbf{M}_4$ and for the mirror copy of $\mathbf{M}_3$. Therefore we will denote $\Theta(\mathbf{M}_4)$ by Θ as well. The sector between the last state letter of B_3 and the first state letter of $(B_3')^{-1}$ is locked by any rule from Θ.

The 'mirror' symmetry of the base will be used in Lemma 7.37.

The S-machine $\mathbf{M}_5$ is a circular analog of $\mathbf{M}_4$. We add one more base letter $\check{t}$ to the hardware of $\mathbf{M}_4$. So the standard base B of $\mathbf{M}_5$ it $\{\check{t}\}B_3(B_3')^{-1}\{\check{t}\}$, where the part $\{\check{t}\}$ has only one letter $\check{t}$ and the first part $\{t\}$ is identified with the last part. For example, $\{\check{t}\}B_3(B_3')^{-1}\{t\}B_3(B_3')^{-1}$ can be a base of an admissible word for $\mathbf{M}_5$. Furthermore, sectors involving $\check{t}^{\pm 1}$ are locked by every rule from Θ. The accordingly modified sets $\mathbf{M}_{3,i}$ are denoted by $\mathbf{M}_{5,i}$.

In particular, for $\mathbf{M}_5$, we have the start and stop words $I_5(\alpha^k, H)$ and $A_5(H)$ similar to the configurations $I_3(\alpha^k, H)$ and $A_3(H)$, and the following analog of Lemma 3.15 can be proved in the same way as Lemma 3.15.

Lemma 3.16. *(1) If a word α^k is accepted by the Turing machine $\mathbf{M}_0$, then for some word H, there is a reduced computation of $I_5(\alpha^k, H) \to \cdots \to A_5(H)$ of the S-machine $\mathbf{M}_5$.*

(2) If there is a computation $\mathcal{C}\colon I_5(\alpha^k, H) \to \cdots \to A_5(H')$ of $\mathbf{M}_5$, then the word α^k is accepted by $\mathbf{M}_0$ and $H' \equiv H$.

Definition 3.17. We call the base of an admissible word of an S-machine *faulty* if

(1) it starts and ends with the same base letter,

(2) only the first and the last letters can occur in the base twice

(3) it is not a reduced word.

Lemma 3.18. *There is a constant $C = C(\mathbf{M}_5)$, such that for every reduced computation $\mathcal{C}\colon W_0 \to \cdots \to W_t$ of $\mathbf{M}_5$ with a faulty base and every $j = 0, 1, \dots, t$, we have $|W_j|_Y \le C\max(|W_0|_Y, |W_t|_Y)$.*

Proof. **Step 1.** One may assume that $|W_r|_Y > \max(|W_0|_Y, |W_t|_Y)$ for every $0 < r < t$ since otherwise it suffices to prove the statement for two shorter computations $W_0 \to \cdots \to W_r$ and $W_r \to \cdots \to W_t$. Since χ-rules do not change the length of configurations, the history H of $\mathcal{C}$ cannot start or end with a χ-rule.

Step 2. If the history H of $\mathcal{C}$ has no χ-rules, then the statement with $C \ge 18$ follows from Lemmas 3.12 (a), 3.4 and 3.10.

Step 3. If there is only one χ-rule χ in H, then $H = H'\chi^{\pm 1}H''$, where H' is a copy of the history of a computation of a copy of $\mathbf{LR}$ or $\mathbf{RL}$ and H'' is the history of a computation of a copy of $\overline{\mathbf{M}}_2$ (or vice versa). For the computation $W_r \to \cdots \to W_0$ with history $(H')^{-1}$, we have $|W_r|_Y \le |W_0|_Y$ by Lemmas 3.12 (a) and 3.4. This contradicts the assumption of Step 1, and so one may assume further that H has at least two χ-rules.

Step 4. The base B of the computation $\mathcal{C}$ has no history sectors PP^{-1}-, $R^{-1}R$-, QQ^{-1}-, or $Q^{-1}Q$-sectors, since every χ-rule locks the PQ- and QR-sectors of the standard base.

The same statement is true for the mirror copies of the above-mentioned sectors, and this stipulation works throughout the remaining part of the proof.

Step 5. Assume that the history $H^{\pm 1}$ is of the form $H_1\chi(i-1,i)H_2\chi(i,i+1)H_3$ for some i, where H_2 is the history of a computation of a copy of $\overline{\mathbf{M}}_2$. Since B is not reduced, there is a 2-letter subword of the base of the form $U^{\pm 1}U^{\mp 1}$ (for some part U of the set of state letters). By Lemma 2.4, then this subword must be a history subword of the form $P^{-1}P$ or RR^{-1} since every sector of the standard base of $\mathbf{M}_3$, except for history RP-sectors is locked either by $\chi(i-1,i)$ or by $\chi(i,i+1)$.

Let us consider the case of $P^{-1}P$ since the second case is similar. Depending on the parity of i either a prefix H_3' of H_3 is the history of a computation of a copy of $\mathbf{LR}$ or the suffix H_1' of H_1 is the history of a computation of a copy of $\mathbf{LR}$. These two cases are similar so we consider only the first one.

Then between the P-letter of the $P^{-1}P$-sector of an admissible word in the subcomputation of $\mathcal{C}$ with the history H_3' and the corresponding R-letter in that admissible word, there is always a Q-letter or a P^{-1}-letter, hence the P-letter never meets the corresponding R-letter during that subcomputation and no transition rules rules can apply to any of the admissible words of that subcomputation. Therefore $H_3' = H_3$ and for the subcomputation $\mathcal{C}'\colon W_s \to \cdots \to W_t$ of $\mathbf{LR}$ with history H_3 we have $|W_s|_Y \le |W_t|_Y$ by Lemmas 3.3 (1) and 3.4. This contradicts Step 1, and so the assumption made in the beginning of Step 5 was false.

Step 6. Assume that there is a history of a subcomputation of $\mathcal{C}$ of the form $H_1\chi H_2\chi^{-1}H_3$, where χ is a χ-rule, H_2 is the history of a computation of a copy of $\overline{\mathbf{M}}_2$. Then we claim that the base of $\mathcal{C}$ has no history $P^{-1}P$- or RR^{-1}-sectors. To prove this, we consider only the former case since the latter one is similar.

If the subcomputation $\mathcal{C}'$ of $\mathcal{C}$ with history H_3 starts with an admissible word W having in the $P^{-1}P$-sector all Y-letters from the right alphabets, then, as in Step 5, H_3 corresponds to the work of $\mathbf{LR}$, which gives a contradiction as in item 5.

If the $P^{-1}P$-sector of W has all Y-letters from the left alphabet, then the subcomputation of $\mathcal{C}^{-1}$ with history χH_2^{-1} will conjugate the Y-projection of that sector by a non-empty reduced word from the right alphabet. Therefore in the last admissible word of that subcomputation, there will still be letters from both left and right alphabets, and so it cannot be in the domain of any χ-rule or its inverse, a contradiction.

Together with Step 4, this implies that the base of $\mathcal{C}$ has no mutually inverse letters from history sectors staying next to each other.

Since the base is faulty, it must contain an input $P_1^{-1}P_1$ or $R_0R_1^{-1}$-sector. This implies that the base does not contain input $(R_0P_1)^{\pm 1}$-sectors since the first and the last letters of the base are equal (say, positive) and the base has no proper subwords with this property. In both cases the configuration W_r corresponding to the transition $\chi\colon W_{r-1} \to W_r$ is the shortest one in $\mathcal{C}$ since the Y-projection of that word is of the form α^k, each rule from $\mathcal{C}$ conjugates the Y-projection from the input sector, and α^k cannot be shortened by any conjugation. This contradicts Step 1.

Step 7. It follows from items 2,3, 5 and 6 that $H = H_1\chi H_2\chi' H_3$, for two χ-rules (or their inverses). Moreover H_2 is the history of a computation $\mathcal{C}_2$ of a copy of **LR** or of **RL** and H_1, H_3 are histories of computations $\mathcal{C}_1, \mathcal{C}_3$ of copies of $\overline{\mathbf{M}}_2$, i.e.,H has exactly two χ-rules (otherwise H has a subword which is ruled out in the previous steps of the proof).

Step 8. We claim that we can assume that the admissible words in the computation $\mathcal{C}$ do not have a history $(PR)^{\pm 1}$-sectors. Indeed, if such a sector exists, then for the subcomputation $\mathcal{C}_1\colon W_0 \to \dots \to W_r$ with history $H_1\chi$, we have $|W_r|_Y \le c|W_0|$ by Lemma 3.11. A similar estimate is true for the subcomputation with history $\chi' H_3$ starting with some W_s. So in order to prove the inequality from the lemma, it suffices to apply Step 2 to the three subcomputations $\mathcal{C}_1, \mathcal{C}_2.\mathcal{C}_3$.

Step 9. Suppose that the base of $\mathcal{C}$ contains a history subword of the form $P^{-1}P$.

If the admissible word from $\mathcal{C}$ in the domain of χ has no letters from the left alphabets, then H_2 is the history of a computation of a copy of **LR** and the state P-letter will never meet the corresponding state R- or Q-letter during the computation $\mathcal{C}_2$, so an application χ' is not possible after $\mathcal{C}_2$ ends, a contradiction.

Thus we can assume that if the base of $\mathcal{C}$ contains a history subword of the form $P^{-1}P$, then the last admissible word of $\mathcal{C}_2$ (which is in the domain of χ) contains letters from the left alphabet.

Similarly, if the base of $\mathcal{C}$ contains a history subword of the form RR^{-1}, then the last admissible word in $\mathcal{C}_2$ contains letters from the right alphabet. This implies, in particular that the base of $\mathcal{C}$ cannot contain both a history subword of the form $P^{-1}P$, and a history subword $R'(R')^{-1}$. Without loss of generality, we will assume that there are no subwords $R'(R')^{-1}$.

Step 10. It follows from Steps 4,8 and 9, that there are no unlocked by χ history sectors of the base except for $P^{-1}P$-sectors, and if there is such a sector UV,then $\mathcal{C}_2$ is a computation of a copy of **RL**. Therefore UV may cotain tape letters from a left alphabet, while every rule θ of $\mathcal{C}_1^{-1}$ multiplies this sector from both sides by letters from a right alphabet. So θ increases the lengths of every history sectors by 2. The rule χ locks working sectors (except for the input one), and so by Lemma 2.3 (**), θ can decrease the lengths of every working sector at most by one. Since working sectors alternate with history ones in any base, we have $||W_r|| \le ||W_0||$, contrary to Step 1.

Step 11. To complete the proof of the lemma, it remains to assume that there are no history sectors in the base of $\mathcal{C}$. Then the faulty base of $\mathcal{C}$ must contain input subwords of the form $R_0R_0^{-1}$ only, because every χ-rule locks all sectors of the standard base except for the input and history sectors. Then any admissible word of $\mathcal{C}$ from the domain of a χ-rule in H is the shortest admissible word in $\mathcal{C}$ since (as in Step 6) every rule of the computation conjugates $R_0R_0^{-1}$-sectors and a word α^k cannot be shortened by any conjugation. The lemma is proved since we can refer to Step 1 again. □

4 The main S-machine M

4.1 The definition of M

We use the S-machine $\mathbf{M}_5$ from Section 3.5, $\mathbf{LR}_m$ from Section 3.1 and three more easy S-machines to compose the main circular S-machine $\mathbf{M}$ needed for this paper. The standard base of $\mathbf{M}$ is the same as the standard base of $\mathbf{M}_5$, i.e.,$\{t\}B_3(B_3')^{-1}$, where B_3 has the form (3.4). However we will use $\check{Q}_0$ instead of Q_0, $\check{R}_1$ instead of R_1 and so on to denote parts of the set of state letters since $\mathbf{M}$ has more state letters in every part of its hardware.

The rules of $\mathbf{M}$ will be partitioned into five sets (S-machines) $\mathbf{\Theta}_i$ $(i = 1, \dots, 5)$ with transition rules $\theta(i, i+1)$ connecting i-th and $i+1$-st sets. The state letters are also disjoint for different sets $\mathbf{\Theta}_i$. It will be clear that $\check{Q}_0$ is the disjoint union of 5 disjoint sets including Q_0, $\check{R}_1$ is the disjoint union of five disjoint sets including R_1, etc.

By default, every transition rule $\theta(i, i+1)$ of $\mathbf{M}$ locks a sector if this sector is locked by all rules from $\mathbf{\Theta}_i$ or if it is locked by all rules from $\mathbf{\Theta}_{i+1}$. It also changes the end state letters of $\mathbf{\Theta}_i$ to the start state letters of $\mathbf{\Theta}_{i+1}$.

Set $\mathbf{\Theta}_1$ inserts input words in the input sectors. The set contains only one positive rule inserting the letter α in the input sector next to the left of a letter p from $\check{P}_1$. It also inserts a copy of α^{-1} next to the right of the corresponding letter $(p')^{-1}$ (the similar mirror symmetry is assumed in the definition of all other rules.) So the positive rule of $\mathbf{\Theta}_1$ has the form

$$[q_0 \xrightarrow{\ell} q_0, r_1 \to r_1, p_1 \xrightarrow{\ell} \alpha p_1, ..., (p_1')^{-1} \to (p_1')^{-1}\alpha^{-1}, (r_1')^{-1} \xrightarrow{\ell} (r_1')^{-1}, t \xrightarrow{\ell} t]$$

The rules of $\mathbf{\Theta}_1$ do not change state letters, so it has one state letter in each part of its hardware.

The connecting rule $\theta(12)$ changes the state letters of $\mathbf{\Theta}_1$ by their copies in a disjoint alphabet. It locks all sectors except for the input sector $\check{R}_0\check{P}_1$ and the mirror copy of this sector.

Set $\mathbf{\Theta}_2$ is a copy of the S-machine $\mathbf{LR}_m$ working in the input sector and its mirror image in parallel, i.e.,we identify the standard base of $\mathbf{LR}_m$ with $\check{R}_0\check{P}_1\check{Q}_1$. The connecting rule $\theta(23)$ locks all sectors except for the input sector $\check{R}_0\check{P}_1$ and its mirror image.

Set $\mathbf{\Theta}_3$ inserts history in the history sectors. This set of rules is a copy of each of the left alphabets $X_{i,l}$ of the S-machine $\mathbf{M}_2$. Every positive rule of $\mathbf{\Theta}_3$ inserts a copy of the corresponding positive letter in every history sector $\check{R}_i\check{P}_{i+1}$ next to the right of a state letter from $\check{R}_i$.

Again, $\mathbf{\Theta}_3$ does not change the state letters, so each part of its hardware contains one letter.

The transition rule $\theta(34)$ changes the state letters by their copies from Set $\mathbf{M}_{5,1}$ of $\mathbf{M}_5$. It locks all sectors except for the input sectors and the history sectors. The history sectors in admissible words from the domain of $\theta(34)$ have Y-letters from the left alphabets $X_{i,l}$ of the S-machine $\mathbf{M}_5$.

Set $\mathbf{\Theta}_4$ is a copy of the S-machine $\mathbf{M}_5$. The transition rule $\theta(45)$ locks all sectors except for history ones. The admissible words in the domain of $\theta(45)$ have no letters from right alphabets.

Set $\mathbf{\Theta}_5$**.** The positive rules from $\mathbf{\Theta}_5$ simultaneously erase the letters of the history sectors from the right of the state letter from $\check{R}_i$. That is, parts of the rules are of the form $r \to ra^{-1}$ where r is a state letter from $\check{R}_i$, a is a letter from the left alphabet of the history sector.

Finally the accept rule θ_0 from $\mathbf{M}$ can be applied when all the sectors are empty, so it locks all the sectors and changes the end state letters of $\mathbf{M}_5$ to the corresponding end state letters of $\mathbf{M}$. Thus the main S-machine $\mathbf{M}$ has unique accept configuration which we will denote by W_{ac}.

For every $i = 1, 2, 3, 4$, we will sometimes denote $\theta(i, i+1)^{-1}$ by $\theta(i+1, i)$.

4.2 Standard computations of M

We say that the history H of a computation of $\mathbf{M}$ (and the computation itself) is *eligible* if it has no neighbor mutually inverse letters except possibly for the subwords $\theta(23)\theta(23)^{-1}$. (The subword $\theta(23)^{-1}\theta(23)$ is not allowed.)

Remark 4.1. Clearly the history H^{-1} is eligible if and only if H is. Every reduced computation is eligible.

Considering eligible computations instead of just reduced computations is necessary for our interpretation of $\mathbf{M}$ in a group.

The history H of an eligible computation of $\mathbf{M}$ can be factorized so that every factor is either a transition rule $\theta(i, i+1)^{\pm 1}$ or a maximal non-empty product of rules of one of the sets $\mathbf{\Theta}_1 - \mathbf{\Theta}_5$. If, for example, $H = H'H''H'''$, where H' is a product of rules from $\mathbf{\Theta}_2$, H'' has only one rule $\theta(23)$ and H''' is a product of rules from $\mathbf{\Theta}_3$, then we say that the *step history* of the computation is (2)(23)(3). Thus the step history of a computation is a word in the alphabet $\{(1), (2), (3), (4), (5), (12), (23), (34), (45), (21), (32), (43), (54)\}$, where (21) is used for the rule $\theta(12)^{-1}$ an so on. For brevity, we can omit some transition symbols, e.g. we may use (2)(3) instead of (2)(23)(3) since the only rule connecting Steps 2 and 3 is $\theta(23)$.

If the step history of a computation consists of only one letter (i), $i = 1, \dots, 5$, then we call it a *one step computation*. The computations with step histories $(i)(i, i \pm 1)$, $(i \pm 1, i)(i)$ and $(i \pm 1, i)(i)(i, i \pm 1)$ are also considered as one step computations. Any eligible one step computation is always reduced by definition.

The step history of any computation cannot contain certain subwords. For example, (1)(3) is not a subword of any step history because domains of rules from $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_3$ are disjoint. In this subsection, we eliminate some less obvious subwords in step histories of eligible computations.

Lemma 4.2. *If the base of a computation $\mathcal{C}$ has at least one history subword UV, then there are no reduced computations $\mathcal{C}$ of $\mathbf{M}$ with step history*

(1) (34)(4)(43) *or* (54)(4)(45)*, provided $UV \equiv (\check{R}_{i-1}\check{P}_i)^{\pm 1}$ for some i,*

(2) (23)(3)(32).

Proof. (1) We consider only the step history (34)(4)(43) since the second case is similar. Let W_0 be the first admissible word of $\mathcal{C}$. Suppose that the history $H = \theta(34)H'\theta(43)$ of $\mathcal{C}$ has χ-letters. By Lemma 3.13 each χ letter $\chi^{\pm 1}$ appears in H_χ only once. Each χ-rule changes the state letters, and words in the domains of different (positive) χ-rules have different state letters. Therefore $W \cdot \theta(34)H'$ has different state letters than W_0, hence $W_0 \cdot \theta(34)H'$ is not in the domain of $\theta(43)$, a contradiction.

If H' has no χ-letters, then it is a history of **RL**, and we obtain a contradiction with Lemma 3.3 (4) (and Remark 3.7).

(2) Suppose the step history of $\mathcal{C}$ is (23)(3)(32). Since the history sectors are locked by $\theta(23)^{\pm 1}$, the history subwords in the base of $\mathcal{C}$ must have the form $(R_{i-1}P_i)^{\pm 1}$ for some i. Every rule of $\mathbf{\Theta}_3$ inserts a letter next to the left of every P_i-letter, different rules insert different letters, same letter for the same rule. Since at the beginning and at the end of the subcomputation with step history (3) all history sectors are empty of Y-letters, the word inserted during the subcomputation must be freely trivial. That contradicts the assumption that this subcomputation is reduced. ☐

By definition, the rule $\theta(23)$ locks all history sectors of the standard base of $\mathbf{M}$ except for the input sector $\check{R}_0\check{P}_1$ and its mirror copy. Hence every admissible word in the domain of $\theta(23)^{-1}$ has the form $W(k,k') \equiv w_1\alpha^k w_2(a')^{-k'}w_3$, where k and k' are integers and w_1, w_2, w_3 are fixed words in state letters; w_1 starts with $\check{t}$. Recall that W_{ac} is the accept word of $\mathbf{M}$.

Lemma 4.3. *There are no reduced computations of* $\mathbf{M}$ *with the standard base whose step history is* (12)(2)(21) *or* (32)(2)(23).

Proof. Consider only the step history (12)(2)(12). Thus the history H of the computation is $\theta(12)H'\theta(21)^{-1}$ and H' is a computation of a copy of $\mathbf{LR}_m$ working in the input sectors of admissible words of $\mathbf{M}$. The applying Lemma 3.3 (4) and Remark 3.3 we can conclude that H' is empty, a contradiction. ☐

Lemma 4.4. *Let a reduced computation* $\mathcal{C}\colon W_0 \to \cdots \to W_t$ *have the history* H *of the form (a)* $\chi(i-1,i)H'\chi(i,i+1)$ *(i.e.,the S-machine works as* $\mathbf{M}_3$ *with step history (4)) or (b)* $\zeta^{(i-1,i)}H'\zeta^{(i,i+1)}$ *(i.e.,it works as* $\mathbf{LR}_m$ *with step history (2)).*

Then the base of the computation $\mathcal{C}$ *is a reduced word, and all configurations of* $\mathcal{C}$ *are uniquely defined by the history* H *and the base of* $\mathcal{C}$*,* $|W_0|_Y = |W_1|_Y = \cdots = |W_t|_Y$*, and* $||H|| < 2||W_0||$.

Moreover, $||H|| = 2s+3$ *(resp.,* $||H|| = s+2$*), where s is the Y-length of every history sector (input sector) in case (a) (in case (b), resp.), and* H' *is the copy of the maximal a-word contained in arbitrary history (resp., input) sector of* W_0.

Proof. (a) Every history sector of the standard base is locked either by one of the rules $\chi(i-1,i), \chi(i,i+1)$, or by a rule of H'. Every non-history sector of the standard base is also locked either by $\chi(i-1,i)$ or by $\chi(i,i+1)$. It follows from Lemma 2.4 that the base of $\mathcal{C}$ is a reduced word. By Lemma 3.3 (3), the histories of the primitive S-machines subsequently restore the tape words in all history sectors. Since one of the rules $\chi(i-1,i), \chi(i,i+1)$ locks all non-history sectors, Lemma 3.12 applied to $\mathcal{C}$ implies equalities $|W_0|_Y = |W_1|_Y = \cdots = |W_t|_Y$, and gives the other statements.

(b) The same proof up to change of the history sectors by the input ones. ☐

Lemma 4.5. *The step history of every eligible computation of* $\mathbf{M}$ *with standard base either*

(A) contains one of the words (34)(4)(45), (54)(4)(43), (12)(2)(23), (32)(2)(21) *as a subword or*

(B) is a subword of one of the words

$$(4)(45)(5)(54)(4), (4)(43)(3)(34)(4), (2)(23)(3)(34)(4),$$

$$(4)(43)(3)(32)(2), (2)(21)(1)(12)(2), (2)(23)(32)(2).$$

Proof. The statement is obvious if there are neither (2) nor (4) in the step history. Lemmas 4.2 (1) (Lemma 4.3) implies that if (4) (resp. (2)) is not the first or the last letter in the step history then it can occur in a subword of the form (34)(4)(45) or (45)(4)(34) (resp., (12)(2)(23) or (23)(2)(12)), i.e.,we have Property (A).

If the first letter in the step history is (2) and Property (A) fails, then the same lemmas give us the longest possible step histories (2)(23)(32)(2), (2)(1)(2) and (2)(23)(3)(34)(4). The assumption that the last letter in the step history is (2) adds one more possible longest step history word (4)(43)(3)(32)(2).

Similarly, we may assume that (4) is either the first or the last letter in the step history and conclude that the step history is a subword of one of the words (4)(5)(4), (4)(3)(4), (2)(3)(4) and (4)(3)(2) provided Property (A) fails. □

Lemma 4.6. *(1) If the word α^k is accepted by the S-machine $\mathbf{M}_0$, then there is a reduced computation of $\mathbf{M}$, $W(k,k) \to \cdots \to W_{ac}$ whose history has no rules of $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$.*

(2) If the history of a computation $\mathcal{C}\colon W(k,k) \to \cdots \to W_{ac}$ of $\mathbf{M}$ has no rules of $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$, then the word α^k is accepted by $\mathbf{M}_0$.

Proof. (1) By Lemma 3.16, there is a computation $I_5(a_k,H) \to \cdots \to A_5(H)$ of the S-machine $\mathbf{M}_5$ for some H. So we have the corresponding computation of $\mathbf{\Theta}_4$:

$$\mathcal{D}\colon I_6(a_k,H) \to \cdots \to A_6(H).$$

Now the computation of $\mathbf{\Theta}_3$ inserting letters in history sectors and a computation of $\mathbf{\Theta}_5$ erasing these letters extend $\mathcal{D}$ and provide us with a computation $W(k,k) \to \cdots \to I_6(a_k,H) \to \cdots \to A_6(H) \to \cdots \to W_{ac}$.

(2) By Lemma 4.2 (1), the step history of $\mathcal{C}$ begins with (3)(4)(5), and so there is a subcomputation of Set 4 of the form $I_5(\alpha^\ell,H) \to \cdots \to A_5(H)$ for some ℓ and H, where according to Lemma 3.16 (2), the word α^ℓ is accepted by $\mathbf{M}_0$. Since the computation of Set 3 does not change the input sector, we have $\ell = k$. □

4.3 The first estimates of computations of M

Lemma 4.7. *Let $\mathcal{C}\colon W_0 \to \cdots \to W_t$ be a computation of $\mathbf{M}$ satisfying Property (B) of Lemma 4.5 or any computation of $\mathbf{M}$ with step history of length at most 2. Then for some constant c_2 (see Section 2.3)*

(a) $|W_j|_Y \le c_2 \max(|W_0|_Y, |W_t|_Y)$ for $j = 0,1,\dots,t$; moreover

$$|W_0|_Y \le c_2 \max(|W_t|_Y$$

if W_0 is a tame configuration.

(b) $t \le c_2^2(||W_0|| + ||W_t||)$; moreover, $t \le c_2^2||W_t||$ if W_0 is a tame configuration.

Proof. (a) If $\mathcal{C}$ is a one-step computation and its step history is (1), (3), or (5), then Statement (a) follows from Lemma 2.5 (c). For step history (2) (resp. (4) it follows from Lemma 3.12 (a) (resp., Lemma 3.14 (c)).

If there is a transition rule of $\mathbf{M}$ in the history H of $\mathcal{C}$, then H can be decomposed in at most three factors $H = H_1H_2H_3$, where H_2 is a one-step computation of step history (1), (3) or (5), or $H_2 = (23)(32)$ and H_1, H_3, if non-empty, are of step history (2) or (4). Respectively, the computation $\mathcal{C}$ is a composition of at most three subcomputations $\mathcal{C}_1\colon W_0 \to \cdots \to W_r$, $\mathcal{C}_2\colon W_r \to \cdots \to W_s$ and $\mathcal{C}_3\colon W_s \to \cdots \to W_t$. Now we can bound $|W_r|_Y$ and $|W_s|_Y$ by $c\max(|W_0|_Y, |W_t|_Y)$ applying either Lemma 3.12 (a) (for step history (2)) or Lemma 3.14 (c) (for step history (4)) to $\mathcal{C}_1$ and $\mathcal{C}_3$. The same lemmas applied to subcomputations $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$ completes the proof since we can assume that $c_2 \gg c$ (see Section 2.3).

If W_0 is a tame configuration, then the same Lemmas linearly bound $||W_0||$ in terms of $||W_t||$, and the required estimates follow.

(b) It suffices to bound the lengths of at most three one step subcomputations $\mathcal{C}'\colon W_j \to \cdots \to W_k$, where $\max(|W_j|_Y, |W_k|_Y) \le c_2\max(|W_0|_Y, |W_t|_Y)$ by (1). For step history (1), (3) or (5), the history lengths are bounded by Lemma 2.5 (b). For (2), we refer to Lemma 3.12 (b). The computation with step history (4) has at most $4m$ χ-rules in the history as follows from Lemma 3.13. So it has at most $4m+1$ maximal subcomputations of the form $W_l \to \cdots \to W_s$, corresponding to one of the $4m+1$ subsets $\mathbf{M}_{3,i}$ of the set of rules of $\mathbf{M}_3$, where $\max(|W_l|_Y, |W_s|) \le c_2\max(|W_0|_Y, |W_t|)$ by part (1) of the lemma. Hence we have the same upper bound for $s-l$ by Lemmas 3.3 (3) (if it is a computation of $\mathbf{LR}$) and 3.9 (if it is acomputation of $\mathbf{M}_2$). This completes the proof of the first inequality since we have $c_2 \gg m$ (Section 2.3). The tame case is treated in the same way (the proof is even shorter). □

4.4 Computations of M with faulty bases

Lemma 4.8. *For every eligible computation $\mathcal{C}\colon W_0 \to \cdots \to W_t$ of $\mathbf{M}$ with a faulty base and every $j = 0, 1, \dots, t$, we have $|W_j|_Y \le c_1\max(|W_0|_Y, |W_t|_Y)$.*

Step 1. As in Step 1 of the proof of Lemma 3.18, one may assume that $|W_j|_Y > \max(|W_0|_Y, |W_t|_Y)$ if $1 < j < t$ and so the history H of $\mathcal{C}$ neither starts nor ends with a transition rule $\theta(i, i+1)^{\pm 1}$.

Step 2. If $\mathcal{C}$ is a one step computation and (i) is its step history, then the statement follows from Lemma 2.5 (c) for $i = 1, 3, 5$, (since $c_1 \ge 2$), Lemma 3.12 (a) for $i = 2$ (since $c_1 \ge 2$) and Lemma 3.18 (since $c_1 \ge C$). Hence one may assume further that H contains a transition rule of $\mathbf{M}$ or its inverse.

Step 3. Assume that $\mathcal{C}$ (or the inverse computation) has a transition rule $\theta(23)$, $W_{j+1} = W_j \cdot \theta(23)$. Recall that the $\theta(23)$ does not lock only the input $\check{R}_0\check{P}_1$-sector and its mirror copy. So by Lemma 2.4, we should have an input subword $\check{R}_0\check{R}_0^{-1}$ or $\check{P}_1^{-1}\check{P}_1$ in the faulty base. Moreover, we must have exactly two such input subwords in the base and no subwords $(\check{R}_0\check{P}_1)^{\pm 1}$ since the first and the last letters of the base are equal (e.g., positive) and the base has no proper subwords with this property (see Definition 3.17).

The input sectors of both W_j and W_{j+1} have Y-projections of the form α^k, and they are not longer than the corresponding Y-words in the input sectors of any other W_i since α^k cannot be shortened by conjugation. It follows that $|W_j|_Y, |W_{j+1}|_Y \le \max(|W_0|_Y, |W_t|_Y)$ contrary to Step 1. Thus, one may assume further that H has no letters $\theta(23)^{\pm 1}$. In particular, $\mathcal{C}$ is a reduced computation.

The same argument eliminates letters $\theta(12)^{\pm 1}$ from H, and so the letter (1) from the step history of $\mathcal{C}$. Hence one can assume that the step history contains neither (1) nor (2).

Step 4. Suppose H (or H^{-1}) contains a subhistory $H'\theta(45)$, where H' is a maximal subword of H which is word in $\mathbf{\Theta}_4$ (which is a copy of the S-machine $\mathbf{M}_5$). By Lemma 2.4, the faulty base of the computation $\mathcal{C}$ contains one of the history subwords $\check{R}_{i-1}\check{R}_{i-1}^{-1}$ or $\check{P}_i^{-1}\check{P}_i$ for some i, because all non-history sectors are locked by $\theta(45)$.

Suppose the base of $\mathcal{C}$ contains a history subword $\check{R}_{j-1}\check{R}_{j-1}^{-1}$ for some j. The word H' must have a suffix which is a word in the alphabet of a copy of **RL** working in parallel in the history sectors (see the definition of $\mathbf{M}_{3,4m+1}$). The state letters from $\check{R}_{j-1}$ in the $\check{R}_{j-1}\check{R}_{j-1}^{-1}$-sector will then never meet a letter from either $\check{Q}_{j-1}$ or $\check{P}_j$. Therefore H' cannot contain the transition rule $\chi(4m,4m+1)^{\pm 1}$ or $\theta(45)^{-1}$. Thus H' is a prefix of H, is a computation of a copy of **RL**, and by Lemma 3.12 (a) applied to the subcomputation of $\mathcal{C}^{-1}$ with history $(H')^{-1}$, we get a contradiction with Step 1 because admissible words in the domain of $\theta(45)^{-1}$ is tame.

Suppose the base of $\mathcal{C}$ contains a subword $(\check{R}_{i-1}\check{P}_i)^{\pm 1}$. Then H has no subword $\theta(45)^{-1}H'\theta(45)$ by Lemma 4.2 (1) If H' has neither transition rules nor χ-rules, then we have a contradiction by Lemma 3.12 (a). Hence H has a subword $\chi(4m,4m+1)H''\theta(45)$, but then by Lemma 3.3 (3), H' has a rule locking all the sectors $\check{R}_{i-1}\check{P}_i$ of the standard base, and we get a contradiction with Lemma 2.4.

Finally suppose all history subwords in the base of $\mathcal{C}$ have the form $\check{P}_i^{-1}\check{P}_i$. Then the rules of a copy of **RL** from H' do not change the history sectors of admissible words in the corresponding subcomputation $\mathcal{C}'$ of $\mathcal{C}$, hence the lengths of all admissible words in $\mathcal{C}'$ stay the same. Moreover since the state letters in the history sectors do not change during the subcomputation $\mathcal{C}$, none of the admissible words in that subcomputation is in the domain of $\chi(4m,4m+1)^{\pm 1}$. Therefore the rules of H' do not change the lengths of admissible words, and either H' is a prefix of H and we get a contradiction with Step 1 or we have the subhistory $\theta(45)^{-1}H'\theta(45)$.

In the latter case, we consider the maximal subhistory H'' of type 5 following after the rule $\theta(45)$ (or before $\theta(45)^{-1}$). All the admissible words of the corresponding subcomputation $\mathcal{C}''$ have equal lengths since the base has no letters $\check{R}_i$. Arguing in this way we see that the history of $\mathcal{C}$ has Steps 4 and 5 only, and all the admissible words in $\mathcal{C}$ have equal length, which proves the inequality of the lemma.

We can conclude that H does not contain $\theta(45)^{\pm 1}$. By Step 2, (5) is not in the step history of $\mathcal{C}$ and the only possible transition rules of $\mathbf{M}$ in H are $\theta(34)^{\pm 1}$.

Step 5. Assume that there is a subhistory of H of the form $H_1\theta(34)H_2\theta(34)^{-1}H_3$, where H_2 is the history of $\mathbf{M}_5$. Then the base of $\mathcal{C}$ has no history sectors of the form $\check{R}_i\check{R}_i^{-1}$ (since, as before, the machine **RL** starting with $\theta(34)$ would never end with $\chi(12)$).

If there is a history subword $\check{R}_{i-1}\check{P}_i$ in the faulty base, then H_2 cannot follow by the transition rule $\theta(34)^{-1}$, by Lemma 3.13 if H_2 contains χ-rules and by Lemma 3.3 (4) otherwise, a contradiction.

Thus the base of $\mathcal{C}$ has no $\check{R}$-letters from history sectors. It also has no $\check{P}_1$-letters from input sectors, because otherwise the base would contain the letter $\check{R}_1$ of the history sector next to the input sector since the sectors $\check{P}_1\check{Q}_1$ and $\check{Q}_1\check{R}_1$ are locked by $\theta(34)$.

Thus, all history sectors have the form $\check{P}_i^{-1}\check{P}_i$ in the faulty base of $\mathcal{C}$, and so H cannot have the rule $\chi(1,2)^{\pm 1}$ (for the same reason the rule $\chi(4m,4m+1)$ was eliminated in Step

4). But without $\chi(1,2)^{\pm 1}$, one cannot get a rule in H changing history sectors $\check{P}_i^{-1}\check{P}_i$ since the rules of $\mathbf{\Theta}_3$ leave such sectors unchanged. The input sectors $\check{R}_0\check{R}_0^{-1}$ of the base of $\mathcal{C}$ (if any) cannot be shorten by a subcomputation since no conjugation shortens a power of one letter in a free group. therefore the rules $\theta(34)^{\pm 1}$ are applied to the shortest admissible word of $\mathcal{C}$, contrary to Step 1.

So our assumption was wrong.

Step 6. If there is only one transition rule $\theta(34)$ in $H^{\pm 1}$, then $H^{\pm 1} = H'\theta(34)H''$, where H'' is the history of $\mathbf{M}_5$. If H'' is the history of a copy of $\mathbf{RL}$, starting with an admissible word W_r, then $|W_r|_Y \le |W_t|_Y$ by Lemmas 3.12 (a) and 3.4, contrary to Step 1. Otherwise we have a subhistory $\theta(34)H_0\chi(1,2)$, and by Lemma 3.3 (3), there are no history subsectors of the form $\check{R}_i\check{R}_i^{-1}$ or $\check{P}_i^{-1}\check{P}_i$ in the base of $\mathcal{C}$. If there is a history sector $\check{R}_{i-1}\check{P}_i$, then one can linearly bound $|W_r|_Y$ in terms of $|W_t|_Y$ applying Lemmas 3.12 (b) and 3.11 several times, namely at most $4m+1$ times by Lemma 3.13. Since $c_1 \gg C, c_1 \gg m$ (see Section 2.3) one consider two subcomputations of $\mathcal{C}$ can divide $\mathcal{C}$: $W_0 \to \dots \to W_r$ and $W_r \to \dots \to W_t$ and reduce the proof to Step 2.

Thus, one may assume that the base of $\mathcal{C}$ has no letters $\check{P}$ and $\check{R}$ from history sectors. This also eliminates the letter $\check{P}_1$ of the input sector and gives the inequality

$$|W_r|_Y \le \max(|W_0|_Y, |W_t|_Y),$$

contrary to Step 1. Therefore the assumption of Step 6 was wrong.

Step 7. It remains to consider the case when $H^{\pm 1}$ is of the form $H_1\theta(34)^{-1}H_2\theta(34)H_3$, where H_2 is the history of $\mathbf{\Theta}_3$, H_1 and H_3 are histories of $\mathbf{\Theta}_4$, and it suffices to repeat the argument of Step 6 with decomposition of $\mathcal{C}$ in the product of three subcomputaions, because we did not use there that the subword $H_1\theta(34)$ was absent.

The lemma is proved.

4.5 Space and length of M-computations with standard base

Let us call a configuration W of $\mathbf{M}$ *accessible* if there is a W*-accessible computation*, i.e., either an accepting computation starting with W or a computation $s_1(\mathbf{M}) \to \dots \to W$, where $s_1(\mathbf{M})$ is the start configuration of $\mathbf{M}$ (i.e., the configuration where all state letters are start state letters of Θ_1 and the Y-projection is empty).

Lemma 4.9. *If W is an accessible configuration, then for a constant $c_3 = c_3(\mathbf{M})$, there is a W-accessible computation $\mathcal{C}$ of length at most $c_3||W||$ whose step history is either a suffix of (4)(5) or a prefix of (1)(2)(3)(4). The Y-length of every configuration of $\mathcal{C}$ does not exceed $c_2|W|_Y$. (Recall that c_2, c_3 are parameters in Section 2.3.)*

Proof. Assume that a W-accessible computation $\mathcal{C}$ has (4) in its step history and its history H has a rule $\chi(i,i+1)$ with $1 < i < 4m$. Since $\mathcal{C}$ is accessible, we have by Lemma 3.14 (b), a subcomputation $W_l \to \dots \to W_r$ with history of the form (a) $\chi(i,i+1)H'\chi(i+1,i+2)$ or (b) $\chi(i,i+1)^{-1}H'\chi(i-1,i)^{-1}$, where H' is a history of a canonical computation of $\mathbf{M}_5$. By Lemma 4.4 we also conclude that every history sector of W_l and of W_r is a copy of H'. It makes possible to accept W_r using erasing rules of Set 5 in case (a) or to construct a computation of type (1)(2)(3) starting with $s_1(\mathbf{M})$ and ending with W_l in case (b).

It follows now from Lemma 3.14 that one can choose a accessible computation $\mathcal{C}$ having no subhistories of type (34)(4)(45) or (45)(4)(34), and so Set 4 can occur only in the beginning or at the end of H. In the first case H has to have type (4)(5), and the required inequalities follow from Lemma 4.7 since $c_3 \gg c_2$.

In the second case, the step history ends with (3)(4), and the connection

$$\theta(34)\colon W_{k-1} \to W_k$$

provides us with copies in all history sectors and in all input sectors since W_k is accessible. Hence one may assume that the step history has the form (1)(2)(3)(4). Here $|W_k|_Y \le c_1|W|_Y$ by Lemma 3.14 (c). The canonical computation with step history (1)(2)(3) does not decrease the lengths of configurations. Now the required estimates follows from Lemma 4.7 for four one-step subcomputations since we chose c_3 after c_2.

One obtains even better estimates than the required iequalities if $\mathcal{C}$ has no Set 4 in the step history since it has one of the types (5) or (1)(2)(3), or (1)(2) or (1). □

For any accessible word W we choose an accessible computation $\mathcal{C}(W)$ according to Lemma 4.9.

Lemma 4.10. *Let W_0 be an accessible word, $\mathcal{C}\colon W_0 \to \dots \to W_t$ be an eligible computation of $\mathbf{M}$ and H_0, H_t be the histories of $\mathcal{C}(W_0)$ and $\mathcal{C}(W_t)$, respectively. Then for some constants c_4, c_5 (see Section 2.3) either*

(a) $t \le c_4 \max(||W_0||, ||W_t||)$ and $||W_j|| \le c_5 \max(||W_0||, ||W_t||)$, for every $j = 0, \dots, t$ or

(b) $||H_0|| + ||H_t|| \le t/500$ and the sum of lengths of all subcomputations of $\mathcal{C}$ with step histories (12)(2)(23), (23)(2)(12), (34)(4)(45) and (45)(4)(34) is at least $0.99t$.

Proof. One may assume that $t > c_4 \max(||W_0||, ||W_t||)$, because otherwise Property (a) holds for sufficiently large c_5 since an application of every rule can increase the length of a configuration by a constant depending on $\mathbf{M}$. Hence by Lemma 4.9,
$||H_0|| + ||H_t|| \le 2c_3 \max(||W_0||, ||W_t||) \le t/500$.

The computation $\mathcal{C}$ is not a B-computation by Lemma 4.7 since $c_2 < c_4$. Therefore it is a computation satisfying Property (A) of Lemma 4.5, and one has a maximal subcomputation $\mathcal{C}''\colon W_r \to \dots \to W_s$ starting and ending with subcomputations with step histories 2 or 4, which where listed in part (A) of that lemma. We have $\mathcal{C} = \mathcal{C}'\mathcal{C}''\mathcal{C}'''$, and applying Lemma 4.7 to $(\mathcal{C}')^{-1}$ and $\mathcal{C}'''$, we have $\max(||W_r||, ||W_s||) \le c_2 \max(||W_0||, ||W_t||)$ and the lengths l' and l''' of $\mathcal{C}'$ and $\mathcal{C}'''$ do not exceed $c_4 \max(||W_0||, ||W_t||)/1000$ since $c_4 > 1000c_2c_3$. Therefore $l' + l''' \le t/500$.

Lemma 4.5 implies that the subcomputation $\mathcal{C}''$ is a product $\mathcal{C}_1\mathcal{D}_1 \dots \mathcal{C}_{k-1}\mathcal{D}_{k-1}\mathcal{C}_k$, where $k \ge 1$, every $\mathcal{C}_i$ has one of the four step histories from item (a) of that lemma, and every $\mathcal{D}_i$ is a subcomputation having type 1 or 3, or 5, or just empty if the history $H(i)$ of $\mathcal{C}_i$ ends with $\theta(23)$ and $H(i+1)$ starts with $\theta(23)^{-1}$. Let $K(i)$ be the history of $\mathcal{D}_i$. Below we will prove that $||K(i)|| \le (||H(i)|| + ||H(i+1)||)/1000$. In turn, this will imply that $\sum_i ||H(i)|| \ge 500 \sum_i ||K(i)||$, and the last claim of the lemma will follow since $l' + l''' \le t/500$.

So, let $\mathcal{D}_i\colon W_x \to \dots \to W_y$. Then on the one hand, $||K_i|| \le |V_x|_Y + |V_y|_Y$ by Lemma 2.5 (b); here $V_x \to \dots \to V_y$ is the restriction $\mathcal{D}_i$ to a sector with base of lengths two, where the rules of $\mathcal{D}_i$ insert/delete letters. On the other hand, $||H(i)|| \ge 2m|V_x|_Y$, as it follows from Remark 3.7 (if $\mathcal{C}_i$ has type 2) and from Lemmas 3.3 (3), 3.13, 3.14 (a) and the definition of Set 4 (if $\mathcal{C}_i$ has type 4). Similarly we have $||H(i+1)|| \ge 2m|V_y|_Y$, whence $||H(i)|| + ||H(i+1)||)/1000 \ge m(|V_x|_Y + |V_y|)/500 \ge ||K(i)||$ by the choice of m. □

We call a base B of an eligible computation (and the computation itself) *revolving* if $B \equiv xvx$ for some letter x and a word v, and B has no proper subword of this form.

If $v \equiv v_1 z v_2$ for some letter z, then the word zv_2xv_1z is also revolving. One can cyclically permute the sectors of revolving computation with base xvx and obtain a uniquely defined computation with the base zv_2xv_1z, which is called a cyclic permutation of the original computation. The history and lengths of configurations do not change when one cyclically permutes a computation.

Lemma 4.11. *Suppose the base B of an eligible computation $\mathcal{C}\colon W_0 \to \cdots \to W_t$ is revolving. Then one of the following statements hold:*

(1) we have inequality $||W_j|| \le c_4 \max(||W_0||, ||W_t||)$, for every $j = 0, \dots, t$ or

(2) we have the following properties:

(a)the word xv or $v^{-1}x^{-1}$ is a cyclic permutation of the standard base of $\mathbf{M}$ and

(b) the corresponding cyclic permutations W_0' and W_t' of the words W_0 and W_t are accessible words, and

(c) the step history of $\mathcal{C}$ (or of the inverse computation) contains a subword $(12)(2)(23)$ *or* $(34)(4)(45)$*; moreover, the sum of lengths of corresponding subwords of the history is at least* $0.99t$ *and*

(d) we have $||H'|| + ||H''|| < t$ for the histories H' and H'' of $\mathcal{C}(W_0)$ and $\mathcal{C}(W_t)$.

Proof. If the computation is faulty, then Property (1) is given by Lemma 4.8 since $c_4 > c_1$. If it is non-faulty, then we have all sectors of the base in the same order as in the standard base (or its inverse), and we obtain Property (2a). Therefore we may assume now that the base xv is standard and Property (1) does not hold.

If $\mathcal{C}$ is a B-computation, we obtain a contradiction with Lemma 4.7 since $c_4 > c_2$. Therefore we assume further that $\mathcal{C}$ is an A-computation. So it (or the inverse one) contains a subcomputation with step history $(12)(2)(23)$ or $(34)(4)(45)$. In case of $(34)(4)(45)$, we consider the transition $\theta(45)\colon W_j \to W_{j+1}$. By Lemma 4.4, the words in the history sectors $\check{R}_{i-1}\check{P}_i$ are copies of each other. Therefore they can be simultaneously erases by the rules of Set 5, and so W_{j+1} and all other configurations are accepted. Similarly one applies Lemma 4.4 in case $(12)(2)(23)$ and concludes that Property (2b) holds.

Now the second part of (2c) and (d) follow from Lemma 4.10. □

4.6 Two more properties of standard computations

Here we prove two lemmas needed for the estimates in Subsection 7.2. The first one says (due to Lemma 4.3 (2)) that if a standard computation $\mathcal{C}$ is very long in comparison with the lengths of the first and the last configuration, then it can be completely restored if one knows the history of $\mathcal{C}$, and the same is true for the long subcomputations of $\mathcal{C}$. This makes the auxiliary parameter $\sigma_\lambda(\Delta)$ useful for some estimates of areas of diagrams Δ. The second lemma is also helpful for the proof of Lemma 7.38 in Subsection 7.2.

Lemma 4.12. *Let $\mathcal{C}\colon W_0 \to \cdots \to W_t$ be a reduced computation with standard base, where $t \ge c_4 \max(||W_0||, ||W_t||)$. Suppose the word W_0 is accessible. Then the history of any subcomputation $\mathcal{D}\colon W_r \to \cdots \to W_s$ of $\mathcal{C}$ (or the inverse for $\mathcal{D}$) of length at least* $0.4t$ *contains a word of the form (a) $\chi(i-1,i)H'\chi(i,i+1)$ (i.e.,the S-machine works as $\mathbf{M}_3$ at Set 4) or (b) $\zeta^{i-1,i}H'\zeta^{i,i+1}$ (i.e.,it works as $\mathbf{LR}_m$ at Set 2).*

Proof. By Lemma 4.10, that the sum of lengths of all subcomputations $\mathcal{C}'$ of $\mathcal{C}$ with step histories (12)(2)(23), (23)(2)(12), (34)(4)(45) and (45)(4)(34) is at least $0.99t$. Therefore $\mathcal{D}$ has to contain a subcomputation $\mathcal{D}'$ of type 2 or 4, which is a subcomputation of some $\mathcal{C}'$, and $||K'|| \geq 0.3||H'||$ for the histories K' and H' of $\mathcal{D}'$ and $\mathcal{C}'$, respectively.

It suffices to show that such a subcomputation $\mathcal{D}'$ of a computation $\mathcal{C}'$ with step history (34)(4)(45) (with (12)(2)(23)) contains a subcomputation of the form (a) (form (b), resp.) For $\mathcal{C}'$ of type (34)(4)(45), this follows from Lemma 3.14 (b) since $m > 10$. For $\mathcal{C}'$ of type (12)(2)(23), the same property holds since the S-machine $\mathbf{LR}_m$ has to repeat the cycles of $\mathbf{LR}$ m times by Lemma 3.3 (3,4). □

Lemma 4.13. *Let a reduced computation $\mathcal{C}\colon W_0 \to \dots \to W_t$ start with a accessible word W_0, have standard base and have step history of length 1. Assume that for some index j, we have $|W_j|_Y > 3|W_0|$. Then there is a sector QQ' such that a state letter from Q or from Q' inserts an a-letter increasing the length of this sector after any transition of the subcomputation $W_j \to \dots \to W_t$.*

Proof. First of all we observe that the a-words in all history sectors (in all input sectors) of any configuration W_i are copies of each other, because W_0 is accessible. Also inducting on t one can assume that $|W_1|_Y > |W_0|_Y$.

I we have one of the Sets 1, 3, 5, then inequality $|W_0|_Y < |W_1|_Y$ implies $|W_1|_Y < |W_2|_Y < \dots$ since the second rule cannot be inverse for the first one, and so on, i.e., we obtain the desired property of any input sector for Set 1 or of any history sector for Sets 3 or 5.

If we have Set 2, then the statement for any imput sector follows from Lemma 3.3 (1) .

Let the step history be (4). Recall that the rules of Set 4 are subdivided in several sets, where each set copies the work of either $\mathbf{LR}$ or $\mathbf{M}_3$. If a $\mathbf{PM}$-rule of the subcomputation $\mathcal{D}\colon W_0 \to \dots \to W_j$ increases the length of a history sector, then we refer to Lemma 3.3 (1) as above. So one may assume that no $\mathbf{PM}$-rules of $\mathcal{D}$ increase the length of history sectors.

Assume now that $\mathcal{D}$ has an $\mathbf{M}_3$-rule increasing the length of history sectors. It has to insert a letter from $X_{i,\ell}$ from the left and a letter from $X_{i,r}$ from the right. Since the obtained word is not a word over one of these alphabets, the work of $\mathbf{M}_3$ is not over, and the next rule has to increase the length of the sector again in the same manner since the computation is reduced. This procedure will repeat until one gets W_t. This proves the statement for any history sector.

It remains to assume that there are no transitions in $\mathcal{D}$ increasing the lengths of history sectors and the first transition $W_0 \to W_1$ is provided by a rule θ of $\mathbf{M}_3$. It cannot shorten history sectors (by 2). Indeed can θ change the length of working sectors of neighbor working sectors at most by 1 (see Lemma 2.3 (3)), which implies $|W_0|_Y \geq |W_1|_Y$, a contradiction. It follows that no further rules of $\mathbf{M}_3$ can shorten history sectors. Then Lemma 3.9 implies that all history sectors in all configurations of $\mathcal{D}$ have equal lengths.

By Lemma 2.6 (b) the lengths of the history of the maximal subcomputation $\mathcal{E}\colon W_0 \to \dots \to W_s$ of $\mathbf{M}_3$ in $\mathcal{D}$ does not exceed h, where h is the Y-length of all history sectors of the configurations from $\mathcal{D}$.

Every rule of the subcomputation $\mathcal{E}$ can change the length of any working sector at most by 1. (See Lemma 2.3 (2)). Hence if its length inW_0 is ℓ, its length in W_s is at most $\ell + h$. It follows that $|W_s|_Y \leq 3|W_0|_Y$, because the working sectors of $\mathbf{M}_2$ and its history sectors alternate in the standard base; and the same inequality $|W_r|_Y \le 3|W_0|_Y$ holds for any configuration W_r of $\mathcal{E}$. Hence $s \ne j$ and the subcomputation $\mathcal{E}$ follows in $\mathcal{D}$ by a subcomputation $\mathcal{F}$ of **PM**, which does not change the length of configurations by Lemma 3.12.

So $\mathcal{F}$ has to follow in $\mathcal{D}$ by a maximal subcomputation $\mathcal{G}$ of $\mathcal{M}_3$ again. Since we have the canonical work of $\mathcal{M}_3$ in history sectors, a prefix of the history of $\mathcal{G}^{-1}$ is a copy of the entire $H(\mathcal{E})^{-1}$, where $H(\mathcal{E})$ is the history of $\mathcal{E}$. ($\mathcal{G}$ cannot be shorter than $\mathcal{E}$ since otherwise the configuration W_j would have a copy in $\mathcal{E}$, whence $|W_j|_Y \le 3|W_0|_Y$, a contradiction.) It follows that a configuration W_l of $\mathcal{G}$ is a copy of W_0, and so $|W_l|_Y = |W_0|_Y$. Since the subcomputation $W_l \to \dots \to W_j \to \dots \to W_t$ is shorter than $\mathcal{C}$, we complete the proof of the lemma inducting on t. $\square$

5 Groups and diagrams

5.1 The groups

Every S-machine can be simulated by a finitely presented group (see [28], [20], [21], etc.). Here we apply a modified construction from [28] to the S-machine **M**. To simplify formulas, it is convenient to change the notation. From now on we shall denote by N the length of the standard base of **M**.

Thus the set of state letters is $Q = \sqcup_{i=0}^{N-1} Q_i$ (we set $Q_N = Q_0 = \{t\}$) $Y = \sqcup_{i=1}^{N} Y_i$, and Θ is the set of rules of the S-machine **M**.

The finite set of generators of the group M consists of *q-letters*, *a-letters* and *θ-letters* defined as follows.

For every letter $q \in Q$ the set of generators of M contains L copies $q^{(i)}$ of it, $i = 1, \dots, L$, if the letter q occurs in the rules of $\mathbf{\Theta}_1$ or $\mathbf{\Theta}_2$. (The number L is one of the parameters from Section 2.3.) Otherwise only the letter q is included in the generating set of M.

For every letter $a \in Y$ the set of generators of M contains a and L copies $\alpha^{(i)}$ of it.

For every $\theta \in \Theta^+$ we have N generators $\theta_0, \dots, \theta_N$ in M (here $\theta_N \equiv \theta_0$) if θ is a rule of Set 3 (excluding $\theta(23)$) or Set 4, or Set 5. For θ of Set 1 or 2 (including $\theta(23)$), we introduce LN generators $\theta_j^{(i)}$, where $j = 0, \dots, N$, $i = 1, \dots, L$ and $\theta_N^{(i)} = \theta_0^{(i+1)}$ (the superscripts are taken modulo L).

The relations of the group M correspond to the rules of the S-machine **M** as follows. For every rule $\theta = [U_0 \to V_0, \dots U_N \to V_N] \in \Theta^+$ of Set 1 or 2, we have

$$U_j^{(i)}\theta_{j+1}^{(i)} = \theta_j^{(i)} V_j^{(i)}, \qquad \theta_j^{(i)} a^{(i)} = a^{(i)}\theta_j^{(i)}, \quad j = 0, ..., N, \quad i = 1, \dots L, \tag{5.6}$$

for all $a \in Y_j(\theta)$, where $U_j^{(i)}$ and $V_j^{(i)}$ are obtained from U_j and V_j by addiing the superscript i to every letter.

For $\theta = \theta(23)$, we introduce relations

$$U_j^{(i)}\theta_{j+1}^{(i)} = \theta_j^{(i)} V_j, \qquad a^{(i)}\theta_j^{(i)} = \theta_j^{(i)} a, \tag{5.7}$$

i.e.,the superscripts are erased in the words $U_j^{(i)}$ and in the a-letters after an application of (5.7).

For every rule $\theta = [U_0 \to V_0, \dots U_N \to V_N] \in \Theta^+$ of Set 3 or 4, or 5, we define

$$U_j\theta_{j+1} = \theta_j V_j, \qquad a\theta_j = \theta_j a \tag{5.8}$$

The first type of relations (5.6 - 5.8) will be called (θ, q)-*relations*, the second type - (θ, a)-*relations*.

Finally, the required group G is given by the generators and relations of the group M and by two more additional relations, namely the *hub*-relations

$$W_{st}^{(1)} \dots W_{st}^{(L)} = 1 \;\; and \;\; (W_{ac})^L = 1, \tag{5.9}$$

where the word $W_{st}^{(i)}$ is a copy with superscript (i) of the start word W_{st} (of length N) of the S-machine $\mathbf{M}$ and W_{ac} is the accept word of $\mathbf{M}$.

Remark 5.1. The main difference of the construction of M and the groups based on S-machines with hubs from our previous papers[28, 20, 21, 18] and others, is that relations (5.6) are defined differently for different rules of the S-machine. We also use two hub relations instead of just one, although it is easy to see that one hub relation follows from the other (and other relations).

Note also that, as usual, M is a multiple HNN extension of the free group generated by all a- and q-letters.

5.2 Van Kampen diagrams

Recall that a van Kampen *diagram* Δ over a presentation $P = \langle A|\mathcal{R}\rangle$ (or just over the group P) is a finite oriented connected and simply–connected planar 2–complex endowed with a *labeling function* $\mathrm{Lab}\colon E(\Delta) \to A^{\pm 1}$, where $E(\Delta)$ denotes the set of oriented edges of Δ, such that $\mathrm{Lab}(e^{-1}) \equiv \mathrm{Lab}(e)^{-1}$. Given a *cell* (that is a 2-cell) Π of Δ, we denote by $\partial\Pi$ the boundary of Π; similarly, $\partial\Delta$ denotes the boundary of Δ. The labels of $\partial\Pi$ and $\partial\Delta$ are defined up to cyclic permutations. An additional requirement is that the label of any cell Π of Δ is equal to (a cyclic permutation of) a word $R^{\pm 1}$, where $R \in \mathcal{R}$. The label and the combinatorial length $||\mathbf{p}||$ of a path $\mathbf{p}$ are defined as for Cayley graphs.

The van Kampen Lemma [12, 14, 27] states that a word W over the alphabet $A^{\pm 1}$ represents the identity in the group P if and only if there exists a diagram Δ over P such that $\mathrm{Lab}(\partial\Delta) \equiv W$, in particular, the combinatorial perimeter $||\partial\Delta||$ of Δ equals $||W||$. ([12], Ch. 5, Theorem 1.1; our formulation is closer to Lemma 11.1 of [14], see also [27, Section 5.1]). The word W representing 1 in P is freely equal to a product of conjugates to the words from $R^{\pm 1}$. The minimal number of factors in such products is called the *area* of the word W. The *area* of a diagram Δ is the number of cells in it. The proof of the van Kampen Lemma [14, 27] shows that $\mathrm{Area}(W)$ is equal to the area of a van Kampen diagram having the smallest number of cells among all van Kampen diagrams with boundary label $\mathrm{Lab}(\partial\Delta) \equiv W$.

We will study diagrams over the group presentations of M and G. The edges labeled by state letters ($=$ *q-letters*) will be called *q-edges*, the edges labeled by tape letters ($=$ *a-letters*) will be called *a-edges*, and the edges labeled by θ-letters are *θ-edges*.

We denote by $|\mathbf{p}|_{\mathbf{Y}}$ (by $|\mathbf{p}|_\theta$, by $|\mathbf{p}|_{\mathbf{q}}$) the *$a$-length* (resp., the *$\theta$-length*, the q-length) of a path/word $\mathbf{p}$, i.e., the number of a-edges/letters (the number of θ-edges/letters, the number of q-edges/letters) in $\mathbf{p}$.

The cells corresponding to relations (5.9) are called *hubs*, the cells corresponding to (θ, q)-relations are called (θ, q)-*cells*, and the cells are called (θ, a)-*cells* if they correspond to (θ, a)-relations.

A Van Kampen diagram is *reduced*, if it does not contain two cells (= closed 2-cells) that have a common edge e such that the boundary labels of these two cells are equal if one reads them starting with e (if such pairs of cells exist, they can be removed to obtain a diagram of smaller area and with the same boundary label).

5.2.1 The superscript shift of a van Kampen diagram over M or G

Remark 5.2. If one changes all superscripts of the generators of M or G by adding the same integer k: $(i) \to (i+k)$ (modulo L) in all letters having a superscript, then one obtains the relations again, as it is clear from formulae (5.6 - 5.9). Therefore similar change $\Delta \to \Delta^{(+k)}$ of the edge labels transforms a (reduced) diagram Δ to a (reduced) diagram $\Delta^{(+k)}$. Let us call such a transformation *superscript shift* (or *k-shift*) of Δ.

5.2.2 Bands

To study (van Kampen) diagrams over the group G we shall use their simpler subdiagrams such as bands and trapezia, as in [16], [28], [1], etc. Here we repeat one more necessary definition.

Definition 5.3. Let $\mathcal{Z}$ be a subset of the set of letters in the set of generators of the group M. A $\mathcal{Z}$-band $\mathcal{B}$ is a sequence of cells $\pi_1, ..., \pi_n$ in a reduced van Kampen diagram Δ such that

- Every two consecutive cells π_i and π_{i+1} in this sequence have a common boundary edge $\mathbf{e}_i$ labeled by a letter from $\mathcal{Z}^{\pm 1}$.
- Each cell π_i, $i = 1, ..., n$ has exactly two $\mathcal{Z}$-edges in the boundary $\partial\pi_i$, $\mathbf{e}_{i-1}^{-1}$ and $\mathbf{e}_i$ (i.e.,edges labeled by a letter from $\mathcal{Z}^{\pm 1}$) with the requirement that either both $\mathrm{Lab}(e_{i-1})$ and $\mathrm{Lab}(e_i)$ are positive letters or both are negative ones.
- If $n = 0$, then $\mathcal{B}$ is just a $\mathcal{Z}$-edge.

The counter-clockwise boundary of the subdiagram formed by the cells $\pi_1, ..., \pi_n$ of $\mathcal{B}$ has the factorization $\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f}\mathbf{q}_2^{-1}$ where $\mathbf{e} = \mathbf{e}_0$ is a $\mathcal{Z}$-edge of π_1 and $\mathbf{f} = \mathbf{e}_n$ is an $\mathcal{Z}$-edge of π_n. 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 $\mathbf{top}(\mathcal{B})$. The *trimmed* top/bottom label are the maximal subwords of the top/bottom labels starting end ending with q-letters.

Top/bottom paths and their inverses are also called the *sides* of the band. The $\mathcal{Z}$-edges $\mathbf{e}$ and $\mathbf{f}$ are called the *start* and *end* edges of the band. If $n \geq 1$ but $\mathbf{e} = \mathbf{f}$, then the $\mathcal{Z}$-band is called a *$\mathcal{Z}$-annulus* .

If $\mathcal{B}$ is a $\mathcal{Z}$-band with $\mathcal{Z}$-edges $e_1, ..., e_n$ (in that order), then we can form a broken line connecting midpoints of the consecutive edges $e_1, ..., e_n$ and laying inside the union of the cells from $\mathcal{B}$ which will be called the *median* of $\mathcal{B}$.

We will consider *q-bands*, where $\mathcal{Z}$ is one of the sets Q_i of state letters for the S-machine $\mathbf{M}$, *θ-bands* for every $\theta \in \Theta$, and *a-bands*, where $\mathcal{Z} = \{a, a^{(1)}, \dots, a^{(L)}\} \subseteq Y$. The convention is that Y-bands do not contain (θ, q)-cells, and so they consist of (θ, a)-cells only.

Lemma 5.4. *Let $\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f}\mathbf{q}_2^{-1}$ be the boundary of a θ-band $\mathcal{B}$ 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 $\mathcal{B}$ has (θ, q)-cells.

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

Proof. (1) If every cell π_i of $\mathcal{B}$ is a (θ, a)-cell, then both θ-edges of the boundary $\partial\pi_i$ have equal labels, as it follows from the definition of (θ, a)-relations. Then the definition of band implies that $\mathrm{Lab}(\mathbf{e}) = \mathrm{Lab}(\mathbf{f})$, a contradiction.

(2) Proving by contradiction, we have that that π_i and π_j $(i \neq j)$ share a boundary q-edge $\mathbf{g}$. We may assume that the difference $j - i > 0$ is minimal, and so the subband formed by $\pi_{i+1}, \dots, \pi_{j-1}$ has no (θ, q)-cells. It folows from (1) that π_i and π_j have the same boundary labels if one read then starting with $\mathrm{Lab}(\mathbf{g})$, contrary to the assumption that the diagram is reduced.

□

Remark 5.5. To construct the top (or bottom) path of a band $\mathcal{B}$, 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). No θ-letter is being canceled in the word $W \equiv \mathrm{Lab}(\mathbf{x}_1) \dots \mathrm{Lab}(\mathbf{x}_n)$ if $\mathcal{B}$ is a q- or Y-band since otherwise two neighbor cells of the band would make the diagram non-reduced. By Lemma 5.4 (2), there are no cancellations of q-letters of W if $\mathcal{B}$ is a θ-band.

If $\mathcal{B}$ is a θ-band then a few cancellations of a-letters (but not q-letters) are possible in W. (This can happen if one of π_i, π_{i+1} is a (θ, q)-cell and another one is a (θ, a)-cell.) We will always assume that the top/bottom label of a θ-band is a reduced form of the word W. This property is easy to achieve: by folding edges with the same labels having the same initial vertex, one can make the boundary label of a subdiagram in a van Kampen diagram reduced (e.g., see [14] or [28]).

We shall 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 they have a common cell and $\mathcal{Z}_1 \cap \mathcal{Z}_2 = \emptyset$.

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

Our previous papers (see [28], [1], etc.) contain the proof of the next lemma in a more general setting. The difference caused by different simulation of the S-machine $\mathbf{M}$ by defining relations of M does not affect the validity of the proof since the proof uses the properties mentioned in Lemma 5.4 and Remark 5.5. To convince the reader, below we recall the proof of one the following claims.

Lemma 5.6. *A reduced van Kampen diagram Δ over M has no q-annuli, no θ-annuli, and no a-annuli. Every θ-band of Δ shares at most one cell with any q-band and with any Y-band.*

Proof. We will prove only the property that a θ-band $\mathcal{T}$ and a q-band $\mathcal{Q}$ cannot cross each other two times. Taking a minimal counter-example, one asuumes that these bands have exactly two common cells π and π', and Δ has no cells outside the region bounded by $\mathcal{T}$ and $\mathcal{Q}$. Then $\mathcal{Q}$ has exactly two cells since otherwise a maximal θ-band starting

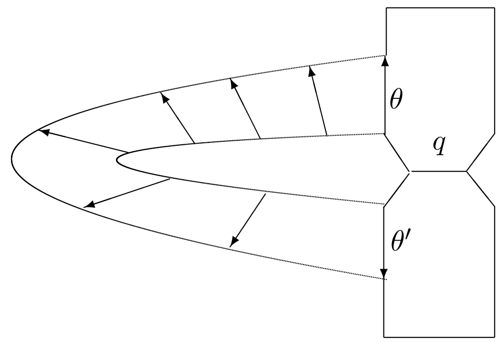


Figure 2: A Q-band intersects a θ-band twice.

with a cell π'' of $\mathcal{Q}$, where $\pi'' \notin \{\pi, \pi'\}$, has to end on $\mathcal{Q}$, bounding with a part of $\mathcal{T}$ a smaller counter-example.

Thus, the boundaries of π and π' share a q-edge.

For the similar reason, $\mathcal{T}$ has no (θ, q)-cells except for π and π', and by Lemma 5.4 (1), these cells have the same pairs of θ-edges in the boundaries. This makes the diagram non-reduced, a contradiction. □

If $W \equiv x_1...x_n$ is a word in an alphabet X, X' is another alphabet, and $\phi\colon X \to X' \cup \{1\}$ (where 1 is the empty word) is a map, then $\phi(W) \equiv \phi(x_1)...\phi(x_n)$ is called the *projection* of W onto X'. We shall consider the projections of words in the generators of M onto Θ (all θ-letters map to the corresponding element of Θ, all other letters map to 1), and the projection onto the alphabet $\{Q_0 \sqcup \cdots \sqcup Q_{N-1}\}$ (every q-letter maps to the corresponding Q_i, all other letters map to 1).

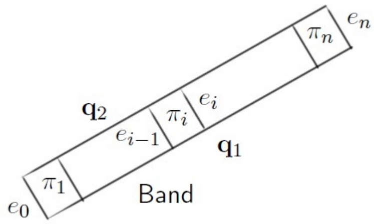
Definition 5.7. The projection of the label of a side of a q-band onto the alphabet Θ is called the *history* of the band. The step history of this projection is the *step history* of the q-band. The projection of the label of a side of a θ-band onto the alphabet $\{Q_0, ..., Q_{N-1}\}$ is called the *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

As in the case of words, we will use representatives of Q_j-s in base words.

If W is a word in the generators of M, then by $W^\emptyset$ we denote the projection of this word onto the alphabet of the S-machine $\mathbf{M}$, we obtain this projection after deleting all superscripts in the letters of W. In particular, $W^\emptyset \equiv W$, if there are no superscripts in the letters of W.

We call a word W in q-generators and a-generators *permissible* if the word $W^\emptyset$ is admissible, and the letters of any 2-letter subword of W have equal superscripts (if any), except for the subwords $(q\check{t})^{\pm 1}$, where the letter q has some superscript (i) and $q^\emptyset \in Q_{N-1}$; in this case the superscript of the letter $\check{t}$ must be $(i+1)$ (modulo L).

Remark 5.8. It follows from the definition that if V is θ-admissible for a rule θ of $\{\theta(23)^{-1}\} \cup \mathbf{\Theta}_3 \cup \{\theta(34)\} \cup \mathbf{\Theta}_4 \cup \{\theta(45)\} \cup \mathbf{\Theta}_5$, then there is exactly one permissible word W such that $W^\emptyset \equiv V$, namely, $W \equiv V$. If θ is a rule of $\mathbf{\Theta}_1 \cup \{\theta(12)\} \cup \mathbf{\Theta}_2 \cup \{\theta(23)\}$, then the permissible word W with property $W^\emptyset \equiv V$ exists and it is uniquely defined if one choose arbitrary superscript for the first letter (or for any particular letter) of W.



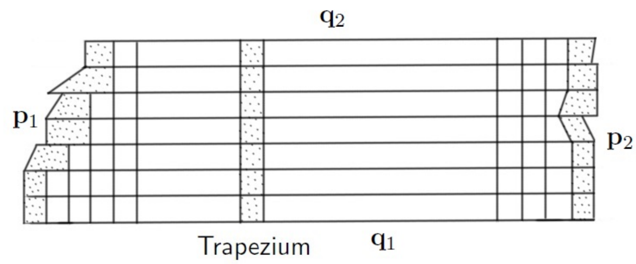


Figure 3: Band and Trapezium

Lemma 5.9. *(1) The trimmed bottom and top labels W_1 and W_2 of any reduced θ-band $\mathcal{T}$ containing at least one $(\theta, q)-$cell are permissible and $W_2^{\emptyset} \equiv W_1^{\emptyset} \cdot \theta$.*

(2) If W is a θ-admissible word, then for a permissible word W_1 such that $W_1^{\emptyset} \equiv W$ (given by Remark 5.8) one can construct a reduced θ-band with the trimmed bottom label W_1 and the trimmed top label W_2, where $W_2^{\emptyset} \equiv W_1^{\emptyset} \cdot \theta$.

Proof. (1) By Lemma 5.4 (2), we have $W_1 \equiv q_1^{\pm 1} u_1 q_2^{\pm 1} \dots u_k q_{k+1}^{\pm 1}$, where $q_j^{\pm 1}$ and $q_{j+1}^{\pm 1}$ 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 5.4 (1), all the θ-edges between $\pi(j)$ and $\pi(j+1)$ have the same labels. It follows from the list of (θ, a)-relations that all a-letters of the word u_j have to belong in the same subalphabet. In particular, if we have the subword $q_j u_j q_{j+1}$, then the projection of this subword is a subword of $W_1^{\emptyset}$ satisfying the first condition from the definition of admissible word. Similarly one obtains other conditions if q_j or/and q_{j+1} occur in W_1 with exponent -1. Hence the word $W_1^{\emptyset}$ (and $W_2^{\emptyset}$) are admissible, and the words W_1, W_2 are permissible since again the condition on 2-letter subwords follows from Lemma 5.4 and the relations (5.6 - 5.8).

If $\mathbf{x} = \mathbf{x}_1 \dots \mathbf{x}_n$ ($\mathbf{y} = \mathbf{y}_1 \dots \mathbf{y}_n$) is the product of the top paths $\mathbf{x}_i$-s (bottom paths $\mathbf{y}_i$-s) of the all cells $\pi_1, \dots, \pi_n$ of $\mathcal{T}$, as in Remark 5.5, then the transition from the trimmed label of $\mathbf{x}$ to the trimmed label of $\mathbf{y}$ with erased superscripts, is the application of θ, as it follows from relations (5.6 - 5.8). Since by definition, the application of θ automatically implies possible cancellations, we have $W_2^{\emptyset} \equiv W_1^{\emptyset} \cdot \theta$ for the reduced words W_1 and W_2, as required.

(2)Since W is θ-admissible, there is an equality $W' \equiv W \cdot \theta$. Therefore we can simulate the application of θ to every letter of W as follows. We draw a path $\mathbf{p} = \mathbf{e_1} \dots \mathbf{e_n}$ labeled by W_1 and attach a cell π_i corresponding to one of the defining relations of M to every edge $\mathbf{e}_i$ of $\mathbf{p}$ from the left. Since the word W_1 is permissible, the θ-edges started with the common vertex of π_i and π_{i+1} must have equal labels, and so these two edges can be identified. Finally, we obtain a required θ-band. It is reduced diagram since the permissible word W_1 is reduced. □

5.2.3 Trapezia

Definition 5.10. Let Δ be a reduced diagram over M, 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.

Then Δ is called a *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 trapezium. The history (step history) of the q-band whose side is $\mathbf{p}_2$ is called the *history* (resp., step history) of the trapezium; the length of the history is called the *height* of the trapezium. The base of $\mathrm{Lab}(\mathbf{q}_1)$ is called the *base* of the trapezium.

Remark 5.11. Notice that the top (bottom) side of a θ-band $\mathcal{T}$ does not necessarily coincides 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. We shall denote the *trimmed* top and bottom sides of $\mathcal{T}$ by $\mathbf{ttop}(\mathcal{T})$ and $\mathbf{tbot}(\mathcal{T})$. By definition, for arbitrary θ-band $\mathcal{T}$, $\mathbf{ttop}(\mathcal{T})$ is obtained by such a trimming only if $\mathcal{T}$ starts and/or ends with a (θ,q)-cell; otherwise $\mathbf{ttop}(\mathcal{T})=\mathbf{top}(\mathcal{T})$. The definition of $\mathbf{tbot}(\mathcal{T})$ is similar.

By Lemma 5.6, any trapezium Δ of height $h\ge 1$ can be decomposed into θ-bands $\mathcal{T}_1,...,\mathcal{T}_h$ connecting the left and the right sides of the trapezium. The analog of the following lemma is false for Turing machines. (See [19] for a discussion.)

Lemma 5.12. *(1) Let Δ be a trapezium with history $H\equiv\theta(1)\dots\theta(d)$ $(d\ge 1)$. 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 the trimmed top labels of $\mathcal{T}_j$, $(j=1,\dots,d)$. Then H is an eligible word, U_j, V_j are permissible words,*

$$V_1^{\emptyset}\equiv U_1^{\emptyset}\cdot\theta(1),\ \ U_2\equiv V_1,\ \ \dots,\ \ U_d\equiv V_{d-1},\ \ V_d^{\emptyset}\equiv U_d^{\emptyset}\cdot\theta(d)$$

Furthemore, if the first and the last q-letters of the word U_j or of the word V_j have some superscripts (i) and (i'), then the difference $i'-i$ (modulo L) does not depend on on the choice of U_j or V_j.

(2) For every eligible computation $U\to\dots\to U\cdot H\equiv V$ of $\mathbf{M}$ with $||H||=d\ge 1$ there exists a trapezium Δ with bottom label U_1 (given by Remark 5.8) such that $U_1^{\emptyset}\equiv U$, top label V_d such that $V_d^{\emptyset}\equiv V$, and with history H.

Proof. (1) The trimmed top side of one of the bands $\mathcal{T}_j$ is the same as trimmed bottom side of $\mathcal{T}_{j+1}$ $(j=1,\dots,d-1)$, and the equalities $U_2\equiv V_1,\dots,U_d\equiv V_{d-1}$ follow. The equalities $V_j^{\emptyset}\equiv U_j$ $(j=1,\dots d)$ are given by Lemma 5.9 (1). By the same lemma the words U_j and V_j are permissible.

Assume that there is a cancellation: $\theta(i+1)\equiv\theta(i)^{-1}$. Since Δ is a reduced diagram, any pair of (θ,q)-cells $\pi\in\mathcal{T}_i$ and $\pi'\in\mathcal{T}_{i+1}$ with a common q-edge $\mathbf{e}$ are not cancellable. Hence the relations given by these cells are not uniquely defined by the q-letter $\mathrm{Lab}(\mathbf{e})$ and the history letter $\theta(i)$. It follows from the list of defining relations (5.6 - 5.8) that $\mathrm{Lab}(\mathbf{e})$ has no superscripts while other labels of the boundary edges of these two cells do have superscripts. Thus, these relations are in the list (5.7) and $\theta(i)\equiv\theta(23)$, which prove that the history H is eligible.

Since by Lemma 5.6 every maximal q-band of Δ connects the top and the bottom of Δ, it suffices to prove the last claim under assumption that the base of Δ is a word $Q^{\pm1}(Q')^{\pm1}$of length 2. Then by definition of permissible word, $i'-i=0$, except for the base $Q_{N-1}Q_N$ (or the inverse one) with $i'-i=1$ modulo L (resp., $i'-i=-1$ modulo L). Since all the words U_j and V_j have equal bases, the last statement of (1) is proved.

(2) We can obtain the $\theta(1)$-band $\mathcal{T}_1$ by Lemma 5.9 (2). By induction, there is a trapezium Δ' of height $d-1$ with bottom label $U_2\equiv U_1$ an top label V such that

$U_2^{\emptyset} \equiv U_1^{\emptyset} \cdot \theta(1)$ and $V_d^{\emptyset} \equiv V$, such that the union Δ of $\mathcal{T}_1$ and Δ' has history H. If Δ is not reduced then we have a pair of cancellable cells $\pi \in \mathcal{T}_1$ and $\pi' \in \mathcal{T}_2$. Then as in item (1) we conlude that $\theta(1) \equiv \theta(23)$, and so the top $\mathbf{q}$ of $\mathcal{T}_1$ has no superscript in the boundary label. Therefore one can replace Δ' with its subscript shift $(\Delta')^{+1}$ in Δ. After such a modification, Δ becomes a reduced diagram since for any pair cells π and π' with common boundary edge from $\mathbf{q}$, the other edges have now different superscripts in their labels. Since $V_d^{\emptyset}$ does not change under the superscript shift, the lemma is proved. □

5.2.4 Big and standard trapezia

Using Lemma 5.12, one can immediately derive properties of trapezia from the properties of computations obtained earlier.

If $H' \equiv \theta(i) \dots \theta(j)$ is a subword of the history H from Lemma 5.12 (1), then the bands $\mathcal{T}_i, \dots, \mathcal{T}_j$ form a subtrapezium Δ' of the trapezium Δ. This subtrapezium is uniquely defined by the subword H' (more precisely, by the occurrence of H' in the word $\theta_1 \dots \theta_d$), and Δ' is called the *H'-part* of Δ.

We say that a trapezium Δ is *standard* if the base of Δ is the standard base $\mathbf{B}$ of $\mathbf{M}$ or $\mathbf{B}^{-1}$, and the history of Δ (or the inverse one) contains one of the words (a) $\chi(i-1,i)H'\chi(i,i+1)$ (i.e.,the S-machine works as $\mathbf{\Theta}_4$) or (b) $\zeta^{i-1,i}H'\zeta^{i,i+1}$ (i.e.,it works as $\mathbf{\Theta}_2$).

Definition 5.13. We say that a trapezium Γ is *big* if

(1) the base of Δ or the inverse word has the form xvx, where xv a cyclic shift of the L-s power of the standard base;

(2) the diagram Γ contains a standard trapezium.

Lemma 5.14. *Let Δ be a trapezium whose base is xvx, where x occurs in v exactly $L-1$ times and other letters occur $< L$ times each. Then either Δ is big or the length of a side of every θ-band of Δ does not exceed $c_5(||W||+||W'||)$, where W, W' are the labels of its top and bottom, respectively.*

Proof. The diagram Δ is covered by L subtrapezia Γ_i $(i = 1, \dots, L)$ with bases xv_ix.

Assume that the the step history of Δ (or inverse step history) contains one of the subwords $\chi(i-1,i)H'\chi(i,i+1)$ or (b) $\zeta^{i-1,i}H'\zeta^{i,i+1}$. Then by Lemma 4.4 (and 5.12), the base of Δ has the form $(xu)^L x$, where xu is a cyclic shift of the standard base (or the inverse one). Since Δ contains a standard subtrapezia, it is is big.

Now, under the assumption that the step history has no subwords mentioned in the previous paragraph, it suffices to bound the the length of a side of every θ-band of arbitrary Γ_i by $\le c_4(||V|_Y + ||V'||)$, where V and V' are the labels of the top and the bottom of Γ_i.

Assume that the word xv_ix has a proper subword yuy, where u has no letters y, and any other letter occurs in u at most once. Then the word yuy is faulty since v_i has no letters x. By Lemma 4.8, we have $|U_j|_Y \le c_1 \max(|U_0|_Y, |U_t|_Y)$ for every configuration U_j of the computation given by Lemma 5.12 (1) restricted to the base yuy. Since $c_4 > c_1$, it suffices to obtain the desired estimate for the computation whose base is obtained by deleting the subword yu from xv_ix. Hence inducting on the length of the base of Γ_i, one may assume that it has no proper subwords yuy, and so the base of Γ_i is revolving. Now the required upper estimate for Γ_i follows from Lemma 4.11 (see (1) and (2c) there). □

6 Diagrams without hubs

6.1 A modified length function

Let us modify the length function on the words and paths. The standard length of a word (a path) will be called its *combinatorial length*. From now on we use the word *'length'* for the modified length. We set the length of every q-letter equal 1, and the length of every a-letter equal a small enough number δ so that $J\delta < 1$.

We also set to 1 the length of every word of length ≤ 2 which contains exactly one θ-letter and no q-letters (such words are called (θ, a)*-syllables*). The length of a decomposition of an arbitrary word in a product of letters and (θ, a)-syllables is the sum of the lengths of the factors. *The length* $|w|$ *of a word* w is the smallest length of such decompositions. *The length* $|\mathbf{p}|$ *of a path* in a diagram is the length of its label. The *perimeter* $|\partial\Delta|$ of a van Kampen diagram is similarly defined by cyclic decompositions of the boundary $\partial\Delta$.

The next statement follows from the above definitions and from the property of (θ, q)-relations and their cyclic shifts: the subword between two q-letters in arbitrary (θ, q)-relation is a syllable.

Lemma 6.1. *Let* $\mathbf{s}$ *be a path in a diagram* Δ *having* c θ*-edges and* d a*-edges. Then*

(a) $|s| \geq \max(c, c + (d - c)\delta)$*;*

(b) $|s| = c$ *if* $\mathbf{s}$ *is a top or a bottom of a* q*-band.*

(c) For any product $\mathbf{s} = s_1 s_2$ *of two paths in a diagram, we have*

$$|s_1| + |s_2| \geq |s| \geq |s_1| + |s_2| - \delta \tag{6.10}$$

(d) Let $\mathcal{T}$ *be a* θ*-band with base of length* l_b*. Let* l_Y *be the number of* a*-edges in the top path* $\mathbf{t}opp(\mathcal{T})$*. Then the length of* $\mathcal{T}$ *(i.e., the numer of cells in* $\mathcal{T}$*) is between* $l_Y - l_b$ *and* $l_Y + 3l_b$*.*

6.1.1 Rim bands

Let $\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f}\mathbf{q}_2^{-1}$the factorization $\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f}\mathbf{q}_2^{-1}$ be the standard factorization of a band. If the path $(\mathbf{e}^{-1}\mathbf{q}_1\mathbf{f})^{\pm 1}$ or the path $(\mathbf{f}\mathbf{q}_2^{-1}\mathbf{e}^{-1})^{\pm 1}$ is the subpath of the boundary path of Δ then the band is called a *rim* band of Δ.

From now on we shall fix a constant K

$$K > 2K_0 = 4LN \tag{6.11}$$

The following easy basic facts about θ-bands are [18, Lemmas 6.2, 6.3] (see also [21, Lemma 4.5]). These facts will allow us to remove short enough rim bands from van Kampen diagrams (see Lemma 6.13 below).

Lemma 6.2. *Let* Δ *be a van Kampen diagram whose rim* θ*-band* $\mathcal{T}$ *has base with at most* K *letters. Denote by* Δ' *the subdiagram* $\Delta\backslash\mathcal{T}$*. Then* $|\partial\Delta| - |\partial\Delta'| > 1$*.*

Proof. Let $\mathbf{s}$ be the top side of $\mathcal{T}$ and $\mathbf{s} \subset \partial\Delta$. Note that the difference between the number of a-edges in the bottom $\mathbf{s}'$ of $\mathcal{T}$ and the number of a-edges in s cannot be greater than $2K$, because every (θ, q)-relator has at most two a-letters. Hence $|\mathbf{s}'| - |\mathbf{s}| \leq 4LN\delta$. However, Δ' is obtained by cutting off $\mathcal{T}$ along $\mathbf{s}'$, and its boundary contains two θ-edges fewer than Δ. Hence we have $|\mathbf{s_0}| - |\mathbf{s_0'}| \geq 2 - 2\delta$ for the complements $\mathbf{s}_0$ and $\mathbf{s'}_0$ of s and s', respectively, in the boundaries $\partial\Delta$ and $\partial\Delta'$. Finally,

$$|\partial\Delta| - |\partial\Delta'| \geq 2 - 2\delta - 2K\delta - 4\delta > 1$$

by (2.3) and (6.10). □

We call a base word w *tight* if

(1) for some letter x the word w has the form $uxvx$, where the letter x does not occur in u and x occurs in v exactly $L-1$ times,

(2) every proper prefix w' of w does not satisfy property (1).

Lemma 6.3. *([18], Lemma 6.3). If a base w of a θ-band has no tight prefixes, then $||w|| \leq K_0$, where $K_0 = 2LN$.*

Proof. The hub base includes every base letter L times. Hence every word in this group alphabet of length $\geq K_0 + 1$ includes one of the letters $L+1$ times. □

6.1.2 Combs

Definition 6.4. We say that a reduced diagram Γ is a *comb* if it has a maximal q-band $\mathcal{Q}$ (the *handle* of the comb), such that

(C_1) $\mathbf{bot}(\mathcal{Q})$ is a part of $\partial\Gamma$, and every maximal θ-band of Γ ends at a cell in $\mathcal{Q}$.

If in addition the following properties hold:

(C_2) one of the maximal θ-bands $\mathcal{T}$ in Γ has a tight base (if one reads the base towards the handle) and

(C_3) other maximal θ-bands in Γ have tight bases or bases without tight prefixes

then the comb is called *tight*.

The number of cells in the handle $\mathcal{Q}$ is the *height* of the comb, and the maximal length of the bases of the θ-bands of a comb is called the *basic width* of the comb.

Notice that every trapezium is a comb.

Lemma 6.5. *([21], Lemma 4.10) Let l and b be the length and the basic width of a comb Γ and let $\mathcal{T}_1, \dots \mathcal{T}_l$ be consecutive θ-bands of Γ (as in Figure 4). We can assume that $\mathbf{b}ot(\mathcal{T}_1)$ and $\mathbf{t}op(\mathcal{T}_l)$ are contained in $\partial\Gamma$. Denote by $\alpha = |\partial\Gamma|_Y$ the number of a-edges in the boundary of Γ, and by α_1 the number of a-edges on $\mathbf{b}ot(\mathcal{T}_1)$. Then $\alpha + 2lb \geq 2\alpha_1$, and the area of Γ does not exceed $c_0bl^2 + 2\alpha l$ for some constant $c_0 = c_0(\mathbf{M})$. (Recall that c_0 is one of the parameters from Section 2.3.)*

We say that a subdiagram Γ of a diagram Δ is a *subcomb* of Δ if Γ is a comb, the handle of Γ divides Δ in two parts, and Γ is one of these parts.

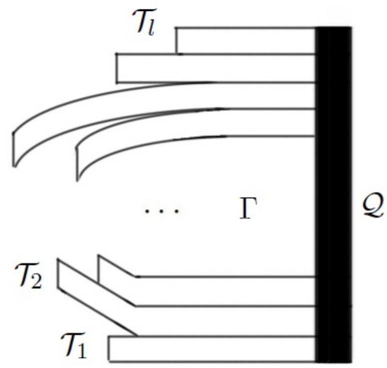


Figure 4: Comb.

Lemma 6.6. *Let Δ be a reduced diagram over G with non-zero area, where every rim θ-band has base of length at least K. Assume that*

(1) Δ is a diagram over the group M or

(2) Δ has a subcomb of basic width at least K_0.

Then there exists a maximal q-band $\mathcal{Q}$ dividing Δ in two parts, where one of the parts is a tight subcomb with handle Q.

Proof. Let $\mathcal{T}_0$ be a rim band of Δ (fig.5). Its base w is of length at least K, and therefore w has disjoint prefix and suffix of lengths K_0 since $K > 2K_0$ by (6.11). The prefix of this base word must have its own tight subprefix w_1, by Lemma 6.3 and the definition of tight words. A q-edge of $\mathcal{T}_0$ corresponding to the last q-letter of w_1 is the start edge of a maximal q-band $\mathcal{Q}'$ which bounds a subdiagram Γ' containing a band $\mathcal{T}$ (a subband of $\mathcal{T}_0$) satisfying property (C_2). It is useful to note that a minimal suffix w_2 of w, such that w_2^{-1} is tight, allows us to construct another band $\mathcal{Q}''$ and a subdiagram Γ'' which satisfies (C_2) and has no cells in common with Γ'.

Thus, there are $\mathcal{Q}$ and Γ satisfying (C_2). Let us choose such a pair with minimal Area(Γ). Assume that there is a θ-band in Γ which does not cross $\mathcal{Q}$. Then there must exist a rim band $\mathcal{T}_1$ which does not cross $\mathcal{Q}$ in Γ. Hence one can apply the construction from the previous paragraph to $\mathcal{T}_1$ and construct two bands $\mathcal{Q}_1$ and $\mathcal{Q}_2$ and two disjoint subdiagrams Γ_1 and Γ_2 satisfying the requirement (C_2) for Γ. Since Γ_1 and Γ_2 are disjoint, one of them, say Γ_1, is inside Γ. But the area of Γ_1 is smaller than the area of Γ, and we come to a contradiction. Hence Γ is a comb and condition (C_1) is satisfied.

Assume that the base of a maximal θ-band $\mathcal{T}$ of Γ has a tight proper prefix (we may assume that $\mathcal{T}$ terminates on $\mathcal{Q}$), and again one obtain a q-band $\mathcal{Q}'$ in Γ, which provides us with a smaller subdiagram Γ' of Δ, satisfying (C_2), a contradiction. Hence Γ satisfies property (C_3) as well.

(2) The proof is shorter since a comb is given in the very beginning. □

6.2 The mixture

We will need a numerical parameter associated with van Kampen diagrams introduced in [17], it was called *mixture*.

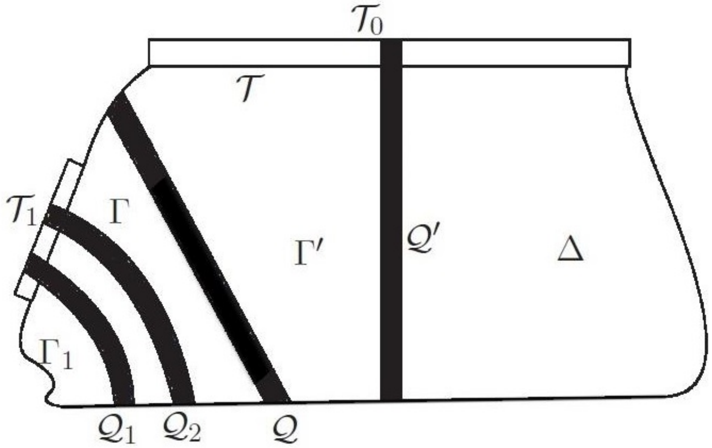


Figure 5: Lemma 6.6.

Let O be a circle with two-colored (black and white) finite set of points (or vertices) on it. We call O a *necklace* with black and white *beads* on it.

Assume that there are n white beads and n' black ones on O. We define sets $\mathbf{P}_j$ of ordered pairs of distinct white beads as follows. A pair (o_1, o_2) $(o_1 \neq o_2)$ belongs to the set $\mathbf{P}_j$ if the simple arc of O drawn from o_1 to o_2 in the clockwise direction has at least j black beads. We denote by $\mu_J(O)$ the sum $\sum_{j=1}^{J} \mathrm{card}(\mathbf{P}_j)$ (the *J-mixture* of O). Below similar sets for another necklace O' are denoted by $\mathbf{P'}_j$. . In this subsection, $J \geq 1$, but later on it will be a fixed large enough number J from the list (2.3).

Lemma 6.7. *([17], Lemma 6.1) (a)* $\mu_J(O) \leq J(n^2 - n)$.

(b) Suppose a necklace O' is obtained from O after removal of a white bead v. Then $\mathrm{card}(\mathbf{P}_j) - n < \mathrm{card}(\mathbf{P'}_j) \leq \mathrm{card}(\mathbf{P}_j)$ *for every j, and* $\mu_J(O) - Jn < \mu_J(O') \leq \mu_J(O)$.

(c) Suppose a necklace O' is obtained from O after removal of a black bead v. Then $\mathrm{card}(\mathbf{P'}_j) \leq \mathrm{card}(\mathbf{P}_j)$ *for every j, and* $\mu_J(O') \leq \mu_J(O)$.

(d) Assume that there are three black beads v_1, v_2, v_3 of a necklace O, such that the clockwise arc $v_1 - v_3$ contains v_2 and has at most J black beads (excluding v_1 and v_3), and the arcs $v_1 - v_2$ and $v_2 - v_3$ have m_1 and m_2 white beads, respectively. If O' is obtained from O by removal of v_2, then $\mu_J(O') \leq \mu_J(O) - m_1 m_2$.

For any diagram Δ over G, we introduce the following invariant $\mu(\Delta) = \mu_J(\partial\Delta)$ depending on the boundary of Δ only. To define it, we consider the boundary $\partial(\Delta)$, as a *necklace*, i.e., we consider a circle O with $||\partial\Delta||$ edges labeled as the boundary path of Δ. By definition, the white beads are the mid-points of the θ-edges of O and black beads are the mid-points of the q-edges O. Then, by definition, the *mixture* of Δ is $\mu(\Delta) = \mu_J(O)$.

6.3 Quadratic upper bound for quasi-areas of diagrams over M.

6.3.1 The G-area of a diagram over M

The Dehn function of the group M is super-quadratic (in fact by [21] it is at least $n^2 \log n$ because M is a mulltiple HNN extension of a free group). However we are going to obtain a quadratic Dehn function of G, and first we want to bound the areas of the words vanishing in M with respect to the presentation of G. For this goal we artificially introduce the concept of G-area, as in [18]. The G-area of a big trapezia can be much less that the real area of it in M. This concept will be justified at the end of this paper, where some big trapezia are replaced by diagrams with hubs, but having lesser areas.

Definition 6.8. The G-area $\mathrm{Area}_G(\Gamma)$ of a big trapezium Γ is, by definition, the minimum of the half of its area (i.e., the number of cells) and the product

$$c_5 h(||\mathbf{t}op(\Gamma)|| + ||\mathbf{b}ot(\Gamma)||),$$

where h is the height of Γ and c_5 is the constant from (2.3).

To define the G-area of a diagram Δ over M, we consider a family $\mathbf{S}$ of big subtrapezia (i.e.,subdiagrams, which are trapezia) and single cells of Δ such that every cell of Δ belongs to a member Σ of this family, and if a cell Π belongs to different Σ_1 and Σ_2 from $\mathbf{S}$, then both Σ_1 and Σ_2 are big subtrapezia of Δ with bases xv_1x, xv_2x, and Π is a (θ, x)-cell. (In the later case, the intersection $\Sigma_1 \cap \Sigma_2$ must be an x-band.) There is such a family 'covering' Δ, e.g. just the family of all cells of Δ.

The G-area of $\mathbf{S}$ is the sum of G-areas of all big trapezia from $\mathbf{S}$ plus the number of single cells from $\mathbf{S}$ (i.e.,the G-area of a cell Π is $area(\Pi) = 1$). Finally, the *G-area* $\mathrm{Area}_G(\Delta)$ is the minimum of the G-areas of all 'covering' $\mathbf{S}$ as above.

It follows from the definition that $\mathrm{Area}_G(\Delta) \le \mathrm{Area}(\Delta)$ since the G-area of a big trapezium does not exceed a half of its area.

Lemma 6.9. *Let Δ be a reduced diagram, and every cell π of Δ belongs in one of subdiagrams $\Delta_1, \dots, \Delta_m$, where any intersection $\Delta_i \cap \Delta_j$ either has no cells or it is a q-band, Then* $\mathrm{Area}_G(\Delta) \le \sum_{i=1}^m \mathrm{Area}_G(\Delta_i)$.

Proof. Consider the families $\mathbf{S}_1, \dots, \mathbf{S}_m$ given by the definition of G-areas for the diagrams $\Delta_1, \dots, \Delta_m$. Then the family $\mathbf{S} = \mathbf{S}_1 \cup \dots \cup \mathbf{S}_m$ 'covers' the entire Δ according to the above definition. This implies the required inequality for G-areas, □

6.3.2 Combs of a potential counterexample

We want to show that for some constants N_1, N_2 the G-area of any reduced diagram Δ over M with perimeter n does not exceed $N_2 n^2 + N_1 \mu(\Delta)$. (Using the quadratic upper bound for $\mu(\Delta)$ from Lemma 6.7 (a), one then deduces that the G-area is bounded by $N' n^2$ for some constant N'.) Roughly speaking, we are doing the following. We use induction on the perimeter of the diagram. First we remove rim θ-bands (those with one side and both ends on the boundary of the diagram) with short bases. This operation decreases the perimeter and preserves the sign of $N_2 n^2 + N_1 \mu(\Delta) - \mathrm{Area}_G(\Delta)$, so we can assume that the diagram does not have such bands. Then we use Lemma 6.6 and find a tight comb inside the diagram with a handle $\mathcal{C}$. We also find a long enough q-band $\mathcal{C}'$ that is close to $\mathcal{C}$. We use a surgery which amounts to removing a part of the diagram between $\mathcal{C}'$ and $\mathcal{C}$ and then gluing the two remaining parts of Δ together. The main difficulty is to show that, as a result of this surgery, the perimeter decreases and the measure and the mixture change in such a way that the expression $N_2 n^2 + N_1 \mu(\Delta) - \mathrm{Area}_G(\Delta)$ does not change its sign. In the proof, we need to consider several cases depending on the shape of the subdiagram between $\mathcal{C}'$ and $\mathcal{C}$. Note that neither $N_2 n^2$ nor $N_1 \mu(\Delta)$

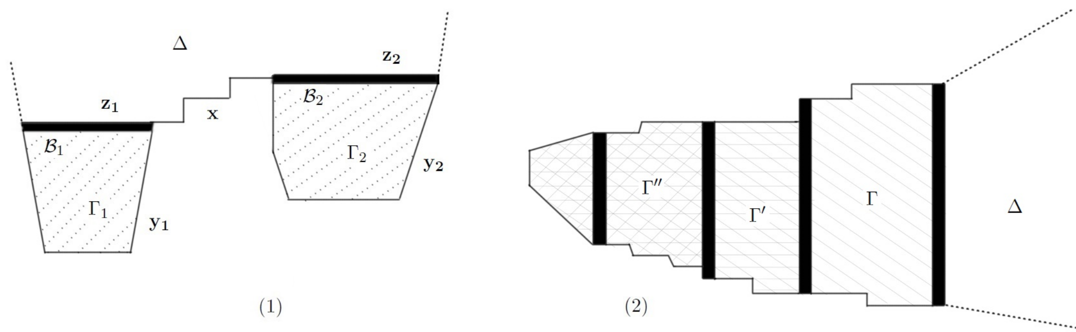


Figure 6: Lemma 6.10

nor $\text{Area}_G(\Delta)$ alone behave in the appropriate way as a result of the surgery, but the expression $N_2 n^2 + N_1\mu(\Delta) - \text{Area}_G(\Delta)$ behaves as needed.

So, we want to prove that the G-area of a reduced diagram Δ over M does not exceed $N_2 n^2 + N_1\mu(\Delta)$, where $n = |\partial\Delta|$. Arguing by contradiction in the remaining part of this section, we consider a **counter-example** Δ with minimal perimeter n. Of course, its G-area is positive, and, by Lemma 5.6, we have at least 2 θ-edges on the boundary $\partial\Delta$, and so $n \geq 2$.

Lemma 6.10. *(1) The diagram Δ has no two disjoint subcombs Γ_1 and Γ_2 of basic widths at most K with handles $\mathcal{B}_1$ and $\mathcal{B}_2$ such that some ends of these handles are connected by a subpath $\mathbf{x}$ of the boundary path of Δ with $|\mathbf{x}|_q \leq N$.*

(2) The boundary of every subcomb Γ with basic width $s \leq K$ has $2s$ q-edges.

Proof. We will prove the statements (1) and (2) using simultaneous induction on $A = \text{Area}(\Gamma_1) + \text{Area}(\Gamma_2)$ (resp., on $A = \text{Area}(\Gamma)$). Arguing by contradiction, we consider a counter-example with minimal A.

(1) Since the area of Γ_i $(i = 1, 2)$ is less than A, we may use Statement (2), and so we have at most $2K$ q edges in $\partial\Gamma_i$.

Let h_1 and h_2 be the lengths of the handles $\mathcal{B}_1$ and $\mathcal{B}_2$ of Γ_1 and Γ_2, resp. Without loss of generality, we assume that $h_1 \leq h_2$. Denote by $\mathbf{y}_i\mathbf{z}_i$ the boundaries of Γ_i $(i = 1, 2)$, where $\mathbf{z}_i$ is the part of $\partial\Delta$ and $\mathbf{y}_i$ is the side of the handle of Γ_i (so $\mathbf{y}_1\mathbf{x}\mathbf{y}_2$ is the part of the boundary path of Δ, see fig. 6 (1)). Then each of the θ-edges $\mathbf{e}$ of $\mathbf{y}_1$ is separated in $\partial\Delta$ from every θ-edge $\mathbf{f}$ of $\mathbf{y}_2$ by less than $4K + N < J$ q-edges. Hence every such pair $(\mathbf{e}, \mathbf{f})$ (or the pair of white beads on these edges) makes a contribution to $\mu(\Delta)$.

Let Δ' be the diagram obtained by deleting the subdiagram Γ_1 from Δ. When passing from $\partial\Delta$ to $\partial\Delta'$, one replaces the θ-edges from $\mathbf{z}_1$ by the θ-edge of $\mathbf{y}_1$ belonging to the same maximal θ-band. The same is true for white beads.

But each of the h_1h_2 pairs in the corresponding set P' of white beads is separated in $\partial\Delta'$ by less number of black beads than the pair defined by Δ. Indeed, since the handle of Γ_1 is removed when one replaces $\partial\Delta$ by $\partial\Delta'$, two black bead at the ends of this handle are removed, and therefore

$$\mu(\Delta) - \mu(\Delta') \geq h_1h_2 \tag{6.12}$$

by Lemma 6.7 (d).

Let α be the number of a-edges in $\partial\Gamma_1$. It follows from Lemma 6.5 that the area, and so the G-area of Γ_1, does not exceed $C_1(h_1)^2 + 2\alpha h_1$, where $C_1 = c_0 K$.

Remark 6.11. The constants C_1, C_2, C_{12}, C_3 are not included in the list (2.3) since their values chosen here make sense only in the present subsection.

Since the boundary of Δ' has at least two q-edges fewer than Δ and $|\mathbf{z}_1| = h_1 \le |\mathbf{y}_1|$, we have $|\partial\Delta'| \le |\partial\Delta| - 2$. Moreover, we have from Lemma 6.1 (a) and Lemma 5.6 that

$$|\partial\Delta| - |\partial\Delta'| \ge \gamma = \max(2, \delta(\alpha - 2h_1)) \tag{6.13}$$

because the top/the bottom of $\mathcal{B}_1$ has at most h_1 a-edges.

This inequality, inequality (6.12), and the inductive assumption related to Δ', imply that the G-area of Δ' is not greater than

$$N_2(n-\gamma)^2 + N_1\mu(\Delta) - N_1 h_1 h_2$$

Adding the G-area of Γ_1 we see that the G-area of Δ does not exceed

$$N_2 n^2 - N_2\gamma n + N_1\mu(\Delta) - N_1 h_1 h_2 + C_1 h_1^2 + 2\alpha h_1.$$

Since $h_1 \le h_2$ and $\phi(n) \ge 1$, this will contradict the choice of the counter-example Δ when we prove that

$$-N_2\gamma n - N_1 h_1^2 + C_1 h_1^2 + 2\alpha h_1 < 0 \tag{6.14}$$

If $\alpha \le 4h_1$, then inequality (6.14) follows from the inequalities $\gamma \ge 2$ and

$$N_1 \ge C_1 + 8 \tag{6.15}$$

Assume that $\alpha > 4h_1$. Then by (6.13), we have $\gamma \ge \frac{1}{2}\delta\alpha$ and $N_2\gamma n > 2\alpha h_1$ since $n \ge 2h_1$ by Lemma 5.6, and

$$N_2 > 2\delta^{-1}. \tag{6.16}$$

Since $N_1 h_1^2 > C_1 h_1^2$ by (6.15), the inequality (6.14) follows.

(2) If there are at least two derivative subcombs of Γ, then one can find two of them satisfying the assumptions of Statement (1), and $\mathrm{Area}(\Gamma_1) + \mathrm{Area}(\Gamma_2) < \mathrm{Area}(\Gamma) = A$, a contradiction. Therefore there is a most one derivative subcomb Γ' in Γ (fig 6 (2)). In turn, Γ' has at most one derivative subcomb Γ'', and so one. It follows that there are no maximal q-bands in Γ except for the handles of $\Gamma', \Gamma'', \dots,$ Since the basic width of Γ is s, we have s maximal q-bands in Γ, and the lemma is proved. □

Lemma 6.12. *There are no pair of subcombs Γ and Γ' in Δ with handles $\mathcal{X}$ and $\mathcal{X}'$ of length ℓ and ℓ' such that Γ' is a subcomb of Γ, the basic width of Γ does not exceed K_0 and $\ell' \le \ell/2$.*

Proof. Proving by contradiction, one can choose Γ' so that ℓ' is minimal for all subcombs in Γ and so Γ' has no proper subcombs, i.e. its basic width is 1 (fig. 7). It follows from Lemma 6.5 that for $\alpha = |\Gamma'|_a$, we have

$$\mathrm{Area}_G(\Gamma') \le \mathrm{Area}(\Gamma') \le c_0(l')^2 + 2\alpha l' \tag{6.17}$$

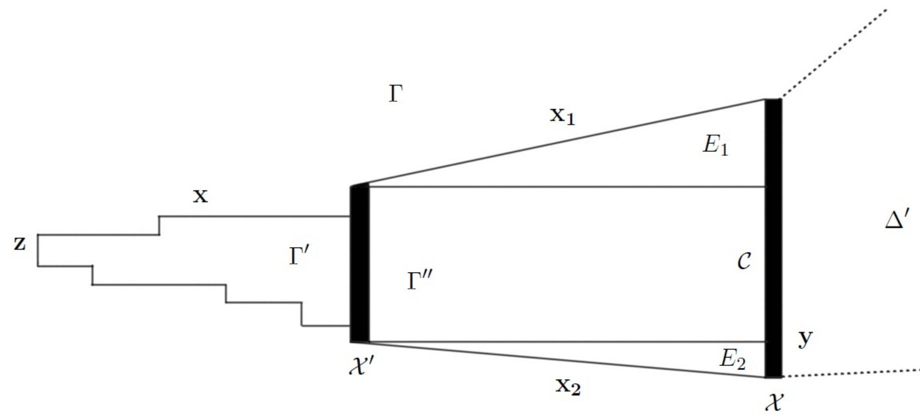


Figure 7: Lemma 6.12

Let Δ' be the diagram obtained after removing the subdiagram Γ' from Δ. The following inequality is the analog of (6.13) (where h_1 is replaced by ℓ')

$$|\partial\Delta| - |\partial\Delta'| \geq \gamma = \max(2, \delta(\alpha - 2l')) \tag{6.18}$$

The q-band $\mathcal{X}$ contains a subband $\mathcal{C}$ of length l'. Moreover one can choose $\mathcal{C}$ so that all maximal θ-bands of Γ crossing the handle $\mathcal{X}'$ of Γ', start from $\mathcal{C}$. These θ-bands form a comb Γ'' contained in Γ, and in turn, Γ'' contains Γ'. The two parts of the compliment $\mathcal{X}\backslash\mathcal{C}$ are the handles of two subcombs E_1 and E_2 formed by maximal θ-bands of Γ, which do not cross $\mathcal{X}'$. Let the length of these two handles be ℓ_1 and ℓ_2, respectively, and so we have $l_1 + l_2 = l - l' > l'$. (E_1 or E_2 can be empty; then l_1 or l_2 equals 0.)

It will be convenient to assume that Γ is drawn from the left of the vertical handle $\mathcal{X}$. Denote by $\mathbf{yz}$ the boundary path of of Γ, where $\mathbf{y}$ is the right side of the band $\mathcal{X}$. Thus, there are l_1 (resp., l_2) θ-edges on the common subpath $\mathbf{x}_1$ (subpath $\mathbf{x}_2$) of $\mathbf{z}$ and ∂E_1 (and ∂E_2).

By Lemma 6.10 (2), the path $\mathbf{z}$ contains at most $2K_0$ q-edges, because the basic width of Γ is at most K_0.

Consider the factorization $\mathbf{z} = \mathbf{x}_2\mathbf{x}\mathbf{x}_1$, where $\mathbf{x}$ is a subpath of $\partial\Gamma'$. It follows that between every white bead on $\mathbf{x}_1$ (i.e. the middle point of the θ-edges on $\mathbf{x}_1$) and a white bead on $\mathbf{x}$ we have at most $2K_0$ black beads (i.e. the middle points of the q-edges of the path $\mathbf{x}$). Since J is greater than $2K_0$, every pair of white beads, where one bead belongs in $\mathbf{x}$ and another one belongs in $\mathbf{x}_1$ (or, similarly, in $\mathbf{x}_2$) contributes 1 to $\mu(\Delta)$. Let P denote the set of such pairs. By the definition of E_1 and E_2, we have $\mathrm{card}(P) = l'(l_1 + l_2) = l'(l - l') > (l')^2$.

When passing from $\partial\Delta$ to $\partial\Delta'$, one replaces the left-most θ-edges of every maximal θ-band from Γ' with the right-most θ-edges lying on the right side of $\mathcal{X}'$. The same is true for white beads. But each of the $l'(l - l')$ pairs in the corresponding set P' of white beads is separated in $\partial\Delta'$ by less number of black beads since the q-band $\mathcal{X}'$ is removed. Therefore every pair from P' gives less by 1 contribution to the mixture, as it follows from the definition of mixture. Hence $\mu(\Delta) - \mu(\Delta') \geq l'(l - l') \geq (l')^2$. This inequality, inequality (6.18), and the inductive assumption related to Δ', imply that the G-area of

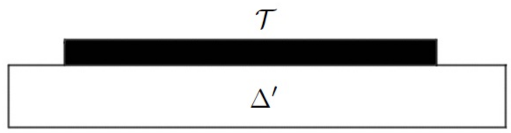


Figure 8: Rim θ-band

Δ' is not greater than

$$N_2(n-\gamma)^n + N_1\mu(\Delta) - N_1(\ell')^2$$

Adding the G-area of Γ' (6.17) we see that the G-area of Δ does not exceed

$$N_2n^2 + N_1\mu(\Delta) - N_2\gamma n - N_1(l')^2 + c_0(l')^2 + 2\alpha l'.$$

This will contradict the choice of the counter-example Δ when we prove that

$$-N_2\gamma n - N_1(l')^2 + c_0(l')^2 + 2\alpha l' < 0, \tag{6.19}$$

Consider two cases.

(a) Let $\alpha \leq 4l'$. Then inequality (6.19) follows from the inequalities $\gamma \geq 2$ and

$$N_1 \geq c_0 + 8. \tag{6.20}$$

(b) Assume that $\alpha > 4l'$. Then by (6.18) we have $\gamma \geq \frac{1}{2}\delta\alpha$ and $N_2\gamma n > 2\alpha l'$ since $n \geq 2l \geq 4l'$ by Lemma 5.6, and (6.16)

Since $N_1(l')^2 > c_0(l')^2$ by (6.20), the inequality (6.19) follows.

Thus, the lemma is proved by contradiction. □

6.3.3 Removing rim θ-bands

Recall that $K > 2K_0 = 4LN$.

Lemma 6.13. *Δ has no rim θ-band whose base has $s \leq K$ letters.*

Proof. Assume by contradiction that such a rim θ-band $\mathcal{T}$ exists, and $\mathbf{top}(\mathcal{T})$ belongs in $\partial(\Delta)$ (fig.8). When deleting $\mathcal{T}$, we obtain, by Lemma 6.2, a diagram Δ' with $|\partial\Delta'| \leq n-1$. Since $\mathbf{top}(\mathcal{T})$ lies on $\partial\Delta$, we have from the definition of the length , that the number of a-edges in $\mathbf{top}(\mathcal{T})$ is less than $\delta^{-1}(n-s)$. By Lemma 6.1, the length of $\mathcal{T}$ is at most $3s + \delta^{-1}(n-s) < \delta^{-1}n$. Thus, by applying the inductive hypothesis to Δ', we have that G-area of Δ is not greater than $N_2(n-1)^2 + N_1\mu(\Delta) + \delta^{-1}n$ because $\mu(\Delta') \leq \mu(\Delta)$ by Lemma 6.7 (b). But the first term of this sum does not exceed $N_2n^2 - N_2n$ and so the entire sum is bounded by $N_2n^2 + N_1\mu(\Delta)$ provided

$$N_2 \geq \delta^{-1}. \tag{6.21}$$

This contradicts to the choice of Δ, and the lemma is proved. □

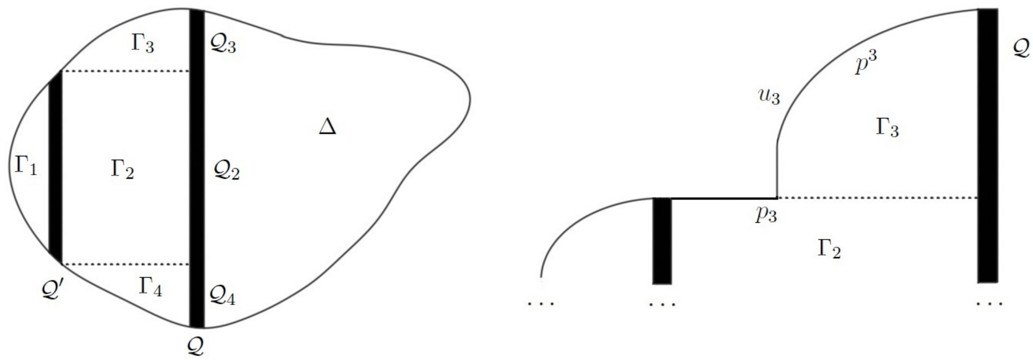


Figure 9: Lemma 6.14.

6.3.4 The quadratic estimate

The next lemma is one of the main ingredient in our proof.

Lemma 6.14. *The G-area of a reduced diagram Δ over* $\mathbf{M}$ *does not exceed $N_2n^2 + N_1\mu(\Delta)$, where $n = |\partial\Delta|$.*

Proof. We continue studying the hypothetical minimal counter-example Δ. By Lemma 6.13, now we can apply Lemma 6.6 (1). By that lemma, there exists a tight subcomb $\Gamma \subset \Delta$. Let $\mathcal{T}$ be a θ-band of Γ with a tight base.

The basic width of Γ is less than K_0 by Lemma 6.3. Since the base of Γ is tight, it is equal to $uxvx$ for some x, where the last occurrence of x corresponds to the handle $\mathcal{Q}$ of Γ, the word u does not contain x, and v has exactly $L-1$ occurrences of x. Let $\mathcal{Q}'$ be the maximal x-band of Γ crossing $\mathcal{T}$ at the cell corresponding to the first occurrence of x in $uxvx$ (fig. 9 (a)).

We consider the smallest subdiagram Γ' of Δ containing all the θ-bands of Γ crossing the x-band $\mathcal{Q}'$. It is a comb with handle $\mathcal{Q}_2 \subset \mathcal{Q}$. The comb Γ' is covered by a trapezium Γ_2 placed between $\mathcal{Q}'$ and $\mathcal{Q}$, and a comb Γ_1 with handle $\mathcal{Q}'$. The band $\mathcal{Q}'$ belongs to both Γ_1 and Γ_2. The remaining part of Γ is a disjoint union of two combs Γ_3 and Γ_4 whose handles $\mathcal{Q}_3$ and $\mathcal{Q}_4$ contain the cells of $\mathcal{Q}$ that do not belong to the trapezium Γ_2. The handle of Γ is the composition of handles $\mathcal{Q}_3$, $\mathcal{Q}_2$, $\mathcal{Q}_4$ of Γ_3, Γ' and Γ_4 in that order.

Let the lengths of $\mathcal{Q}_3$ and $\mathcal{Q}_4$ be l_3 and l_4, respectively. Let l' be the length of the handle of Γ'. Then by Lemma 6.12, we have

$$l' > l/2 \quad \text{and} \quad l = l' + l_3 + l_4 \tag{6.22}$$

For $i \in \{3,4\}$ and $\alpha_i = |\partial\Gamma_i|_a$, Lemma 6.5 gives inequalities

$$A_i \le C_1 l_i^2 + 2\alpha_i l_i, \tag{6.23}$$

where A_i is the G-area of Γ_i. (We take into account that G-area cannot exceed area.)

Let $\mathbf{p}_3, \mathbf{p}_4$ be the top and the bottom of the trapezium Γ_2. Here $\mathbf{p}_3^{-1}$ (resp. $\mathbf{p}_4^{-1}$) shares some initial edges with $\partial\Gamma_3$ (with $\partial\Gamma_4$), the rest of these paths belong to the boundary of Δ. We denote by d_3 the number of a-edges of $\mathbf{p}_3$ and by d_3' the number of the a edges of $\mathbf{p}_3$ which do not belong to Γ_3. Similarly, we introduce d_4 and d_4'.

Let A_2 be the G-area of Γ_2. Then by Lemma 5.14 and the definition of the G-area for big trapezia, we have

$$A_2 \leq C_2 l'(d_3 + d_4 + 1) \tag{6.24}$$

for some constant $C_2 < \delta^{-1}$, because the basic width of Γ_2 is less than K.

The first and the last base letters of the base of the trapezium Γ_2 differ by their superscripts. Therefore the last statement of Lemma 5.9 (1) implies that the q-band $\mathcal{Q}_2$ is a superscript shift of $\mathcal{Q}'$ since the superscripts of two q-letters of every (θ, q)-relation determine the superscripts of other letters. Therefore there is a superscript shift $\Gamma_1^{(+k)}$ of the comb Γ_1 such that the handle $(\mathcal{Q}')^{(+k)}$ of Γ_1^{+k} is a copy of $\mathcal{Q}_2$. This makes possible to identify $(\mathcal{Q}')^{(+k)}$ and Q_2 when one construct the diagram Δ_0.

This makes the following surgery possible. The diagram Δ is covered by two subdiagrams: Γ and another subdiagram Δ_1, having only the band in common. We construct a new auxiliary diagram by attaching of $\Gamma_1^{(+k)}$ to Δ_1 with identification of the of the band $\mathcal{Q}'$ of Γ_1 and the band $\mathcal{Q}_2$. We denote the constructed diagram by Δ_0.

Δ_0 is a reduced diagram because every pair of its cells having a common edge, has a copy either in Γ_1 or in $\Delta_1 \cup \mathcal{Q}$. Now we need the following claim.

Lemma 6.15. *The G-area A_0 of Δ_0 is at least the sum of the G-areas of Γ_1 and Δ_1 minus l'.*

Proof. Consider a minimal covering $\mathbf{S}$ of Δ_0 from the definition of G-area, and assume that there is a big trapezium $E \in \mathbf{S}$, such that neither $\Gamma_1^{(+k)}$ nor Δ_1 contains it. Then E has a base ywy, where $(yw)^{\pm 1}$ is a cyclic permutation of the L-th power of the standard base, and the first y-band of E is in $\Gamma_1^{(+k)}$, but it is not a subband of $\mathcal{Q}'$.

Since the history H of the big trapezium E is a subhistory of the history of Γ_2, and H uniquely determines the base starting with given letter by Lemma 4.4, we conlude that Γ_2 is a big trapezium itself, and therefore $(xv)^{\pm 1}$ is an L-th power of the standard base. Since the first y occurs in uxvx before the first x it follows that we have the $(L+1)-th$ occurrence of y before the last occurrence of x in the word $uxvx$. But this contradicts the definition of tight comb Γ.

Hence every big trapezium from $\mathbf{S}$ entirely belongs either in $\Gamma_1^{(+k)}$ or in Δ_1. Therefore one can obtain 'coverings' $\mathbf{S}'$ and $\mathbf{S}''$ of these two diagrams if (1) every Σ from $\mathbf{S}$ is assigned either to $\mathbf{S}'$ or to $\mathbf{S}''$ and then (2) one add at most l' single cells since the common band $\mathcal{Q}'$ in Δ_0 should be covered twice in disjoint diagrams $\Gamma_1^{(+k)}$ and Δ_1. These construction complete the proof of the lemma. □

Let us continue the proof of Lemma 6.14.

By Lemma 6.9, the G-area of Δ does not exceed the sum of G-areas of the five subdiagrams Γ_1, Γ_2, Γ_3, Γ_4 and Δ_1. But the direct estimate of each of these values is not efficient. Therefore we will use Lemma 6.15 to bound the G-area of the auxiliary diagram Δ_0 built of two pieces Γ_1 and Δ_1.

It follows from our constructions and lemmas 6.9, 6.15, that

$$\mathrm{Area}_G(\Delta) \leq A_2 + A_3 + A_4 + A_0 + l' \tag{6.25}$$

Let $\mathbf{p}^3$ be the segment of the boundary $\partial\Gamma_3$ that joins $\mathcal{Q}$ and Γ_2 along the boundary of Δ (fig. 9 (b)). It follows from the definition of d_3, d_3', l_3 and α_3, that the number of a-edges lying on $\mathbf{p}^3$ is at least $\alpha_3 - (d_3 - d_3') - l_3$.

Let $\mathbf{u}_3$ be the part of $\partial\Delta$ that contains $\mathbf{p}^3$ and connects $\mathcal{Q}$ with $\mathcal{Q}'$. It has l_3 θ-edges. Hence we have, by Lemma 6.1, that

$$|\mathbf{u}_3| \geq \max(l_3, l_3 + \delta(|p^3|_a - l_3)) \geq \max(l_3, l_3 + \delta(\alpha_3 - (d_3 - d_3') - 2l_3)).$$

Since $\mathbf{u}_3$ includes a subpath of length d_3' having no θ-edges, we also have by Lemma 6.1 (c) that $|\mathbf{u}_3| \geq l_3 + \delta(d_3' - 1)$.

One can similarly define $\mathbf{p}^4$ and $\mathbf{u}_4$ for Γ_4. When passing from $\partial\Delta$ to $\partial\Delta_0$ we replace the end edges of $\mathcal{Q}'$, $\mathbf{u}_3$ and $\mathbf{u}_4$ by two subpaths of $\partial\mathcal{Q}$ having lengths l_3 and l_4. Let $n_0 = |\partial\Delta_0|$. Then it follows from the previous paragraph that

$$n - n_0 \geq 2 + \delta(\max(0, d_3' - 1, \alpha_3 - (d_3 - d_3') - 2l_3) + \max(0, d_4' - 1, \alpha_4 - (d_4 - d_4') - 2l_4)) \quad (6.26)$$

In particular, $n_0 \leq n - 2$. By the inductive hypothesis,

$$A_0 \leq N_2 n_0^2 + N_1 \mu(\Delta_0) \quad (6.27)$$

We note that the mixture $\mu(\Delta_0)$ of Δ_0 is not greater than $\mu(\Delta) - l'(l - l')$. Indeed, by Lemma 6.12 (2), one can use the same trick as in Lemma 6.12 as follows. For every pair of white beads, where one bead corresponds to a θ-band of Γ_2 and another one to a θ-band of Γ_3 or Γ_4, the contribution of this pair to $\mu(\Delta_0)$ is less than the contribution to Δ. It remains to count the number of such pairs: $l'(l_3 + l_4) = l(l - l')$.

Therefore, by inequality (6.27), the G-area of Δ is not greater than

$$N_2 n^2 + N_1 \mu(\Delta) - N_2 n(n - n_0) - N_1 l'(l - l') + A_2 + A_3 + A_4 + l' \quad (6.28)$$

In view of inequalities (6.24), (6.23) and (6.25), to obtain the desired contradiction, it suffices to prove that

$$N_2 n(n - n_0) + N_1 l'(l - l') \geq C_{12} l'(d_3 + d_4 + 1) + C_{12}(l_3^2 + l_4^2) + 2\alpha_3 l_3 + 2\alpha_4 l_4 + l' \quad (6.29)$$

where $C_{12} = \max(C_1, C_2)$.

Since $l = l' + l_3 + l_4$, it suffices to prove that

$$N_2 n(n - n_0) + N_1 l'(l - l') \geq C_3 l'(d_3 + d_4 + 1) + C_3(l_3^2 + l_4^2) + 2\alpha_3 l_3 + 2\alpha_4 l_4 \quad (6.30)$$

where $C_3 = C_{12} + 1$.

Note that we can assume that

$$C_3 >> 1. \quad (6.31)$$

First we can choose N_1 big enough so that $N_1 l'(l - l')/3 \geq C_3(l_3 + l_4)^2 \geq C_3(l_3^2 + l_4^2)$. Indeed, by (6.22), we obtain $\frac{N_1}{3} l'(l - l') \geq \frac{N_1}{3}(l_3 + l_4)(l_3 + l_4)$, so it is enough to assume that

$$N_1 > 3C_3. \quad (6.32)$$

We also have that

$$\frac{N_2}{2} n(n - n_0) \geq C_3 l' \quad (6.33)$$

because $n - n_0 \geq 2$, $n \geq 2l'$ and $N_2 \geq C_3$ by (6.32).

It remains to prove that

$$\frac{N_2}{2} n(n - n_0) + \frac{2N_1}{3} l'(l - l') > C_3 l'(d_3 + d_4) + 2\alpha_3 l_3 + 2\alpha_4 l_4. \quad (6.34)$$

We assume without loss of generality that $\alpha_3 \geq \alpha_4$, and consider two cases.

(a) Suppose $\alpha_3 \leq 2C_3(l-l')$.
Since $d_i \leq \alpha_i + d_i'$ for $i = 3, 4$, by inequality (6.26), we have

$$d_3+d_4 \leq \alpha_3+\alpha_4+d_3'+d_4' < 4C_3(l-l')+\delta^{-1}(n-n_0)+2-2\delta^{-1} < 4C_3(l-l')+\delta^{-1}(n-n_0).$$

Therefore

$$\frac{N_1}{3}l'(l-l') + \frac{N_2}{2}n(n-n_0) \geq 4C_3^2 l'(l-l') + C_3\delta^{-1}(n-n_0)l' > C_3 l'(d_3+d_4) \tag{6.35}$$

since we can assume that

$$N_1 > 12C_3^2, \qquad N_2/2 > C_3\delta^{-1}. \tag{6.36}$$

We also have by (6.22):

$$\frac{N_1}{3}l'(l-l') \geq \frac{N_1}{3}(l_3+l_4)(l_3+l_4) \geq \frac{N_1}{3}\frac{\alpha_3+\alpha_4}{4C_3}(l_3+l_4) > 2\alpha_3 l_3 + 2\alpha_4 l_4 \tag{6.37}$$

since we can assume that

$$N_1 > 24C_3. \tag{6.38}$$

The sum of inequalities (6.35) and (6.37) gives us the desired inequality (6.34).

(b) Assume now that $\alpha_3 > 2C_3(l-l')$. Then, applying Lemma 6.5 to the comb Γ_3, we obtain

$$d_3 - d_3' < \frac{1}{2}\alpha_3 + K_0 l_3 \leq \frac{5}{6}\alpha_3 \tag{6.39}$$

since $l_3 \leq l - l' < \frac{\alpha_3}{2C_3}$ and

$$C_3 > 3K_0. \tag{6.40}$$

We also have $d_4 - d_4' < \frac{1}{2}\alpha_4 + K_0 l_4 \leq \frac{5}{6}\alpha_3$. These two inequalities and inequality (6.26) lead to

$$d_3 + d_4 \leq \frac{5}{3}\alpha_3 + \delta^{-1}(n-n_0) \tag{6.41}$$

It follows from (6.39) that

$$\alpha_3 - (d_3 - d_3') - 2l_3 \geq \frac{1}{6}\alpha_3 - \frac{2}{2C_3}\alpha_3 \geq \frac{1}{7}\alpha_3,$$

since $l_3 \leq l - l' < \frac{\alpha_3}{2C_3}$ and $C_3 > 42$ by (6.31). Therefore, by (6.26),

$$n - n_0 \geq \frac{1}{7}\delta\alpha_3. \tag{6.42}$$

Thus, by (6.41),

$$d_3 + d_4 < 13\delta^{-1}(n-n_0). \tag{6.43}$$

Since $2l' < n$ and $n - n_0 \geq 2$, inequality (6.43) implies

$$\frac{N_2}{3}n(n-n_0) > C_3 l'(d_3+d_4) \tag{6.44}$$

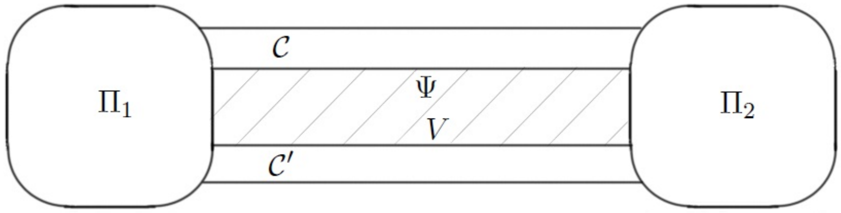


Figure 10: Cancellation of two hubs

because we can assume that

$$N_2 >> C_3\delta^{-1} \tag{6.45}$$

($N_2 > 21C_3\delta^{-1}$ is enough).

Inequalities (6.42), (6.45), $\alpha_3 \geq \alpha_4$, and $4(l_3 + l_4) \leq n$ give us

$$\frac{N_2}{6}n(n-n_0) \geq \frac{7}{2}C_3\delta^{-1}(n-n_0)n \geq 2\alpha_3(l_3+l_4) \geq 2\alpha_3 l_3 + 2\alpha_4 l_4 \tag{6.46}$$

The inequality (6.34) follows now from inequalities (6.44), and (6.46). □

7 Minimal diagrams over G

7.1 Diagrams with hubs

Given a reduced diagram Δ over the group G, one can construct a planar graph whose vertices are the hubs of this diagram plus one improper vertex outside Δ, and the edges are maximal $\check{t}$-bands of Δ.

7.1.1 Eliminating pairs of hubs connected by two $\check{t}$-bands

Let us consider two hubs Π_1 and Π_2 in a reduced diagram, connected by two neighbor $\check{t}$-bands $\mathcal{C}_i$ and $\mathcal{C}_{i+1}$, where there are no other hubs between these $\check{t}$-bands. These bands, together with parts of $\partial\Pi_1$ and $\partial\Pi_2$, bound either a subdiagram having no cells, or a trapezium Ψ of height ≥ 1 (fig. 10).

The former case is impossible. Indeed, in this case the hubs have to correspond the same hub relation since the relations (5.9) have no common letters. Hence the diagram is not reduced since a cyclic shift of a hub relation starting with a fixed copy of the letter $\check{t}$ is unique.

We want to show that the latter case is not possible either if the diagram Δ is chosen with minimal number of hubs among the diagrams with the same boundary label.

Indeed, by Lemma 5.9 (1), the $\check{t}$-band $\mathcal{C}_{i+1}$ is a k-shift of $\mathcal{C}_i$ In fact, $k = \pm 1$ since the superscripts of the letters in W_{st}^L change by one after every $\check{t}$-letter. One may assume that $k = 1$. So if we construct a 1-shift Ψ_2 of $\Psi_1 = \Psi$, then the first maximal $\check{t}$-band of Ψ_2 is a copy of $\mathcal{C}_{i+1}$ (the second $\check{t}$-band in Ψ_1). Similarly one can construct $\Psi_3 = \Psi_2^{(+1)} = \Psi_1^{(+2)}, \dots, \Psi_L = \Psi_1^{(+L)}$. Let us separately construct an auxiliary diagram Δ_1 consequently attaching the bottoms of $\Psi_1, \Psi_2, \dots, \Psi_L$ to Π_1 and identifying the second $\check{t}$-band of Ψ_i with the first $\check{t}$-band of Ψ_{i+1} (indices modulo L). This is possible since the

L-shift of any diagram is equal to itself. Now we can attach Π_2 to the tops of Ψ_i-s in Δ_1 and obtain a spherical diagram Δ_2. The diagram Δ_2 contains a copy of the subdiagram Γ of Δ formed by Π_1, Π_2 and Ψ. Hence the boundary label of Γ is equal to the boundary label of the compliment Γ' of (the copy of) the subdiagram Γ in Δ_2. Thus, one can replace Γ with Γ' in Δ decreasing the number of hubs, contrary to the choice of Δ.

7.1.2 Disks

Definition 7.1. A permissible word V is called a *disk word* if $V^{\emptyset} \equiv W^L$ for some accessible word W. The cyclic permutations of W and W^{-1} are also disk words by definition.

Lemma 7.2. *Every disk word V is equal to* 1 *in the group G.*

Proof. Assume there is an eligible computation $W_{st} \to \dots \to W$, where $V^{\emptyset} \equiv W^L$. Then the computation $W_{st}^L \to \dots \to W^L$ with the same history is eligible too. By Lemma 5.12 (2), one can construct a trapezium Δ with bottom label $W_{st}^{(1)} \dots W_{st}^{(L)}$ and top label V' such that $(V')^{\emptyset} \equiv V^{\emptyset}$, and so V' is a cyclic shift of the word V. The two sides of Δ have equal labels since the L-shift preserves superscripts. So one can identify these sides and attach the obtained annulus to the hub cell labeled by $W_{st}^{(1)} \dots W_{st}^{(L)}$. Since V' is the boundary label of the obtained disk diagram, we have $V' = 1$ in G, and so $V = 1$, as required. If there there is an eligible computation $W \to \dots \to W_{ac}$, then the proof is similar with bottom label of Δ equal to W_{ac}^L. □

Remark 7.3. In fact, for the disk word W, we have built a van Kampen diagram using one hub and L trapezia corresponding to an accessible computation for W.

We will increase the set of relations of G by adding the (infinite) set of *disk relation* V for every disk word V. So we will consider diagrams with *disks*, where every disk cell (or just *disk*) is labeled by such a word V. (In particular, a hub is a disk.)

Again, if two disks are connected by two $\check{t}$-bands and there are no other disks between these $\check{t}$-bands, then one can reduce the number of disks in the diagram. For this aid, it suffices to apply the trick exploited above for a pair of hubs.

Definition 7.4. We will call a reduced diagram Δ *minimal* if

(1) the number of disks is minimal for the diagrams with the same boundary label as Δ and

(2) Δ has minimal number of (θ, t)-cells among the diagrams with the same boundary label and with minimal number of disks.

Clearly, a subdiagram of a minimal diagram is minimal itself.

Thus, no two disks of a minimal diagram are connected by two $\check{t}$-bands, such that the subdiagram bounded by them contains no other disks. This property makes the disk graph of a reduced diagram hyperbolic, in a sense, if the degree L of every proper vertex (=disk) is high ($L \gg 1$). Below we give a more precise formulation (proved for diagrams with such a disk graph, in particular, in [28], Lemma 11.4 and in [16], Lemma 3.2).

Lemma 7.5. *If a a minimal diagram contains a least one disk, then there is a disk Π in Δ such that $L-3$ consecutive maximal $\check{t}$-bands $\mathcal{B}_1, \dots \mathcal{B}_{L-3}$ start on $\partial\Pi$, end on the boundary $\partial\Delta$, and for any $i \in [1, L-4]$, there are no disks in the subdiagram Γ_i bounded by $\mathcal{B}_i$, $\mathcal{B}_{i+1}$, $\partial\Pi$, and $\partial\Delta$ (fig. 11).*

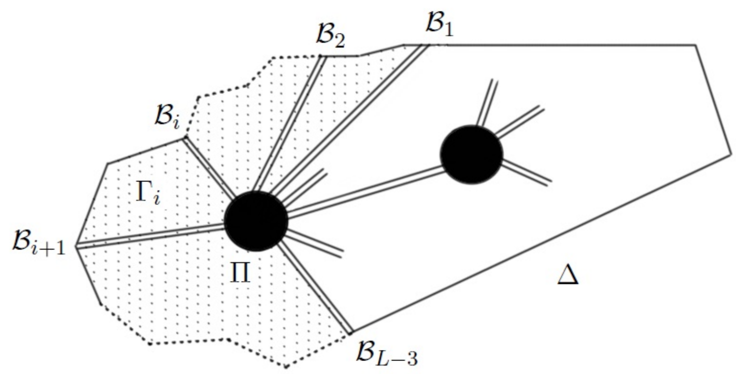


Figure 11: Lemma 7.5

A maximal q-band starting on a disk of a diagram is called a *spoke*.

Lemma 7.5 implies by induction on the number of hubs:

Lemma 7.6. *(see [17], Lemma 5.19) If a reduced diagram Δ has $m \geq 1$ hubs, then the number of $\check{t}$-spokes of Δ ending on the boundary $\partial\Delta$, and therefore the number of $\check{t}$-edges in the boundary path of Δ, is greater than $mLN/2$.*

7.1.3 The band moving transformation

Recall the following band moving transformation for diagrams with disks, exploited earlier in [16], [28]. Assume there is a disk Π and a θ-band $\mathcal{T}$ subsequently crossing some spokes $\mathcal{B}_1, \dots, \mathcal{B}_k$ which start (say, counter-clockwise) from Π. Assume that $k \geq 2$ and there are no other cells between Π and the bottom of $\mathcal{T}$, and so there is a subdiagram Γ formed by Π and $\mathcal{T}$.

We describe the (see, e.g., [28]) as follows. By Lemma 5.9 (1), we have a word $V \equiv (tW)(tW)^{(+1)} \dots (tW)^{(+(k-2))} t^{(+(k-1))}$ written on the top of the subband $\mathcal{T}'$ of $\mathcal{T}$, that starts on $\mathcal{B}_1$ and ends on $\mathcal{B}_k$. The bottom $\mathbf{q}_2$ of $\mathcal{T}'$ is the subpath of the boundary path $\mathbf{q}_2\mathbf{q}_3$ of Π (fig. 12), its label is a part of a disk word, and so is V by Lemma 5.9.

Therefore one can construct a new disk $\overline{\Pi}$ with boundary label

$$(tW)(tW)^{(+1)} \dots (tW)^{(+(L-1))}$$

and boundary $\mathbf{s}_1\mathbf{s}_2$, where $\mathrm{Lab}(\mathbf{s}_1) \equiv V$. Also one construct an auxiliary band $\mathcal{T}''$ with top label $(W^{-1}t^{-1})^{(+(L-1))} \dots (W^{-1}t^{-1})^{(+k)}(W^{-1})^{(+(k-1))}$, and attach it to $\mathbf{s}_2^{-1}$, which has the same label. Finally we replace the subband $\mathcal{T}'$ by $\mathcal{T}''$ (and make cancellations in the new θ-band $\overline{\mathcal{T}}$ if any appear). The new diagram $\overline{\Gamma}$ formed by $\overline{\Pi}$ and $\overline{\mathcal{T}}$ has the same boundary label as Γ.

Remark 7.7. After the band moving, the first (θ, t)-cells of $\check{t}$-spokes $\mathcal{B}_1, \dots, \mathcal{B}_k$ are removed and the total number of common (θ, t)-cells of the new spokes $\overline{\mathcal{B}}_1, \dots, \overline{\mathcal{B}}_k$ of $\overline{\Pi}$ and $\overline{\mathcal{T}}$ is less than the number of common (θ, t)-cells of $\mathcal{B}_1, \dots, \mathcal{B}_k$ and $\mathcal{T}$ at least by k. In particular, if $k > L - k$, then the number of (θ, t)-cells in $\overline{\Gamma}$ is less than the number of (θ, t)-cells in Γ. This observation implies

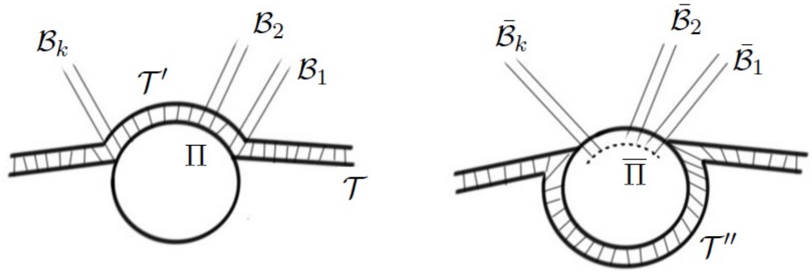


Figure 12: The band moving transformation of a θ-band and a disk

Lemma 7.8. *Let Δ be a minimal diagram.*

(1) Assume that a θ-band $\mathcal{T}_0$ crosses k $\check{t}$-spokes $\mathcal{B}_1, \dots, \mathcal{B}_k$ starting on a disk Π, and there are no disks in the subdiagram Δ_0, bounded by these spokes, by $\mathcal{T}_0$ and by Π. Then $k \le L/2$.

(2) Assume that there are two disjoint θ-bands $\mathcal{T}$ and $\mathcal{S}$ whose bottoms are parts of the boundary of a disk Π and these bands correspond to the same rule θ (if their histories are read towards the disk) and $\theta \ne \theta(23)$. Suppose $\mathcal{T}$ crosses $k \ge 2$ $\check{t}$-spokes starting on $\partial\Pi$ and $\mathcal{S}$ crosses $\ell \ge 2$ $\check{t}$-spokes starting on $\partial\Pi$. Then $k + \ell \le L/2$.

(3) Δ contains no θ-annuli.

(4) A θ-band cannot cross a maximal q-band (in particular, a spoke) twice.

Proof. (1) Since every cell, except for disks, belongs to a maximal θ-band, it follows from Lemma 5.6 that there is a θ-band $\mathcal{T}$ such that $\mathcal{T}$ crosses all $\mathcal{B}_1, \dots, \mathcal{B}_k$ and Δ_0 has no cells between $\mathcal{T}$ and Π. If $k > L/2$, then by Remark 7.7, the band moving $\mathcal{T}$ around Π would decrease the number of (θ, t)-cells in Δ, a contradiction, since Δ is a minimal diagram.

(2) As above, let us move the band $\mathcal{T}$ aroud Π. This operation removes k (θ, t)-cells but add $L - k$ new (θ, t)-cells in $\overline{\mathcal{T}}$. However ℓ (θ, t)-cells of $\mathcal{S}$ and ℓ (θ, t)-cells of $\overline{\mathcal{T}}$ will form mirror pairs, because for $\theta \ne \theta(23)$, the boundary label of a (θ, q)-cell π, considered as a θ-band, is uniquely determined by the history θ and the label of the top q-edge of π. So after cancellations one will have at most $L - k - 2\ell$ new (θ, t)-cells. This number is less than k if $k + \ell > L/2$ contrary to the minimality of the original diagram. Therefore $k + \ell \le L/2$.

(3) Proving by contradiction, consider the subdiagram Δ' bounded by a θ-annulus. It has to contain disks by Lemma 5.6. Hence it must contain spokes $\mathcal{B}_1, \dots, \mathcal{B}_{L-3}$ introduced in Lemma 7.5. But this contradits to item (1) of the lemma since $L - 3 > L/2$.

(4) The argument of item (3) works if there is a subdiagram Δ' of Δ bounded by an q-band and a θ-band. □

The band moving will be used for removing disks from quasi-trapezia.

7.1.4 Quasi-trapezia

Definition 7.9. A *quasi-trapezium* is the same as trapezium (Definition 5.10), but may contain disks. (So a quasi-trapezium without disks is a trapezium.)

Lemma 7.10. *Let a minimal diagram Γ be a quasi-trapezium with standard factorization of the boundary as $\mathbf{p}_1^{-1}\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2^{-1}$. Then there is a diagram Γ' such that*

(1) the boundary of Γ' is $(\mathbf{p}_1')^{-1}\mathbf{q}_1'\mathbf{p}_2'(\mathbf{q}_2')^{-1}$, where $\mathrm{Lab}(\mathbf{p'}_j)\equiv\mathrm{Lab}(\mathbf{p}_j)$ and $\mathrm{Lab}(\mathbf{q'}_j)\equiv\mathrm{Lab}(\mathbf{q}_j)$ for $j=1,2$;

(2) the numbers of hubs and (θ,q)-cells in Γ' are the same as in Γ;

(3) the vertices $(\mathbf{p'}_1)_-$ and $(\mathbf{p'}_2)_-$ (the vertices $(\mathbf{p}_1')_+$ and $(\mathbf{p'}_2)_+$) are connected by a simple path $\mathbf{s}_1$ (by $\mathbf{s}_2$, resp.) such that we have three subdiagrams $\Gamma_1,\Gamma_2,\Gamma_3$ of Γ', where Γ_2 is a trapezium with standard factorization of the boundary $(p_1')^{-1}\mathbf{s}_1\mathbf{p}_2'\mathbf{s}_2^{-1}$ and all cells of the subdiagrams Γ_1 and Γ_3 with boundaries $\mathbf{q_1's_1^{-1}}$ and $\mathbf{s_2(q_1')^{-1}}$ are disks;

(4) All maximal θ-bands of Γ and all maximal θ-bands of Γ_2 have the same number ot (θ,t)-cells (equal for Γ and Γ_2) .

Proof. Every maximal θ-band of Γ must connect an edge of $\mathbf{p_1}$ with an edge of $\mathbf{p}_2$; this follows from Lemma 7.8 (3). Hence we can enumerate these bands from bottom to top: $\mathcal{T}_1,\dots,\mathcal{T}_h$, where $h=|\mathbf{p}_1|=|\mathbf{p}_2|$.

If Γ has a disk, then by Lemma 7.5, there is a disk Π such that at least $L-3$ $\check t$-spokes of it end on $\mathbf{q}_1$ and $\mathbf{q}_2$, and there are no disks between the spokes ending on $\mathbf{q}_1$ (and on $\mathbf{q}_2$). By Lemma 7.8 (2), at least $L-3-L/2\ge 2$ of these spokes must end on $\mathbf{q}_1$ (resp., on $\mathbf{q}_2$).

If Π lies between $\mathcal{T}_j$ and $\mathcal{T}_{j+1}$, then the number of its $\check t$-spokes crossing $\mathcal{T}_j$ (crossing $\mathcal{T}_{j+1}$) is at least 2. So one can move each of the two θ-bands around Π. So we can move the disk toward $\mathbf{q_1}$ (or toward $\mathbf{q_2}$) until the disk is removed from the quasi-trapezium. (We use the fact that if k $\check t$-spokes $\mathcal{B}_1,\dots,\mathcal{B}_k$ of Π end on $\mathbf{q_1}$, then after band moving, we again have k $\check t$-spokes $\overline{\mathcal{B}}_1,\dots,\overline{\mathcal{B}}_k$ of $\overline{\Pi}$ ending on $\mathbf{q_1}$. - See the notation of Remark 7.7.)

No pair $\mathcal{T}_j$ and $\mathcal{T}_{j+1}$ corresponds to two mutual inverse letters $\theta\theta^{-1}$ of the history if $\theta\ne\theta(23)$. This follows from Lemma 5.12 (1) if there are no disks between these θ-bands. If there is a disk, then this is impossible too by Lemma 7.10 (2) since one could choose a disk Π as in the previous paragraph. So the projection of the label of $\mathbf{p}_1$ on the history is eligible.

Let us choose i such that the number m of (θ,t)-cells in $\mathcal{T}_i$ is minimal. It follows that Γ has at least hm (θ,t)-cells.

If the disk Π lies above $\mathcal{T}_i$, we will move it upwards using the band moving transformation. So after a number of iterations all such (modified) disks will be placed above the θ-band number h and form the subdiagram Γ_1. Similarly we can form Γ_3 moving other disks downwards.

In the resulting diagram Γ_2 lying between Γ_1 and Γ_3, every θ-band is reduced by the definition of band moving. The neighbor maximal θ-band of Γ_2 cannot be mirror copies of each other since the labels of $\mathbf{p}_1$ and $\mathbf{p}_1'$ are equal and $\mathrm{Lab}(\mathbf{p}_1)$ is a reduced word by Remark 5.5. It follows that the diagram Γ_2 (without disks) is a reduced diagram, and so it is a trapezium of height h.

The θ-band $\mathcal{T}_i$ did not participate in the series of band moving transformations above. Therefore it is a maximal θ-band of Γ_2. Hence the trapezium Γ_2 contains exactly mh (θ,t)-cells, which does not exceed the number of (θ,t)-cells in Γ. In fact these two numbers are equal since Γ is a minimal diagram. So every maximal θ-band of Γ and every maximal θ-band of Γ_2 has m (θ,t)-cells.

□

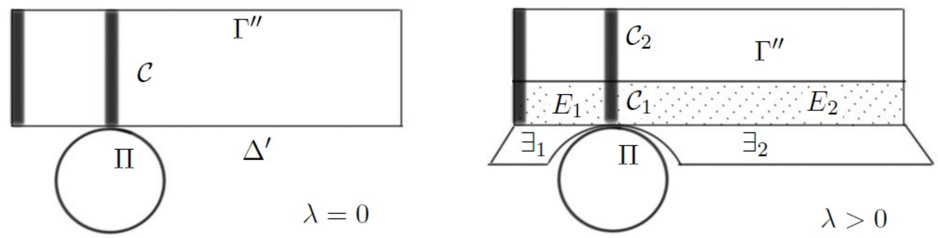


Figure 13: Lemma 7.12.

7.1.5 Shafts

We say that a history word H is *standard* if there is a standard trapezium with history H.

Definition 7.11. Suppose we have a disk Π with boundary label $(tW)^L$ and $\mathcal{B}$ be a t-spoke starting on Π. Suppose there is a subband $\mathcal{C}$ of $\mathcal{B}$, which also starts on Π and has a standard history H, for which the word tW is H-admissible. Then we call the $\check{t}$-band $\mathcal{C}$ a *shaft.*

For a constant $\lambda \in [0; 1/2)$ we also define a stronger concept of λ-shaft at Π as follows. A shaft $\mathcal{C}$ with history H is a λ-*shaft* if for every factorization of the history $H \equiv H_1H_2H_3$, where $||H_1|| + ||H_3|| < \lambda||H||$, the middle part H_2 is still a standard history. (So a shaft is a 0-shaft).

Lemma 7.12. *Let Π be a disk in a minimal diagram Δ and $\mathcal{C}$ be a λ-shaft at Π with history H. Then $\mathcal{C}$ has no factorizations $\mathcal{C} = \mathcal{C}_1\mathcal{C}_2\mathcal{C}_3$ such that*

(a) the sum of lengths of $\mathcal{C}_1$ and $\mathcal{C}_3$ do not exceed $\lambda||H||$ and

(b) Δ has a quasi-trapezium Γ such that top (or bottom) label of Γ has $L+1$ occurrences of $\check{t}$-letters and $\mathcal{C}_2$ starts on the bottom and ends on the top of Γ.

Proof. Proving by contradiction, we first replace Γ by a trapezium Γ' according to Lemma 7.10. The transpositions used for this goal do not affect neither Π nor $\mathcal{C}$ since $\mathcal{C}$ crosses all the maximal θ-bands of Γ. Also one can replace Γ' by a trapezium with shorter base and so we assume that the base of it starts and ends with letter t.

For the beginning, we assume that $\mathcal{C}$ is a shaft (i.e.,$\lambda = 0$). Then it follows from the definition of shaft and Lemma 4.4 that $\mathbf{b}ot(\Gamma')$ is labeled by a word Vt such that $V^{\emptyset} \equiv (tW)^L$, where the word tW has standard base. Now it follows from Remark 5.8 and Lemma 5.12that V is the boundary label of Π. One can remove the last maximal $\check{t}$-band from Γ' and obtain a subtrapezium Γ'' whose boundary label coincides with the label of $\partial\Pi$ (up to cyclic permutation), and $\partial\Gamma''$ shares a t-edge with $\partial\Pi$ (fig.13 with $\lambda = 0$). It follows that the subdiagram $\Delta' = \Pi \cup \Gamma''$ has boundary label freely equal to Lab($\mathbf{top}(\Gamma'')$). However Lab($\mathbf{top}(\Gamma'') \equiv V'$, where $V' = V \cdot H$ by Lemma 5.12, and so there is a disk Π' with boundary label V'. Therefore the subdiagram Δ' can be replaced by a single disk. So we decrease the number of (θ, t)-cells contrary to the minimality of Δ.

Now we consider the general case, where $\mathcal{C} = \mathcal{C}_1\mathcal{C}_2\mathcal{C}_3$. As above, we replace Γ by a trapezium Γ' and obtain a trapezium Γ'' after removing of one $\check{t}$-band in Γ'. To obtain

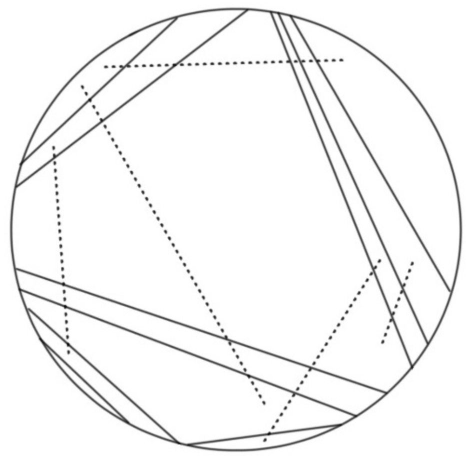


Figure 14: Design

a contradiction, it suffices to consider the diagram $\Delta' = \Pi \cup \mathcal{C}_1\mathcal{C}_2 \cup \Gamma''$ (forgetting of the complement of Δ' in Δ) and find another diagram Δ'' with one disk and fewer (θ, t)-cells such that $\mathrm{Lab}(\partial\Delta'') = \mathrm{Lab}(\partial\Delta')$ in the free group.

Since both histories H and H_2 (and so H_1H_2) are standard, one can enlarge Γ'' and construct a trapezium Γ''' with history H_1H_2. (The added parts E_1 and E_2 are dashed in figure 13 with $\lambda > 0$). Note that we add $< \lambda||H||L$ new (θ, t)-cells since every maximal θ-band of Γ''' has L such cells. As in case $\lambda = 0$, this trapezium Γ''' and the disk Π can be replaced by one disk Π'. However to obtain the boundary label equal to $\mathrm{Lab}(\partial\Delta')$, we should attach the mirror copies $\exists_1$ and $\exists_2$ of E_1 and E_2 to Π'. The obtained diagram Δ'' has at most $\lambda||H_1||L$ (θ, t)-cells, while Δ' has at least $||H_2||L \geq (1-\lambda)||H||$ (θ, t)-cells. Since $\lambda < 1 - \lambda$, we have the desired contradiction. □

7.1.6 Designs

As in [18], we are going to use *designs*.

Let $\mathcal{D}$ be the Euclidean unit disk and $\mathbf{T}$ be a finite set of disjoint *chords* (plain lines in fig. 14) and $\mathbf{Q}$ a finite set of disjoint simple curves in $\mathcal{D}$ (dotted lines in fig. 14). We assume that a curve is a non-oriented broken line, i.e., it is built from finitely many finite line segments. To distinguish the elements from $\mathbf{T}$ and $\mathbf{Q}$, we will say that the elements of $\mathbf{Q}$ are *arcs*.

We shall assume that the arcs belong to the open disk D^o, an arc may cross a chord transversally at most once, and the intersection point cannot coincide with one of the two ends of an arc.

Under these assumptions, we shall say that the pair $(\mathbf{T}, Q)$ is a *design*. The number of elements in $\mathbf{T}$ and $\mathbf{Q}$ are denoted by $\#\mathbf{T}$ and $\#\mathbf{Q}$.

By definition, the *length* $|C|$ of an arc C is the number of the chords crossing C. The term *subarc* will be used in the natural way. Oviously one has $|D| \leq |C|$ if D is a subarc of an arc C.

We say that an arc C_1 is *parallel* to an arc C_2 and write $C_1 \parallel C_2$ if every chord (from $\mathbf{T}$) crossing C_1 also crosses C_2. So the relation $\parallel$ is transitive (it is not necessarily symmetric). For example, the arc of length 2 is parallel to the arc of length 5 in fig. 14.

Definition 7.13. Given $\lambda \in (0;1)$ and an integer $n \geq 1$, the property $P(\lambda, n)$ of a design says that for any n different ars $C_1, \dots, C_n$, there exist no subarcs $D_1, \dots, D_n$, respectively, such that $|D_i| > (1-\lambda)|C_i|$ for every $i = 1, \dots, n$ and $D_1 \parallel D_2 \parallel \dots \parallel D_n$.

By definition, the length $\ell(\mathbf{Q})$ of the set of arcs $\mathbf{Q}$ is defined by the equality

$$\ell(\mathbf{Q}) = \sum_{C \in \mathbf{Q}} |C| \tag{7.47}$$

The number of chords will be denoted by $\#\mathbf{T}$. Here is the main statement about designs from [18].

Theorem 7.14 (Theorem 8.2 [18])**.** *There is a constant $c = c(\lambda, n)$ such that for any design $(\mathbf{T}, \mathbf{Q})$ with property $P(\lambda, n)$, we have*

$$\ell(\mathbf{Q}) \leq c(\#\mathbf{T}) \tag{7.48}$$

7.1.7 Designs and the σ_λ invariant

Let $\lambda \in [0, 1/2)$. For every $\check{t}$-spoke $\mathcal{B}$ of a minimal diagram Δ, we choose the λ-shaft of maximal length in it (if a λ-shaft exists). If $\mathcal{B}$ connects two disks Π_1 and Π_2, then there can be two maximal λ-shafts: at Π_1 and at Π_2. We denote by $\sigma_\lambda(\Delta)$ the sum of lengths of all λ-shafts in this family.

Lemma 7.15. *There is a constant $c = c(\lambda)$ such that $\sigma_\lambda(\Delta) \leq c|\partial\Delta|$ for every minimal diagram Δ over the group G.*

Proof. Let us associate the following design with Δ. We say that the median lines of the maximal θ-bands are the chords and the median lines of the maximal λ-shafts are the arcs. Here we use two disjoint median lines if two maximal λ-shafts share a $(\theta, \check{t})$-cell. By Lemma 7.8 (3), (4), we indeed obtain a design.

Observe that the length $|C|$ of an arc is the number of cells in the λ-shaft and $\#\mathbf{T} \leq |\partial\Delta|/2$ since every maximal θ-band has two θ-edges on $\partial\Delta$.

Thus, by Theorem 7.14, it suffices to show that the constructed design satisfies the condition $P(\lambda, n)$, where n does not depend on Δ.

Let $n = 2L + 1$. If the property $P(\lambda, n)$ does not hold, then we have n maximal λ-shafts $\mathcal{C}_1, \dots, \mathcal{C}_n$ and a subband $\mathcal{D}$ of $\mathcal{C}_1$, such that $|\mathcal{D}| > (1-\lambda)|\mathcal{C}_1|$, and every maximal θ-band crossing $\mathcal{D}$ must cross each of $\mathcal{C}_2, \dots, \mathcal{C}_n$. (Here $|\mathcal{B}|$ is the length of a $\check{t}$-band $\mathcal{B}$.) It follows that each of these θ-band crosses at least $L+1$ maximal $\check{t}$-bands. (See Lemma 7.8 (3,4). Here we take into account that the same $\check{t}$-spoke can generate two arcs in the design.) Hence using the λ-shaft $\mathcal{C}_1$ one can construct a quasi-trapezium of height $|\mathcal{D}|$, which contradicts Lemma 7.12. □

7.2 Upper bound for G-areas of diagrams over the group G.

7.2.1 The area of a disk is quadratic

By definition, the *G-area of a disk* Π is just the minimum of areas of diagrams over the presentation (5.6,5.9) of G having the same boundary label as Π.

Lemma 7.16. *There is a constant c_6 such that both area and the G-area of any disk does not exceed $c_6|\partial\Pi|^2$.*

Proof. By Remark 7.3, a disk with boundary label V can be built of one hub and L trapezia corresponding to a accessible computation $\mathcal{C}$ for W, where $W^L \equiv V^\emptyset$. By Lemma 4.9, the length of $\mathcal{C}$ can be bound by $c_2||W||$ and the length of every configuration of $\mathcal{C}$ does not exceed $c_1||W||$ Hence by Lemma 6.1, the area and the G-area of the disk is bounded by $c_6|\partial\Pi|^2$ since the constant c_6 can be chosen after c_1, c_2 and δ. □

By definition, the G-area of a minimal diagram Δ over G is the sum of G-areas of its disks plus the G-area of the compliment Γ. For the compliment, as in subsection 6.3, we consider a family **S** of big subtrapezia and single cells of Γ such that every cell of Γ belongs to a member Σ of this family, and if a cell Π belongs to different Σ_1 and Σ_2 from Σ, then both Σ_1 and Σ_2 are big subtrapezia of Γ with bases xv_1x, xv_2x, and Π is an (θ, x)-cell.) Hence the statement of Lemma 6.9 holds for minimal diagrams over G as well.

We want to prove that for big enough constants N_3 and N_4, Area$(\Delta) \le N_4(n + \sigma_\lambda(\Delta))^2 + N_3\mu(\Delta)$ for every minimal diagram Δ with perimeter n. For this goal, we will argue by contradiction in this section and study a **counter-example** Δ with minimal $n + \sigma_\lambda(\Delta)$.

7.2.2 Getting rid of rim bands with short base

Lemma 7.17. *The diagram Δ has no rim θ-bands with base of length at most K.*

Proof. The proof of Lemma 6.13 works for the minimal counter-example over G. It suffices to replace N_2 and N_1 by N_4 and N_3, replace n by $n+\sigma_\lambda(\Delta)$, and notice that the value of σ_λ does nor increase when passing from Δ to Δ'. □

7.2.3 The cloves

By Lemma 6.14, Δ has at least one disk. Using Lemma 7.5, we fix a disk Π in Δ such that $L-3$ consecutive maximal $\check{t}$-bands $\mathcal{B}_1, \dots \mathcal{B}_{L-3}$ start on $\partial\Delta$, end on the boundary $\partial\Pi$, and for any $i \in [1, L-4]$, there are no disks in the subdiagram bounded by $\mathcal{B}_i$, $\mathcal{B}_{i+1}$, $\partial\Pi$, and $\partial\Delta$. (See fig. 11.)

We denote by $\Psi = cl(\Pi, \mathcal{B}_1, \mathcal{B}_{L-3})$ the subdiagram without disks bounded by the spokes $\mathcal{B}_1$, $\mathcal{B}_{L-3}$ (and including them) and by subpaths of the boundaries of Δ and Π, and call this subdiagram a *clove*. Similarly one can define the cloves $\Psi_{ij} = cl(\Pi, \mathcal{B}_i, \mathcal{B}_j)$ if $1 \le i < j \le L-3$.

7.2.4 A clove cannot contain certain subcombs

Lemma 7.18. *The clove $\Psi = cl(\Pi, \mathcal{B}_1, \mathcal{B}_{L-3})$ has no subcombs of basic width at least K_0.*

Proof. Proving by contradiction, we may assume that there is a tight subcomb Γ by Lemma 6.6 (2). Then contradiction appears exactly as in the proofs of Lemmas 6.10–6.14. It suffices to replace N_2 and N_1 by N_4 and N_3, replace n by $n+\sigma_\lambda(\Delta)$, and notice that the value of σ_λ does nor increase when passing from Δ to Δ_0 since no $\check{t}$-band of Γ is a part of a spoke. □

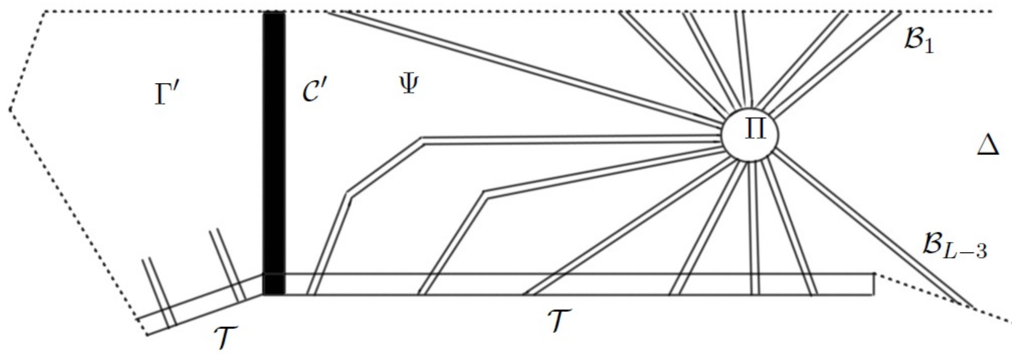


Figure 15: Lemma 7.20

Below we use the following analog of Lemma 6.10 (with identical proof):

Lemma 7.19. *(1) The counter-example Δ has no two disjoint subcombs Γ_1 and Γ_2 of basic widths at most K with handles $\mathcal{C}_1$ and $\mathcal{C}_2$ such that some ends of these handles are connected by a subpath $\mathbf{x}$ of the boundary path of Δ with $|\mathbf{x}|_q \le N$.*

(2) The boundary of every subcomb Γ with basic width $s \le K$ has $2s$ q-edges.

□

7.2.5 θ-bands in a clove

Lemma 7.20. *(1) Every maximal θ-band of Ψ crosses either $\mathcal{B}_1$ or $\mathcal{B}_{L-1}$. (2) There exists r, $L/2-3 \le r \le L/2$, such that the θ-bands of Ψ crossing $\mathcal{B}_{L-3}$ do not cross $\mathcal{B}_r$, and the θ-bands of Ψ crossing $\mathcal{B}_1$ do not cross $\mathcal{B}_{r+1}$;*

Proof. (1) If the claim were wrong, then one could find a rim θ-band $\mathcal{T}$ in Ψ, which crosses neither $\mathcal{B}_1$ nor $\mathcal{B}_{L-3}$. By Lemma 7.17, the basic width of $\mathcal{T}$ is greater than K. Since (1) a disk has LN spokes, (2) no q-band of Ψ intersects $\mathcal{T}$ twice by Lemma 5.6 (3), $\mathcal{T}$ has at least K q-cells, and (4) $K > 2K_0 + LN$, there exists a maximal q-band $\mathcal{C}'$ such that a subdiagram Γ' separated from Ψ by $\mathcal{C}'$ contains no edges of the spokes of Π and the part of $\mathcal{T}$ belonging to Γ' has at least K_0 q-cells (fig. 15).

If Γ' is not a comb, and so a maximal θ-band of it does not cross $\mathcal{C}'$, then Γ' must contain another rim band $\mathcal{T}'$ having at least K q-cells. This makes possible to find a subdiagram Γ'' of Γ' such that a part of $\mathcal{T}'$ is a rim band of Γ'' containing at least K_0 q-cells, and Γ'' does not contain $\mathcal{C}'$. Since Area$(\Gamma') >$ Area$(\Gamma'') > \dots$, such a procedure must stop. Hence, for some i, we obtain a subcomb $\Gamma^{(i)}$ of basic width $\ge K_0$, contrary to Lemma 7.18.

(2) Assume there is a maximal θ-band $\mathcal{T}$ of Ψ crossing the spoke $\mathcal{B}_1$. Then assume that $\mathcal{T}$ is the closest to the disk Π, i.e. the intersection of $\mathcal{T}$ and $\mathcal{B}_1$ is the first cell of the spoke $\mathcal{B}_1$. If $\mathcal{B}_1, \dots, \mathcal{B}_r$ are all the spokes crossed by $\mathcal{T}$, then $r \le L/2$ by Lemma 6.3. Since the band $\mathcal{T}$ does not cross the spoke $\mathcal{B}_{r+1}$, no other θ-band of Ψ crossing $\mathcal{B}_1$ can cross $\mathcal{B}_{r+1}$. and no θ-band crossing the spoke $\mathcal{B}_{L-3}$ can cross $\mathcal{B}_r$. The same argument shows that $r+1 \ge L/2-2$ if there is a θ-band of Ψ crossing the spoke $\mathcal{B}_{L-3}$. □

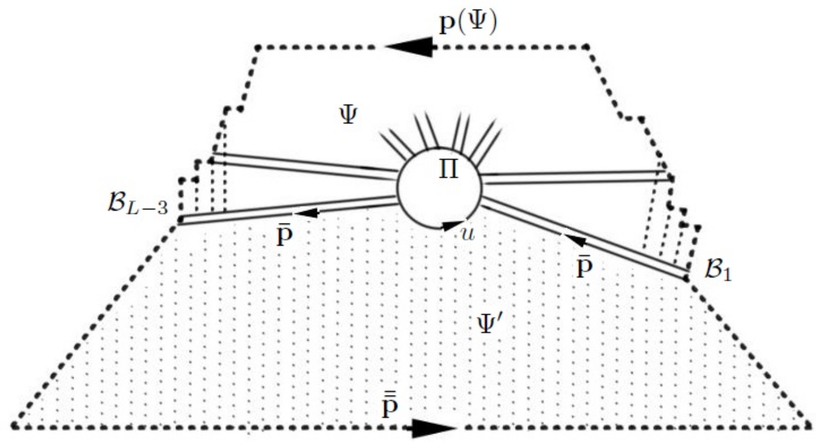


Figure 16: Boundaries of Ψ and Ψ'

For the clove $\Psi = cl(\pi, \mathcal{B}_1, \mathcal{B}_{L-3})$ in Δ, we denote by $\mathbf{p}(\Psi)$ the common subpath of $\partial\Psi$ and $\partial\Delta$ starting with the $\check{t}$-edge of $\mathcal{B}_1$ and ending with the $\check{t}$-edge of $\mathcal{B}_{L-3}$. Similarly we define the (outer) path $\mathbf{p}_{ij} = \mathbf{p}(\Psi)_{ij}$ for every smaller clove Ψ_{ij}.

7.2.6 The clove Ψ and related subdiagrams.

Lemma 7.21. *Every path $\mathbf{p}_{i,i+1}$ $(i = 1, \dots, L-4)$ has fewer than $3K_0$ q-edges.*

Proof. Let a maximal q-band $\mathcal{C}$ of Ψ starts on $\mathbf{p}_{i,i+1}$ and does not end on Π. Then is has to end on $\mathbf{p}_{i,i+1}$ too. If Γ is the subdiagram without disks separated by $\mathcal{C}$, then every maximal θ-band of Γ has to cross the q-band $\mathcal{C}$ since its extension in Ψ must cross either $\mathcal{B}_1$ or $\mathcal{B}_{L-3}$ by Lemma 7.20. Therefore Γ is a comb with handle $\mathcal{C}$.

Consider the q-bands of this kind defining maximal subcombs $\Gamma_1, \Gamma_2, \dots \Gamma_k$ in $\Psi_{i,i+1}$. The basic width of each of them is less than K_0 by Lemma 7.18. Therefore $k \le 1$ since otherwise one can get two subcombs contradicting to Lemma 7.19 (1), because there are at most $N+1$ maximal q-bands starting on $\partial\Pi$ in $\Psi_{i,i+1}$. By Lemma 7.19 (2), such a subcomb has at most $2K_0$ q-edges in the boundary. Hence there are at most $2K_0 + N < 3K_0$ q-edges in the path $\mathbf{p}_{i,i+1}$. □

We denote by $\overline{\Delta}$ the subdiagram formed by Π and Ψ, and denote by $\overline{\mathbf{p}}$ the path $\mathbf{top}(\mathcal{B}_1)\mathbf{u}^{-1}\mathbf{bot}(\mathcal{B})_{L-3}^{-1}$, where $\mathbf{u}$ is a subpath of $\partial\Pi$, such that $\overline{\mathbf{p}}$ separates $\overline{\Delta}$ from the remaining subdiagram Ψ' of Δ (fig. 16).

Similarly we define subdiagrams $\overline{\Delta}_{ij}$, paths $\overline{\mathbf{p}}_{i,j} = \mathbf{top}(\mathcal{B}_i)\mathbf{u}_{ij}^{-1}\mathbf{bot}(\mathcal{B})_j^{-1}$, where $\mathbf{u}_{ij}$ is a subpath of $\partial\Pi$, and the subdiagrams Ψ'_{ij}.

We denote by $H_1, \dots, H_{L-3}$ the histories of the spokes $\mathcal{B}_1, \dots, \mathcal{B}_{L-3}$ (read starting from the disk Π) and by $h_1, \dots, h_{L-3}$ their lengths, i.e., the numbers of (θ, t)-cells. By Lemma 7.20, these lengths non-increase and then non-decrease as follows:

$$h_1 \ge h_2 \ge \dots \ge h_r; \;\; h_{r+1} \le \dots \le h_{L-3} \;\; (L/2 - 3 \le r \le L/2), \tag{7.49}$$

and therefore H_{i+1} is a prefix of H_i (H_j is a prefix H_{j+1}) for $i = 1, \dots, r-1$ (resp., for $j = r+1, \dots, L-4$).

Recall that by Definition 7.1 the boundary label of $\partial\Pi$ is a disk word V, where $V^{\emptyset} \equiv W^L$, where W is a accessible word.

Lemma 7.22. *We have the following inequalities*

$$|\bar{\mathbf{p}}_{ij}| \le h_i + h_j + (L - j + i)|W| - 1$$

and, if $i \le r$ and $j \ge r+1$, then

$$|\mathbf{p}_{ij}| \ge |\mathbf{p}_{ij}|_\theta + |\mathbf{p}_{ij}|_q \ge h_i + h_j + (j-i)N + 1$$

Proof. The first iequality follows from Lemma 6.1 (b) since the path $\mathbf{u}_{ij}$ has $L-j+i-1$ t-edges. To prove the second inequality, we observe that the path $|\mathbf{p}_{ij}|$ has $(j-i)N+1$ q-edges and it has $h_i + h_j$ θ-edges by Lemma 7.20. □

Lemma 7.23. *If $j-i > L/2$, then we have $\mu(\Delta) - \mu(\Psi'_{ij}) > -2Jn(h_i+h_j) \ge -2Jn|\mathbf{p}_{ij}|$.*

Proof. The number of q-edges in the path $\bar{\mathbf{p}}_{ij}$ (or in the path $\mathbf{u}_{ij}$) does not exceed the similar number for $\mathbf{p}_{ij}$ provided $j - i \ge L/2$. Therefore any two white beads o, o' of the necklace on $\partial\Delta$, provided they both do not belong to $\mathbf{p}_{ij}$, are separated by at least the same number of black beads in the necklace for Δ as in the necklace for Ψ'_{ij} (either the clockwise arc $o - o'$ includes $\mathbf{p}_{ij}$ or not). So such a pair contributes to $\mu(\Delta)$ at least the amount it contributes to $\mu(\Psi'_{ij})$. Thus, to estimate $\mu(\Delta) - \mu(\Psi'_{ij})$ from below, it suffices to consider the contribution to $\mu(\Psi')$ for the pairs o, o', where one of the two beads lies on $\mathbf{p}_{ij}$. The number of such (unordered) pairs is bounded by $n(h_i+h_j)$ Taking into account the definition of μ for diagrams and inequalities (7.49), we get the required inequality. □

Lemma 7.24. *If $j - i > L/2$, then the following inequality holds: $|\mathbf{p}_{ij}| < (1+\varepsilon)|\bar{\mathbf{p}}_{ij}|$, where $\varepsilon = N_4^{-\frac{1}{2}}$. Moreover, we have $|\mathbf{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij}) < (1+\varepsilon)|\bar{\mathbf{p}}_{ij}|$.*

Proof. It suffices to prove the second statement. Let d be the difference $|\mathbf{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij}) - |\bar{\mathbf{p}}_{ij}|$ and assume to contradiction that $d \ge \varepsilon|\bar{\mathbf{p}}_{ij}|$. Then $d \ge |\mathbf{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij}) - \varepsilon^{-1}d$, whence

$$d \ge (1+\varepsilon^{-1})^{-1}(|\mathbf{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij})) \ge \frac{\varepsilon}{2}(|\mathbf{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij})) \ge \frac{\varepsilon y}{2}, \tag{7.50}$$

where by definition, $y = |\mathbf{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij})$.

We have $(|\partial\Delta| + \sigma_\lambda(\Delta)) - (|\partial\Psi'_{ij}| + \sigma_\lambda(\Psi'_{ij})) \ge d > 0$, because $|\partial\Delta| - |\partial\Psi'_{ij}| \ge |\mathbf{p}_{ij}| - |\bar{\mathbf{p}}_{ij}|$ and $\sigma_\lambda(\bar{\Delta}_{ij}) + \sigma_\lambda(\bar{\Psi}'_{ij}) \le \sigma_\lambda(\Delta)$. Therefore for $x = n + \sigma_\lambda(\Delta)$, we obtain from the minimality of the counter-example Δ that

$$\begin{aligned}\mathrm{Area}_G(\Psi'_{ij}) &\le N_4(x-d)^2 + N_3\mu(\Psi'_{ij}) \le N_4x^2 - N_4xd + N_3\mu(\Delta)\\ &+ 2N_3Jn|\mathbf{p}_{ij}| \le N_4x^2 + N_3\mu(\Delta) - N_4xd + 2N_3Jny\end{aligned} \tag{7.51}$$

by Lemma 7.23 and inequality $\sigma_\lambda(\Psi'_{ij})) \le \sigma_\lambda(\Delta)$. By Lemma 7.22, $|\bar{\mathbf{p}}_{ij}| < |\mathbf{p}_{ij}| + |\partial\Pi|$, and so the perimeter $|\partial\Psi_{ij}|$ is less than $2|\mathbf{p}_{ij}| + |\partial\Pi|$. Since $|\partial\Pi| \le L|\bar{\mathbf{p}}_{ij}| < L(|\mathbf{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij}))$, we have

$$|\partial\Psi_{ij}| < (2+L)|\mathbf{p}_{ij}| + L\sigma_\lambda(\bar{\Delta}_{ij}) \le (L+2)y \tag{7.52}$$

By the inequalities $N_2 > N_1$, (7.52), lemmas 6.14 and 6.7 (a), the G-area of Ψ_{ij} does not exceed

$$N_2(2+L)^2y^2 + N_1\mu(\Psi_{ij}) \le N_2(J+1)(2+L)^2y^2 \qquad (7.53)$$

By Lemma 7.16, the G-area of Π does not exceed $c_6|\partial\Pi|^2 \le c_6(L+2)^2y^2$, and so there is a constant $c_7 = c_7(L)$ such that $\mathrm{Area}_G(\Pi) \le c_7y^2$.

This estimate and (7.53) give the inequality $\mathrm{Area}_G(\bar\Delta_{ij}) \le N_2(J+1)(2+L)^2y^2 + c_7y^2$, and we obtain with (7.51) that

$$\mathrm{Area}_G(\Delta) \le N_4x^2 + N_3\mu(\Delta) - N_4xd + 2N_3Jny + N_2(J+1)(2+L)^2y^2 + c_7y^2$$

To obtain the desired contradiction, it suffices to show that here, the number $T = N_4xd/3$ is greater than each of the last three summands. Recall that $x \ge n$, $d > \varepsilon y/2$ by (7.50), $\varepsilon = N_4^{-1/2}$, and so $T > 2N_3Jny$ if N_4 is large enough in comparison with N_3 and other constant chosen earlier. Also we have $T > N_2(J+1)(2+L)^2y^2$, because $x = n+\sigma_\lambda(\Delta) > \mathbf{p}_{ij}|+\sigma_\lambda(\bar\Delta_{ij}) = y$, and so $xd > x\varepsilon y/2 \ge \varepsilon y^2/2$. Finally, $T > c_7y^2$ since

$$xd > x\varepsilon y/2 \ge y^2\varepsilon/2$$

□

For every path $\mathbf{p}_{i,i+1}$ we will fix a shortest path $\mathbf{q}_{i,i+1}$ homotopic to $\mathbf{p}_{i,i+1}$ in the subdiagram Ψ_{ij}, such that the first and the last $\check t$-edges of $\mathbf{q}_{i,i+1}$ coincide with the first and the last $\check t$-edges of $\mathbf{p}_{i,i+1}$. For $j > i+1$ the path $\mathbf{q}_{i,j}$ is formed by $\mathbf{q}_{i,i+1}, \dots \mathbf{q}_{j-1,j}$.

Lemma 7.25. *If $i \le r$ and $j \ge r+1$, then*

$$|\mathbf{q}_{ij}| \ge |\mathbf{q}_{ij}|_\theta + |\mathbf{q}_{ij}|_q \ge h_i + h_j + (j-i)N + 1$$

The proof is similar to the second part of Lemma 7.22.

Let Ψ^0_{ij} (let Ψ^0, Δ^0) be the subdiagram of Ψ_{ij} (of Ψ, of Δ) obtained after replacement of the subpath $\mathbf{p}_{ij}$ (of $\mathbf{p}$) by $\mathbf{q}_{ij}$ (by $\mathbf{q} = \mathbf{q}_{1,L-3}$, resp.) in the boundary.

Lemma 7.26. *(1) The subdiagram Ψ^0_{ij} has no maximal q-bands except for the q-spokes starting from $\partial\Pi$. (2) Every θ-band of Ψ^0_{ij} is crossed by the path $\mathbf{q}_{ij}$ at most once.*

Proof. Assume there is a q-band $\mathcal{Q}$ of Ψ^0_{ij} starting and ending on $\mathbf{q}_{ij}$ Then $j = i+1$ and $\mathbf{q}_{i,i+1} = \mathbf{uevfw}$, where $\mathcal{Q}$ starts with the q-edge $\mathbf{e}$ and ends with the q-edge $\mathbf{f}$. Suppose that $\mathcal{Q}$ has length ℓ. Then $|v| \ge \ell$ since every maximal θ-band of $\Psi^0_{i,i+1}$ crossing $\mathcal{Q}$ has to end on the subpath $\mathbf{v}$. So one has $|\mathbf{evf}| = \ell + 2$, and replacing the subpath $\mathbf{evf}$ by a side of $\mathcal{Q}$ of length ℓ one replaces the path $\mathbf{q}_{i,i+1}$ with a shorter homotopic path by Lemma 6.1. This contradicts the choice of $\mathbf{q}_{i,i+1}$, and so statement (1) is proved. The proof (2) is similar. □

It follows from Lemma 7.20 that between the spokes $\mathcal{B}_j$ and $\mathcal{B}_{j+1}$ $(1 \le j \le r-1)$, there is a trapezium Γ_j of height h_{j+1} with the side $\check t$-bands . Similarly, we have trapezia Γ_j for $r+1 \le j \le L-4$. By Lemma 7.26 (2), every trapezium Γ_j is contained in both $\Psi_{j,j+1}$ and $\Psi^0_{j,j+1}$.

The bottoms $\mathbf{y}_j$ of all trapezia Γ_j belong to $\partial\Pi$ and have the same label $W\check t$. We will use $\mathbf{z}_j$ for the tops of these trapezia. Since Γ_j and Γ_{j-1} $(2 \le j \le r-1)$ have the same bottom labels and the history H_j is a prefix of H_{j-1}, by Lemma 5.12, h_j different θ-bands

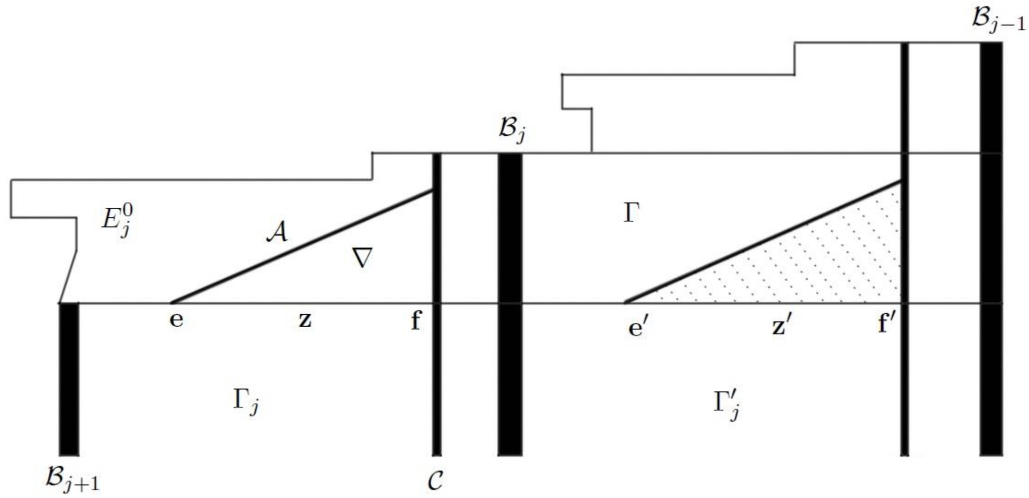


Figure 17: Lemma 7.27

of Γ_{j-1} form the copy Γ'_j of the trapezium Γ_j (more precisely, a copy of a superscript shift $\Gamma_j^{(+(\pm 1))}$) with top and bottom paths $\mathbf{z}'_j$ and $\mathbf{y}'_j = \mathbf{y}_{j-1}$.

‘ We denote by E_j (by E_j^0) the comb formed by the maximal θ-bands of $\Psi_{j,j+1}$ (of $\Psi^0_{j,j+1}$, resp.) crossing the $\tilde{t}$-spoke $\mathcal{B}_j$ but not crossing $\mathcal{B}_{j+1}$ ($1 \le j \le r-1$, see fig. 17). Its handle $\mathcal{C}_j$ of height $h_j - h_{j+1}$ is contained in $\mathcal{B}_j$. The boundary ∂E_j (resp., ∂E_j^0) consists of the side of this handle, the path $\mathbf{z}_j$ and the path $\mathbf{p}_{j,j+1}$ (the path $\mathbf{q}_{j,j+1}$, respectively).

Assume that a maximal Y-band $\mathcal{A}$ of E_j^0 ($2 \le j \le r-1$) starts on the path $\mathbf{z}_j$ and ends on a side a-edge of a maximal q-band $\mathcal{C}$ of E_j^0. Then $\mathcal{A}$, a part of $\mathcal{C}$ and a part $\mathbf{z}$ of $\mathbf{z}_j$ bound a comb ∇.

Lemma 7.27. *There is a copy of the comb ∇ in the trapezium $\Gamma = \Gamma_{j-1} \backslash \Gamma'_j$. It is a superscript shift of ∇.*

Proof. The subpath $\mathbf{z}$ of $\mathbf{z}_j$ starts with an a-edge $\mathbf{e}$ and ends with a q-edge $\mathbf{f}$. There is a copy $\mathbf{z}'$ of $\mathbf{z}$ in $\mathbf{z}'_j$ starting with $\mathbf{e}'$ and ending with $\mathbf{f}'$. Note that the θ-cells π and π' attached to $\mathbf{f}$ and to $\mathbf{f}'$ in ∇ and in Γ are copies of each other since they correspond to the same letter of the history. Now moving from $\mathbf{f}$ to $\mathbf{e}$, we see that the whole maximal θ-band $\mathcal{T}_1$ of ∇ containing π has a copy in Γ. Similarly we obtain a copy of the next maximal θ-band $\mathcal{T}_2$ of ∇, and so on. □

7.2.7 Bounding the number of Y-bands in a sector of a clove

Lemma 7.28. *At most N Y-bands starting on the path $\mathbf{y}_j$ can end on the (θ, q)-cells of the same θ-band. This property holds for the Y-bands starting on $\mathbf{z}_j$ too.*

Proof. We will prove the second claim only since the proof of the first one is similar. Assume that the a-bands $\mathcal{A}_1, \dots, \mathcal{A}_s$ start from $\mathbf{z}_j$ and end on some (θ, q)-cells of a θ-band $\mathcal{T}$. Let $\mathcal{T}_0$ be the minimal subband of $\mathcal{T}$, where the a-bands $\mathcal{A}_2, \dots, \mathcal{A}_{s-1}$ end and

$\bar{\mathbf{z}}_j$ be the minimal subpath of $\mathbf{z}_j$, where they start. Then by Lemma 5.6, every maximal q-band starting on $\bar{\mathbf{z}}_j$ has to cross the band $\mathcal{T}_0$ and vice versa. Hence the base of $\mathcal{T}_0$ is a subbase of the standard base (or of its inverse). Since every rule of $\mathbf{M}$ can change at most $N-2$ a-letters in a word with standard base, all (θ,q)-cells of $\mathcal{T}_0$ have at most $N-2$ a-edges, and the statement of the lemma follows. □

Without loss of generality, we assume that

$$h = h_{L_0+1} \geq h_{L-L_0-3}. \tag{7.54}$$

(Recall that L_0 is one of the parameters used in the paper, a number between c_5 and L, Section 2.3.)

7.2.8 Estimating the sizes of trapezia Γ_j

Lemma 7.29. *If $h \leq L_0^2|W|_Y$, then the number of trapezia Γ_j with the properties $|\mathbf{z}_j|_Y \geq |W|_Y/c_5N$ for $j \in [L_0+1, r-1]$ or $j \in [r+1, L-L_0-5]$, is less than $L/5$.*

Proof. Consider Γ_j as in the assumption of the lemma with $j \in [L_0+1, r-1]$. The subcomb E_j^0 has at most N maximal q-bands by Lemma 7.26. So there are at most N maximal a-bands starting on $\mathbf{z}_j$ and ending on each of the θ-bands of E_j^0. Proving by contradiction, we have at least $L|W|_a/5c_5N$ such a-bands for all $j \in S$, where S the set of integers in $[L_0+1, r-1] \cup [r+1, L-L_0-5]$; denote this set of a-bands by $\mathbf{A}$. But the number of maximal θ-bands in all such subcombs E_j^0 does not exceed $2h$. Therefore at least $L|W|_a/5c_5N - 2hN$ bands from $\mathbf{A}$ end on the subpaths $\mathbf{q}_{j,j+1}$ for $j \in S$. Therefore by Lemmas 7.25 and 6.1, we have

$$\begin{aligned} |\mathbf{p}_{L_0+1,L-L_0-5}| \geq |\mathbf{q}_{L_0+1,L-L_0-5}| &\geq h_{L_0+1} + h_{L-L_0-5} + LN/2 + \delta(L|W|_a/5c_5N - 2hN) \\ &\geq h_{L_0+1} + h_{L-L_0-5} + LN/2 + \delta L|W|_a/10c_5N \end{aligned} \tag{7.55}$$

since $2hN \leq 2L_0^2N|W|_a < L_0^3|W|_a/10c_5N \leq L|W|_a/10c_5N$ by the choice of L_0 and L.

Also by Lemma 7.22, we have

$$\begin{aligned} |\bar{\mathbf{p}}_{L_0+1,L-L_0-5}| &\leq h_{L_0+1} + h_{L-L_0-5} + 3L_0N + 3L_0\delta|W|_a \\ &\leq h_{L_0+1} + h_{L-L_0-5} + 3L_0N + \delta L|W|_a/20c_5N, \end{aligned} \tag{7.56}$$

because by the choice of L, $3L_0 < L/20c_5N$. Since $h_{L_0+1} + h_{L-L_0-5} \leq 2h \leq 2L_0^2|W|_a < L|W|_a$, L is chosen after c_5N, and $\varepsilon = N_4^{-1/2}$ is chosen after L, the inequality

$$\frac{|\mathbf{p}_{L_0+1,L-L_0-5}|}{|\bar{\mathbf{p}}_{L_0+1,L-L_0-5}|} \geq 1 + \frac{\delta}{20c_5N} > 1 + \varepsilon$$

follows from (7.55, 7.56), contrary to Lemma 7.24. The lemma is proved by contradiction. □

Lemma 7.30. *If $h \leq L_0^2|W_Y|$, then the histories H_1 and H_{L-3} have different first letters unless these letters are $\theta(23)^{-1}$.*

Proof. Let $\mathcal{T}$ and $\mathcal{S}$ be the maximal θ-bands of Ψ crossing $\mathcal{B}_1$ and $\mathcal{B}_{L-3}$, respectively, and the closest to the disk Π. Let they cross k and ℓ spokes of Π, respectively. By Lemma 7.29, $k+\ell > L - L/5 - 3L_0 > 2L/3$, and also $k, \ell \geq 2$ since $L/2 - 3 \leq r \leq L/2$. It follows from Lemma 7.8 (2) that the first letters of H_1 and H_{L-3} are different. □

Lemma 7.31. *We have $h > L_0^2|W|_Y$.*

Proof. If this inequality is wrong, then by Lemma 7.29, there are at least $L - L/5 - 3L_0 > 0.7L$ trapezia Γ_j with $|\mathbf{z}_j|_Y < |W|_Y/c_5N$, and one can choose two such trapesia Γ_k and Γ_ℓ such that $k < r$, $\ell \ge r+1$ and $\ell - k > 0.6L$. Since H_{k+1} (resp. H_ℓ) is a prefix of H_1 (of H_{L-3}), it follows from Lemma 7.30 that the first letters of H_{k+1} and H_ℓ are different unless they are equal to $\theta(23)^{-1}$.

Since the bottoms of Γ_k and Γ_ℓ (which belong in $\partial\Delta$) have the same label, up to a superscript shift, one can construct an auxiliary trapezium E identifying the bottom of a copy of Γ_k and the bottom of a mirror copy of Γ_ℓ. The history of E is $H_\ell^{-1}H_{k+1}$, which is an eligible word if the first letters of H_k and H_ℓ are different.

If both first letters are $\theta(23)^{-1}$, then the word $H_\ell^{-1}H_{k+1}$ also eligible by definition. If the bottom θ-bands of Γ_k and Γ_l are just copies of each other then the above constructed diagram E is not reduced. However one can modify the construction replacing Γ_k by an auxiliary superscript shift $\Gamma_k^{(+1)}$. By the definition of relations (5.7), the bottom labels of $\Gamma_k^{(+1)}$, Γ_k and Γ_l are all equal, but the top labels of the first θ-bands of $\Gamma_k^{(+1)}$ and Γ_l are not mirror copies of each other (they differ by 1-shift), and so the diagram E obtained by identifying the bottom of a copy of $\Gamma_k^{(+1)}$ and the bottom of a mirror copy of Γ_ℓ is reduced, i.e.,we can obtain the trapezium E in any case.

The top and the bottom of E have Y-lengths less than $|W|_Y/c_5N$. Without loss of generality, one may assume that $h_{k+1} \ge h_\ell$, and so $h_{k+1} \ge t/2$, where t is the height of E.

Note that the difference of Y-lengths $|W|_Y - |W|_Y/c_5N > |W|_Y/2$, and so

$$h_{k+1}, h_\ell > |W|_Y/2N \tag{7.57}$$

since the difference of Y-lengths for the top and the bottom of every maximal θ-band of E does not exceed N. Therefore $t > |W|_Y/N$, and the computation corresponding E satisfies the assumption of Lemma 4.12.

So for every factorization $H'H''H'''$ of the history of Γ_k, where $||H'|| + ||H''|| \le \lambda||H'H''H'''||$, we have $||H''|| > 0.4t$, since $\lambda < 1/5$. Therefore by Lemma 4.12, the spoke $\mathcal{B}_{k+1}$ is a λ-shaft.

Using Lemma 7.22, we obtain:

$$|\mathbf{p}_{k+1,\ell}| + \sigma_\lambda(\overline{\Delta}_{k+1,\ell}) \ge h_{k+1} + h_\ell + 0.6LN + h_{k+1} \tag{7.58}$$

By inequality (7.57), we have $\delta L|W|_Y \le 2LN\delta h_{k+1} < h_{k+1}$ by the choice of δ and by Lemma 7.22,

$$|\bar{\mathbf{p}}_{k+1,\ell}| \le h_{k+1} + h_\ell + 0.4LN + 0.4L\delta|W|_Y \le h_{k+1} + h_\ell + h_{k+1}/2 \tag{7.59}$$

The right-hand side of the inequality (7.58) divided by the right-hand side of (7.59) is greater than 1.1 (because $h_{k+1} \ge h_\ell$), which contradicts to Lemma 7.24. Thus, the lemma is proved. □

Lemma 7.32. *We have $h_i > \delta^{-1}$ for every $i = 1, \dots, L_0$.*

Proof. By inequalities (7.54) and (7.49), we have $h_i \ge h_{L-L_0-3}$. Proving by contradiction, we obtain $|W|_a < h_i \le \delta^{-1}$ by Lemma 7.31. Then

$$|\bar{\mathbf{p}}_{i,L-L_0-3}| < h_i + h_{L-L_0-3} + 3L_0(N + \delta^{-1}\delta) \le h_i + h_{L-L_0-3} + 4L_0N$$

by Lemma 7.22, and $|\mathbf{p}_{i,L-L_0-3}| \geq h_i + h_{L-L_0-3} + LN/2$. Since $h_i + h_{L-L_0-3} \leq 2\delta^{-1}$ and $4L_0N < LN/4$, we see that $\frac{|\mathbf{p}_{i,L-L_0-3}|}{|\bar{\mathbf{p}}_{i,L-L_0-3}|} > 1+\delta > 1+\varepsilon$ contrary to Lemma 7.24. The lemma is proved by contradiction. □

.

7.2.9 Bounding shafts in a clove and obtaining corollaries.

Lemma 7.33. *None of the spokes $\mathcal{B}_1, ..., \mathcal{B}_{L_0}$ contains a λ-shaft at Π of length at least δh.*

Proof. On the one hand, by Lemmas 7.22 and 7.31,

$$|\bar{\mathbf{p}}_{L_0+1,L-L_0-3}| < h_{L_0+1}+h_{L-L_0-3}+3L_0(N+\delta|W|_a) < h_{L_0+1}+h_{L-L_0-3}+3L_0(N+\delta L_0^{-2}h) \tag{7.60}$$

On the other hand, by Lemma 7.22,

$$|\mathbf{p}_{L_0+1,L-L_0-3}| > h_{L_0+1} + h_{L-L_0-3} + (L-3L_0)N \tag{7.61}$$

If the statement of the lemma were wrong, then we would have $\sigma_\lambda(\bar{\Delta}) \geq \delta h$, and inequalities (7.60) and (7.61) would imply that

$$|\mathbf{p}_{L_0+1,L-L_0-3}| - |\bar{\mathbf{p}}_{L_0+1,L-L_0-3}| + \sigma_\lambda(\bar{\Delta}) \geq (L-6L_0)N - 3L_0^{-1}\delta h + \delta h \geq LN/2 + \delta h/2$$

The right-hand side of the last inequality divided by the right-hand side of (7.60) is greater than $\varepsilon = N_4^{-\frac{1}{2}}$, because $h \geq h_{L_0+1}, h_{L-L_0-3}$, which would contradict to Lemma 7.24. Thus, the lemma is proved. □

Lemma 7.34. *For every $j \in [1, L_0-1]$, we have $|\mathbf{z}_j|_Y > h_{j+1}/c_5$.*

Proof. If $|\mathbf{z}_j|_a \leq h_{j+1}/c_5$, then the computation $\mathcal{C}: W_0 \to \cdots \to W_t$ corresponding to the trapezium Γ_j satisfies the assumption of Lemma 4.12, since $t = h_{j+1} > c_5|W_t|_a = c_5|\mathbf{z}_j|_a$ and by Lemma 7.31, $t = h_{j+1} \geq L_0^2|W_0|_a \geq c_5|W|_a$ since $L_0 > c_5$. Hence $\mathcal{B}_{j+1}$ is a λ-shaft by Lemma 4.12 since $\lambda < 1/2$. We obtain a contradiction with Lemma 7.33 since $\delta h \leq h \leq h_{j+1}$. Thus, the lemma is proved. □

Lemma 7.35. *For every $j \in [1, L_0-1]$, we have $h_{j+1} < (1-\frac{1}{10c_5N})h_j$.*

Proof. By Lemma 7.34, we have $|\mathbf{z}_j|_a \geq h_{j+1}/c_5$. Let us assume that $h_{j+1} \geq (1-\frac{1}{10c_5N})h_j$, that is the handle $\mathcal{C}_j$ of E_j has height at most $h_j/10c_5N$. By Lemma 7.28, at most $h_j/10c_5$ maximal a-bands of E_j starting on $\mathbf{z}_j$ can end on the (θ,q)-cells of E_j. Hence at least

$$|\mathbf{z}_j|_a - h_j/10c_5 \geq |\mathbf{z}_j|_a - 2h_{j+1}/10c_5 \geq 0.8h_{j+1}/c_5 > 0.7h_j/c_5$$

of them have to end on the path $\mathbf{p}_{j,j+1}$.

The path $\mathbf{p}_{j,j+1}$ has at most $\frac{h_j}{10c_5N}$ θ-edges. Hence by Lemma 6.1,

$$|\mathbf{p}_{j,j+1}| \geq h_j - h_{j+1} + \delta(0.7h_j/c_5 - h_j/10c_5N) \geq h_j - h_{j+1} + 0.6\delta h_j/c_5,$$

and therefore by Lemma 7.22, $|\mathbf{p}_{j,L-L_0-3}| \geq LN/2 + h_j + h_{L-L_0-3} + 0.6\delta h_j/c_5$. On the other hand by Lemma 7.22, we have

$$|\bar{\mathbf{p}}_{j,L-L_0-3}| \leq h_j + h_{L-L_0-3} + 3NL_0 + 3L_0\delta|W|_a \leq h_j + h_{L-L_0-3} + 3NL_0 + 3L_0^{-1}\delta h_{j+1}$$

by Lemma 7.31 and inequality $h \leq h_{j+1}$. Hence $\frac{|\mathbf{p}_{j,L-L_0-3}|}{|\bar{\mathbf{p}}_{j,L-L_0-3}|} \geq (1 + \delta/10c_5)$ since $h_{L-L_0-3} \leq h_{L_0+1} \leq h_{j+1} \leq h_j$ and $L_0 >> c_5$. We have a contradiction with Lemma 7.24 since $\delta/10c_5 > \varepsilon$. The lemma is proved by contradiction. □

The proof of the next lemma is similar.

Lemma 7.36. *For every $j \in [2, L_0 - 1]$. we have $|\mathbf{z}_j|_Y \leq 2Nh_j$,*

Proof. Assume that $|\mathbf{z}_j|_a \geq 2Nh_j$. By Lemma 7.28, at most Nh_j maximal a-bands of E_j starting on $\mathbf{z}_j$ can end on the (θ, q)-cells of E_j. Hence at least $|\mathbf{z}_j|_a - Nh_j \geq Nh_j$ of them has to end on the path $\mathbf{p}_{j,j+1}$. The path $\mathbf{p}_{j,j+1}$ has at most h_j θ-edges. Hence by Lemma 6.1, $|\mathbf{p}_{j,j+1}| \geq h_j - h_{j+1} + \delta(Nh_j - h_j) = h_j - h_{j+1} + \delta(N-1)h_j$ and therefore by Lemma 7.22, $|\mathbf{p}_{j,L-L_0-3}| \geq LN/2 + h_j + h_{L-L_0-3} + \delta(N-1)h_j$. On the other hand by Lemmas 7.22 and 7.31, we have

$$|\bar{\mathbf{p}}_{j,L-L_0-3}| \leq h_j + h_{L-L_0-3} + 3NL_0 + 3L_0\delta|W|_a \leq h_j + h_{L-L_0-3} + 3NL_0 + \frac{3\delta h_j}{L_0}$$

because $h \leq h_j$. Since $h_j \geq h \geq h_{L-L_0-3}$, we have $\frac{|\mathbf{p}_{j,L-L_0-3}|}{|\bar{\mathbf{p}}_{j,L-L_0-3}|} \geq (1+\varepsilon)$, a contradiction by Lemma 7.24. □

7.2.10 Certain subtrapezia with one step history do not exist

Lemma 7.37. *There is no $i \in [2, L_0 - 3]$ such that the histories $H_{i-1} = H_iH' = H_{i+1}H''H' = H_{i+2}H'''H''H'$ and the computation $\mathcal{C}$ with history H_i corresponding to the trapezium Γ_{i-1} satisfy the following condition:*

(*) The history $H'''H''H'$ has only one step, and for the subcomputation $\mathcal{D}$ with this history, there is a sectors $Q'Q$ such that a state letter from Q or from Q' inserts a letter increasing the length of this sector after every transition of $\mathcal{D}$.

Proof. Recall that the standard base of $\mathbf{M}$ is built of the standard base B of $\mathbf{M}_4$ and its inverse copy $(B')^{-1}$ (plus letter t). Due to this mirror symmetry of the standard base, we have mirror symmetry for any accepting computation, in particular, for $\mathcal{C}$ and $\mathcal{D}$. Therefore proving by contradiction, we may assume that the a-letters are inserted from the left of Q.

Let $\mathcal{Q}$ be the maximal q-spoke of the subdiagram $E_i^0 \subset \Gamma_i$ corresponding to the base letter Q. If $\mathcal{Q}'$ is the neighbor from the left q-spoke for $\mathcal{Q}$ (the spokes are directed from the disk Π), then the subpath $\mathbf{x}$ of $\mathbf{z}_i$ between these two q-spokes has at least $h_{i+1} - h_{i+2} = ||H'''||$ a-letters. Indeed, Γ_i contains a copy Γ'_{i+1} of Γ_{i+1}, the bottom of the trapezium $\Gamma_i \backslash \Gamma'_{i+1}$ is the copy $\mathbf{z}'_{i+1}$ of $\mathbf{z}_{i+1}$ and the top of it iz $\mathbf{z}_i$, and so the subcomputation with history H''' has already increased the length of the $Q'Q$-sector. Thus, by lemmas 7.35, 7.31 and the choice of $L_0 > 100c_5N$, we have

$$|\mathbf{x}|_a \geq h_{i+1} - h_{i+2} \geq \frac{1}{10c_5N}h_{i+1} \geq 10L_0|W|_a \tag{7.62}$$

Note that an a-band $\mathcal{A}$ starting on $\mathbf{x}$ cannot end on a (θ,q)-cell from $\mathcal{Q}$. Indeed, otherwise by Lemma 7.27, there is a copy of this configuration in the diagram Γ_{i-1}, i.e. the copy of $\mathcal{A}$ ends on the copy of $\mathcal{Q}$ contrary the assumption that the rules of computation with history $H'''H''H'$ do not delete a-letters.

Let us consider the comb bounded by $\mathcal{Q}$, $\mathcal{Q}'$, $\mathbf{x}$ and the boundary path of Δ^0 (without the cells from $\mathcal{Q}'$). If the lengths of $\mathcal{Q}$ and $\mathcal{Q}'$ are s and s', repectively, then there are $|\mathbf{x}|+s$ maximal a-bands starting on $\mathbf{x}$ and $\mathcal{Q}$ and ending either on $\mathcal{Q}'$ or on $\partial\Delta^0$ since the comb has no maximal q-bands by Lemma 7.26. At most $s'<s$ of these a-bands can end on $\mathcal{Q}'$. Therefore at least $|\mathbf{x}|+s-s'$ of them end on the segment of the boundary path of Δ^0 lying between the ends of $\mathcal{Q}'$ and $\mathcal{Q}$.

Since this segment has $s-s'$ θ-edges, its length is at least $s-s'+\delta|\mathbf{x}|_a$ by Lemma 6.1. This inequality, Lemma 7.25 and inequality (7.62) imply

$$|\mathbf{p}_{i,L-L_0-3}|\ge|\mathbf{q}_{i,L-L_0-3}|\ge LN/2+h_i+h_{L-L_0-3}+\frac{\delta}{10c_5N}h_{i+1}$$

$$\ge LN/2+h_i+h_{L-L_0-3}+10\delta L_0|W|_a$$

Therefore

$$|\mathbf{p}_{i,L-L_0-3}|-\frac{7\delta}{100c_5N}h_{i+1}>3L_0N+h_i+h_{L-L_0-3}+3\delta L_0|W|_a\ge|\bar{\mathbf{p}}_{i,L-L_0-3}|, \quad (7.63)$$

by Lemma 7.22, and since Δ is a minimal counter-example, we have

$$\mathrm{Area}_G(\Psi'_{i+1,L-L_0-3})\le N_4(n+\sigma_\lambda(\Delta)-\frac{7\delta}{100c_5N}h_{i+1})^2+N_3\mu(\Psi'_{i,L-L_0-3})$$

$$\le N_4(n+\sigma_\lambda(\Delta))^2-N_4\frac{7\delta n}{100c_5N}h_{i+1}+N_3\mu(\Psi'_{i,L-L_0-3}) \quad (7.64)$$

By Lemma 7.31, $|W|_a\le L_0^{-2}h_i$, and by Lemma 7.32, $h_i>\delta^{-1}>100L_0N$, whence $|\bar{\mathbf{p}}_{i,L-L_0-3}|\le 2h_i+3L_0N+3\delta L_0|W|_a\le(2+0.03+\frac{3\delta}{L_0})h_i\le 2.1h_i$ and by Lemma 7.24, we have

$$|\mathbf{p}_{i,L-L_0-3}|\le(1+\varepsilon)|\bar{\mathbf{p}}_{i,L-L_0-3}|<2.2h_i \quad (7.65)$$

By lemmas 6.14 and 6.7 (a) and inequalities (7.65) and (7.63), the G-area of $\Psi_{i,L-L_0-3}$ does not exceed

$$N_2(2|\mathbf{p}_{i,L-L_0-3}|)^2+N_1\mu(\Psi_{i,L-L_0-3})\le N_2(4J+4)|\mathbf{p}_{i,L-L_0-3}|^2\le 5N_2(4J+4)h_i^2 \quad (7.66)$$

By Lemma 7.16, the G-area of Π is bounded by $c_6F(|\partial\Pi|)$. The inequalities (7.63) and (7.65) implies the inequality $|\partial\Pi|<L|\bar{\mathbf{p}}_{i,L-L_0-3}|<L|\mathbf{p}_{i,L-L_0-3}|<3Lh_i$. Therefore one may assume that the constant c_7 is chosen so that

$$\mathrm{Area}_G(\Pi)<c_6|\partial\Pi|^2<c_7h_i^2=\le c_7h_i^2 \quad (7.67)$$

(Recall that $h_i\le n/2$ here since h_i is the number θ-bands crossing $\mathcal{B}_i$; they start and end on $\partial\Delta$.) It follows from (7.66) and (7.67) that

$$\mathrm{Area}_G(\bar\Delta_{i,L-L_0-3})\le 5N_2(4J+4)h_i^2+c_7h_i^2g(n) \quad (7.68)$$

We need an estimate for $\mu(\Psi'_{i+1,L-L_0-3})-\mu(\Psi'_{i,L-L_0-3})$ now. To obtain it, we observe that by Lemma 7.20, the common q-edge $\mathbf{f}$ of the spoke $\mathcal{B}_i$ and $\partial\Delta$ separates at least

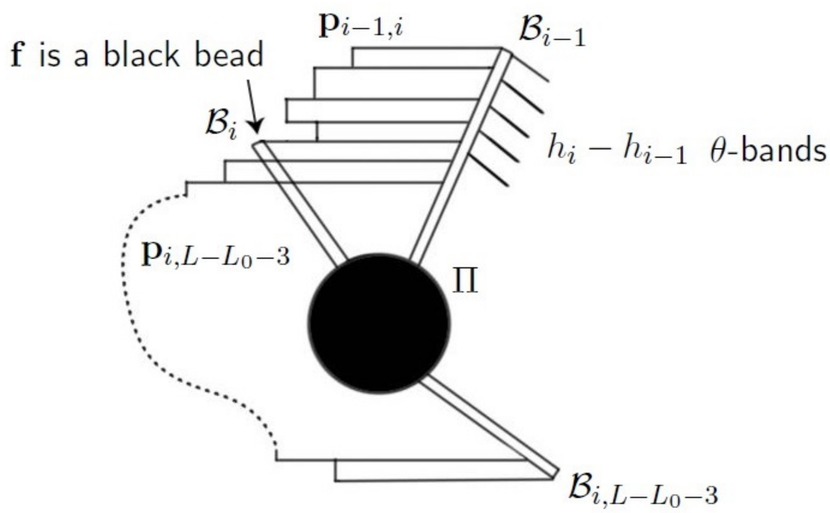


Figure 18: $\mu(\Psi'_{i+1,L-L_0-3}) - \mu(\Psi'_{i,L-L_0-3})$

$h_{i-1} - h_i = m_1$ θ-edges of the path $\mathbf{p}_{i-1,i}$ and m_2 ones lying on $\mathbf{p}_{i,L-L_0-3}$, where $m_2 = h_i + h_{i,L-L_0-3}$ (see fig. 18). Since the number of q-edges of $\mathbf{p}$ is less than $3K_0L < J$ by Lemma 7.21, one decreases $\mu(\Psi'_{i+1,L-L_0-3})$ at least by m_1m_2 when erasing the black bead on $\mathbf{f}$ in the necklace on $\partial\Psi'_{i+1,L-L_0-3}$ by Lemma 6.7 (d,b,c). Hence

$$\mu(\Psi'_{i+1,L-L_0-3})) - \mu(\Psi'_{i,L-L_0-3}) \geq m_1m_2$$

$$= (h_{i-1} - h_i)(h_i + h_{L-L_0-3}) \geq \frac{1}{c_5N}h_{i-1}(h_i + h_{L-L_0-3})$$

by Lemma 7.35. This inequality and Lemma 7.23 applied to $\Psi_{i+1,L-L_0-3}$, imply

$$\mu(\Delta) - \mu(\Psi'_{i,L-L_0-3}) \geq -2Jn(h_{i+1} + h_{L-L_0-3}) + \frac{1}{10c_5N}h_{i-1}(h_i + h_{L-L_0-3})$$

Note that $(h_{i+1} + h_{L-L_0-3}) \leq 2h_{i+1}$ by (7.49) and (7.54). Hence

$$N_3\mu(\Delta) - N_3\mu(\Psi'_{i,L-L_0-3}) \geq -4N_3Jnh_{i+1} + \frac{N_3}{10c_5N}h_{i-1}(h_i + h_{L-L_0-3}) \qquad (7.69)$$

It follows from (7.68, 7.69, 7.64) that

$$\mathrm{Area}_G(\Delta) \leq \mathrm{Area}_G(\Psi'_{i,L-L_0-3}) + \mathrm{Area}_G(\bar\Delta_{i,L-L_0-3}) \leq$$

$$\leq N_4(n + \sigma_\lambda(\Delta))^2 - N_4\frac{7\delta n}{100c_5N}h_{i+1} + N_3\mu(\Psi'_{i,L-L_0-3}) + 5N_2(4J+4)h_i^2 + c_7h_i^2$$

$$\leq N_4(n + \sigma_\lambda(\Delta))^2 + N_3\mu(\Delta) - N_4\frac{7\delta n}{100c_5N}h_{i+1} - \frac{N_3}{10c_5N}h_{i-1}(h_i + h_{L-L_0-3})$$

$$+4N_3Jnh_{i+1} + 5N_2(4J+4)h_i^2 + c_7h_i^2$$

We come to a contradiction since we obtain inequality $\mathrm{Area}_G(\Delta) \leq N_4(n+\sigma_\lambda(\Delta))^2 + N_3\mu(\Delta)$, because $N_4\frac{7\delta}{100c_5N} > 4N_3J$ and $\frac{N_3}{10c_5N} > 5N_2(4J+4) + c_7$. □

7.2.11 A clove with a disk can be removed and no counterexample exists

Lemma 7.38. *There exists no counter-example* Δ, *and therefore* $\mathrm{Area}_G(\Delta) \le N_4(n + \sigma_\lambda(\Delta))^2 + N_3\mu(\Delta)$ *for any minimal diagram* Δ *with* $|\partial\Delta| = n$.

Proof. Recall that for $j = 1, \dots, L_0 - 1$, we have $h_{j+1} < (1 - \frac{1}{10c_5N})h_j$ by Lemma 7.35, and by lemmas 7.34 and 7.36, we have inequalities $|\mathbf{z}_j|_Y \ge h_{j+1}/c_5$ and $|\mathbf{z}_k|_Y \le 2Nh_k$. One can choose an integer $\rho = \rho(\mathbf{M})$ (it depends on the S-machine $\mathbf{M}$ only as c_5 and N) so that $(1 - \frac{1}{10c_5N})^\rho < \frac{1}{6Nc_5}$, and so $h_{j+1} > 6Nc_5h_k$ if $k - j - 1 \ge \rho$. Hence

$$|\mathbf{z}_j|_Y \ge h_{j+1}/c_5 \ge 6Nh_k > 3|\mathbf{z}_k|$$

If L_0 is large enough, say $L_0 > 2000\rho$, one can obtain 1000 indices $j_1 < j_2 < \dots < j_{1000} < L_0$ such that for $i = 2, \dots, 1000$, one obtains inequalities

$$|\mathbf{z}_{j_{i-1}}| > 3|\mathbf{z}_{j_i}| \;\; and \;\; h_{j_{i-1}} \ge h_{j_{i-1}+1} > 6c_5Nh_{j_i} \tag{7.70}$$

Let $\mathcal{C}\colon\ W \equiv W_0 \to \dots \to W_t$ be the computation corresponding to the trapezium Γ_{j_2}. Since it contains the copy Γ'_{j_2+1} of Γ_{j_2+1}, which in turn contains a copy of Γ_{j_2+2} and so on, we have some configurations $W(k)$ in $\mathcal{C}$ ($k = 1, \dots, 999$), that are the labels of some $\mathbf{z}_{i_k}$ (but without superscripts) and $|W(k+1)|_Y > 3|W(k)|_Y$ for $k = 1, \dots, 998$. If for some k we were obtain one-step subcomputation $W(k) \to \dots \to W(k+4)$, then the statement of Lemma 4.13 would give a subcomputation $W(k+1) \to \dots \to W(k+4)$ contradicting to the statement of Lemma 7.37. Hence no five consecutive words $W(k)$-s are configuration of a one-step subcomputation, and so the number of steps in $W(1) \to \dots \to W(999)$ in at least 100.

It follows now from Lemma 4.5 that the step history of $\Gamma_{j_2}\backslash\Gamma$, where Γ is the copy of Γ_{L_0} in Γ_{j_2}, has a subword (34)(4)(45) or (45)(4)(34), or (12)(2)(23), or (23)(2)(12).

Let us consider the case (34)(4)(45) (or (45)(4)(34)). Then the history H_{j_2+1} of Γ_{j_2} can be decomposed as $H'H''H'''$, where H'' has form $\chi(i-1,i)H'\chi(i,i+1)$ (the S-machine works as $\mathbf{M}_3$) and $||H'|| \ge h$ since the height of Γ is at least h. Moreover, by Lemma 3.14 (b), one can choose i so that $||H'|| \ge ||H''||$ since the number of cycles m is large enough.

Since $h_{j_1+1} > 2h_{j_2}$ by (7.70), the history H_{j_1+1} of Γ_{j_1} has a prefix $H'H''H^*$, where $||H^*|| = ||H'|| \ge ||H''||$, and so the $\check{t}$-spoke $\mathcal{B}_{j_1+1}$ has a $\check{t}$-subband $\mathcal{C}$ starting with $\partial\Pi$ and having the history $H'H''H^*$.

For any factorization $\mathcal{C} = \mathcal{C}_1\mathcal{C}_2\mathcal{C}_3$ with $||\mathcal{C}_1|| + ||\mathcal{C}_2|| \le ||\mathcal{C}||/3$, the history of $\mathcal{C}_2$ contains the subhistory H'', since $||H^*|| = ||H'||| \ge ||H''||$. It follows that $\mathcal{C}$ is a λ-shaft, because $\lambda < 1/3$. The shaft has length at least $||H'|| \ge h$ contrary to Lemma 7.33.

The case of (12)(2)(23) (of (23)(2)(12)) is similar but $H'' = \zeta^{i-1,i}H'\zeta^{i,i+1}$ (the S-machine works as $\mathbf{LR}_m$) and the cycles of $\mathbf{LR}_m$ have equal lengths by Lemma 3.3 (3). We come to the final contradiction in this section. $\square$

8 Proof of Theorem 1.2

8.1 The Dehn function of the group G

Lemma 8.1. *For every big trapezia* Δ, *there is a diagram* $\tilde{\Delta}$ *over* G *with the same boundary label, such that the area of* $\tilde{\Delta}$ *does not exceed* $2\mathrm{Area}_G(\Delta)$.

Proof. Consider the computation $\mathcal{C}\colon\ V_0 \to \cdots \to V_t$ corresponding to Δ. According to Definition 6.8, one may assume that $\mathrm{Area}_G(\Delta) = c_5 h(||V_0|| + ||V_t||)$ since otherwise $\tilde\Delta = \Delta$.

Δ is the covered by L trapezia $\Delta_1, \dots, \Delta_L$ with base xvx, where xv (or the inverse word) is a cyclic shift ot the standard base of $\mathbf{M}$. By Lemmas 4.4 and 5.6, all $\Delta_1, \dots, \Delta_L$ are superscript shifts of each other. Let us apply Lemma 4.11 to any of them, say to Δ_1, whose top and bottom have labels W_0 and W_t. If we have Property (1) of that lemma, then the area of Δ_1 does not exceed $c_4 h(||V_0|| + ||V_t||)$ since every maximal θ-band of Δ_1 has at most $c_4(||V_0|| + ||V_t||)$ cells in this case. Hence area of Δ does not exceed $Lc_4h(||V_0||+||V_t||) \le 2c_4h(||W_0||+||W_t||) < c_5h(||W_0||+||W_t||) = \mathrm{Area}_G(\Delta)$, i.e. $\tilde\Delta = \Delta$ in this case too.

Hence one may assume that Property (2) of Lemma 4.11 holds for Δ_1. By that Lemma, items (b,d), the corresponding cyclic shifts $(W_0')^{\emptyset}$ and $(W_t')^{\emptyset}$ are accessible, and so removing the last letters x from V_0 and V_t we obtain disk words V_0' and V_t'. For the histories H' and H'' of $\mathcal{C}((W_0')^{\emptyset})$ and $\mathcal{C}((W_t')^{\emptyset})$, Lemma 4.11 gives inequality $||H'|| + ||H''|| \le t$.

Denote by Δ_- the diagram Δ without maximal rim x-band. So Δ_- has the boundary $\mathbf{p}_1\mathbf{q}_1\mathbf{p}_2^{-1}\mathbf{q}_2^{-1}$, where $\mathrm{Lab}(\mathbf{p}_1)$ and $\mathrm{Lab}(\mathbf{p}_2)$ are disk words and $\mathrm{Lab}(\mathbf{q})_{\mathbf{1}}) \equiv \mathrm{Lab}(\mathbf{q})_{\mathbf{2}})$ since the first and the last maximal two x-bands of Δ are L-shifts of each other.

If we attach disks Π_1 and Π_2 (of radius $\le t$ each) along their boundaries to the top and the bottom of Δ_-, we obtain a diagram, whose boundary label is trivial in the free group. Hence there is a diagram E with two disks whose boundary label is equal to the boundary label of Δ_-, and the area is less than $\le 3c_1t(||V_0'|| + ||V_t'||$ by Lemma 4.9. If we attach one x-band of length t to E, we construct the required diagram $\tilde\Delta(i)$ of area at most $\le 3c_1t(||V_0|| + ||V_t||) < c_5t(||V(1)|| + ||V(2)||)$ ☐

Lemma 8.2. *The Dehn function $d(n)$ of the group G is $O(n^2)$.*

Proof. To obtain the quadratic upper bound for $d(n)$ (with respect to the finite presentation of G given in Section 5), it suffices, for every word W vanishing in G with $||W|| \le n$, to find a diagram over G of area $O(n^2)$ with boundary label W. Since $|W| \le ||W||$, van Kampen's lemma and Lemma 7.31 provide us with a minimal diagram Δ such that $\mathrm{Area}_G(\Delta) \le N_4(n + \sigma_\lambda(\Delta))^2 + N_3\mu(\Delta)$ for some costants N_3 and N_4 depending on the presentation of G. By Lemmas 7.15, 6.7 (a) and the definition of $\mu(\Delta)$, the right-hand side does not exceed $N_4((1+c)n)^2 + N_3Jn^2$, and we conclude that $\mathrm{Area}_G(\Delta) \le C_0n^2$ for some constant C_0.

Recall that in the definition of G-area, the subdiagrams, which are big trapezia $\Gamma, \Gamma', \dots,$ can have common cells in their rim q-bands only. By Lemma 8.1, any big trapezia Γ from this list with top $\mathbf{p}_1$ and bottom $\mathbf{p}_2$ can be replaced by a diagram $\tilde\Gamma$ with (combinatorial) area at most $2\mathrm{Area}_G(\Gamma)$. When replacing all big trapezia $\Gamma, \Gamma', \dots,$ in this way, we should add q-bands for the possible intersection of big trapezia, but for every Γ of height h, we add at most $2h$ new cells. So the area of the modified diagram E is at most $3\mathrm{Area}_G(\Delta) \le 3C_0n^2$. Hence the required diagram is found for given word W. ☐

8.2 The conjugacy problem in G

Recall that the rule $\theta(23)$ locks all sectors of the standard base of $\mathbf{M}$ except for the input sector $\check R_0 \check P_1$ and its mirror copy. Hence every $\theta(23)^{-1}$-admissible word has the form

$W(k,k') \equiv w_1\alpha^k w_2(a')^{-k'}w_3$, where k and k' are integers and w_1, w_2, w_3 are fixed word in state letters; w_1 starts with $\check{t}$.

Lemma 8.3. *A word $W(k,k)$ is a conjugate of the word W_{ac} in the group G (and in the group M) if and only if the input α^k is accepted by the Turing machine $\mathbf{M}_0$.*

Proof. Let the S-machine $\mathbf{M}_0$ accept α^k. Then by Lemma 4.6, we have an accepting computation $\mathcal{C}$ of $\mathbf{M}$ starting with $W(k,k)$ and ending with W_{ac}. By Lemma 5.12, one can construct a corresponding trapezium Δ. Since the computation $\mathcal{C}$ uses neither the rules of Step 1, nor the rules of Step 2, nor the rules $\theta(23)^{\pm 1}$, the labels of the edges of Δ have no superscripts. Hence the bottom of Δ is labeled by $W(k,k)$, the top label is W_{ac} and the sides of Δ have equal labels since the S-machine $\mathbf{M}$ have cyclic standard bases. It follows from van Kampen Lemma that the words $W(k,k)$ and W_{ac} are conjugate in the group M, as required.

For the converse statement, we assume that the words $W(k,k)$ and W_{ac} are conjugate in G. Recall that the definition of annular diagram Δ over a group G is similar to the definition of van Kampen diagram, but the compliment of Δ in the plane has two connected components. So Δ has two boundary components. By van Kampen-Schupp lemma (see [12], Lemma 5.2 or [14], Lemma 11.2) there is a annular diagram Δ whose boundary components $\mathbf{p}_1$ and $\mathbf{p}_2$ have clockwise labels $W(k,k)$ and W_{ac}. As for van Kampen diagrams (see Subsection 7.1.2), one may assume that Δ is a minimal diagram and there are no two disks in Δ connected by two $\check{t}$-spokes $\mathcal{B}$ and $\mathcal{C}$ provided there are neither disks nor boundary components of Δ between $\mathcal{B}$ and $\mathcal{C}$. This property makes the disk graph of Δ hyperbolic as in Subsection 7.1.2: if Δ has a disk, then there is a disk with at least $L/2$ spokes ending on $\partial\Delta$ (see Corollary 10.1 in [14]).

However each of $\mathbf{p}_1$ and $\mathbf{p}_2$ has only one $\check{t}$-edge, and it follows that Δ has no disks since $L/2 > 2$. Hence the unique maximal $\check{t}$-band $\mathcal{B}$ of Δ has to connect these $\check{t}$-edges. Cutting Δ along a side $\mathbf{q}$ of $\mathcal{B}$, we obtain a reduced van Kampen diagram Γ over the group M. Its boundary path is $\mathbf{p}_1\mathbf{q}\mathbf{p}_2^{-1}\mathbf{q}'^{-1}$, where $\mathrm{Lab}(\mathbf{q}') \equiv \mathrm{Lab}(\mathbf{q})$. The maximal θ-bands of Γ connect $\mathbf{q}$ and $\mathbf{q}'$ since they cannot cross a q-band twice by Lemma 5.6. Hence Γ is a trapezium with top $\mathbf{p}_1$ and bottom $\mathbf{p}_2$. The base of Γ is standard since the top/bottom labels have standard base.

The equality $\mathrm{Lab}(\mathbf{q}') \equiv \mathrm{Lab}(\mathbf{q})$ implies that the side edges have no superscripts because $\mathrm{Lab}(\mathbf{q}')$ has to be a ± 1-shift of $\mathrm{Lab}(\mathbf{q})$. It follows from Lemma 5.12 and the definition of (θ, q)-relations that Γ corresponds to a reduced computation $\mathcal{C}\colon W(k,k) \to \cdots \to W_{ac}$ having no rules of Steps 1,2 and no $\theta(23)^{\pm 1}$. Therefore the word α^k is accepted by $\mathbf{M}_0$ by Lemma 4.6 (2).

$\square$

Proof of Theorem 1.2. Since the Turing machine $\mathbf{M}_0$ accepts a non-recursive language, the conjugacy problem is undecidable for the group G by Lemma 8.3. The Dehn function of G is at most quadratic by Lemma 8.2. To obtain a lower quadratic estimate, it suffices to verify that if a θ-letter θ and a-letter a commute, then by Lemmas 7.5 and 5.6, the area of the word $a^n\theta^n a^{-n}\theta^{-n}$ is equal to n^2 (or to use [2]: every non-hyperbolic finitely presented group has at least quadratic Dehn function). The theorem is proved. $\square$

References

[1] 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. (Cited on pages 2, 4, 36, and 37).

[2] B. H. Bowditch, Notes on Gromov's hyperbolicity criterion for path-metric spaces, from: "Group theory from a geometrical viewpoint (Trieste, 1990)", (E Ghys, A Haefliger, A Verjovsky, editors), World Sci. Publ., River Edge, NJ (1991) 64–167. (Cited on pages 2 and 77).

[3] M. R. Bridson, D. Groves, The quadratic isoperimetric inequality for mapping tori of free group automorphisms. Mem. Amer. Math. Soc. 203 (2010), no. 955, xii+152 pp. (Cited on page 3).

[4] M. R. Bridson and A. Haefliger.Metric spaces of non-positive curvature, volume 319 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999. (Cited on page 3).

[5] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V.F. Levy, M.S. Paterson, and W.P. Thurston. Word processing in groups. Jones and Bartlett Publishers. Boston, MA, 1992. (Cited on page 3).

[6] M.Gromov, Hyperbolic groups, in: Essays in Group Theory (S.M.Gersten, ed.), M.S.R.I. Pub. 8, Springer, 1987, 75–263. (Cited on page 2).

[7] M.Gromov, Asymptotic invariants of infinite groups, in: Geometric Group Theory. Vol. 2 (G.A.Niblo and M.A.Roller, eds.), London Math. Soc. Lecture Notes Ser., 182 (1993), 1–295. (Cited on page 2).

[8] F. Grunewald, D. Segal, Some general algorithms. I. Arithmetic groups. Ann. of Math. (2) 112 (1980), no. 3, 531–583. (Cited on page 3).

[9] V. Guba, The Dehn function of Richard Thompson's group F is quadratic. Invent. Math. 163 (2006), no. 2, 313–342. (Cited on page 3).

[10] V. Guba, M. Sapir, Diagram groups, Memoirs of the AMS, November, 1997. (Cited on page 3).

[11] V. S. Guba, M.V. Sapir, On Dehn functions of free products of groups. Proc. Amer. Math. Soc. 127 (1999), no. 7, 1885–1891. (Not cited).

[12] R. C. Lyndon and P. E. Schupp. *Combinatorial group theory.* Springer-Verlag, 1977. (Cited on pages 2, 35, and 77).

[13] C.F. Miller III. On group-theoretic decision problems and their classification. Annalls of Mathematics Studies, no. 68. Princeton University Press, Princeton, N.J., University of Tokio Press, Tokio, 1971. (Cited on page 5).

[14] A. Yu. Olshanskii, Geometry of defining relations in groups, Moscow, Nauka, 1989, 448 p.p. (in Russian, English translation by Kluwer Academic Publ., 1991) (Cited on pages 2, 12, 35, 37, and 77).

[15] A. Yu. Olshanskii, Hyperbolicity of groups with subquadratic isoperimetric inequality. Internat. J. Algebra Comput. 1 (1991), no. 3, 281–289. (Cited on page 2).

[16] A. Yu. Olshanskii, On the subgroup distortion in finitely presented groups, Matem. Sbornik, 188 (1997), N 11, 73–120 (Cited on pages 7, 36, 56, and 57).

[17] A. Yu. Olshanskii, Groups with undecidable word problem and almost quadratic Dehn function, Journal of Topology, 5, 4(2012), 785–886 (with Appendix by M.V.Sapir). (Cited on pages 6, 11, 44, 45, and 57).

[18] A. Yu. Olshanskii, Polynomially-bounded Dehn functions of groups, arXiv: 1710.00550, 2017, 92 pp. (Cited on pages 4, 5, 6, 7, 10, 35, 42, 43, 46, 61, and 62).

[19] 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. (Cited on page 40).

[20] A.Yu. Olshanskii, M.V. Sapir, Conjugacy problem and Higman embeddings, Memoirs of the AMS 170(2004), no. 804 p.p. vii+131. (Cited on pages 4, 11, 34, and 35).

[21] A.Yu. Olshanskii, M.V. Sapir, Groups with small Dehn functions and bipartite chord diagrams, Geometric and Functional Analysis, 16 (2006), 1324–1376. (Cited on pages 3, 4, 5, 6, 7, 10, 34, 35, 42, 43, and 45).

[22] A. Yu. Olshanskii, M.V. Sapir, Groups with non-simply connected asymptotic cones. Topological and asymptotic aspects of group theory, 203–208, Contemp. Math., 394, Amer. Math. Soc., Providence, RI, 2006. (Cited on page 4).

[23] A. Yu. Olshanskii, M. Sapir, On flat submaps of maps of non-positive curvature, arXiv:1702.08205 (2017), Accepted in the Bulletin of Mathematical Sciences, 2018. (Cited on page 5).

[24] P. Papasoglu, On the asymptotic cone of groups satisfying a quadratic isoperimetric inequality, J. Differential Geometry, Vol. 44 (1996), 789–806. (Cited on page 2).

[25] J.Rotman, An introduction to the theory of groups, 3d edition, Allyn and Bacon Inc., Boston, Mass, 1984. (Not cited).

[26] M. V. Sapir, Algorithmic and asymptotic properties of groups. International Congress of Mathematicians. Vol. II, 223–244, Eur. Math. Soc., Zürich, 2006. (Cited on pages 4 and 7).

[27] M. V. Sapir, Combinatorial algebra: syntax and semantics. With contributions by Victor S. Guba and Mikhail V. Volkov. Springer Monographs in Mathematics. Springer, Cham, 2014. xvi+355 pp. (Cited on pages 2, 4, and 35).

[28] M. V. Sapir, J. C. Birget, E. Rips, Isoperimetric and isodiametric functions of groups, Annals of Mathematics, 157, 2(2002), 345-466. (Cited on pages 2, 4, 6, 7, 10, 34, 35, 36, 37, 56, and 57).

[29] Robert Young, The Dehn function of $SL(n,\mathbb{Z})$, Ann. of Math. (2) 177 (2013), no. 3, 969–1027. (Cited on page 3).

Subject index

Alexander Yu. Ol'shanskii
Department of Mathematics
Vanderbilt University and
Department of Higher Algebra,
MEHMAT, Moscow State University
alexander.olshanskiy@vanderbilt.edu

Mark Sapir,
Department of Mathematics
Vanderbilt University
mark.sapir@vanderbilt.edu