Torsion Subgroups of Groups with Quadratic Dehn Function

By

Francis G Wagner

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Mathematics

August 31st, 2021

Nashville, Tennessee

Approved:

Alexander Ol'shanskii, Ph.D.

Michael Mihalik, Ph.D.

Denis Osin, Ph.D.

Mark Sapir, Ph.D.

Robert Scherrer, Ph.D.

ACKNOWLEDGEMENTS

I thank my advisor, Professor Alexander Ol'shanskii, for his eminent patience and guidance. I can only aspire to his vast knowledge and insight, and his approach to sharing just a bit of it with me will be something I strive to emulate. The influence of Professor Ol'shanskii on this work cannot be overstated.

I thank Professors Denis Osin and Michael Mihalik for the extensive knowledge they shared with me and for the advice that they have always been willing to provide.

I thank Professor Mark Sapir, whose brilliant work laid the foundation for my progress.

I thank my Master's advisor, Professor Lisa Carbone, for introducing me to the field and providing continuous counsel, without which I can confidently say none of this would be possible.

I thank Nancy and John Hood for introducing me to new methods of critical thinking and systematic approaches to learning.

I thank Professor Robert Scherrer for spending time to review my qualifying paper and my dissertation and to serve on my committee.

I thank the faculty of the mathematics department at Vanderbilt for the spectacular graduate courses they provided that taught me so much, for the aid they offered to make my teaching duties run as smoothly as possible, and for the events and committees they sponsored to make my time so enjoyable.

I thank my friends and colleagues for their camaraderie and ability to endure absurd jokes. In particular, I thank Diana Sordillo for her treasured support and companionship.

Finally, I thank my family for their endless love, help, and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

In this thesis, we construct the first examples of finitely presented groups with quadratic Dehn function which contain a finitely generated infinite torsion subgroup, answering a problem of Ol'shanskii. These examples are "optimal" in the sense that the Dehn function of any such finitely presented group must be at least quadratic. Moreover, we show that for any $n \geq 2^{48}$ such that n is either odd or divisible by 2^9, any infinite free Burnside group with exponent n is a quasi-isometrically embedded subgroup of a finitely presented group with quadratic Dehn function satisfying the Congruence Extension Property.

1.1 Dehn Functions

Given a set $\mathcal{A}$, a *word over* $\mathcal{A} \cup \mathcal{A}^{-1}$ is a finite sequence of elements of $\mathcal{A} \cup \mathcal{A}^{-1}$ called the *letters* of the word. Given such a word w, the (combinatorial) *length* of w, denoted $\|w\|$ or $|w|_{\mathcal{A}}$, is the number of letters that comprise its defining sequence.

The free group with basis $\mathcal{A}$ is denoted $F(\mathcal{A})$. The notation $F(\mathcal{A})$ is also used to denote the set of *reduced* words over $\mathcal{A} \cup \mathcal{A}^{-1}$, i.e the set of all words which contain no subsequence of the form aa^{-1} or $a^{-1}a$ for some $a \in \mathcal{A}$.

If G is a group generated by $\mathcal{A}$, i.e $G = \langle \mathcal{A} \rangle$, and w and v are two words over $\mathcal{A} \cup \mathcal{A}^{-1}$, then the notation $w =_G v$ is used to indicate that the words w and v represent the same element of G. If further w and v are exactly the same word, then they are said to be *visually equal*, denoted $w \equiv v$. Note that if w and v are both reduced, then $w \equiv v$ if and only if $w =_{F(\mathcal{A})} v$.

For any subset $\mathcal{R} \subset F(\mathcal{A})$, the *normal closure* of $\mathcal{R}$ in $F(\mathcal{A})$, denoted $\langle\langle \mathcal{R} \rangle\rangle$, is the smallest normal subgroup of $F(\mathcal{A})$ containing $\mathcal{R}$. Observe that a nontrivial element $w \in F(\mathcal{A})$ satisfies $w \in \langle\langle \mathcal{R} \rangle\rangle$ if and only if there exist $k \in \mathbb{N}$, $f_1, \dots, f_k \in F(\mathcal{A})$, $\varepsilon_1, \dots, \varepsilon_k \in \{\pm 1\}$, and

$R_1, \dots, R_k \in \mathcal{R}$ such that

$$w =_{F(\mathcal{A})} \prod_{i=1}^{k} f_i R_i^{\varepsilon_i} f_i^{-1} \tag{1.1}$$

If a group G is isomorphic to the quotient group $F(\mathcal{A})/\langle\langle\mathcal{R}\rangle\rangle$, then $\langle \mathcal{A} \mid \mathcal{R} \rangle$ is said to be a *presentation* of G. In this case, it is convenient to view G as being generated by $\mathcal{A}$. As such, $\mathcal{A}$ is called the *set of generators* of the presentation, while $\mathcal{R}$ is called the *set of relators*. If $\mathcal{A}$ is a finite set, then G is called *finitely generated*; if in addition $\mathcal{R}$ is finite, then G is called *finitely presented* and $\langle \mathcal{A} \mid \mathcal{R} \rangle$ is called a *finite presentation*.

Note that it is a consequence of the universal property of free groups that every group admits a presentation. Indeed, letting $\mathcal{A}$ be any generating set of G (one may take $\mathcal{A} = G$, for example), then there exists an epimorphism $\varphi : F(\mathcal{A}) \to G$; so, letting $\mathcal{R} = \ker \varphi$, it follows that G has presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$.

Given a nontrivial word w over $\mathcal{A} \cup \mathcal{A}^{-1}$, it is an immediate consequence of the above definitions that $w =_G 1$ if and only if there exists a representation of w as in (1.1). The minimal natural number k in such a representation is called the *area of w with respect to the presentation* $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$, denoted $\text{Area}_{\mathcal{P}}(w)$. This definition is extended to the case where $w =_{F(\mathcal{A})} 1$, in which case it is taken that $\text{Area}_{\mathcal{P}}(w) = 0$.

If $\mathcal{A}$ is a finite set and $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ is a presentation of the group G, then the *Dehn function of $\mathcal{P}$* is the function $\delta_{\mathcal{P}} : \mathbb{N} \to \mathbb{N}$ defined by

$$\delta_{\mathcal{P}}(n) = \max\{\text{Area}_{\mathcal{P}}(w) : w =_G 1,\ \|w\| \leq n\}$$

Dehn functions are taken up to the asymptotic equivalence relation $\sim$ on functions $\mathbb{N} \to \mathbb{N}$ induced by the preorder $\preccurlyeq$ given by $f \preccurlyeq g$ if and only if there exists $C > 0$ such that for all $n \in \mathbb{N}$,

$$f(n) \leq Cg(Cn) + Cn + C$$

Observe that $n \preccurlyeq f$ for any function $f : \mathbb{N} \to \mathbb{N}$. What's more, if f is a polynomial of degree d, then $f \sim n^d$.

Dehn functions were first defined and investigated by Gromov in [8]. There, it was shown that the Dehn function of a finite presentation tells one a good deal about the finitely presented group to which it is associated. For one thing, it is shown that the function is invariant (up to $\sim$) with respect to choice of finite presentation:

Theorem 1.1. *(Gromov [8]) If $\mathcal{P}$ and $\mathcal{S}$ are finite presentations for the finitely presented group G, then $\delta_{\mathcal{P}} \sim \delta_{\mathcal{S}}$.*

Therefore, it is well-founded to speak of 'the' *Dehn function of a finitely presented group G*, denoted δ_G, in reference to the Dehn function of any of its finite presentations.

The Dehn function proves to be a useful invariant for the study of finitely presented groups. For example, it is shown in [5] and [26] that the Dehn function is closely related to the solvability of the Word Problem of the group, i.e the ability to create an algorithm to determine whether or not a given word over the generators represents the trivial element in the group. Moreover, the Dehn function gives a characterization of the geometric quality of the hyperbolicity of a group:

Theorem 1.2. *(Gromov [8]) A finitely presented group G is word hyperbolic if and only if $\delta_G \sim n$*

Theorem 1.3. *(Gromov [8], Bowditch [4], Ol'shanskii [15]) A finitely presented group G is word hyperbolic if and only if $\delta_G \prec n^2$.*

Theorems 1.2 and 1.3 imply the existence of an 'isoperimetric gap', i.e a jump in possible Dehn functions between word hyperbolic groups (with linear Dehn function) and finitely presented groups with quadratic Dehn functions. This begs a class of questions:

Question 1: *While properties of hyperbolic groups abound, which such properties are satisfied by all groups with quadratic Dehn function?*

1.2 Infinite Torsion Groups

Given a group G, an element g is called *torsion* if there exists $n \in \mathbb{N} \setminus \{0\}$ such that $g^n = 1$. The minimal number n satisfying this equation is called the *order* of g and is denoted $|g|$. The group G is called *torsion* if each of its elements is itself torsion. If in addition the orders of the elements of G are bounded, then G is said to have *bounded exponent*.

The Burnside problem, perhaps the oldest problem in group theory dating back to 1902, asked whether all finitely generated torsion groups are finite. A variant of the problem, the Bounded Burnside problem, asked whether all finitely generated torsion groups of bounded exponent are finite.

For any $n > 1$ and any set $\mathcal{A}$, let $F(\mathcal{A})^n$ be the normal subgroup of $F(\mathcal{A})$ generated by all words of the form w^n. Then the group $F(\mathcal{A})/F(\mathcal{A})^n$ is the *free group relative the class of groups of exponent* n (this class is also known as the Burnside variety $\mathcal{B}_n$). This terminology is justified by the universal property of relatively free groups: If G is a group such that $g^n = 1$ for all $g \in G$ and G is generated by $\{g_i\}_{i \in I}$, then for $\mathcal{A} = \{a_i\}_{i \in I}$, there exists an epimorphism $\phi : F(\mathcal{A})/F(\mathcal{A})^n \to G$ such that $\phi(a_i F(\mathcal{A})^n) = g_i$ for all $i \in I$. For convenience, the group $F(\mathcal{A})/F(\mathcal{A})^n$ is called a *free Burnside group* and is denoted $B(\mathcal{A}, n)$, or simply $B(m, n)$ if $|\mathcal{A}| = m$.

Hence, the Bounded Burnside problem essentially asks whether there exists $m, n \in \mathbb{N}$ such that $B(m, n)$ is infinite (and, if so, for which choices of m, n).

The first solution to the Burnside problem was provided by Golod and Shafarevich in 1964 [7], when they constructed examples of infinite finitely generated torsion subgroups. However, the groups constructed for this purpose were shown to have unbounded exponent, and so do not provide a solution to the Bounded Burnside problem.

In 1968, Novikov and Adian [13] provided the first solutions to the Bounded Burnside problem when they showed that $B(m, n)$ is infinite for all $m > 1$ and $n \geq 4381$ odd. Adian later improved the bound placed on the odd integer n to $n \geq 665$ in 1978 [1], and then $n \geq 101$ in 2015 [2]. In 1982, Ol'shanskii provided a simpler geometric proof that $B(m, n)$ is infinite for $m > 1$ and sufficiently large odd n (say $n > 10^{10}$), as well as proving the existence of the so-called Tarski monster groups [14].

The seminal result for even values of n came in 1994 due to S.V. Ivanov [9], who proved that $B(m,n)$ is infinite for $m > 1$ and $n \geq 2^{48}$ that is either odd or divisible by 2^9. For simple reference in what follows, the set of natural numbers n in Ivanov's setting is denoted $\mathbb{N}^*$. The bound was later improved by Lysenok in 1996 [12], where the analogous statement was proved for $n \geq 8000$ and divisible by 2^4.

Though the infinite torsion groups arising from these constructions were all shown to have solvable Word Problem, it was also shown that they cannot be finitely presented. As such, one cannot speak of the Dehn function of these free Burnside groups.

However, an infinite free Burnside group may be isomorphic to a proper subgroup of a finitely presented group, begging the following question of Ol'shanskii:

Question 2: *If G is a finitely presented group containing a finitely generated infinite torsion subgroup H, then what can be said about δ_G? What about the specific case where $H = B(m,n)$ for some m, n?*

1.3 Main Results

The first partial answer to Question 2 is due to Ghys and de la Harpe in 1991 [6], and shows that Question 2 can be thought of as one belonging to the class of questions detailed in Question 1:

Theorem 1.4. *(Ghys-de la Harpe [6])*

If H is word hyperbolic and T is a torsion subgroup of H, then T is finite.

Conversely, in 2000 Ol'shanskii and Sapir [19] constructed a finitely presented group G satisfying $\delta_G \preccurlyeq n^{10}$ and containing a subgroup isomorphic to the finitely generated infinite torsion subgroup $B(m,n)$ for any $m > 1$ and n sufficiently large and odd.

The following provides an optimal such example:

Theorem A. *For $m > 1$ and $n \in \mathbb{N}^*$, there exists a finitely presented group $G_{m,n}$ with quadratic Dehn function into which the free Burnside group $B(m,n)$ embeds. In particular, there exists a* *finitely presented group G with quadratic Dehn function containing a finitely generated infinite torsion subgroup.*

If $\mathcal{A} = \{a_i\}_{i=1}^{\infty}$, then denote the free Burnside group $B(\mathcal{A}, n)$ simply by $B(\infty, n)$. For n sufficiently large and odd, Shirvanyan exhibited an embedding of $B(\infty, n)$ in $B(2, n)$ [29], while Ivanov and Ol'shanskii exhibited such an embedding for $n \geq 2^{48}$ and divisible by 2^9 [10].

Thus, taking $G_n = G_{2,n}$, Theorem A immediately implies the following corollary:

Corollary A. *For all $n \in \mathbb{N}^*$, there exists a finitely presented group G_n with quadratic Dehn function into which the free Burnside group $B(\infty, n)$ embeds. In particular, for all $m \in \mathbb{N}$, G_n contains a subgroup isomorphic to $B(m, n)$.*

Given metric spaces (X, d_X) and (Y, d_Y), a map $f : X \to Y$ is called a *quasi-isometric embedding* if there exist $C \geq 1$ and $K \geq 0$ such that for all $x_1, x_2 \in X$,

$$\frac{1}{C} d_X(x_1, x_2) - K \leq d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2) + K$$

If $K = 0$, then f is called a *bi-Lipschitz embedding*. Note that, unlike a quasi-isometric embedding, a bi-Lipschitz embedding is necessarily an embedding (as quasi-isometric embeddings need not be injective).

Let G be a finitely generated group with finite generating set X and define the function $d_X : G \times G \to \mathbb{N}$ by $d_X(g_1, g_2) = |g_1^{-1} g_2|_X$. Then d_X is a metric induced by the norm $|\cdot|_X$.

Now, suppose G and H are two finitely generated groups with finite generating sets X and Y, respectively. Further, suppose there exists a monomorphism $\varphi : G \to H$. Then, it is clear that φ is a bi-Lipschitz embedding if and only if there exists a $C \geq 1$ such that for all $g \in G$,

$$\frac{1}{C} |g|_X \leq |\varphi(g)|_Y \leq C |g|_X$$

Letting $C_1 = \max\{|\varphi(x)|_Y : x \in X\}$, note that $|\varphi(g)|_Y \leq C_1 |g|_X$ for all $g \in G$. Hence, φ is a bi-Lipschitz embedding if and only if there exists $C_2 \geq 1$ such that for any $g \in G$, $|g|_X \leq C_2 |\varphi(g)|_Y$.

The concept of quasi-isometry, and so of bi-Lipschitz maps, is especially important and useful in geometric group theory. For example, Theorem 1.1 is originally stated in the more general setting of two finite presentations that are quasi-isometric with respect to the induced word metrics.

Theorem B. *The embedding given in Theorem A is a bi-Lipschitz embedding (and so a quasi-isometric embedding) of the free Burnside group $B(m, n)$ into the finitely presented group $G_{m,n}$.*

A subgroup G of a group H satisfies the *Congruence Extension Property* (CEP) if for any epimorphism $\varepsilon : G \to G_1$, there exists an epimorphism $\bar{\varepsilon} : H \to H_1$ for some group H_1 containing G_1 as a subgroup and such that the restriction of $\bar{\varepsilon}$ to G is ε. In this case, we write $G \leq_{CEP} H$ and say that G is a CEP-subgroup of H or that G is CEP-embedded in H.

There are two convenient reformulations of the definition of CEP:

(1) G is a CEP-subgroup of H if and only if for any normal subgroup $N \triangleleft G$, there exists a normal subgroup $M \triangleleft H$ such that $M \cap G = N$

(2) G is a CEP-subgroup of H if and only if for any subset $S \subseteq G$, $G \cap \langle\langle S\rangle\rangle^G = \langle\langle S\rangle\rangle^H$ (where the normal closure of a subset T in a group K is denoted $\langle\langle T\rangle\rangle^K$).

It is clear from (1) that any retract of a group is a CEP-subgroup and that $\leq_{CEP}$ is a transitive relation. However, some examples are less obvious. For example, Sonkin proved that for sufficiently large odd n, there exists a CEP-embedding of $B(\infty, n)$ into the group $B(2, n)$ [30].

Theorem C. *The embedding given in Theorem A is a CEP-embedding of the free Burnside group $B(m, n)$ into the finitely presented group $G_{m,n}$.*

As $B(m, n)$ is a retract of $B(\infty, n)$, Theorem C immediately implies the following corollary:

Corollary C. *For all odd $n \in \mathbb{N}^*$, there exists a finitely presented group G_n with quadratic Dehn function such that for any $m \geq 2$, there is a CEP-embedding of $B(m, n)$ into G_n. Moreover, G_n contains a CEP-subgroup isomorphic to $B(\infty, n)$.*

1.4 Summary of Contents

Chapter 2 functions to introduce perhaps the most useful tool in the study of Dehn functions: The van Kampen diagram. This leads naturally to the more advanced study of graded small-cancellation theory in Chapter 3, culminating with the proof of Lemma 3.8 which proves vital to the proofs of the main theorems.

As in [18] and [25], the construction of the groups of interest is through S-machines. S-machines were first introduced by Sapir in [26]; for a formal definition, see Chapter 4. Similar to the construction in those settings, we first create several auxiliary machines $\mathbf{M}_1 - \mathbf{M}_4$ satisfying some desirable properties (see Chapter 5). The chief properties on which we base this construction are the following:

(a) The language of accepted inputs is a set of relators for a presentation of the free Burnside group $B(m, n)$ (see Lemma 5.33);

(b) The length of any accepting computation of a word u^n is linearly bounded by $\|u\|$ (see Lemma 5.32);

(c) The majority of an accepting computation is spent on one particular step (see Lemma 5.34); and

(d) The length of a computation in a specific class of bases is linearly bounded by the length of the initial or terminal admissible words (see Lemma 5.37).

Many copies of the machine $\mathbf{M}_4$ are then 'concatenated' to form our main machine $\mathbf{M}$, a process that resembles the construction of the groups of interest in [18] and [25]. However, unlike in those sources, one copy of $\mathbf{M}_4$ is deemed 'special' and is operated upon in a different manner as the other copies, causing a distinct non-uniformity.

The purpose of this lack of symmetry is to allow our machine to accept two configurations which differ only in the insertion/deletion of an accepted input. As all accepted configurations are trivial in the group $G(\mathbf{M})$ associated to the machine, this implies the relation $w = 1$ for any word w over $\mathcal{A} \cup \mathcal{A}^{-1}$ of the input sector that represents the trivial element of $B(\mathcal{A}, n)$.

Conversely, this asymmetry is also the source of several new obstacles not faced in [18] or [25]. For example, many statements in Chapter 6 are devoted to understanding the relationship between computations of one copy of $\mathbf{M}_4$ and computations of the standard base of $\mathbf{M}$ (for example, see Lemmas 6.12 and 6.13), a relationship that would be trivial had the rules operated with symmetry.

In Chapters 7-12, we study diagrams over the groups associated to the S-machine $\mathbf{M}$, culminating in the proof of Theorem A. The general method of study follows a similar path to those followed in [18] and [25], but with one major change: The consideration of the groups $M_\Omega(\mathbf{M})$ and $G_\Omega(\mathbf{M})$ constructed by the addition of extra relations, called a-relations, to the groups $M(\mathbf{M})$ and $G(\mathbf{M})$, respectively. The set of relators Ω corresponding to the a-relations consist of words over the alphabet of the input sector and contains the set $\mathcal{S}$ of all words that represent the trivial element of $B(\mathcal{A}, n)$.

The cells of a diagram over $M_\Omega(\mathbf{M})$ or $G_\Omega(\mathbf{M})$ corresponding to elements of Ω, referred to as a-cells, are invaluable to the proof that φ is an embedding (see Lemma 12.3) but cause a new obstacle in virtually every diagrammatic consideration. For example, the consideration of rim θ-bands of a diagram must be replaced with the consideration of quasi-rim θ-bands, i.e a band that may have a-cells between it and the boundary (see Lemma 9.16).

The proof of Theorem B is presented Chapter 13. Its makeup is similar to the diagrammatic arguments presented in Section 10, but is unique to this setting in that it has no analogue in [18] or [25].

We conclude with the proof of Theorem C in Chapter 14. The proof is a consequence of the arguments pertaining to minimal diagrams introduced in Chapters 7-10.

Finally, we mention here the importance of the group $B(m, n)$ to this construction and proof. In the context of the proof of Theorem A, it is clear that the following two properties of the embedded group were necessary for the proof to follow: The existence of a presentation of the group whose relators satisfy some linear bound as in (b) above and the existence of another presentation of the group such that any van Kampen diagram over this presentation satisfies some quadratic bound as in Lemma 3.8. However, there is a third, more subtle requirement: In the proof of Lemma 12.6, it is essential that the relators are periodic. Due to this demand, that we are studying a group in the Burnside variety is crucial to our construction.

Chapter 2

van Kampen diagrams

A vital tool for many of the arguments to come is the concept of van Kampen diagrams over group presentations, a notion introduced by its namesake in 1933 [31].

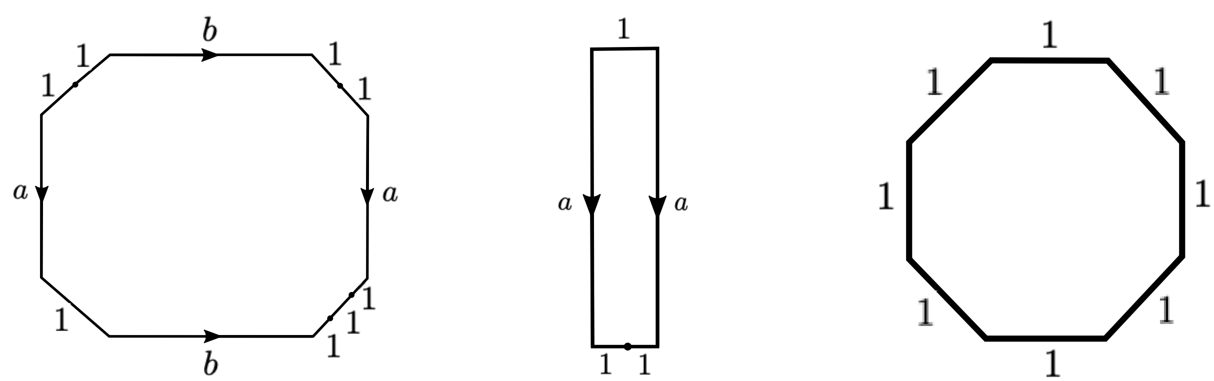
Let G be a group with presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$. Suppose Δ is an oriented 2-complex homeomorphic to a disk equipped with a *labelling function*, i.e a function $\text{Lab} : E(\Delta) \to \mathcal{A} \cup \mathcal{A}^{-1} \cup \{1\}$ which satisfies $\text{Lab}(e^{-1}) \equiv \text{Lab}(e)^{-1}$ for any edge $e \in E(\Delta)$ (with, of course, $1^{-1} \equiv 1$). The label of a path in Δ is defined in the obvious way, that is, $\text{Lab}(e_1 \dots e_n) \equiv \text{Lab}(e_1) \dots \text{Lab}(e_n)$. For any edge e in Δ, e is called a 0-edge if $\text{Lab}(e) \equiv 1$; otherwise, e is called an $\mathcal{A}$-edge.

Suppose that for each cell Π of Δ, one of the following is true:

(1) omitting the label of any zero edges, $\text{Lab}(\partial\Pi)$ (defined by reading starting at some vertex) is visually equal to a cyclic permutation of $R^{\pm 1}$ for some $R \in \mathcal{R}$

(2) $\partial\Pi$ consists of 0-edges and exactly two $\mathcal{A}$-edges e and f, with $\text{Lab}(e) \equiv \text{Lab}(f^{-1})$

(3) $\partial\Pi$ consists only of 0-edges.



(a) $\mathcal{R}$-cell corresponding to the relator $R = aba^{-1}b^{-1}$.
(b) 0-cell of type (2), $a \in \mathcal{A}$.
(c) 0-cell of type (3).

Figure 2.1: Cells in van Kampen diagrams

Then Δ is called a *(disk) van Kampen diagram* (or simply a *disk diagram*) over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$. The cells satisfying condition (1) above are called $\mathcal{R}$-cells, while the others are called 0-cells.

Given a disk diagram Δ, note that $\text{Lab}(\partial\Delta)$ (or indeed the label of any closed path in Δ) is not well-defined. However, it is convenient to understand that such labels are always considered up to cyclic permutation.

The following statement, called van Kampen's lemma, will serve as a fundamental tool for the combinatorial study throughout the rest of this manuscript:

Lemma 2.1. *(van Kampen, 1933 [31]) A word W over $\mathcal{A} \cup \mathcal{A}^{-1}$ satisfies $W =_G 1$ if and only if there exists a disk diagram Δ over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ such that Lab$(\partial\Delta) \equiv W$.*

Proof. Suppose Δ is a disk diagram over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ satisfying $W \equiv \text{Lab}(\partial\Delta)$. If Δ consists of one cell, then either $W =_{F(\mathcal{A})} 1$ (if the cell is a 0-cell) or W is a cyclic permutation of $R^{\pm 1}$ for some $R \in \mathcal{R}$ (if the cell is an $\mathcal{R}$-cell). In either case, $W =_G 1$. Otherwise, supposing Δ consists of more than one cell, let t be some path that cuts Δ into two disk diagrams Δ_1 and Δ_2 with fewer cells. Letting $\partial\Delta_1 = p_1 t$ and $\partial\Delta_2 = t^{-1}p_2$ such that p_1 and p_2 are subpaths of $\partial\Delta$, it follows that $\text{Lab}(p_1p_2)$ is a cyclic permutation of $\text{Lab}(\partial\Delta)$. Inducting on the number of cells in a disk diagram, one can assume that $\text{Lab}(\partial\Delta_j) =_G 1$ for $j = 1, 2$, so that

$$\text{Lab}(p_1p_2) =_{F(\mathcal{A})} \text{Lab}(p_1tt^{-1}p_2) \equiv \text{Lab}(\partial\Delta_1)\text{Lab}(\partial\Delta_2) =_G 1$$

Hence, $W \equiv \text{Lab}(\partial\Delta) =_G 1$.

Now suppose $W =_G 1$ and let $W =_{F(\mathcal{A})} \prod_{i=1}^{k} f_i R_i^{\varepsilon_i} f_i^{-1}$ be some representation of W as in (1.1).

Let s_i be the diagram obtained by taking a polygonal arc labelled by f_i and attaching to its end a polygon labelled by $R_i^{\varepsilon_i}$. Then, consider the 'lollipop' diagram Δ_0 obtained by attaching all s_i at a single vertex so that they are read consecutively (see Figure 2.2). However, this diagram is not homeomorphic to a disk; in order to achieve this, one must add 0-cells to 'thicken' the arcs representing the sticks of the lollipop, yielding a disk diagram Δ_0. By construction,

$\mathrm{Lab}(\partial\Delta_0) =_{F(\mathcal{A})} \prod_{i=1}^{k} f_i R_i^{\varepsilon_i} f_i^{-1}.$

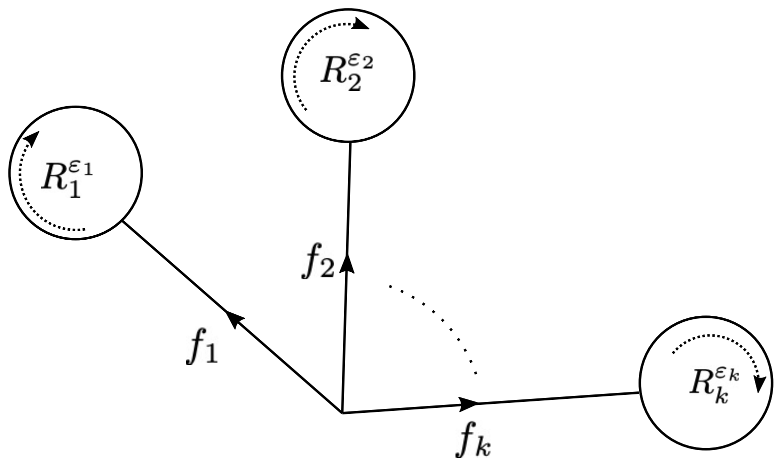


Figure 2.2: Lollipop diagram

Suppose $\partial\Delta_0$ contains a subpath of 0-edges. Then the addition of one 0-cell results in a disk diagram with this segment removed (see Figure 2.3). Doing this for all such subpaths of $\partial\Delta_0$ results in a disk diagram Δ_1 such that $\mathrm{Lab}(\partial\Delta_1)$ is a word over $\mathcal{A}\cup\mathcal{A}^{-1}$ satisfying $\mathrm{Lab}(\partial\Delta_1) =_{F(\mathcal{A})} \prod_{i=1}^{k} f_i R_i^{\varepsilon_i} f_i^{-1}$

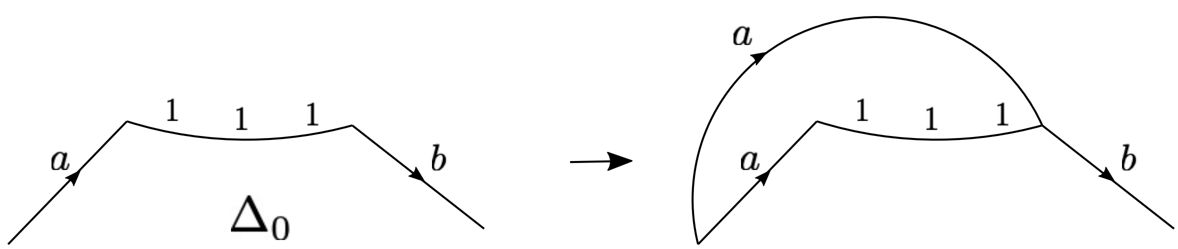


Figure 2.3: The removal of 0-edges

Thus, it suffices to show that mutually inverse pairs of adjacent letters can be removed or inserted into the label of the boundary. But this can also be done by the addition of 0-cells, as evidenced by Figures 2.4 and 2.5:

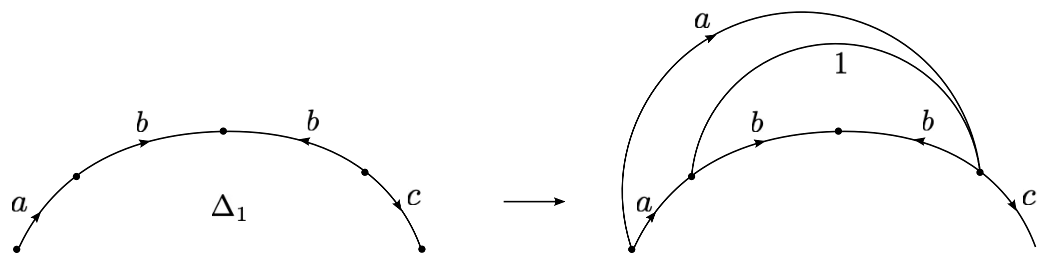


Figure 2.4: The cancellation of mutually inverse letters

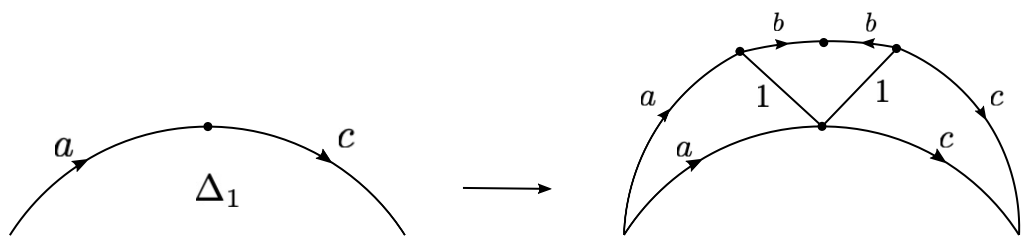


Figure 2.5: The insertion of mutually inverse letters

□

The *area* of a disk diagram Δ, denoted Area(Δ), is the number of $\mathcal{R}$-cells it contains. Further, for any word W satisfying $W = 1$ in G, the *area* of W, Area(W), is the minimal area of a diagram Δ satisfying Lab$(\partial\Delta) \equiv W$.

A *0-refinement* of a disk diagram Δ is a disk diagram Δ' obtained from Δ by the insertion of 0-edges or 0-cells. Note that a 0-refinement has the same area as the diagram from which it arises.

Note that each of the operations performed to the 'lollipop' diagram in the proof of van Kampen's Lemma is a 0-refinement. As a result, the following is an immediate consequence:

Lemma 2.2. *Letting $\mathcal{P} = \langle \mathcal{A} \mid \mathcal{R} \rangle$ be the presentation of G, for any word W over $\mathcal{A} \cup \mathcal{A}^{-1}$ satisfying $W =_G 1$, Area$(W) =$ Area$_{\mathcal{P}}(W)$.*

Let Δ be a disk diagram and Π_1, Π_2 be two $\mathcal{R}$-cells in Δ. Suppose there exists a simple path t between the vertices O_1, O_2 of Π_1, Π_2, respectively, such that:

- Lab$(t) = 1$ in $F(\mathcal{A})$, and
- Lab$(\partial\Pi_1)$ read starting at O_1 is mutually inverse to Lab$(\partial\Pi_2)$ read starting at O_2

Then Π_1 and Π_2 are called *cancellable* in Δ.

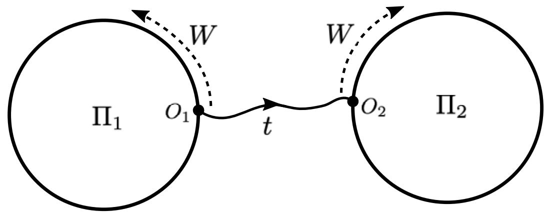


Figure 2.6: Cancellable cells

This term is justified by the ability to 'remove' the cells Π_1 and Π_2 from Δ without affecting its contour label, yielding a disk diagram Δ' satisfying Lab$(\partial\Delta') \equiv$ Lab$(\partial\Delta)$ and such that Area$(\Delta') <$ Area(Δ).

Naturally, a disk diagram is called *reduced* if it has no pair of cancellable cells. By simply removing pairs of cancellable cells, any disk diagram over a presentation can be made reduced. This immediately implies a strengthened version of van Kampen's lemma:

Lemma 2.3. *A word W over $\mathcal{A} \cup \mathcal{A}^{-1}$ satisfies $W =_G 1$ if and only if there exists a reduced disk diagram Δ over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ with Lab$(\partial\Delta) \equiv W$.*

An annular (Schupp) diagram over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ is defined in the analogous way. It is then an immediate consequence of van Kampen's lemma that two words W and V are conjugate in G if and only if there exists a reduced annular diagram Δ with contour components p and q satsifying Lab$(p) \equiv W$ and Lab$(q) \equiv V^{-1}$.

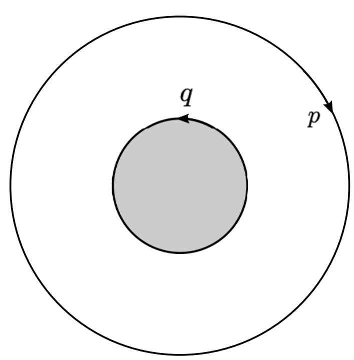


Figure 2.7: Annular diagram

Chapter 3

Graded small-cancellation

3.1 Graded maps on a disk or annulus

The definitions and statements presented in this chapter can be found in [16] and [9]. Those relevant to the proof of Lemma 3.8 are restated here, with reference given in place of proofs.

A *map* Δ is a finite oriented planar graph on a disk which subdivides the surface into polygonal cells. In particular, by 'forgetting' the labelling, one can interpret a van Kampen diagram as a map.

A map Δ is called *graded* if each cell Π in Δ is assigned a nonnegative integer $r(\Pi)$ called its *rank*. The map Δ is called a *map of rank at most k* if all its cells have rank $\leq k$. The minimal k for which Δ is a map of rank at most k is called the *rank* of Δ and denoted $r(\Delta)$.

For $r(\Delta) = k$, the *type* of Δ, $\tau(\Delta)$, is the $(k+2)$-vector $(r(\Delta), \tau_0, \dots, \tau_k)$, where τ_i is the number of cells of rank $k-i$ in Δ. The types of maps are ordered lexicographically, i.e for two maps Δ and Γ with $\tau(\Delta) = (r(\Delta), \tau_0, \dots, \tau_k)$ and $\tau(\Gamma) = (r(\Gamma), \sigma_0, \dots, \sigma_\ell)$, $\tau(\Delta) \leq \tau(\Gamma)$ if the following three conditions hold:

- $r(\Delta) \leq r(\Gamma)$;
- if $r(\Delta) = r(\Gamma)$, then $\tau_0 \leq \sigma_0$;
- for $1 \leq i \leq r(\Delta)$, if $r(\Delta) = r(\Gamma)$ and $\tau_j = \sigma_j$ for all $j < i$, then $\tau_i \leq \sigma_i$.

For simplicity, the cells of rank 0 in a graded map are called *0-cells*. All other cells are called $\mathcal{R}$*-cells* (even though an alphabet $\mathcal{R}$ is not specified).

The edges of the graph are divided into two disjoint sets, called the *0-edges* and the $\mathcal{A}$*-edges*. The *length* of a path p in a graded map Δ, denoted $|p|$, is the number of $\mathcal{A}$-edges that comprise it. In particular, for $\partial\Pi$ the contour of a cell, $|\partial\Pi|$ is called the *perimeter* of Π.

Motivated by the definition of van Kampen diagrams, the following three facts are assumed about graded maps:

(1) the inverse edge of a 0-edge is also a 0-edge

(2) the contour of a 0-cell either consists entirely of 0-edges or of exactly two $\mathcal{A}$-edges in addition to a number of 0-edges

(3) if Π is an $\mathcal{R}$-cell, then $|\partial\Pi| > 0$

If Δ is a graded map and Γ is a subspace homeomorphic to a disk bounded by some edgepath of Δ, then Γ is called a *submap* of Δ.

It is assumed that the contour of a graded map has a fixed decomposition into at most eight distinct parts. In particular, if Δ is a graded map, then there is a standard factorization $p_1 \dots p_k$ of $\partial\Delta$, with $k \leq 8$ and each p_i called a *section* of the contour.

3.2 0-Bonds and 0-contiguity submaps

Let Δ be a graded map and Π be a 0-cell whose contour contains exactly two $\mathcal{A}$-edges, e_1 and e_2. Then the pair of edges e_1, e_2^{-1} are called *immediately adjacent* (as is the pair e_1^{-1}, e_2). Two edges e and f of Δ are then said to be *adjacent* if there exists a sequence of edges $e = e_1, e_2, \dots, e_{k+1} = f$ such that e_i and e_{i+1} are immediately adjacent for $i = 1, \dots, k$.

Let Δ be a graded map with adjacent edges e and f. Suppose e belongs to the contour of the $\mathcal{R}$-cell Π_1 and f^{-1} to the contour of some $\mathcal{R}$-cell Π_2. Per the definition, set $e = e_1, \dots, e_{k+1} = f$ with 0-cells $\pi_1, \dots, \pi_k$ such that the only two $\mathcal{A}$-edges of $\partial\pi_i$ are e_i^{-1} and e_{i+1}.

We can then write $\partial\pi_i = e_i^{-1} p_i e_{i+1} s_i$ for $i = 1, \dots, k$ such that $|p_i| = |s_i| = 0$. With the aid of 0-refinement, we can assume that $p = p_1 \dots p_k$ and $s = s_k \dots s_1$ are simple paths such that each intersects Π_1, Π_2 only on its endpoints.

Then, the submap Γ with contour $p^{-1} e s^{-1} f^{-1}$ consisting of the cells $\pi_1, \dots, \pi_k$ is called a *0-bond* between Π_1 and Π_2. The edges e and f^{-1} are called the *contiguity arcs* of the 0-bond Γ and p and s the *side arcs*.

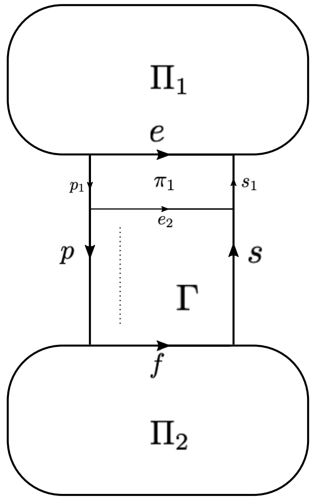


Figure 3.1: A 0-bond between two $\mathcal{R}$-cells

Similarly, if e and f are adjacent edges with e belonging to the contour of some $\mathcal{R}$-cell Π and f^{-1} belonging to some section q of the contour, then a 0-bond between Π and q is defined. A 0-bond between two sections of the contour is defined analogously.

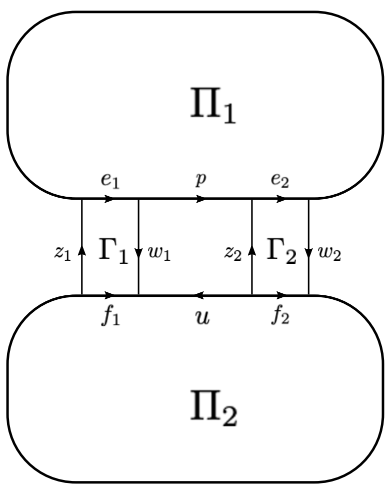


Figure 3.2: A 0-contiguity submap between two $\mathcal{R}$-cells

Now suppose e_1, f_1 and e_2, f_2 are two pairs of adjacent edges such that e_1 and e_2 belong to the contour of some $\mathcal{R}$-cell Π_1 and f_1^{-1}, f_2^{-1} to some $\mathcal{R}$-cell Π_2. Then, construct two 0-bonds, Γ_1 and Γ_2, between the two pairs, with $\partial\Gamma_i = z_i e_i w_i f_i^{-1}$. If $\Gamma_1 = \Gamma_2$, set $\Gamma = \Gamma_1$. Otherwise, there exist subpaths y_1 and y_2 of $\partial\Pi_1$ and $\partial\Pi_2$, respectively, such that $y_1 = e_1 p e_2$ and $y_2 = f_2^{-1} u f_1^{-1}$ (or $y_1 = e_2 p e_1$ and $y_2 = f_1^{-1} u f_2^{-1}$). Then let Γ be the submap with contour $z_1 y_1 w_2 y_2$ (or $z_2 y_1 w_1 y_2$).

If Γ does not contain Π_1 or Π_2, then Γ is called a *0-contiguity submap* of Π_1 to Π_2. In this case, y_1 and y_2 are called the *contiguity arcs* of Γ, denoted $y_i = \Gamma \wedge \Pi_i$. The paths z_1 and w_2 (or z_2 and w_1) are called the *side arcs* of Γ. Note that both side arcs have zero length. The ratio $|y_1|/|\partial\Pi_1|$ is called the *degree of contiguity* of Π_1 to Π_2 with respect to Γ and is denoted (Π_1, Γ, Π_2). Similarly, $(\Pi_2, \Gamma, \Pi_1) = |y_2|/|\partial\Pi_2|$ is the degree of contiguity of Π_2 to Π_1.

Note, however, that if $\Pi_1 = \Pi_2$, then $\Gamma \wedge \Pi_1$ represents two distinct arcs of $\partial\Pi_1$ and (Π_1, Γ, Π_1) a pair of numbers.

As with 0-bonds, a 0-contiguity submap between an $\mathcal{R}$-cell and a section of $\partial\Delta$ is similarly defined, as is a 0-contiguity submap between two sections of $\partial\Delta$. The contiguity arcs, side arcs, and degree of contiguity of such 0-contiguity submaps are defined in the same way as as above; for example, if Γ is a 0-contiguity submap between an $\mathcal{R}$-cell Π and a section q of the contour of $\partial\Delta$, then the degree of contiguity of q to Π is $(q, \Gamma, \Pi) = |\Gamma \wedge q|/|q|$.

Two 0-contiguity submaps Γ_1 and Γ_2 are *disjoint* if they have no common cells, their contiguity arcs have no common points, and their side arcs have no common points.

3.3 Bonds and contiguity submaps

In this section, $\varepsilon \in (0, 1)$ is taken to be a fixed constant. For the moment, one can think of this number as 'sufficiently small', with this interpretation made precise in the Section 3.5.

Set $k > 0$ and suppose the terms j-bond and j-contiguity submap have been defined for all $0 \leq j < k$. Assume further that contiguity arcs, side arcs, and degrees of contiguity are defined for j-contiguity submaps in a way similar to how they were defined for 0-contiguity submaps.

Two submaps Γ_1, Γ_2 such that Γ_i is a j_i-contiguity submap for $j_i < k$ are called *disjoint* if they have no common cells, their contiguity arcs have no common points, and their side arcs have no common points. Note that this definition agrees with that given for the case $j_1 = j_2 = 0$.

Let π, Π_1, and Π_2 be cells of a graded map Δ, $\Pi_1 \neq \Pi_2$, satisfying the following:

(1) $r(\pi) = k$, $r(\Pi_i) > k$ for $i = 1, 2$,

(2) there are disjoint submaps Γ_1, Γ_2 such that Γ_i is a j_i-contiguity submap of π to Π_i for $j_i < k$, Π_1 is not contained in Γ_2, and Π_2 is not contained in Γ_1,

(3) $(\pi, \Gamma_i, \Pi_i) \geq \varepsilon$ for $i = 1, 2$.

For $i = 1, 2$, let $\partial\Gamma_i = v_i s_i$ for $v_i = \Gamma_{i\wedge}\pi$ and $\partial\pi = u_1 v_1 u_2 v_2$. Letting Γ be the submap with contour $s_1 u_1^{-1} s_2 u_2^{-1}$, Γ is the *k-bond* between Π_1 and Π_2 defined by the contiguity submaps Γ_1 and Γ_2 with *principal cell* π. The *contiguity arc* of Γ to Π_i is defined to be $\Gamma_{i\wedge}\Pi_i$ and denoted $\Gamma{\wedge}\Pi_i$. The *side arcs* of Γ are defined in the obvious way.

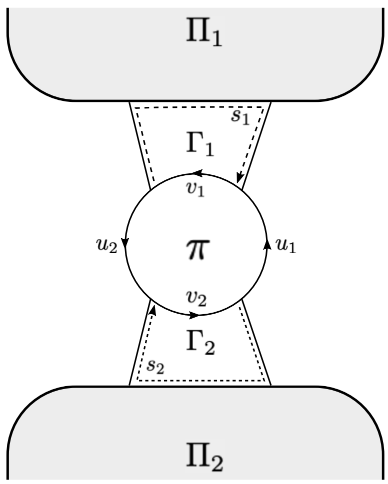


Figure 3.3: A k-bond between two $\mathcal{R}$-cells

A k-bond between an $\mathcal{R}$-cell and a section of the contour or between two distinct sections of the contour is defined similarly.

Suppose Γ_1 is a k-bond between two cells Π_1 and Π_2 and Γ_2 is a j-bond between Π_1 and Π_2 for $j \leq k$. If $\Gamma_1 = \Gamma_2$, then set $\Gamma = \Gamma_1$. Otherwise, if Γ_1 and Γ_2 are disjoint, then set $\partial\Gamma_i = z_i v_i w_i s_i$ for $v_i = \Gamma_{i\wedge}\Pi_1$ and $s_i = \Gamma_{i\wedge}\Pi_2$. Then set y_1 as a subpath of $\partial\Pi_1$ of the form $v_1 v v_2$ (or $v_2 v v_1$) and y_2 as a subpath of $\partial\Pi_2$ of the form $s_2 s s_1$ (or $s_1 s s_2$). Setting Γ as the submap with contour $z_1 y_1 w_2 y_2$ (or $z_2 y_1 w_1 y_2$), if Γ does not contain Π_1 or Π_2, then it is called the *k-contiguity submap* of Π_1 to Π_2 defined by the bonds Γ_1 and Γ_2. As with previous definitions, $y_i = \Gamma{\wedge}\Pi_i$ is called the *contiguity arc* of Γ to Π_i, z_1 and w_2 (or w_1 and z_2) are called the *side arcs* of Γ, $(\Pi_1, \Gamma, \Pi_2) = |y_1|/|\Pi_1|$ is called the *degree of contiguity* of Π_1 to Π_2 with respect to Γ.

A k-contiguity submap between an $\mathcal{R}$-cell and a section of the contour is defined similarly, as is a k-contiguity submap between two sections of the contour.

The number k is often omitted when referring to k-contiguity submaps, so that there will be reference merely to a contiguity submap. Further, if Γ is a contiguity submap between Π_1 and Π_2 and $\partial\Gamma = p_1q_1p_2q_2$ with $q_i = \Gamma \wedge \Pi_i$, then $\partial(\Pi_1, \Gamma, \Pi_2)$ denotes the *standard decomposition* $p_1q_1p_2q_2$.

3.4 Graded Presentations

Given an alphabet $\mathcal{A}$, let $\{\mathcal{S}_i\}_{i=1}^{\infty}$ be a collection of subsets of $F(\mathcal{A})$ such that if $W \in \mathcal{S}_i$ and V is a cyclic permutation of W or W^{-1}, then $V \notin \mathcal{S}_j$ for any $j \neq i$.

Set $\mathcal{R}_j = \cup_{i=1}^{j}\mathcal{S}_i$ for $j \geq 1$, $\mathcal{R}_0 = \emptyset$, and $\mathcal{R} = \cup_{i=1}^{\infty}\mathcal{S}_i$. Further, define $G(j) = \langle \mathcal{A} \mid \mathcal{R}_j \rangle$ for all $j \geq 0$. Note that $G(0) \cong F(\mathcal{A})$.

Then $\langle \mathcal{A} \mid \mathcal{R} \rangle$ is called a *graded presentation* for the group $G = G(\infty)$.

The words in $\mathcal{S}_i$ are called the *relators of rank* i. For words X, Y over $\mathcal{A}$, if $X = Y$ in $G(i)$, then X and Y are said to be *equal in rank* i, with this relation denoted $X \overset{i}{=} Y$.

Given a disk diagram Δ over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$, let Π be an $\mathcal{R}$-cell such that $\text{Lab}(\partial\Pi)$ is a cyclic permutation of a relator of rank i (or the inverse of such a relator). Then Π is called a *cell of rank* i, denoted by the representative notation $r(\Pi) = i$. Naturally, the 0-cells of Δ are called cells of rank 0.

Note that if one forgets the labelling function of a disk diagram Δ over a graded presentation, then Δ is a graded map (with the ranks of cells assigned in the same way). A diagram satisfying this property is called a *graded disk diagram*. It is then natural to define the *rank* and *type* of a graded disk diagram as the rank and type of the underlying map.

Let Δ be a graded disk diagram over $\langle \mathcal{A} \mid \mathcal{R} \rangle$ containing $\mathcal{R}$-cells Π_1, Π_2 with $r(\Pi_1) = r(\Pi_2) = j$. Suppose there exists a 0-refinement Δ' of Δ with copies Π_1', Π_2' of Π_1, Π_2, respectively, and a simple path t in Δ' between vertices O_1, O_2 of Π_1', Π_2', respectively, such that:

- $\text{Lab}(t) \overset{j-1}{=} 1$ and

- $\text{Lab}(\partial\Pi_1')$ read starting at O_1 is mutually inverse to $\text{Lab}(\partial\Pi_2')$ read starting at O_2.

Then Π_1 and Π_2 are called a *j-pair* in Δ.

This generalizes the concept of cancellable cells in a disk diagram over a presentation: If Δ is a graded disk diagram over a graded presentation with a j-pair Π_1, Π_2, then one can 'remove' Π_1 and Π_2 from Δ at the cost of cells of rank $\leq j-1$, producing a graded disk diagram Δ'' over the same presentation with $\text{Lab}(\partial\Delta'') \equiv \text{Lab}(\partial\Delta)$ and $\tau(\Delta'') < \tau(\Delta)$.

A graded disk diagram Δ over $\langle \mathcal{A} \mid \mathcal{R} \rangle$ is called *reduced* if for any graded disk diagram Γ over $\langle \mathcal{A} \mid \mathcal{R} \rangle$ satisfying $\text{Lab}(\partial\Delta) \equiv \text{Lab}(\partial\Gamma)$, the inequality $\tau(\Delta) \leq \tau(\Gamma)$ is satisfied. Similar to reduced disk diagrams over general presentations, one can make any graded disk diagram reduced simply by the removal of j-pairs (for varying j). As a result, van Kampen's Lemma can again be strengthened: Given a graded presentation $G = \langle \mathcal{A} \mid \mathcal{R} \rangle$, a word W over $\mathcal{A}$ represents the identity in G if and only if there exists a reduced graded disk diagram Δ over the presentation with $\text{Lab}(\partial\Delta) \equiv W$.

Graded annular diagrams are defined similarly.

3.5 Auxiliary parameters

The arguments presented through the rest of this section rely on the *lowest parameter principle* introduced in [16]. For this, we introduce the relation $>>$ on parameters defined as follows.

If $\alpha_1, \alpha_2, \dots, \alpha_k$ are (positive) parameters with $\alpha_1 >> \alpha_2 >> \cdots >> \alpha_k$, then for all indices $2 \leq i \leq k$, it is understood that $\alpha_1, \dots, \alpha_{i-1}$ are assigned prior to the assignment of α_i and that the assignment of α_i is dependent on the assignment of its predecessors. The resulting inequalities are then understood as '$\alpha_i \leq$(any positive-valued expression involving $\alpha_1, \dots, \alpha_{i-1}$)'.

The principle makes the sequence of inequalities used throughout the rest of the section consistent without muddling the matter with the arithmetic of particular infinitesimals.

Specifically, the assignment of parameters used in this section is:

$$\beta >> \gamma >> \delta >> \varepsilon >> \zeta >> \iota$$

Note that ε is the parameter used to define contiguity submaps in the previous section.

Further, one more restriction is imposed on the assignment of ι, specifically that its inverse $n = 1/\iota$ is an integer that is either odd or divisible by 2^9 (and so, for small enough ι, $n \in \mathbb{N}^*$).

In [9], these constants are labelled differently. Indeed, exact values are even given there, with $\beta = 0.05$, $\gamma = 0.01$, $\delta = 0.005$, $\varepsilon = 2^{-14}$, $\zeta = 2^{-33}$, $\iota \leq 2^{-48}$.

3.6 The graded presentation of $B(\mathcal{A}, n)$

Letting $\mathcal{A}$ be a finite alphabet, fix a total order $\prec$ on the set of words over $\mathcal{A}$ such that if $|X| < |Y|$, then $X \prec Y$. Also, set $\mathcal{R}_0 = \emptyset$ and $\mathbf{B}(0) = F(\mathcal{A})$.

With this terminology, inductively define $\mathcal{R}_i = \mathcal{R}_{i-1} \cup \{A_i^n\}$ where A_i is the first (relative to $\prec$) with infinite order in $\mathbf{B}(i-1) = \langle \mathcal{A} \mid \mathcal{R}_{i-1} \rangle$.

The following is a main result of [9]:

Lemma 3.1. *(Theorem B of [9]). For each i, A_i exists, with $|A_i| \leq |A_{i-1}|+1$. Moreover, $\mathcal{R} = \cup \mathcal{R}_i$ can be taken as an independent set of defining relations of the free Burnside group $B(\mathcal{A}, n)$ (and so defines a graded presentation $\mathbf{B}(\infty)$ of $B(\mathcal{A}, n)$).*

Let A be a freely cyclically reduced word over $\mathcal{A}$. Then a word W over $\mathcal{A}$ is A-periodic if W is a subword of a power A^k for $k > 0$, i.e $A^k \equiv Z_1 W Z_2$ for some (perhaps empty) words Z_1 and Z_2. A decomposition $W \equiv W_1 W_2$ is called *phase* if there exist positive integers k_1, k_2 with $k_1 + k_2 = k$ such that $A^{k_1} \equiv Z_1 W_1$ and $A^{k_2} \equiv W_2 Z_2$.

If Δ is a graded diagram over $\mathbf{B}(i)$, a section q of $\partial\Delta$ is called A-periodic if Lab(q) is an A-periodic word. In this case, a vertex of q is called phase if the natural decomposition of Lab(q) it defines is phase. Similarly, if Π is an $\mathcal{R}$-cell with $r(\Pi) = j$, then a vertex O of $\partial\Pi$ is called phase if Lab$(\partial\Pi)$ when read starting at O is visually $A_j^{\pm n}$.

Denote $\mathcal{F}(A_i)$ as a finite subgroup of $\mathbf{B}(i-1)$ that is maximal with respect to the property that A_i normalizes $\mathcal{F}(A_i)$. Clearly, such a subgroup must exist as A_i normalizes the trivial group. The following statement establishes the well-defined nature of this construction:

Lemma 3.2. *(Lemma 18.5(a) of [9]). The subgroup $\mathcal{F}(A_i)$ is uniquely defined and is a 2-group.*

A word J is called an $\mathcal{F}(A_i)$-*involution* provided J normalizes $\mathcal{F}(A_i)$ in $\mathbf{B}(i-1)$, $J^2 \in \mathcal{F}(A_i)$ in rank $i-1$, and $J^{-1}A_iJ \overset{i-1}{=} A_i^{-1}F$ with a word $F \in \mathcal{F}(A_i)$.

Let Δ be a diagram over $\mathbf{B}(i)$ and q_1, q_2 be sections of $\partial\Delta$ or the contours of a cell in Δ. Suppose Lab(q_k) is $A_j^{\varepsilon_k}$-periodic for $k \leq i$ and $\varepsilon_k \in \{\pm 1\}$. Then q_1 and q_2 are called j-*compatible* in Δ if either:

(A1) If $\varepsilon_1\varepsilon_2 = -1$, then there are phase vertices $O_k \in q_k$ and a simple path t between O_1 and O_2 such that $|t| < (1+\delta)|A_j|$ and Lab$(t) \overset{j-1}{=} T$ for some $T \in \mathcal{F}(A_j)$.

(A2) If $\varepsilon_1\varepsilon_2 = 1$, then there are phase vertices $O_k \in q_k$ and a simple path t between O_1 and O_2 such that $|t| < (1+\delta)|A_j|$ and Lab$(t) \overset{j-1}{=} T$ for some $\mathcal{F}(A_j)$-involution T.

A disk diagram Δ over $\mathbf{B}(i)$ is called *strictly reduced* if $\partial\Pi_1$ and $\partial\Pi_2$ are not j-compatible for cells Π_1 and Π_2 of rank j (perhaps with $\Pi_1 = \Pi_2$). (Note that in [9], such a diagram is simply called reduced).

This terminology is justified by Theorem C of [9], which essentially assures that cells whose contours are j-compatible can be removed from Δ and replaced with cells of rank $\leq j-1$, reducing the type of Δ (this is done in much the same way as it is done for j-pairs). As such, we may again strengthen van Kampen's Lemma, so that a word W over $\mathcal{A} \cup \mathcal{A}^{-1}$ satisfies $W =_{B(\mathcal{A},n)} 1$ if and only if there exists $i \geq 0$ and a strictly reduced diagram Δ over $\mathbf{B}(i)$ with Lab$(\partial\Delta) \equiv W$.

3.7 Tame diagrams

A word X is *cyclically reduced in rank* i if for any word Y such that $X \overset{i}{=} ZYZ^{-1}$ (i.e X and Y are conjugate in rank i), then $|X| \leq |Y|$. The word A is called *simple in rank* i if A is not conjugate in rank i to $A_k^{\ell}F$ for any $k \leq i$, any integer ℓ, and any $F \in \mathcal{F}(A_k)$.

For Δ a diagram over $\mathbf{B}(i)$, an A-periodic section q of $\partial\Delta$ is called *smooth* if either:

(S1) $A \equiv A_j^{\pm 1}$ and there is no cell Π in Δ such that $r(\Pi) = j$ and $\partial\Pi$ is j-compatible with q

(S2) A is simple in rank i.

If q satisfies (S1), then the rank of q is defined to be $r(q) = j$. Otherwise, we write $r(q) = \infty$.

Note that if Δ is a strictly reduced diagram containing an $\mathcal{R}$-cell Π, Γ is a subdiagram of Δ such that a section q of $\partial\Gamma$ is a subpath of $\partial\Pi$ in Δ, and Π is not contained in Γ, then Γ is strictly reduced and q is a smooth section of $\partial\Gamma$ with $r(q) = r(\Pi)$.

Finally, a strictly reduced diagram Δ over $\mathbf{B}(i)$ is called *tame* if it satisfies both:

(T1) Let $p = \partial\Pi_1$ for some cell Π_1 in Δ and q be a smooth section of $\partial\Delta$ or $q = \partial\Pi_2$ for some cell Π_2 in Δ. If Γ is a contiguity subdiagram between p and q, then $r(\Gamma) < \min(r(p), r(q))$.

(T2) For any $\mathcal{R}$-cell Π in Δ, there is no 0-bond in Δ from $\partial\Pi$ to itself.

The following statements are proved in [9] and listed here for reference.

Lemma 3.3. *(Lemma 9.2 of [9]). Every strictly reduced diagram Δ over $\boldsymbol{B}(i)$ is tame.*

Lemma 3.4. *(Lemma 5.7 of [9]) Let Δ be a tame disk diagram over $\boldsymbol{B}(i)$ whose contour is decomposed into the sections $q_1, \ldots, q_m$ with $m \leq 8$. Then, in Δ, there exists an $\mathcal{R}$-cell π and disjoint contiguity submaps $\{\Gamma_j\}_{j=1}^k$ of π to these sections such that $\sum_{j=1}^{k} |\Gamma_{j\wedge}\pi| > (1-\gamma)|\partial\pi|$.*

The cell π guaranteed by Lemma 3.4 is called a γ-*cell*.

Lemma 3.5. *(Lemma 6.1 of [9]) Let Δ be a tame disk diagram over $\boldsymbol{B}(i)$ with contour qt. If q is a smooth section, then $(1-\beta)|q| \leq |t|$.*

Lemma 3.6. *(Lemma 6.2 of [9]) If Δ is a tame disk diagram over $\boldsymbol{B}(i)$ satisfying the relation $|\partial\Delta| \leq (1-\beta)n|A_k|$ for some $k \leq i$, then $r(\Delta) < k$.*

Lemma 3.7. *(Lemma 3.1 of [9]) Let Δ be a strictly reduced disk diagram over $\boldsymbol{B}(i)$ and Γ be a contiguity submap of a cell Π to a section q of $\partial\Delta$. If Γ is a tame diagram with $\partial\Gamma = d_1p_1d_2q_1$ where $p_1 = \Gamma_\wedge\partial\Pi$ and $q_1 = \Gamma_\wedge q$. Then for $j = r(\Pi)$, $\max(|d_1|, |d_2|) < 2\varepsilon^{-1}|A_j| \leq \zeta n|A_j|$.*

3.8 Mass of a diagram

We now introduce a weighting on diagrams over the presentation $\mathbf{B}(\infty) = \langle \mathcal{A} \mid \mathcal{R} \rangle$ of $B(\mathcal{A}, n)$, generalizing the concept of the area of such a diagram.

For Π an $\mathcal{R}$-cell in a reduced graded disk diagram Δ over the presentation $\mathbf{B}(\infty)$, let A^n be the relator associated to Lab$(\partial\Pi)$. Then, define the *mass* of Π as $\rho(\Pi) = |A|^2$. This definition is extended naturally to the mass of the entire diagram, taking $\rho(\Delta)$ to be the sum of the masses of its $\mathcal{R}$-cells.

Lemma 3.8. *If Δ is a strictly reduced graded disk diagram over the presentation $\boldsymbol{B}(\infty)$ of $B(m, n)$, then $\rho(\Delta) \leq |\partial\Delta|^2$.*

Proof. The proof inducts on $|\partial\Delta|$, with the base case $|\partial\Delta| \leq (1-\beta)n$. In this case, Lemma 3.6 implies that $r(\Delta) = 0$, i.e Δ contains no $\mathcal{R}$-cells. But then $\rho(\Delta) = 0$. Hence, we may assume that $|\partial\Delta| > (1-\beta)n$ and Δ is a 'minimal counterexample' to the lemma.

Partition $\partial\Delta$ into 8 sections, $q_1 \dots q_8$, any two of which differ in length by at most 1. By Lemma 3.3, Δ is a tame diagram. Applying Lemma 3.4, there exists a γ-cell π in Δ together with contiguity submaps Γ_j of π to $q_{\ell(j)}$ for $j = 1, \dots, k$.

As $|\partial\Delta| > (1-\beta)n$ and $\beta < \frac{1}{2}$, $(\frac{1}{8} - \frac{2}{n})|\partial\Delta| < |q_j| < (\frac{1}{8} + \frac{2}{n})|\partial\Delta|$ for all $1 \leq j \leq 8$.

We now proceed in two cases.

1. Suppose there exists $1 \leq m \leq 8$ such that $\ell(j) \neq m$ for all $1 \leq j \leq k$. Without loss of generality, say $m = 1$, i.e no Γ_j is a contiguity submap between π and q_1.

For any $1 \leq j \leq k$, write $\partial(\pi, \Gamma_j, q_{\ell(j)}) = d_1^j p_1^j d_2^j q_1^j$. By Lemma 3.7, $\max(|d_1^j|, |d_2^j|) < \zeta n|A_r| = \zeta|\partial\pi|$ for $r = r(\pi)$.

Now, let Γ be the smallest subdiagram of Δ containing each Γ_j. Then, there exists a decomposition of the contour $\partial\Gamma = s_1 t_1 s_2 t_2$ where t_1 is a subpath of $\partial\pi$, t_2 is a subpath of $\partial\Delta$, and each s_ℓ is a side arc of some Γ_j. Note that each p_1^j is a subpath of t_1, so that $|t_1| \geq \sum |p_1^j| > (1-\gamma)|\partial\pi|$.

As Γ is a tame disk diagram and t_1 is a smooth section of $\partial\Gamma$, Lemma 3.5 implies that

$$(1-\beta)|t_1| \leq |t_2| + |s_1| + |s_2| \leq |t_2| + 2\zeta|\partial\pi|$$

It follows that $|t_2| \geq ((1-\beta)(1-\gamma) - 2\zeta)|\partial\pi|$.

Let t_1' be the complement of t_1 in $\partial\pi$ so that $\partial\pi = t_1 t_1'$. Further, let t_2' be the complement of t_2 in $\partial\Delta$ so that $\partial\Delta = t_2 t_2'$. Letting $u = s_2^{-1} t_1' s_1^{-1}$, then $|t_1'| < \gamma|\partial\pi|$ and $|u| < (\gamma + 2\zeta)|\partial\pi|$.

Cutting Δ along u yields two tame subdiagrams Δ_1 and Δ_2 with contours $u^{-1}t_2$ and ut_2', respectively.

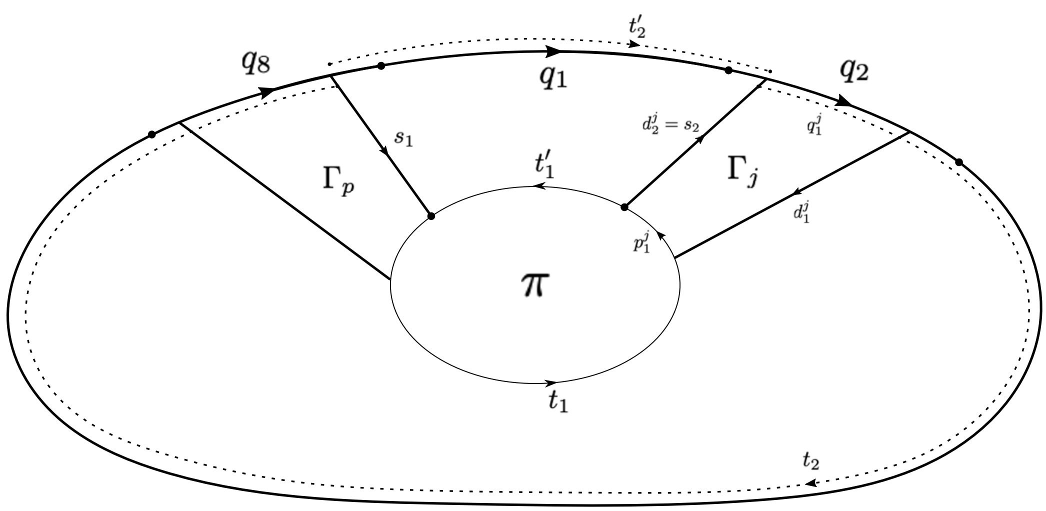


Figure 3.4: Γ_1 is absent, Γ_2 and Γ_{10} are present

Then, by the parameter assignments, we have

$$|u| < (\gamma + 2\zeta)|\partial\pi| < 2\gamma((1-\beta)(1-\gamma) - 2\zeta)^{-1}|t_2| < 3\gamma|t_2| < 3\gamma|\partial\Delta|$$

and $|t_2| < 7(\frac{1}{8} + \frac{2}{n})|\partial\Delta| < \frac{9}{10}|\partial\Delta|$.

Hence, by the parameter assignment for γ,

$$|\partial\Delta_1| = |t_2| + |u| < (1 + 3\gamma)|t_2| < \frac{9}{10}(1 + 3\gamma)|\partial\Delta| < |\partial\Delta|$$

$$|\partial\Delta_2| = |u| + |t_2'| = |u| + |\partial\Delta| - |t_2| < |\partial\Delta| - (1-3\gamma)|t_2| < |\partial\Delta|$$

Applying the inductive hypothesis on both diagrams then yields

$$\rho(\Delta_1) < (1+3\gamma)^2|t_2|^2$$

$$\rho(\Delta_2) < (|\partial\Delta| - (1-3\gamma)|t_2|)^2$$

As γ is sufficiently small, note that $\frac{20}{9}(1-3\gamma) \geq (1+3\gamma)^2 + (1-3\gamma)^2$. So,

$$|t_2|((1+3\gamma)^2 + (1-3\gamma)^2) \leq \frac{20}{9}|t_2|(1-3\gamma) \leq 2|\partial\Delta|(1-3\gamma)$$

Hence, $|t_2|^2(1+3\gamma)^2 + |t_2|^2(1-3\gamma)^2 - 2|\partial\Delta||t_2|(1-3\gamma) \leq 0$, and so

$$(|\partial\Delta| - (1-3\gamma)|t_2|)^2 + (1+3\gamma)^2|t_2|^2 \leq |\partial\Delta|^2$$

This final inequality yields

$$\rho(\Delta) = \rho(\Delta_1) + \rho(\Delta_2) < |\partial\Delta|^2$$

2. Suppose that for every $1 \leq m \leq 8$, there is a j such that Γ_j is a contiguity submap between π and q_m.

For any $m \in \{1, \ldots, 8\}$, let $\Lambda_1^m, \ldots, \Lambda_k^m$ be the collection of Γ_j that are contiguity submaps between π and q_m. Then, let Λ_m be the smallest subdiagram of Δ containing each Λ_j^m.

It follows that we may set $\partial\Lambda_m = s_1^m t_1^m s_2^m t_2^m$ for all m, $\partial\pi = t_1^8 v_8 t_1^7 v_7 \ldots t_1^1 v_1$, and $\partial\Delta = t_2^1 w_1 t_2^2 w_2 \ldots t_2^8 w_8$. Further, for $m = 1, \ldots, 8$, let Δ_m be the subdiagram with contour

$$w_m (s_2^{m+1})^{-1} v_{m+1} (s_1^m)^{-1}$$

(with indices m counted mod 8).

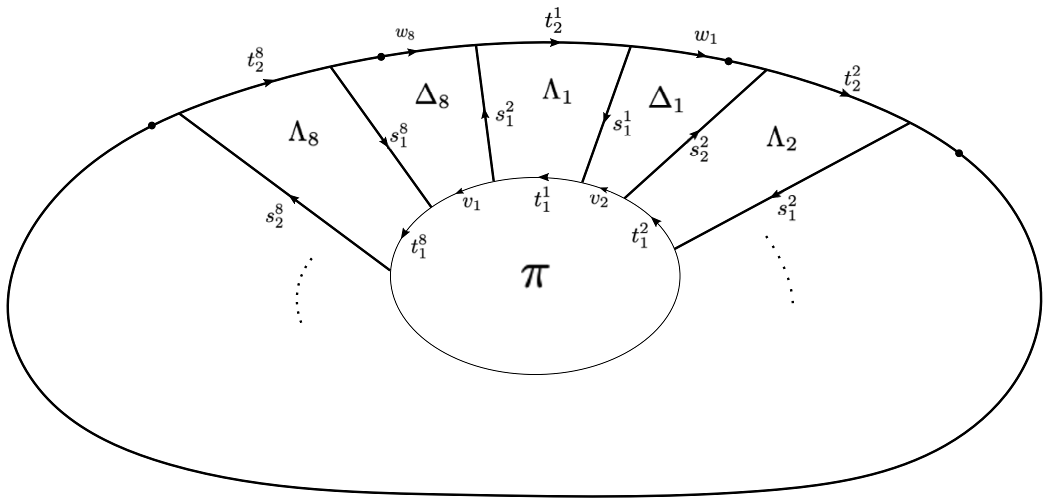


Figure 3.5: All Γ_i are present

As in the previous case, Lemma 3.7 implies that $|s_j^m| < \zeta|\partial\pi|$. Also, as $\beta < \frac{1}{2}$, Lemma 3.6 implies $|\partial\pi| < 2|\partial\Delta|$.

As Λ_m is a tame subdiagram and t_1^m is a smooth section of $\partial\Lambda_m$, Lemma 3.5 implies

$$(1-\beta)|t_1^m| < |s_1^m| + |s_2^m| + |t_2^m|$$

for all m. Further, since we also have $|t_2^m| \leq |q_m| < (\frac{1}{8} + \frac{2}{n})|\partial\Delta|$, it follows from the parameter choices that

$$|s_1^m| + |s_2^m| + |t_2^m| < (\frac{1}{8} + \frac{2}{n} + 4\zeta)|\partial\Delta| < \frac{1}{7}|\partial\Delta|$$

for all m. Hence, for all m,

$$|\partial\Lambda_m| = |t_1^m| + |s_1^m| + |s_2^m| + |t_2^m| < \frac{1}{7}\left(1 + \frac{1}{1-\beta}\right)|\partial\Delta| < \frac{1}{6}|\partial\Delta|$$

So, applying the inductive hypothesis, $\rho(\Lambda_m) < \frac{1}{36}|\partial\Delta|^2$ for all m.

Further, $|w_m| < |q_m| + |q_{m+1}| < (\frac{1}{4} + \frac{4}{n})|\partial\Delta|$ and $|v_m| < \gamma|\partial\pi| < 2\gamma|\partial\Delta|$, so that the parameter assignments yield

$$|\partial\Delta_m| < \left(\frac{1}{4} + \frac{4}{n} + 4\zeta + 2\gamma\right)|\partial\Delta| < \frac{2}{7}|\partial\Delta|$$

for all m. So, applying the inductive hypothesis yields $\rho(\Delta_m) < \frac{4}{49}|\partial\Delta|^2$ for all m.

Finally, note that since π is an $\mathcal{R}$-cell, $\rho(\pi) = (\iota|\partial\pi|)^2 < 4\iota^2|\partial\Delta|^2 < \frac{1}{9}|\partial\Delta|^2$. Thus,

$$\rho(\Delta) = \sum_{m=1}^{8} \rho(\Gamma_m) + \sum_{m=1}^{8} \rho(\Delta_m) + \rho(\pi) < \frac{2}{9}|\partial\Delta|^2 + \frac{32}{49}|\partial\Delta|^2 + \frac{1}{9}|\partial\Delta|^2 < |\partial\Delta|^2$$

$\square$

Chapter 4

S-Machines

4.1 Definition of S-machine as a Rewriting System

There are many equivalent interpretations of S-machines [27]. Following the conventions of [3], [18], [20], [21], [23], [25], [26], and others, we approach them here as a rewriting system for words over group alphabets.

Let (Y, Q) be a pair of sets with $Q = \sqcup_{i=0}^{N} Q_i$ and $Y = \sqcup_{i=1}^{N} Y_i$ for some positive integer N. For convenience of notation, set $Y_0 = Y_{N+1} = \emptyset$ in this setting.

The elements of $Q \cup Q^{-1}$ are called *state letters* or q*-letters*, while those of $Y \cup Y^{-1}$ are *tape letters* or a*-letters*. The sets Q_i and Y_i are called the *parts* of Q and Y, respectively. Note that the parts of the state letters are typically represented by capital letters, while their elements are represented by lowercase.

The *language of admissible words* for (Y, Q) is the collection of reduced words of the form $q_0^{\varepsilon_0} u_1 q_1^{\varepsilon_1} \dots u_k q_k^{\varepsilon_k}$ where $\varepsilon_i \in \{\pm 1\}$ and each subword $q_{i-1}^{\varepsilon_{i-1}} u_i q_i^{\varepsilon_i}$ either:

(1) belongs to $(Q_{j-1} F(Y_j) Q_j)^{\pm 1}$;

(2) has the form quq^{-1} for $q \in Q_j$ and $u \in F(Y_{j+1})$; or

(3) has the form $q^{-1}uq$ for $q \in Q_j$ and $u \in F(Y_j)$

For a reduced word $W \in F(Y \cup Q)$, define its a*-length* $|W|_a$ as the number of a-letters that comprise it. The q-length of W is defined similarly and is denoted $|W|_q$.

Let $W \equiv q_1 u_1 q_2 u_2 q_3 \dots q_s$ be an admissible word with $q_i \in Q_{j(i)}^{\varepsilon_i}$ for $\varepsilon_i \in \{\pm 1\}$ and $u_i \in F(Y)$. Then the *base* of W is $\text{base}(W) \equiv Q_{j(1)}^{\varepsilon_1} Q_{j(2)}^{\varepsilon_2} \dots Q_{j(s)}^{\varepsilon_s}$, where these letters are merely representatives of their corresponding parts, and u_i is called the $Q_{j(i)}^{\varepsilon_i} Q_{j(i+1)}^{\varepsilon_{i+1}}$-sector of W. Note that the base of an admissible word W need not be a reduced word and that W is permitted to have many sectors of the same name (for example, W may contain many Q_0Q_1-sectors).

The base $Q_0 \dots Q_N$ is called the *standard base*. An admissible word with the standard base is called a *configuration*.

Now, set $U_0, \dots, U_m$ and $V_0, \dots, V_m$ as a collection of reduced words over $Y \cup Q$ satisfying:

(1) U_i and V_i have base $Q_{\ell(i)}Q_{\ell(i)+1} \dots Q_{r(i)}$ with $\ell(i) \leq r(i)$ and such that both are subwords of admissible words

(2) $\ell(i+1) = r(i) + 1$ for all i

(3) U_0 and V_0 start with letters from Q_0, while U_m and V_m end with letters from Q_N

Define $Q(\theta)$ as the set of state letters appearing in some U_i. Note that $Q(\theta)$ contains exactly one state letter from each part.

Also, let $Y(\theta) = \cup Y_j(\theta)$ with $Y_j(\theta) \subseteq Y_j$ be some subset of the tape alphabet with the requirement that the set of tape letters appearing in U_i or V_i is a subset of $Y(\theta)^{\pm 1}$. Each $Y_j(\theta)$ is called the *domain* of θ in the corresponding sector of the standard base.

If W is an admissible word with all its state letters contained in $Q(\theta) \cup Q(\theta)^{-1}$ and all its tape letters contained in $Y(\theta) \cup Y(\theta)^{-1}$, then define $W \cdot \theta$ as the result of simultaneously replacing every subword $U_i^{\pm 1} \equiv (u_{\ell(i)}q_{\ell(i)}u_{\ell(i)+1}q_{\ell(i)+1} \dots q_{r(i)}u_{r(i)+1})^{\pm 1}$ of W by the subword $V_i^{\pm 1} \equiv (v_{\ell(i)}q'_{\ell(i)}v_{\ell(i)+1} \dots q'_{r(i)}v_{r(i)+1})^{\pm 1}$, followed by the necessary reduction to make the resulting word again admissible.

In this case, θ is called an S-*rule* of (Y, Q) and is denoted $\theta = [U_0 \to V_0, \dots, U_m \to V_m]$. This notation fully describes the rule θ except for the corresponding sets $Y_j(\theta)$. Henceforth, $Y_j(\theta)$ is assumed to be either Y_j or $\emptyset$ unless otherwise stated, with context making it clear which is chosen.

For any S-rule θ, if θ is applicable to an admissible word W, then W is called θ-*admissible*.

An important note to stress is that the application of an S-rule results in a reduced word, i.e reduction is not a separate step in the application of the S-rule.

If the i-th part of the S-rule θ is $U_i \to V_i$, U_i and V_i have base $Q_{\ell(i)} \dots Q_{r(i)}$, and $Y_{r(i)+1}(\theta) = \emptyset$, then this part of the rule is denoted $U_i \xrightarrow{\ell} V_i$ and θ is said to *lock* the $Q_{r(i)}Q_{r(i)+1}$-sector.

Note that every S-rule θ has a natural inverse, namely $\theta^{-1} = [V_0 \to U_0, \dots, V_m \to U_m]$ with $Y_j(\theta^{-1}) = Y_j(\theta)$ for all j.

An *S-machine* **S** with hardware (Y, Q) is the defined to be the rewriting system whose software is a symmetric set of S-rules $\Theta(\mathbf{S}) = \Theta$, i.e $\theta \in \Theta$ if and only if $\theta^{-1} \in \Theta$.

It is convenient to partition Θ into two disjoint sets, Θ^+ and Θ^-, such that $\theta \in \Theta^+$ if and only if $\theta^{-1} \in \Theta^-$. The elements of Θ^+ are called the *positive rules* and those of Θ^- the *negative rules*.

For $t \geq 0$, suppose $W_0, \dots, W_t$ are admissible words with the same base such that there exist $\theta_1, \dots, \theta_t \in \Theta$ satisfying $W_{i-1} \cdot \theta_i \equiv W_i$ for all $1 \leq i \leq t$. Then the sequence of applications of rules $\mathcal{C} : W_0 \to \cdots \to W_t$ is called a *computation* of *length* or *time* $t \geq 0$ of **S**. The word $H = \theta_1 \dots \theta_t$ is called the *history* of $\mathcal{C}$ and the notation $W_t \equiv W_0 \cdot H$ is used to represent the computation.

A computation is called *reduced* if its history is a reduced word in $F(\Theta^+)$. Every computation can be made reduced without changing the initial and final admissible words of the computation simply by removing consecutive mutually inverse rules.

Typically, it is assumed that each part of the state letters contains two (perhaps the same) fixed elements, called the *start* and *end* state letters. A configuration is called a *start* (or *end*) configuration if all of its state letters are start (or end) letters.

A *recognizing* S-machine is one with specified sectors called the *input sectors*. If a start configuration has all sectors empty except for the input sectors, then it is called an *input configuration* and its projection onto $Y \cup Y^{-1}$ is called its *input*. The end configuration with every sector empty is called the *accept configuration*.

A configuration W is *accepted* by a recognizing S-machine if there is an *accepting computation*, i.e a computation whose initial configuration is W and whose final configuration is the accept configuration. If W is an accepted input configuration with input u, then u is also said to be *accepted*.

If the configuration W is accepted by the S-machine **S**, then $T(W)$ is the minimal time of its accepting computations. For a recognizing S-machine **S**, its *time function* is

$$T_{\mathbf{S}}(n) = \max\{T(W) : W \text{ is an accepted input configuration of } \mathbf{S},\ |W|_a \leq n\}$$

If two recognizing S-machines have the same language of accepted words and Θ-equivalent time functions, then they are said to be *equivalent*.

The following simplifies how one approaches the rules of a recognizing S-machine.

Lemma 4.1. *(Lemma 2.1 of [18]) Every recognizing S-machine* **S** *is equivalent to a recognizing S-machine that satisfies:*

(1) Every part of every rule has a 1-letter base (i.e if $U_i \to V_i$ is a part of a rule θ, then $U_i \equiv u_i q_i u_{i+1}$ and $V_i \equiv v_i q_i' v_{i+1}$ for q_i, q_i' state letters in Q_i)

(2) In every part $u_i q_i u_{i+1} \to v_i q_i' v_{i+1}$ of every rule, $\|u_i\| + \|v_i\| \leq 1$ and $\|u_{i+1}\| + \|v_{i+1}\| \leq 1$.

(3) Moreover, with the terminology of (2), $\|u_i\| + \|v_i\| + \|u_{i+1}\| + \|v_{i+1}\| \leq 1$.

As a result of Lemma 4.1, we may assume that each part of every rule of an S-machine is of the form $q_i \to a q_i' b$ with $\|a\| + \|b\| \leq 1$. However, it will be convenient to allow $\|a\| = \|b\| = 1$ in the defining rules of some of the S-machines we construct.

4.2 Some elementary properties of S-machines

The following is an immediate consequence of the definition of admissible words.

Lemma 4.2. *If the rule θ locks the $Q_i Q_{i+1}$-sector, i.e it has a part $q_i \xrightarrow{\ell} a q_i'$ for some $q_i, q_i' \in Q_i$, then the base of any θ-admissible word has no subword of the form $Q_i Q_i^{-1}$ or $Q_{i+1}^{-1} Q_{i+1}$.*

Through the rest of our discussion of S-machines, we will often use copies of words over disjoint alphabets. To be precise, let A and B be disjoint alphabets, $W \equiv a_1^{\varepsilon_1} \dots a_k^{\varepsilon_k}$ with $a_i \in A$

and $\varepsilon_i \in \{\pm 1\}$, and $\varphi : \{a_1, \dots, a_k\} \to B$ be an injection. Then the *copy* of W over the alphabet B formed by φ is the word $W' \equiv \varphi(a_1)^{\varepsilon_1} \dots \varphi(a_k)^{\varepsilon_k}$. Typically, the injection defining the copy will be contextually clear.

Alternatively, a copy of an alphabet A is a disjoint alphabet A' which is in one-to-one correspondence with A. For a word over A, its copy over A' is defined by the correspondence between the alphabets.

The following are properties of some simple computations in S-machines that are fundamental to the proofs presented in the next two chapters. They are stated here without proof, with a reference provided for their proofs in previous literature.

Lemma 4.3. *(Lemma 2.7 of [18]) Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation, where W_0 is an admissible word with the two-letter base Q_iQ_{i+1}. Denote the tape word of W_j as u_j for each $0 \leq j \leq t$. Suppose that each rule of $\mathcal{C}$ multiplies the Q_iQ_{i+1}-sector by a letter on the left (respectively right). Suppose further that different rules multiply this sector by different letters. Then:*

(a) the history H of $\mathcal{C}$ is a copy of the reduced form of $u_tu_0^{-1}$ read from right to left (respectively $u_0^{-1}u_t$ read left to right). In particular, if $u_0 \equiv u_t$, then $\mathcal{C}$ is empty.

(b) $\|H\| \leq \|u_0\| + \|u_t\|$

(c) if $\|u_{j-1}\| < \|u_j\|$ for some $1 \leq j \leq t-1$, then $\|u_j\| < \|u_{j+1}\|$

(d) $\|u_j\| \leq \max(\|u_0\|, \|u_t\|)$

Lemma 4.4. *(Lemma 3.6 of [24]) Suppose $\mathcal{C} : W_0 \to \cdots \to W_t$ is a reduced computation of an S-machine with base $Q_iQ_i^{-1}$ (respectively $Q_i^{-1}Q_i$). For $0 \leq j \leq t$, let u_j be the tape word of W_j. Suppose each rule of $\mathcal{C}$ multiplies the Q_iQ_{i+1}-sector (respectively the $Q_{i-1}Q_i$-sector) by a letter from the left (respectively from the right), with different rules corresponding to different letters. Then $\|u_j\| \leq \max(\|u_0\|, \|u_t\|)$ for all j and the history of $\mathcal{C}$ has the form $H_1H_2^\ell H_3$, where $\ell \geq 0$, $\|H_2\| \leq \min(\|u_0\|, \|u_t\|)$, $\|H_1\| \leq \|u_0\|/2$, and $\|H_3\| \leq \|u_t\|/2$.*

4.3 Parameters

The arguments spanning the rest of this manuscript are reliant on the *highest parameter principle*, the obvious dual to the lowest parameter principle described in Chapter 3.5. In particular, we introduce the relation $<<$ on parameters defined as follows.

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are parameters with $\alpha_1 << \alpha_2 << \cdots << \alpha_n$, then for all $2 \leq i \leq n$, it is understood that $\alpha_1, \ldots, \alpha_{i-1}$ are assigned prior to the assignment of α_i and that the assignment of α_i is dependent on the assignment of its predecessors. The resulting inequalities are then understood as '$\alpha_i \geq$(any expression involving $\alpha_1, \ldots, \alpha_{i-1}$)'

Specifically, the assignment of parameters we use here is:

$$\begin{aligned} n << \lambda^{-1} << c_0 << k << c_1 << c_2 << c_3 << c_4 << c_5 << L_0 << L << K_0 \\ << K << J << \delta^{-1} << C_1 << C_2 << C_3 << N_1 << N_2 << N_3 << N_4 << N_5 \end{aligned}$$

Chapter 5

Auxiliary Machines

5.1 The machine $\mathbf{M}_1$

Let n be a positive integer and $\mathcal{A}$ be a finite set. Define the *language of defining relations of* $B(\mathcal{A}, n)$ as the set $\mathcal{L} = \{u^n : u \in F(\mathcal{A})\}$.

For $0 \leq i \leq 4$, let $Q_i = \{q_i(j) : j = 1, \dots, 2n\}$. Further, for $1 \leq i \leq 4$, let $Y_i = \{a_i : a \in \mathcal{A}\}$ be a copy of $\mathcal{A}$.

The recognizing S-machine $\mathbf{M}_1$ has hardware $(\sqcup_{i=1}^4 Y_i, \sqcup_{i=0}^4 Q_i)$ and software the set of rules Φ defined below. The input sector is taken to be the Q_0Q_1-sector, while the letters $q_i(1)$ (respectively $q_i(2n)$) are the start (respectively end) letters.

The idea of the function of $\mathbf{M}_1$ is the following. Consider an input configuration with input $u^n \in \mathcal{L}$. The machine removes one copy of u and replaces it in the Q_2Q_3-sector. Next, this copy is moved to the Q_1Q_2-sector. It is then moved back to the Q_2Q_3-sector while another copy of u is simultaneously erased from the input sector. The second and third steps are then repeated $n-2$ more times until the input sector is empty. In the final step of this iteration, though, the natural copy of u^{-1} is written in the Q_3Q_4-sector. Finally, the copies of u and u^{-1} are erased from the Q_2Q_3- and Q_3Q_4-sectors, respectively.

The set of positive rules Φ^+ is decomposed into $2n+1$ subsets, which are denoted $\Phi_1^+, \dots, \Phi_{2n}^+$, and $\{\sigma(i, i+1) : i = 1, \dots, 2n-1\}$.

For each i, the rules of Φ_i^+ are in correspondence with $\mathcal{A}$, with the rule corresponding to $a \in \mathcal{A}$ denoted $\tau_i(a)$.

For simplicity, if a rule θ of $\mathbf{M}_1$ does not lock the $Q_{i-1}Q_i$-sector, then we take $Y_i(\theta) = Y_i$.

The definitions of the positive rules are detailed below, with a comment provided to detail the purpose of each rule:

• The rules of Φ_1^+ are of the form

$$\tau_1(a) = \left[q_0(1) \to q_0(1),\ q_1(1) \xrightarrow{\ell} a_1^{-1} q_1(1),\ q_2(1) \to q_2(1) a_3,\ q_3(1) \xrightarrow{\ell} q_3(1),\ q_4(1) \to q_4(1) \right]$$

Comment: The state letter $q_1(1)$ moves left, removing the copy of a from the Q_0Q_1-sector and replacing its copy in the Q_2Q_3-sector.

• $$\sigma(12) = \left[q_0(1) \to q_0(2),\ q_1(1) \xrightarrow{\ell} q_1(2),\ q_2(1) \to q_2(2),\ q_3(1) \xrightarrow{\ell} q_3(2),\ q_4(1) \to q_4(2) \right]$$

Comment: The state letters are changed while the Q_1Q_2- and Q_3Q_4-sectors are locked.

• For $1 \leq i \leq n-1$, the rules of Φ_{2i}^+ are of the form

$$\tau_{2i}(a) = \left[\begin{array}{l} q_0(2i) \to q_0(2i),\ q_1(2i) \to q_1(2i),\ q_2(2i) \to a_2 q_2(2i) a_3^{-1}, \\ q_3(2i) \xrightarrow{\ell} q_3(2i),\ q_4(2i) \to q_4(2i) \end{array} \right]$$

Comment: The state letter $q_2(2i)$ moves right, removing the copy of a from the Q_2Q_3-sector and replacing its copy in the Q_1Q_2-sector.

• For $1 \leq i \leq n-1$,

$$\sigma(2i, 2i+1) = \left[\begin{array}{l} q_0(2i) \to q_0(2i+1),\ q_1(2i) \to q_1(2i+1),\ q_2(2i) \xrightarrow{\ell} q_2(2i+1), \\ q_3(2i) \xrightarrow{\ell} q_3(2i+1),\ q_4(2i) \to q_4(2i+1) \end{array} \right]$$

Comment: When $q_2(2i)$ reaches $q_3(2i)$, the state letters are changed.

• For $1 \leq i \leq n-2$, the rules of Φ_{2i+1}^+ are of the form

$$\tau_{2i+1}(a) = \left[\begin{array}{l} q_0(2i+1) \to q_0(2i+1),\ q_1(2i+1) \to a_1^{-1} q_1(2i+1), \\ q_2(2i+1) \to a_2^{-1} q_2(2i+1) a_3,\ q_3(2i+1) \xrightarrow{\ell} q_3(2i+1), \\ q_4(2i+1) \to q_4(2i+1) \end{array} \right]$$

Comment: The state letter $q_2(2i+1)$ moves left, removing the copy of a from the Q_1Q_2-sector and replacing its copy in the Q_2Q_3-sector. Simultaneously, the copy of a is removed from the right of the Q_0Q_1-sector.

• For $1 \leq i \leq n-2$,

$$\sigma(2i+1, 2i+2) = \left[\begin{array}{l} q_0(2i+1) \to q_0(2i+2),\ q_1(2i+1) \xrightarrow{\ell} q_1(2i+2), \\ q_2(2i+1) \to q_2(2i+2),\ q_3(2i+1) \xrightarrow{\ell} q_3(2i+2), \\ q_4(2i+1) \to q_4(2i+2) \end{array} \right]$$

Comment: When $q_2(2i+1)$ reaches $q_1(2i+1)$, the state letters are changed.

• The rules of Φ^+_{2n-1} are of the form

$$\tau_{2n-1}(a) = \left[\begin{array}{l} q_0(2i+1)\to q_0(2i+1),\ q_1(2i+1)\to a_1^{-1}q_1(2i+1), \\ q_2(2i+1)\to a_2^{-1}q_2(2i+1)a_3,\ q_3(2i+1)\to q_3(2i+1), \\ q_4(2i+1)\to a_4^{-1}q_4(2i+1) \end{array}\right]$$

Comment: This rule functions similar to $\tau_{2i+1}(a)$ for $1\leq i\leq n-2$, but also inserts a copy of a^{-1} in the Q_3Q_4-sector.

$$\bullet\ \sigma(2n-1,2n) = \left[\begin{array}{l} q_0(2n-1)\xrightarrow{\ell} q_0(2n),\ q_1(2n-1)\xrightarrow{\ell} q_1(2n), \\ q_2(2n-1)\to q_2(2n),\ q_3(2n-1)\to q_3(2n), \\ q_4(2n-1)\to q_4(2n) \end{array}\right]$$

Comment: When both $q_2(2n-1)$ reaches $q_1(2n-1)$ and $q_1(2n-1)$ reaches $q_0(2n-1)$, the state letters are changed.

• The rules of Φ^+_{2n} are of the form

$$\tau_{2n}(a) = \left[\begin{array}{l} q_0(2n)\xrightarrow{\ell} q_0(2n),\ q_1(2n)\xrightarrow{\ell} q_1(2n),\ q_2(2n)\to q_2(2n)a_3^{-1}, \\ q_3(2n)\to q_3(2n),\ q_4(2n)\to a_4q_4(2n) \end{array}\right]$$

Comment: The letters in the Q_2Q_3- and Q_3Q_4-sectors are removed.

5.2 Standard computations of $\mathbf{M}_1$

The machine $\mathbf{M}_1$ can be viewed as the *composition* of $2n$ submachines, which are denoted $\mathbf{M}_1(1),\dots,\mathbf{M}_1(2n)$, and which are concatenated by the rules $\sigma(i,i+1)^{\pm1}$. The set of positive rules of the machine $\mathbf{M}_1(i)$ is Φ_i^+ and each machine has a disjoint set of state letters. So, each $Q_j = Q_{j,1}\sqcup\cdots\sqcup Q_{j,2n}$ where each $Q_{j,i}$ is the corresponding part of the hardware of $\mathbf{M}_1(i)$ (in this machine, each such subset is a singleton).

Accordingly, the rules of the form $\sigma(i,i+1)^{\pm1}$ are called *transition rules*, as their function is to force the steps to be carried out in the correct order. For clarity in later chapters, these rules are henceforth referred to as *σ-rules*.

For simplicity of notation, denote the inverse of each σ-rule by switching the indices, so that $\sigma(i,i+1)^{-1} \equiv \sigma(i+1,i)$.

The history of a reduced computation of $\mathbf{M}_1$ can be factored so that each factor is either a σ-rule or the history of a maximal subcomputation of $\mathbf{M}_1(i)$ for some i. The *step history* of a reduced computation is then defined so as to capture the order of the types of these factors. To do this, we denote the σ-rule $\sigma(i,j)$ by the pair (ij) and a factor that is an element of $F(\Phi_i^+)$ simply by (i).

For example, if $H \equiv H'H''H'''$ where $H' \in F(\Phi_2^+)$, $H'' \equiv \sigma(23)$, and $H''' \in F(\Phi_3^+)$, then the step history of a computation with history H is $(2)(23)(3)$. So, the step history of a reduced computation is some concatenation of the letters

$$\{(1),\ (2),\ \dots,\ (2n),\ (12),\ (23),\ \dots,\ (2n-1,2n),\ (21),\ (32),\ \dots,\ (2n,2n-1)\}$$

It is convenient to allow the omission of a letter representing a σ-rule in a step history when its existence is clear from its necessity. For example, given a reduced computation with step history (2)(23)(3), one can instead write the step history as $(2)(3)$, as the rule $\sigma(23)$ must occur for the maximal subcomputation with step history (3) to be possible.

If the step history of a reduced computation is $(i-1,i)(i,i+1)$, it is also permitted for the step history to be written as $(i-1,i)(i)(i,i+1)$ even though the 'maximal subcomputation' with step history (i) is empty.

A *one-step computation* is a reduced computation of $\mathbf{M}_1$ whose step history has exactly one factor corresponding to a maximal subcomputation of a submachine $\mathbf{M}_1(i)$.

Certain subwords cannot appear in the step history of a reduced computation of $\mathbf{M}_1$. For example, it is clear that it is impossible for the step history of a reduced computation to contain the subword $(1)(3)$. The next statement displays the impossibility of some less obvious potential subwords.

Lemma 5.1. *Let $\mathcal{C}$ be a reduced computation with base B.*

(a) If B contains a subword B' of the form $(Q_2Q_3)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be

$(2i, 2i+1)(2i+1)(2i+1, 2i)$ *or* $(2i+1, 2i)(2i)(2i, 2i+1)$ *for* $1 \le i \le n-1$.

(b) If B contains a subword B' of the form $(Q_1Q_2)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be $(2i-1, 2i)(2i)(2i, 2i-1)$ *or* $(2i+2, 2i+1)(2i+1)(2i+1, 2i+2)$ *for* $1 \le i \le n-1$.

(c) If B contains a subword B' of the form $(Q_3Q_4)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be $(2n-2, 2n-1)(2n-1)(2n-1, 2n-2)$.

Proof. Assuming to the contrary, let $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be the restriction of $\mathcal{C}$ to the subword B'. In each case, $|W_0'|_a = |W_t'|_a = 0$ and the subcomputation $W_1' \to \cdots \to W_{t-1}'$ satisfies the hypotheses of Lemma 4.3. But then this subcomputation must be empty, contradicting the assumption that $\mathcal{C}$ is reduced.

□

For an admissible word W, there is a natural projection of W onto $F(\mathcal{A})$ given by sending each tape letter to its natural copy and each state letter to the identity.

Note that any application of a rule of Φ_1 preserves the projection of a configuration. Similarly, for $i = 2, \dots, 2n-1$, any application of a rule of Φ_i preserves the projection of an admissible word with base $Q_1Q_2Q_3$.

An application of these useful facts (or those similar in nature) is referred to as a *projection argument*.

Lemma 5.2. *For $w \in F(\mathcal{A})$ and $i \in \{2 \dots, 2n-1\}$, there exists a unique reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ with base $Q_1Q_2Q_3$ such that*

(a) the step history of $\mathcal{C}$ is $(i-1, i)(i)(i, i+1)$, and

(b) the projection of W_0 onto $F(\mathcal{A})$ is w.

Moreover, for H_i the history of the maximal subcomputation of $\mathcal{C}$ with step history (i), H_i is a copy of w read left to right (resp read right to left) if i is even (resp odd) and $|W_j|_a = \|w\|$ for all $0 \le j \le t$.

Proof. Let $\mathcal{C}$ be such a computation and suppose i is even.

As W_0 is $\sigma(i-1,i)$-admissible, its Q_1Q_2-sector must be empty. So, since w is reduced, W_0 must have the natural copy of w written in its Q_2Q_3-sector, i.e $W_0 \equiv q_1(i-1)q_2(i-1)w_3q_3(i-1)$ for w_3 the natural copy of w in $F(Y_3)$.

Further, since W_t is $\sigma(i+1,i)$-admissible, its Q_2Q_3-sector must be empty.

But the restriction of the subcomputation $W_1 \to \cdots \to W_{t-1}$ to the Q_2Q_3-sector satisfies the hypotheses of Lemma 4.3, so that its history must be the natural copy of w.

If i is odd, then an analogous argument applies.

□

Lemma 5.3. *(a) Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_1$. Suppose W_0 is a start configuration and W_t is an end configuration. Then there exists $u \in F(\mathcal{A})$ such that the projection of W_0 onto $F(\mathcal{A})$ is $u^n \in \mathcal{L}$.*

(b) For all $u \in F(\mathcal{A})$, there exists a unique reduced computation $\mathcal{D}_1(u) : W_0 \to \cdots \to W_t$ of $\boldsymbol{M}_1$ with step history $(12)(2)(3)\dots(2n-1)(2n-1,2n)$ and such that the projection of W_0 onto $F(\mathcal{A})$ is u^n. This computation has length $(2n-2)\|u\|+(2n-1)$ and W_t has the natural copy of u (respectively u^{-1}) written in its Q_2Q_3-sector (respectively Q_3Q_4-sector).

Proof. (a) By a projection argument, it suffices to assume that the first letter of the step history is (12). Lemma 5.1 then implies that the step history of $\mathcal{C}$ must have prefix

$$(12)(2)(3)\dots(2n-1)(2n-1,2n)$$

Let $\mathcal{D} : W_0 \to \cdots \to W_s$ be the subcomputation with this step history. Further, let $u, v \in F(\mathcal{A})$ be the reduced words such that W_0 has the natural copy of v written in its input sector and the natural copy of u written in its Q_2Q_3-sector.

After restricting $\mathcal{D}$ to the subword $Q_1Q_2Q_3$ of the standard base, Lemma 5.2 implies that the history of $\mathcal{D}$ must be

$$\sigma(12)H_2\sigma(23)H_3\dots H_{2n-1}\sigma(2n-1,2n)$$

where $H_i \in F(\Phi_i^+)$ is the natural copy of u read left to right (resp right to left) for i even (resp i odd).

Then W_{s-1} has the copy of $vu^{-(n-1)}$ written in its input sector. But W_{s-1} is $\sigma(2n-1,2n)$-admissible, so that $vu^{-(n-1)}$ is freely trivial.

Hence, the projection of W_0 onto $F(\mathcal{A})$ is $vu = u^n \in \mathcal{L}$.

(b) Let $u \in F(\mathcal{A})$ and $\mathcal{D} : W_0 \to \cdots \to W_s$ be the computation described in (a).

Let $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be a computation with step history $(12)(2)(3)\ldots(2n-1)(2n-1,2n)$ such that the projection of W_0' onto $F(\mathcal{A})$ is u^n.

Set $x, w \in F(\mathcal{A})$ as the reduced words such that W_0' has the natural copy of x written in its input sector and the natural copy of w written in its Q_2Q_3-sector.

As in (a), applications of Lemma 5.2 then imply that the history of $\mathcal{C}'$ must be

$$\sigma(12)H_2'\sigma(23)H_3'\ldots H_{2n-1}'\sigma(2n-1,2n)$$

where $H_i' \in F(\Phi_i^+)$ is the natural copy of w read left to right (resp right to left) for i even (resp i odd).

Then, W_{t-1}' has the natural copy of $xw^{-(n-1)}$ written in its input sector and is $\sigma(2n-1,2n)$-admissible, so that $x = w^{n-1}$. But the projection of W_0' onto $F(\mathcal{A})$ is $xw = w^n$, so that $w \equiv u$.

Hence, $H_i' \equiv H_i$ for all i and $W_0 \equiv W_0'$, so that $\mathcal{C}' = \mathcal{D}$.

As H_i is a copy of u for each i, the length of $\mathcal{D}$ is $(2n-2)\|u\| + (2n-1)$.

□

Lemma 5.4. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_1$ such that W_0 is an end configuration. If the history H of $\mathcal{C}$ contains a σ-rule, then W_t is not an end configuration.*

Proof. Assuming to the contrary, Lemma 5.1 implies that H can be factored as $H'H''$ such that

$$H' \equiv H_{2n}'\sigma(2n,2n-1)H_{2n-1}'\ldots\sigma(32)H_2'\sigma(21)$$

where $H_i' \in F(\Phi_i^+)$ for each i.

Then, for $W' \equiv W_0 \cdot H'$, Lemma 5.3(a) implies there exists $u \in F(\mathcal{A})$ such that the projection of W' onto $F(\mathcal{A})$ is u^n.

As we assume that W_t is an end configuration, H'' must contain a σ-rule. By Lemma 5.1, H'' must then have a prefix of the form

$$H_1''\sigma(12)H_2''\sigma(23)\dots H_{2n-1}''\sigma(2n-1,2n)$$

where $H_i'' \in F(\Phi_i^+)$ for each i.

By a projection argument, W' and $W' \cdot H_1''$ have the same projection onto $F(\mathcal{A})$ so that Lemma 5.3(b) implies $W' \cdot H_1'' \equiv W'$. But then H_1'' must be empty by Lemma 4.3, contradicting the assumption that $\mathcal{C}$ is reduced.

□

Lemma 5.5. *The language of accepted inputs of* $\boldsymbol{M}_1$ *is* $\mathcal{L}$*. Moreover, for any* $u^n \in \mathcal{L}$*, there exists a unique accepting computation* $\mathcal{C}_1(u)$*.*

Proof. Suppose $\mathcal{C}$ is an accepting computation of some input configuration W with input w.

Lemmas 5.1 and 5.4 then imply that the history H of $\mathcal{C}$ is of the form

$$H_1\sigma(12)H_2\sigma(23)\dots H_{2n-1}\sigma(2n-1,2n)H_{2n}$$

where $H_i \in F(\Phi_i^+)$ for all i.

As $W \cdot H_1$ is $\sigma(12)H_2\sigma(23)\dots H_{2n-1}\sigma(2n-1,2n)$-admissible, Lemma 5.3(a) implies that its projection onto $F(\mathcal{A})$ is u^n for some $u \in F(\mathcal{A})$. A projection argument then implies $w \equiv u^n \in \mathcal{L}$.

Conversely, for any $u \in F(\mathcal{A})$, let $H_0(u)$ be the history of $\mathcal{D}_1(u)$ (see Lemma 5.3(b)). Further, let $H_1(u)$ be the natural copy of u read right to left in $F(\Phi_1^+)$ and $H_{2n}(u)$ be the natural copy of u read left to right in $F(\Phi_{2n}^+)$.

Then, for W the input configuration with input u^n and $H(u) \equiv H_1(u)H_0(u)H_{2n}(u)$, W is

$H(u)$-admissible with $W \cdot H(u)$ the accept configuration. Let $\mathcal{C}_1(u)$ be the reduced computation with history $H(u)$ accepting u^n.

Suppose $\mathcal{C}'$ is an arbitrary accepting computation of u^n. Again, Lemmas 5.1 and 5.4 imply that the history H' of $\mathcal{C}'$ can be factored as

$$H' \equiv H_1'\sigma(12)H_2'\sigma(23)\dots H_{2n-1}'\sigma(2n-1,2n)H_{2n}'$$

where $H_i' \in F(\Phi_i^+)$ for all i.

Then the projection of $W \cdot H_1'$ onto $F(\mathcal{A})$ is u^n, so that Lemma 5.3(b) yields

(i) $W \cdot H_1' \equiv W \cdot H_1 = q_0(1)u_1^{n-1}q_1(1)q_2(1)u_3q_3(1)q_4(1)$,

(ii) $\sigma(12)H_2'\sigma(23)\dots H_{2n-1}'\sigma(2n-1,2n) \equiv H_0(u)$, and

(iii) $W \cdot (H_1'H_0(u)) \equiv q_0(2n)q_1(2n)q_2(2n)u_3q_3(2n)u_4^{-1}q_4(2n)$.

where u_i is the natural copy of u in $F(Y_i)$.

Applications of Lemma 4.3 to the restriction of the subcomputations with history H_1' and H_{2n}' to the Q_2Q_3-sector then imply that $H_1' \equiv H_1(u)$ and $H_{2n}' \equiv H_{2n}(u)$.

Thus, $H' \equiv H(u)$, and so $\mathcal{C}' = \mathcal{C}_1(u)$.

□

As $H_1(u)$ and $H_{2n}(u)$ are copies of u (read in different directions), Lemma 5.3(b) implies that the length of $\mathcal{C}_1(u)$ is $2n\|u\|+2n-1$.

Lemma 5.6. *Let $\mathcal{C} : W_0 \to \dots \to W_t$ be a reduced computation with base $Q_1Q_2Q_3$. Suppose the step history of $\mathcal{C}$ does not contain the letter (1) or $(2n)$. Then for $m = \max(|W_0|_a, |W_t|_a)$, $t \leq 2n(m+1)$.*

Proof. By Lemma 5.1, the step history of $\mathcal{C}$ (or its inverse) is a subword of

$$(12)(2)(3)\dots(2n-1)(2n-1,2n)$$

Suppose the history of $\mathcal{C}$ contains no σ-rule. Then the restriction of $\mathcal{C}$ to the Q_1Q_2-sector satisfies the hypotheses of Lemma 4.3, so that $t \leq |W_0|_a + |W_t|_a \leq 2m$.

So, we may factor the history H of $\mathcal{C}$ as $H_1H_2H_3$, where H_1 and H_3 contain no σ-rules and H_2 starts and ends with a σ-rule. Note that we may have $\|H_2\| = 1$ or $\|H_i\| = 0$ for $i = 1, 3$.

Let w be the projection of W_0 onto $F(\mathcal{A})$. Then a projection argument implies that the projection of W_i onto $F(\mathcal{A})$ is w for all $0 \leq i \leq t$. Hence, $|W_i|_a \geq \|w\|$ for all i.

Lemma 5.2 applies to any subcomputation whose step history is of the form $(j-1, j)(j)(j, j+1)$. So, for the subcomputation $W_r \to \cdots \to W_s$ with history H_2, we have $s - r \leq (2n-2)\|w\| + (2n-1)$ and $|W_i|_a = \|w\|$ for all $r \leq i \leq s$.

Since W_s is σ-admissible for some σ-rule, one of its sectors must be empty. The restriction of the subcomputation $W_s \to \cdots \to W_t$ to this sector then satisfies the hypotheses of Lemma 4.3, so that $t - s \leq |W_t|_a$. An analogous argument implies $r \leq |W_0|_a$.

Hence, $t \leq |W_0|_a + |W_t|_a + (2n-2)\|w\| + (2n-1) \leq 2n(m+1)$.

□

Lemma 5.7. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_1$ in the standard base. Suppose the step history of $\mathcal{C}$ is (i) for some $i \in \{2, \dots, 2n-1\}$ and W_0 is σ-admissible for some σ-rule. Then $|W_0|_a \leq 3|W_t|_a$.*

Proof. Let $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be the restriction of $\mathcal{C}$ to the base $Q_1Q_2Q_3$.

Let w be the projection of W_0' onto $F(\mathcal{A})$. As W_0 is σ-admissible, one sector of W_0' must be empty. The restriction of $\mathcal{C}'$ to this sector then satisfies the hypotheses of Lemma 4.3, so that $t \leq |W_t'|_a$.

Further, a projection argument implies $|W_t'|_a \geq \|w\| = |W_0'|_a$.

Now let $\mathcal{C}'' : W_0'' \to \cdots \to W_t''$ be the restriction of $\mathcal{C}$ to the input sector. As the application of any rule inserts/deletes at most one letter from the input sector, $|W_0''|_a \leq |W_t''|_a + t$.

Similarly, for $\mathcal{C}''' : W_0''' \to \cdots \to W_t'''$ the restriction to the Q_3Q_4-sector, $|W_0'''|_a \leq |W_t'''|_a + t$.

Hence, $|W_0|_a = |W_0'|_a + |W_0''|_a + |W_0'''|_a \leq |W_t'|_a + |W_t''|_a + |W_t'''|_a + 2t \leq 3|W_t|_a$.

□

Lemma 5.8. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_1$ in the standard base. Suppose W_0 is an input configuration and the step history of $\mathcal{C}$ does not contain the letter $(2n)$. Then $|W_0|_a \leq 9n|W_t|_a$.*

Proof. Let $W_0 \to \cdots \to W_r$ be the maximal subcomputation with step history (1). A projection argument implies $|W_0|_a \leq |W_r|_a$. So, it suffices to assume that $r < t$ and prove $|W_r|_a \leq 6n|W_t|_a$.

By Lemma 5.1, the step history of the subcomputation $\mathcal{C}' : W_r \to \cdots \to W_t$ must be a subword of $(12)(2)(3)\dots(2n-1)(2n-1,2n)$.

Let $j \leq 2n-1$ be the maximal index for which $\mathcal{C}'$ has a nonempty maximal subcomputation with step history (j). As σ-rules do not alter the a-length of an admissible word, we may then assume that the step history of $\mathcal{C}'$ is $(12)(2)\dots(j)$.

Let H' be the history of the subcomputation of $\mathcal{C}'$ with step history $(12)\dots(j-1,j)$.

Let $u, v \in F(\mathcal{A})$ be the reduced words such that W_r has the natural copy of u written in its Q_2Q_3-sector and the natural copy of v in its input sector.

By Lemma 5.2, $W_s \equiv W_r \cdot H'$ has the natural copy of u written in either its Q_1Q_2-sector (if j is odd) or its Q_2Q_3-sector (if j is even) and the natural copy of (the reduced form of) $vu^{-\ell}$ written in its input sector for some $\ell \leq n$.

If $2n\|u\| \geq \|v\|$, then $|W_r|_a = \|u\| + \|v\| \leq (2n+1)\|u\| \leq 3n|W_s|_a$.

Otherwise, $\|vu^{-\ell}\| \geq \|v\| - \ell\|u\| \geq \|v\| - n\|u\| \geq \frac{1}{2}\|v\|$. So,

$$|W_r|_a = \|u\| + \|v\| \leq 2(\|u\| + \|vu^{-\ell}\|) \leq 2|W_s|_a$$

As Lemma 5.7 implies $|W_s|_a \leq 3|W_t|_a$, wee have $|W_r|_a \leq 9n|W_t|_a$.

□

Lemma 5.9. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_1$ in the standard base. Suppose the first letter of the step history of $\mathcal{C}$ is $(2n, 2n-1)$. Then $t \leq 15n^2(|W_t|_a + 1)$ and $|W_0|_a \leq 12n|W_t|_a$.*

Proof. Let $u \in F(\mathcal{A})$ be the reduced word such that W_0 has the natural copy of u written in its Q_2Q_3-sector.

Suppose the step history of $\mathcal{C}$ is $(2n, 2n-1)(2n-1)$. Then for $\mathcal{C}' : W_0' \to \cdots \to W_t'$ the restriction of $\mathcal{C}$ to the Q_0Q_1- or Q_1Q_2-sector, Lemma 4.3 implies $|W_t'|_a = |W_0'|_a + t - 1$ and $t \leq |W_t'|_a + 1$. As each of the rules of Φ_{2n-1} can decrease the length of an admissible word with base Q_2Q_3 or Q_3Q_4 by at most one, it then follows that $|W_0|_a \leq |W_t|_a$.

So, Lemma 5.1 allows us to assume that $(2n, 2n-1)(2n-1)(2n-1, 2n-2)$ is a prefix of the step history of $\mathcal{C}$. As a result, W_0 has the natural copy of u^{-1} written in its Q_3Q_4-sector, so that $|W_0|_a = 2\|u\|$.

Next, suppose the step history of $\mathcal{C}$ is a subword of $(2n, 2n-1)(2n-1)(2n-2)\dots(2)(21)$ and let $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be the restriction to the base $Q_1Q_2Q_3$. Then the projection of W_i' onto $F(\mathcal{A})$ is u for all i, so that $\|u\| \leq |W_t'|_a$. Hence, $|W_0|_a \leq 2|W_t'|_a$ and, by Lemma 5.6, $t \leq 2n(|W_t'|_a + 1)$.

So, by Lemma 5.1, we may assume that the step history of $\mathcal{C}$ has prefix

$$(2n, 2n-1)(2n-1)(2n-2)\dots(2)(21)(1)$$

Let $W_0 \to \cdots \to W_s$ be the maximal subcomputation with this step history and $W_r \to \cdots \to W_s$ be the maximal subcomputation with step history (1).

By Lemma 5.3, the projection of W_r onto $F(\mathcal{A})$ is u^n and $r = (2n-2)\|u\| + (2n-1)$.

Let $v \in F(\mathcal{A})$ be the reduced word such that W_s has the natural copy of v written in its Q_2Q_3-sector. The restriction of $W_r \to \cdots \to W_s$ to the Q_2Q_3-sector satisfies the hypotheses of Lemma 4.3, so that $s - r \leq \|u\| + \|v\|$.

By Lemmas 5.1 and 5.4, we may apply Lemma 5.6 to the restriction of the subcomputation $W_s \to \cdots \to W_t$ to the subword $Q_1Q_2Q_3$. So, $t - s \leq 2n(|W_t|_a + 1)$ and $\|v\| \leq |W_t|_a$.

Hence, $t \leq 2n\|u\| + 3n|W_t|_a + 4n$.

Suppose $\|u\| \leq 2n\|v\|$. Then $|W_0|_a \leq 4n|W_t|_a$ and $t \leq (4n^2+3n)|W_t|_a+4n \leq 7n^2(|W_t|_a+1)$.

Otherwise, let $W_x \to \cdots \to W_t$ be the maximal suffix whose history contains no σ-rule. Lemma 5.2 implies that for some $0 \leq \ell \leq n$, W_x has the natural copy of (the reduced form of) $u^n v^{-\ell}$ written in its input sector. Note that $\|u^n v^{-\ell}\| \geq \|u^n\| - \ell\|v\| \geq \|u\| - n\|v\|$, so that $|W_x|_a \geq \frac{1}{2}\|u\|$.

Lemma 5.7 then implies that $\frac{1}{2}|W_0|_a = \|u\| \leq 6|W_t|_a$, so that $t \leq 15n|W_t|_a + 4n \leq 15n(|W_t|_a + 1)$.

□

Lemma 5.10. *For any reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ of $\boldsymbol{M}_1$ in the standard base, $t \leq c_0 \max(\|W_0\|, \|W_t\|)$.*

Proof. First, suppose the step history of $\mathcal{C}$ has no occurrence of $(2n)$.

By the parameter choice $c_0 >> n$ and Lemma 5.6, we may assume that the step history has an occurrence of (1). Lemma 5.1 then implies that there is exactly one occurrence of (1). Let $\mathcal{C}_1 : W_r \to \cdots \to W_s$ be the maximal subcomputation with step history (1).

Further, let $\mathcal{C}_1' : W_r' \to \cdots \to W_s'$ be the restriction of $\mathcal{C}_1$ to the Q_2Q_3-sector and let u and v be the projections of W_r' and W_s', respectively, onto $F(\mathcal{A})$. Lemma 4.3 implies that $s-r \leq \|u\|+\|v\|$.

Next, let $\mathcal{C}_0'' : W_0'' \to \cdots \to W_r''$ and $\mathcal{C}_t'' : W_s'' \to \cdots \to W_t''$ be the restrictions of the corresponding subcomputations to the base $Q_1Q_2Q_3$.

If $\mathcal{C}_0''$ is nonempty, then W_r'' must be $\sigma(12)$-admissible, so that $|W_r''|_a = \|u\|$. A projection argument then implies $|W_0''|_a \geq \|u\| = |W_r''|_a$, so that Lemma 5.6 yields $r \leq 2n(|W_0''|_a + 1)$.

Similarly, if $\mathcal{C}_t''$ is nonempty, then $|W_t''|_a \geq \|v\| = |W_s''|_a$ and $t - s \leq 2n(|W_t''|_a + 1)$.

Hence, $t \leq (2n+1)(|W_0''|_a + |W_t''|_a) + 4n \leq (4n+2)\max(|W_0|_a, |W_t|_a) + 4n \leq 6n\max(\|W_0\|, \|W_t\|)$.

Thus, we may assume that the step history of $\mathcal{C}$ has an occurrence of $(2n)$. Then, Lemma 5.4 implies that there is exactly one occurrence of $(2n)$. Let $\mathcal{C}_{2n} : W_x \to \cdots \to W_y$ be the maximal subcomputation with step history $(2n)$.

Similar to above, applying Lemma 4.3 to the restriction of $\mathcal{C}_{2n}$ to the Q_2Q_3-sector implies $s - r \leq |W_x|_a + |W_y|_a$. But then Lemma 5.9 implies that $r \leq 15n^2(|W_0|_a + 1)$, $|W_r|_a \leq 12n|W_0|_a$, $t - s \leq 15n^2(|W_t|_a + 1)$, and $|W_s|_a \leq 12n|W_t|_a$.

Hence, $t \leq 15n^2(|W_0|_a + |W_t|_a + 2) + |W_r|_a + |W_s|_a \leq (30n^2 + 24n)\max(\|W_0\|, \|W_t\|)$.

□

Lemma 5.11. *For any reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ of $\boldsymbol{M}_1$ in the standard base, $\|W_i\| \leq 3c_0 \max(\|W_0\|, \|W_t\|)$ for all $i = 0, \dots, t$.*

Proof. Note that the application of any rule of $\mathbf{M}_1$ changes the length of a configuration by at most four.

For $i \leq t/2$, Lemma 5.10 implies that the subcomputation $W_0 \to \cdots \to W_i$ has length at most $\frac{1}{2}c_0 \max(\|W_0\|, \|W_t\|)$, so that $\|W_i\| \leq \|W_0\| + 2c_0 \max(\|W_0\|, \|W_t\|)$.

For $i \geq t/2$, the analogous argument applies to the subcomputation $W_i \to \cdots \to W_t$, so that $\|W_i\| \leq \|W_t\| + 2c_0 \max(\|W_0\|, \|W_t\|)$.

Hence, for any i, $\|W_i\| \leq (2c_0 + 1)\max(\|W_0\|, \|W_t\|)$.

□

For $1 \leq i \leq 2n$ and $1 \leq j \leq 4$, suppose there exists an admissible word W with base $Q_{j-1}Q_j$ and $\theta \in \Phi_i$ such that the tape word of $W \cdot \theta$ differs from that of W. Then the application of any rule of $\mathbf{M}_1(i)$ to any admissible word with base $Q_{j-1}Q_j$ inserts/deletes one tape letter. Moreover, this insertion/deletion occurs on the same side of the tape word for fixed i and j.

If the insertion/deletion occurs on the left (resp right) of the tape word, then the subword $Q_{j-1}Q_j$ of the standard base of $\mathbf{M}_1$ is called *left-active* (resp *right-active*) for $\mathbf{M}_1(i)$.

Lemma 5.12. *For $i \in \{1, \dots, 2n\}$, let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_1(i)$ in the standard base. Assume that for some index j, $|W_j|_a > 4|W_0|_a$. Then there are $\ell, r \in \{1, 2, 3, 4\}$ such that $Q_{\ell-1}Q_\ell$ is left-active, $Q_{r-1}Q_r$ is right-active, and for the restriction $W_0' \to \cdots \to W_t'$ to either sector, $|W_j'|_a < |W_{j+1}'|_a < \cdots < |W_t'|_a$.*

Proof. Let $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be the restriction to any sector. Then by the definition of the rules, $\big||W_k'|_a - |W_{k-1}'|_a\big| = 1$ for all $1 \leq k \leq t$. Moreover, if $|W_k'|_a - |W_{k-1}'|_a = 1$, then Lemma 4.3 implies that $|W_m'|_a - |W_{m-1}'|_a = 1$ for all $m \geq k$.

Note that for each i, there exists some sector of the standard base that is left-active and another that is right-active. So, assuming the statement is false, there must exist a sector for which the restriction $\mathcal{C}'$ satisfies $|W'_j|_a = |W'_{j+1}|_a + 1$.

Then, we must have $|W'_{k-1}|_a = |W'_k|_a + 1$ for all $k \leq j$, so that $j \leq |W'_0|_a$.

For any other sector, the restriction $\mathcal{C}'' : W''_0 \to \cdots \to W''_t$ satisfies the hypotheses of Lemma 4.3, so that $|W''_j|_a \leq |W''_0|_a + j$.

But as there are three such sectors, we have $|W_j|_a \leq |W_0|_a + 3j \leq 4|W_0|_a$, yielding a contradiction.

□

5.3 Primitive Machines

As in the constructions of [18] and [25], we introduce two machines, $\mathbf{LR}(Y)$ and $\mathbf{RL}(Y)$ for an alphabet Y, that will be used to alter $\mathbf{M}_1$. These machines are called *primitive machines*.

The standard base of $\mathbf{LR}(Y)$ is $Q^{(1)}PQ^{(2)}$ with $Q^{(1)} = \{q^{(1)}\}$, $P = \{p^{(1)}, p^{(2)}\}$, and $Q^{(2)} = \{q^{(2)}\}$. The letter $p^{(1)}$ is the start letter of P, while $p^{(2)}$ is the end letter.

The tape alphabets are two disjoint copies of Y, denoted $Y^{(1)}$ and $Y^{(2)}$ and assigned in the natural way.

The positive rules of $\mathbf{LR}(Y)$ come in the following three forms:

• $\zeta^{(1)}(a) = [q^{(1)} \to q^{(1)},\ p^{(1)} \to a_1^{-1}p^{(1)}a_2,\ q^{(2)} \to q^{(2)}]$ for all $a \in Y$, where a_i is its copy in $Y^{(i)}$.

Comment. The state letter $p^{(1)}$ moves left, replacing a letter from the $Q^{(1)}P$-sector with its copy in the $PQ^{(2)}$-sector.

• $\zeta^{(12)} = [q^{(1)} \xrightarrow{\ell} q^{(1)},\ p^{(1)} \to p^{(2)},\ q^{(2)} \to q^{(2)}]$

Comment. When $p^{(1)}$ meets $q^{(1)}$, it switches to $p^{(2)}$. This is called the *connecting rule* of the machine.

• $\zeta^{(2)}(a) = [q^{(1)} \to q^{(1)},\ p^{(2)} \to a_1p^{(2)}a_2^{-1},\ q^{(2)} \to q^{(2)}]$ for all $a \in Y$, where a_i is its copy in $Y^{(i)}$.

Comment. The state letter $p^{(2)}$ moves right towards $q^{(2)}$ and replaces a letter in the $PQ^{(2)}$-sector with its copy in the $Q^{(1)}P$-sector.

The state letters of P are called *running state letters*. In practice, they 'run' left to the adjacent state letter and then right to the other, as is indicated by the name of the machine.

Lemma 5.13. *(Lemma 3.1 of [18]) Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of* ***LR***(Y) *in the standard base. Then:*

(1) if $|W_{i-1}|_a < |W_i|_a$ for some $1 \leq i \leq t-1$, then $|W_i|_a < |W_{i+1}|_a$

(2) $|W_i|_a \leq \max(|W_0|_a, |W_t|_a)$ for each i

(3) if $W_0 \equiv q^{(1)}up^{(1)}q^{(2)}$ and $W_t \equiv q^{(1)}vp^{(2)}q^{(2)}$ for some $u, v \in F(Y^{(1)})$, then $u \equiv v$, $|W_i|_a = \|u\| := \ell$ for each i, $t = 2\ell + 1$, and the $Q^{(1)}P$-sector is locked in the rule $W_\ell \to W_{\ell+1}$. Moreover, letting $\bar{u}$ be u read right to left, the history H of $\mathcal{C}$ is a copy of $\bar{u}\zeta^{(12)}u$

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

(5) if W_0 is of the form $q^{(1)}up^{(1)}q^{(2)}$, $q^{(1)}p^{(1)}uq^{(2)}$, $q^{(1)}up^{(2)}q^{(2)}$, or $q^{(1)}p^{(2)}uq^{(2)}$ for some word u, then $|W_i|_a \geq |W_0|_a$ for every i.

Lemma 5.14. *(Lemma 3.4 of [25]) Suppose $W_0 \to \cdots \to W_t$ is a reduced computation of* ***LR***(Y) *with base $Q^{(1)}PP^{-1}(Q^{(1)})^{-1}$ (or $(Q^{(2)})^{-1}P^{-1}PQ^{(2)}$) such that $W_0 \equiv q^{(1)}p^{(i)}u(p^{(i)})^{-1}(q^{(1)})^{-1}$ (or $W_0 \equiv (q^{(2)})^{-1}(p^{(i)})^{-1}vp^{(i)}q^{(2)}$) for $i = 1, 2$ and some word u (or v). Then $|W_0|_a \leq \cdots \leq |W_t|_a$.*

Lemma 5.15. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computuation of* ***LR***(Y) *in the standard base. Then $t \leq |W_0|_a + |W_t|_a + 1$.*

Proof. If the history H of $\mathcal{C}$ contains no connecting rule, then the restriction of $\mathcal{C}$ to the $Q^{(1)}P$-sector satisfies the hypotheses of Lemma 4.3. So, $t \leq |W_0|_a + |W_t|_a$.

By Lemma 5.13(4), we then assume that H contains exactly one connecting rule. Let $W_r \to W_{r+1}$ be the transition corresponding to this connecting rule. Then both W_r and W_{r+1} have empty $Q^{(1)}P$-sector, so that Lemma 4.3 yields $r \leq |W_0|_a$ and $t-r-1 \leq |W_t|_a$.

□

The machine **RL**(Y) is the right analogue of **LR**(Y). To be precise, the standard base of **RL**(Y) is $Q^{(1)}RQ^{(2)}$ with $R = \{r^{(1)}, r^{(2)}\}$, the tape alphabets are again two copies of Y denoted $Y^{(1)}$ and $Y^{(2)}$, and the positive rules are:

- $\xi^{(1)}(a) = [q^{(1)} \to q^{(1)},\ r^{(1)} \to a_1 r^{(1)} a_2^{-1},\ q^{(2)} \to q^{(2)}]$ for all $a \in Y$, where a_i is its copy in $Y^{(i)}$
- $\xi^{(12)} = [q^{(1)} \to q^{(1)},\ r^{(1)} \xrightarrow{\ell} r^{(2)},\ q^{(2)} \to q^{(2)}]$
- $\xi^{(2)}(a) = [q^{(1)} \to q^{(1)},\ r^{(2)} \to a_1^{-1} r^{(2)} a_2,\ q^{(2)} \to q^{(2)}]$ for all $a \in Y$, where a_i is its copy in $Y^{(i)}$.

There are obvious analogues of Lemmas 5.13-5.15 in the setting of **RL**(Y), which can be verified in much the same ways.

When the alphabet Y is contextually clear, it is convenient to omit it from the names of these machines. So, there will be reference in subsequent constructions to the machines **LR** and **RL**.

5.4 The machine $\mathbf{M}_2$

The next machine in our construction, $\mathbf{M}_2$, is the composition of copies of the submachines $\mathbf{M}_1(i)$ with copies of the primitive machines **LR** and **RL**.

Four new parts are added to the standard base of $\mathbf{M}_1$, producing the standard base

$$Q_0 P_1 Q_1 R_1 Q_2 R_2 Q_3 P_4 Q_4$$

However, the parts of the form Q_i have more letters than the corresponding parts of the hardware of $\mathbf{M}_1$. The makeup of each part is contextually clear from the definition of the positive rules below.

The tape alphabets of:

- the Q_0P_1- and P_1Q_1-sectors are copies of Y_1,
- the Q_1R_1- and R_1Q_2-sectors are copies of Y_2,
- the Q_2R_2- and R_2Q_3-sectors are copies of Y_3, and
- the Q_3P_4- and P_4Q_4-sectors are copies of Y_4.

The Q_0P_1-sector functions as the machine's input sector.

The idea of the function of $\mathbf{M}_2$ is the following. The Q_0P_1-, R_1Q_2-, R_2Q_3-, and Q_3P_4-sectors are identified with the sectors of the standard base of $\mathbf{M}_1$, so that a computation of $\mathbf{M}_1$ may be carried out while all other sectors are locked. However, before a transition between two steps of such a computation can take place, every unlocked sector must take part in at least one copy of a standard computation of a primitive machine.

To be precise, $\mathbf{M}_2$ is the concatenation of $4n-1$ submachines, which are denoted $\mathbf{M}_2(2), \dots, \mathbf{M}_2(4n)$. Consequently, each part of the state letters is the disjoint union of $4n-1$ sets corresponding to the hardware of these submachines.

The submachines are concatenated in the natural order. As such, the state letters of $\mathbf{M}_2(2)$ and $\mathbf{M}_2(4n)$ function as the start and end letters, respectively.

To force the correct order of this concatenation, we introduce transition rules $\theta(i, i+1)^{\pm 1}$ for $2 \leq i \leq 4n-1$. The rule $\theta(i, i+1)$ changes the state letters from the end letters of $\mathbf{M}_2(i)$ to the start letters of $\mathbf{M}_2(i+1)$. Further, a sector of the standard base is locked by $\theta(i, i+1)$ if and only if it is locked by every rule of $\mathbf{M}_2(i)$ or every rule of $\mathbf{M}_2(i+1)$. The domain in a sector that is unlocked is the entire corresponding tape alphabet.

For $1 \leq i \leq 2n$, the submachine $\mathbf{M}_2(2i)$ corresponds to the submachine $\mathbf{M}_1(i)$, with each part of the state letters consisting of a singleton. For any $\theta \in \Phi_i^+$, the corresponding positive rule of $\mathbf{M}_2(2i)$ locks the P_1Q_1-, Q_1R_1-, Q_2R_2-, and P_4Q_4-sectors and operates in the remaining sectors as θ, identifying these sectors with the standard base of $\mathbf{M}_1$ in the obvious way. As such, the positive rules of $\mathbf{M}_2(2i)$ are identified with Φ_i^+.

For $3 \leq i \leq 4n-3$ odd, the submachine $\mathbf{M}_2(i)$ is the concatenation of two submachines, which are denoted $\mathbf{M}_2(i^-)$ and $\mathbf{M}_2(i^+)$ and concatenated in this order. To achieve this concatentation, we introduce more transition rules, denoted $\chi_i^{\pm 1}$. To differentiate these transition rules for clarity, we henceforth refer to them as *χ-rules* rather than transition rules. The rule χ_i changes the state letters from the end letters of $\mathbf{M}_2(i^-)$ to the start letters of $\mathbf{M}_2(i^+)$ and has the same domain as the rule $\theta(i-1, i)$.

The submachine $\mathbf{M}_2(i^-)$ operates as the machine $\mathbf{LR} = \mathbf{LR}(Y_1)$. The subword $Q_0P_1Q_1$ is identified with the standard base of $\mathbf{LR}$, with each rule operating on this subword as its corresponding rule. Each other part of the standard base consists of a singleton. Additionally, the R_2Q_3-sector (respectively R_1Q_2-sector) remains unlocked by every rule if i is of the form $4\ell - 1$ (respectively $4\ell + 1$). Every remaining sector of the standard base is locked by every rule.

If i is of the form $4\ell - 1$ (with $\ell < n$), then the submachine $\mathbf{M}_2(i^+)$ operates as the machine $\mathbf{RL} = \mathbf{RL}(Y_3)$. The subword $Q_2R_2Q_3$ is identified with the standard base of $\mathbf{RL}$, with each rule operating on this subword as its corresponding rule. Each other part of the standard base consists of a singleton. Additionally, the Q_0P_1-sector is unlocked by every rule, while all remaining sectors of the standard base are locked by every rule.

If i is of the form $4\ell + 1$, then the submachine $\mathbf{M}_2(i^+)$ operates as the machine $\mathbf{RL} = \mathbf{RL}(Y_2)$. The subword $Q_1R_1Q_2$ is identified with the standard base of $\mathbf{RL}$, while each other part of the standard base consists of a singleton. Again, the Q_0P_1-sector is unlocked by every rule, while all remaining sectors of the standard base are locked by every rule.

Finally, the submachine $\mathbf{M}_2(4n-1)$ is the concatenation of k submachines, where k is the parameter specified in Chapter 4.3. These submachines are denoted $\mathbf{M}_2((4n-1)_1), \dots, \mathbf{M}_2((4n-1)_k)$ and are concatenated in the natural way. So, the start letters of $\mathbf{M}_2((4n-1)_1)$ and the end letters of $\mathbf{M}_2((4n-1)_k)$ function as the start and end letters of $\mathbf{M}_2(4n-1)$, respectively.

To force the correct order of this concatenation, we introduce more transition rules, denoted $\chi(j, j+1)^{\pm 1}$. The rule $\chi(j, j+1)$ changes the state letters from the end letters of $\mathbf{M}_2((4n-1)_j)$ to the start letters of $\mathbf{M}_2((4n-1)_{j+1})$, locking all sectors of the standard base except for the R_2Q_3-

and Q_3P_4-sectors.

As with the transition rules within other submachines, the rules $\chi(j,j+1)^{\pm 1}$ are called χ-rules. As a result, forthcoming references to 'transition rules' of $\mathbf{M}_2$ are implicitly restricted to those of the form $\theta(i,i+1)^{\pm 1}$.

Each submachine $\mathbf{M}_2((4n-1)_j)$ operates in parallel as a copy of $\mathbf{RL}=\mathbf{RL}(Y_3)$ on the subword $Q_2R_2Q_3$ and a copy of $\mathbf{LR}=\mathbf{LR}(Y_4^{-1})$ on the subword $Q_3P_4Q_4$. As such, for every $a\in\mathcal{A}$ and $i\in\{1,2\}$, there exists a positive rule of this submachine that simultaneously acts as $\xi^{(i)}(a)$ on the subword $Q_2R_2Q_3$ and as $\zeta^{(i)}(a^{-1})$ on the subword $Q_3P_4Q_4$. The subsets of Q_0, P_1, Q_1, and R_1 corresponding to this submachine are singletons, while the remaining sectors are locked by every rule.

Note that we may interpret $\mathbf{M}_2$ as the concatenation of $6n-4+k$ submachines, which are concatenated in the following order:

$$\mathbf{M}_2(2),\ \mathbf{M}_2(3^-),\ \mathbf{M}_2(3^+),\ \dots,\ \mathbf{M}_2(4n-4),\ \mathbf{M}_2((4n-3)^-),\ \mathbf{M}_2((4n-3)^+),$$
$$\mathbf{M}_2(4n-2),\ \mathbf{M}_2((4n-1)_1),\ \dots,\ \mathbf{M}_2((4n-1)_k),\ \mathbf{M}_2(4n)$$

5.5 Standard computations of $\mathbf{M}_2$

The step history of a reduced computation of $\mathbf{M}_2$ is defined in much the same way as it is defined for a reduced computation of $\mathbf{M}_1$. As such, we first factor the computation's history so that each factor is either the history of a maximal subcomputation of one of the defining submachines $\mathbf{M}_2(i)$ or a transition rule $\theta(i,i+1)^{\pm 1}$. Then, a factor corresponding to a computation of $\mathbf{M}_2(i)$ is represented by (i) and a factor corresponding to a transition rule $\theta(i,j)$ is represented by (ij), where we take $\theta(i,i+1)^{-1}\equiv\theta(i+1,i)$.

The notational conventions described in Chapter 5.2 are used for step histories of this machine. For example, we may write the step history of a reduced computation of $\mathbf{M}_2$ as $(2)(3)$, omitting reference to the rule $\theta(23)$ as its presence is clear from its necessity.

Similarly, a one-step computation of $\mathbf{M}_2$ is a reduced computation whose step history has exactly one letter corresponding to a computation of $\mathbf{M}_2(i)$.

The following statement is an immediate consequence of Lemma 5.1.

Lemma 5.16. *Let $\mathcal{C}$ be a reduced computation of $\boldsymbol{M}_2$ with base B.*

(a) If B contains a subword of the form $(R_2Q_3)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be of the form $(4i+1, 4i+2)(4i+2)(4i+2, 4i+1)$ or $(4i+1, 4i)(4i)(4i, 4i+1)$ for $1 \leq i \leq n-1$.

(b) If B contains a subword of the form $(R_1Q_2)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be of the form $(4i-1, 4i)(4i)(4i, 4i-1)$ or $(4i+3, 4i+2)(4i+2)(4i+2, 4i+3)$ for $1 \leq i \leq n-1$.

(c) If B contains a subword of the form $(Q_3P_4)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be of the form $(4n-3, 4n-2)(4n-2)(4n-2, 4n-3)$.

Further, the following statement is an immediate consequence of Lemma 5.13(4).

Lemma 5.17. *Let $\mathcal{C}$ be a reduced computation of $\boldsymbol{M}_2$ in the standard base. Then the step history of $\mathcal{C}$ cannot be $(2i, 2i+1)(2i+1)(2i+1, 2i)$ or $(2i+2, 2i+1)(2i+1)(2i+1, 2i+2)$ for any $1 \leq i \leq 2n-1$.*

Suppose $\mathcal{C}$ is a one-step computation of $\mathbf{M}_2$ with step $(4n-1)$. Then the history of $\mathcal{C}$ is called *controlled* if it (or its inverse) is of the form

$$\chi(j-1, j)H'\chi(j, j+1)$$

for $1 \leq j \leq k$, where H' contains no χ-rule, $\chi(0,1)$ is taken to be $\theta(4n-2, 4n-1)$, and $\chi(k, k+1)$ is taken to be $\theta(4n-1, 4n)$.

Lemma 5.18. *Let $\mathcal{C}: W_0 \to \dots \to W_t$ be a reduced computation of $\boldsymbol{M}_2$ with controlled history H. Then the base B of the computation is reduced and all configurations are uniquely defined by H and B. Moreover, if $\mathcal{C}$ is a computation in the standard base, then $|W_i|_a = |W_0|_a$ for all $0 \leq i \leq t$, $\|H\| = |W_0|_a + 3$, and W_0 is accepted by $\boldsymbol{M}_2$.*

Proof. Without loss of generality, suppose $H \equiv \chi(j-1,j)H'\chi(j,j+1)$. Then, any sector of the standard base not locked by $\chi(j-1,j)$ is locked by the connecting rule of $\mathbf{M}_2((4n-1)_j)$. So, Lemma 4.2 implies that the base must be reduced. Lemma 5.13(3) then implies that $\mathcal{C}$ is uniquely defined by H and B.

If $\mathcal{C}$ is a computation in the standard base, then the parallel nature of the rules implies that there exists $w \in F(\mathcal{A})$ such that W_0 has the natural copy of w written in its R_2Q_3-sector and the natural copy of w^{-1} written in its Q_3P_4-sector. By Lemma 5.13(3), we then have $|W_i|_a = |W_0|_a$ for all i and $\|H\| = 2\|w\| + 3 = |W_0|_a + 3$.

Using Lemma 5.13(3), we may construct a reduced computation $W_0 \to \cdots \to W_h$ of the submachine $\mathbf{M}_2(4n-1)$ such that W_h is $\theta(4n-1,4n)$-admissible. Then, $W \equiv W_h \cdot \theta(4n-1,4n)$ is the end configuration with the natural copy of w written in its R_2Q_3-sector and the natural copy of w^{-1} written in its Q_3P_4-sector. Setting H'' as the natural copy of w in $F(\Phi_{2n}^+)$, $W \cdot H''$ is the accept configuration. Hence, W_0 is accepted.

□

The following statement is a similar consequence of Lemma 5.13(3).

Lemma 5.19. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\mathbf{M}_2$ with history H. Suppose the step history of $\mathcal{C}$ is $(2i,2i+1)(2i+1)(2i+1,2i+2)$ for some $1 \le i \le 2n-2$. Then the base B of the computation is reduced and all configurations are uniquely defined by H and B.*
Moreover, if $\mathcal{C}$ is a computation in the standard base, then $|W_i|_a = |W_0|_a$ for all $0 \le i \le t$ and $\|H\| = 2|W_0|_a + 5$.

A configuration of $\mathbf{M}_2$ is called *tame* if its P_1Q_1-, Q_1R_1-, Q_2R_2-, and P_4Q_4-sectors are all empty. Note that for $i = 1, \dots, 2n$, a tame configuration W of $\mathbf{M}_2(2i)$ corresponds naturally to a configuration W' of $\mathbf{M}_1(i)$. Moreover, if W is admissible for a rule of $\mathbf{M}_2(2i)$, then W' is admissible for the corresponding rule of $\mathbf{M}_1(i)$. Similarly, if W is $\theta(2i,2i+1)$-admissible (respectively $\theta(2i,2i-1)$-admissible), then W' is $\sigma(i,i+1)$-admissible (respectively $\sigma(i,i-1)$-admissible).

Suppose $\mathcal{C} : W_0 \to \cdots \to W_t$ is a reduced computation of $\mathbf{M}_2$ in the standard base such that neither the first nor last letter of its step history is of the form $(2i+1)$. Then by Lemma 5.17, any occurrence of $(2i+1)$ in the step history of $\mathcal{C}$ must be part of a subword that is of the form

(a) $(2i, 2i+1)(2i+1)(2i+1, 2i+2)$ or

(b) $(2i+2, 2i+1)(2i+1)(2i+1, 2i)$.

Let $W_r \to \cdots \to W_s$ be a subcomputation of $\mathcal{C}$ with such a step history. Then W_r and W_s are both tame configurations. Moreover, Lemma 5.13(3) implies that the configurations W_r' and W_s' of $\mathbf{M}_1$ corresponding to W_r and W_s, respectively, satisfy $W_r' \cdot \sigma(i, i+1) \equiv W_s'$ if the step history is of the form (a) or $W_r' \cdot \sigma(i+1, i) \equiv W_s'$ if the step history is of the form (b).

So, we can associate to $\mathcal{C}$ a reduced computation $\mathcal{C}'$ of $\mathbf{M}_1$ by doing the following:

- replace each subcomputation $W_r \to \cdots \to W_s$ whose step history is of the form (a) with the single transition $W_r' \to W_r' \cdot \sigma(i, i+1)$,
- replace each subcomputation $W_r \to \cdots \to W_s$ whose step history is of the form (b) with the single transition $W_r' \to W_r' \cdot \sigma(i+1, i)$,
- if the first letter of the step history is of the form $(2i+1, 2i)$ (respectively $(2i-1, 2i)$), then replace the transition $W_0 \to W_1$ with the transition $W_1' \cdot \sigma(i, i+1) \to W_1'$ (respectively $W_1' \cdot \sigma(i, i-1) \to W_1'$),
- if the last letter of the step history is $(2i, 2i+1)$ (respectively $(2i, 2i-1)$), then replace the transition $W_{t-1} \to W_t$ with the transition $W_{t-1}' \to W_{t-1}' \cdot \sigma(i, i+1)$ (respectively $W_{t-1}' \to W_{t-1}' \cdot \sigma(i, i-1)$), and
- replace all other transitions $W_{j-1} \to W_j$ with the corresponding transition $W_{j-1}' \to W_j'$.

In this case, the reduced computation $\mathcal{C}'$ is called the $\boldsymbol{M}_1$ *computation associated to* $\mathcal{C}$.

Note that any subcomputation of $\mathcal{C}$ that is 'removed' to construct $\mathcal{C}'$ corresponds to an occurrence of a σ-rule in the history of $\mathcal{C}'$. Hence, Lemmas 5.1 and 5.4 imply that at most $8n$ distinct subcomputations are removed.

Lemma 5.20. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation such that W_0 is an end configuration. If the history H of $\mathcal{C}$ contains a transition rule, then W_t is not an end configuration.*

Proof. Assuming toward contradiction, neither the first nor the last letter of the step history of $\mathcal{C}$ can be of the form $(2i+1)$. So, we may construct $\mathcal{C}'$ the $\mathbf{M}_1$ computation associated to $\mathcal{C}$.

Then $\mathcal{C}'$ is a reduced computation of $\mathbf{M}_1$ starting and ending with an end configuration. However, the existence of a transition rule in H necessitates the existence of a σ-rule in the history of $\mathcal{C}'$, so that $\mathcal{C}'$ contradicts Lemma 5.4.

□

Lemma 5.21. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_2$ in the standard base. Suppose W_0 is an input configuration and the step history of $\mathcal{C}$ does not contain the letter $(4n)$. Then $|W_0|_a \leq 9n|W_t|_a$.*

Proof. Let $\mathcal{D} : W_0 \to \cdots \to W_s$ be the maximal subcomputation such that the last letter of the step history of $\mathcal{D}$ is not of the form $(2i+1)$. Then, we may construct $\mathcal{D}' : W_0' \to \cdots \to W_s'$ the $\mathbf{M}_1$ computation associated to $\mathcal{D}$.

Lemma 5.8 then implies that $|W_0|_a = |W_0'|_a \leq 9n|W_s'|_a = 9n|W_s|_a$.

If the subcomputation $W_s \to \cdots \to W_t$ is nonempty, then its step history is of the form $(2i+1)$. But then Lemma 5.13 implies $|W_s|_a \leq |W_t|_a$, so that $|W_0|_a \leq 9n|W_t|_a$.

□

The next statement follows from an analogous proof, using Lemma 5.9 in place of 5.8.

Lemma 5.22. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_2$ in the standard base. Suppose the first letter of the step history of $\mathcal{C}$ is $(4n, 4n-1)$. Then $|W_0|_a \leq 12n|W_t|_a$.*

Lemma 5.23.

(a) Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_2$. Suppose W_0 is a start configuration and W_t is an end configuration. Then there exists $u \in F(\mathcal{A})$ such that the projection of W_0 onto $F(\mathcal{A})$ is $u^n \in \mathcal{L}$.

(b) For all $u \in F(\mathcal{A})$, there exists a unique reduced computation $\mathcal{D}_2(u) : W_0 \to \cdots \to W_t$ of $\boldsymbol{M}_2$ in the standard base with step history $(23)(3)\dots(4n-1)(4n-1,4n)$ and such that the projection of W_0 onto $F(\mathcal{A})$ is u^n.

(c) Let t be the length of $\mathcal{D}_2(u)$ and ℓ be the length of the subcomputation with step history $(4n-2,4n-1)(4n-1)(4n-1,4n)$. Then $\ell = 2k\|u\| + 2k + 1$ and $t - \ell \leq c_0(\|u\| + 1)$.

Proof. (a) Note that neither the first nor the last letter of the step history of $\mathcal{C}$ can be of the form $(2i+1)$. So, we may construct $\mathcal{C}' : W_0' \to \cdots \to W_t'$ the $\mathbf{M}_1$ computation associated to $\mathcal{C}$.

Then W_0' is a start configuration and W_t' is an end configuration, so that Lemma 5.3 implies that there exists $u \in F(\mathcal{A})$ such that the projection of W_0' onto $F(\mathcal{A})$ is u^n. As W_0 is tame, its projection must also be u^n.

(b) For any reduced computation $\mathcal{C}$ satisfying the statement, the $\mathbf{M}_1$ computation associated to $\mathcal{C}$ must be the computation $\mathcal{D}_1(u)$ in Lemma 5.3(b).

The removed computations correspond to primitive computations at the σ-rules. As the tape words of the terminal configuration of $\mathcal{D}_1(u)$ are mutually inverse, the statement follows from Lemma 5.13(3).

(c) By Lemma 5.3(b), the subcomputation with step history $(4n-2,4n-1)(4n-1)(4n-1,4n)$ operates on the base $Q_2R_2Q_3$ as k copies of the standard computation of **RL** with tape word u. So, Lemma 5.13(3) implies $\ell = 2k\|u\| + 2k + 1$.

Let $\mathcal{E}$ be the maximal subcomputation of $\mathcal{D}_2(u)$ with step history $(23)(3)\dots(4n-2)$. So, the length of $\mathcal{E}$ is $t - \ell$.

The $\mathbf{M}_1$ computation associated to $\mathcal{E}$ is the maximal subcomputation of $\mathcal{D}_1(u)$ with step history $(12)(2)(3)\dots(2n-1)$. So, its length is $(2n-2)(\|u\|+1)$.

The subcomputations removed from $\mathcal{E}$ correspond to the subcomputations with step history of the form $(2i, 2i+1)(2i+1)(2i+1, 2i+2)$. Let $W_r \to \cdots \to W_s$ be such a subcomputation. Then W_r has the natural copy of u written in its R_2Q_3- or R_1Q_2-sector, depending on the parity of i, and the natural copy of u^{n-m} written in its Q_0P_1-sector for some $1 \leq m \leq n-1$. Lemma 5.19 then implies that $s - r = 2\|u\| + 2\|u^{n-m}\| + 5 \leq 2(n+1)\|u\| + 5$.

As there are $2n-2$ such subcomputations of $\mathcal{E}$, $t - \ell \leq (2n-2)\left((2n+3)\|u\| + 6\right)$. So, the statement follows by a parameter choice of c_0.

□

Lemma 5.24. *The language of accepted inputs of* $\boldsymbol{M}_2$ *is* $\mathcal{L}$*. Moreover, for any* $u^n \in \mathcal{L}$*, there exists a unique accepting computation* $\mathcal{C}_2(u)$*.*

Proof. By Lemma 5.23(a), any accepted input must be an element of $\mathcal{L}$.

Conversely, for any $u^n \in \mathcal{L}$, Lemma 5.5 provides a unique accepting computation $\mathcal{C}_1(u)$. Concatenating the steps of $\mathcal{C}_1(u)$ with primitive computations corresponding to Lemma 5.13(3) then yields an accepting computation $\mathcal{C}_2(u)$.

Uniqueness of this computation follows from Lemmas 5.16, 5.17, 5.20, 5.23(b), and 4.3.

□

Lemma 5.25. *For* i *odd, any reduced computation* $\mathcal{C} : W_0 \to \cdots \to W_t$ *of* $\boldsymbol{M}_2(i)$ *in the standard base satisfies* $t \leq 2k \max(\|W_0\|, \|W_t\|)$*.*

Proof. If the history contains no χ-rules, then the statement follows from Lemma 5.15. So, we may assume there exists a maximal subcomputation $W_r \to \cdots \to W_s$ starting and ending with χ-rules.

Suppose $i \neq 4n-1$. Then Lemma 5.13(4) implies that there is exactly one χ-rule in the history H of $\mathcal{C}$, so that $s = r+1$. Lemmas 5.13(5) and 5.15 then imply that $t - s \leq 2|W_t|_a + 1$, $|W_s|_a \leq |W_t|_a$, $r \leq 2|W_0|_a + 1$, and $|W_r|_a \leq |W_0|_a$. Hence, $t \leq 4\max(\|W_0\|, \|W_t\|)$.

If $i = 4n-1$, then there are at most k copies of primitive computations in the subcomputation $W_r \to \cdots \to W_s$. Lemma 5.18 then implies that $s - r \leq k\|W_r\| = k\|W_s\|$. Again, Lemmas 5.13(5) and 5.15 imply that $t - s \leq 2\|W_t\|$, $\|W_s\| \leq \|W_t\|$, $r \leq 2\|W_0\|$, and $\|W_r\| \leq \|W_0\|$.

Hence, $t \leq (k+4)\max(\|W_0\|, \|W_t\|)$, so that the statement is implied by the parameter choice $k \geq 4$.

□

Lemma 5.26. *For any reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ of $\boldsymbol{M}_2$ in the standard base, $t \leq c_1 \max(\|W_0\|, \|W_t\|)$.*

Proof. By Lemma 5.25 and the parameter choice $c_1 >> k$, we may assume that $\mathcal{C}$ is not a one-step computation with step $(2i+1)$.

Let $\mathcal{D} : W_x \to \cdots \to W_y$ be the maximal subcomputation of $\mathcal{C}$ such that neither the first nor the last letter of the step history of $\mathcal{D}$ is of the form $(2i+1)$.

Then, let $\mathcal{D}' : W_x' \to \cdots \to W_y'$ be the $\mathbf{M}_1$ computation associated to $\mathcal{D}$ and ℓ be the length of $\mathcal{D}'$ (note that ℓ may be less than $y-x$ if subcomputations are removed).

By Lemma 5.10, $\ell \leq c_0 \max(\|W_x'\|, \|W_y'\|) \leq c_0 \max(\|W_x\|, \|W_y\|)$. Moreover, by Lemma 5.11, $\|W_i'\| \leq 3c_0 \max(\|W_x'\|, \|W_y'\|) \leq 3c_0 \max(\|W_x\|, \|W_y\|)$ for all $x \leq i \leq y$ for which W_i' is part of the computation $\mathcal{D}'$.

The difference between $y-x$ and ℓ arises from removed subcomputations $W_r \to \cdots \to W_s$. By Lemma 5.25, the length of such a subcomputation is at most $2k\|W_r\|$. So, this removed subcomputation has length at most $6c_0 k \max(\|W_x\|, \|W_y\|)$.

As there are at most $8n$ such removed subcomputations, we have $y-x-\ell \leq k^2 \max(\|W_x\|, \|W_y\|)$ as $k >> c_0$.

Hence, $y-x \leq (k^2+c_0)\max(\|W_x\|, \|W_y\|)$.

If the subcomputation $W_y \to \cdots \to W_t$ is nonempty, then its step history is of the form $(2i+1)$. Lemma 5.25 then implies that $t-y \leq 2k\max(\|W_y\|, \|W_t\|)$. But W_y is tame, so that Lemma 5.13 implies $|W_y|_a \leq |W_t|_a$. Hence, $t-y \leq 2k\|W_t\|$.

By the analogous arguments, $|W_x|_a \leq |W_0|_a$ and $x \leq 2k\|W_0\|$.

Thus, $t \leq (k^2+4k+c_0)\max(\|W_0\|, \|W_t\|)$, so that the statement is implied by the parameter choices $c_1 >> k >> c_0$.

□

For $2 \leq i \leq 4n$, let UV be a two-letter subword of the standard base of $\mathbf{M}_2$. Suppose that the application of any rule of $\mathbf{M}_2(i)$ to an admissible word with base UV either leaves the tape word fixed or inserts/deletes one letter on the left of the tape word. Then UV is called *left-active* in $\mathbf{M}_2(i)$. A *right-active* two-letter subword is defined analogously.

For example, the subword Q_0P_1 is right-active for the submachine $\mathbf{M}_2(3)$ even though applications of the rules of $\mathbf{M}_2(3^+)$ do not alter an admissible word with base Q_0P_1.

Note that a two-letter subword of the standard base of $\mathbf{M}_2$ is left-active (resp right-active) in $\mathbf{M}_2(2i)$ if and only if it is operated upon as a sector of $\mathbf{M}_1(i)$ and the corresponding two-letter subword of the standard base of $\mathbf{M}_1$ is left-active (resp right-active) in $\mathbf{M}_1(i)$.

Further, a two-letter subword that is neither left-active nor right-active in $\mathbf{M}_2(i)$ is locked by every rule of the submachine.

Lemma 5.27. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation in the standard base of $\mathbf{M}_2(i)$ for some i. Assume that for some index j, $|W_j|_a > 4|W_0|_a$. Then there are subwords $U_\ell V_\ell$ and $U_r V_r$ of the standard base such that $U_\ell V_\ell$ is left-active in $\mathbf{M}_2(i)$, $U_r V_r$ is right-active in $\mathbf{M}_2(i)$, and for $W_0' \to \cdots \to W_t'$ the restriction of $\mathcal{C}$ to either sector, $|W_j'|_a < |W_{j+1}'|_a < \cdots < |W_t'|_a$.*

Proof. If i is even, then the statement is an immediate consequence of Lemma 5.12. So, we may assume i is odd.

Further, inducting on t, we may assume that $|W_1|_a > |W_0|_a$.

If $i \neq 4n-1$, then Lemma 5.13(1) implies that $\mathcal{C}$ is a computation of either $\mathbf{M}_2(i^-)$ or $\mathbf{M}_2(i^+)$, i.e there exists a three-letter subword of the standard base on which $\mathcal{C}$ operates as a primitive machine. As one of the corresponding sectors is left-active while the other is right-active, the statement follows.

If $i = 4n-1$, then the R_2Q_3- and P_4Q_4-sectors are left-active, the Q_2R_2- and Q_3P_4-sectors are right-active, and all other sectors are locked. As any rule changes the a-length of any of the four sectors above by one, at least three must have their length increased at the first transition. Two of these three must then be operated upon by $\mathcal{C}$ as a copy of a primitive machine.

Hence, Lemma 5.13(1) implies that there exists $1 \leq m \leq k$ such that $\mathcal{C}$ is a computation of $\mathbf{M}_2((4n-1)_m)$. The statement then follows as above.

$\square$

5.6 The machine $\mathbf{M}_3$

The recognizing S-machine $\mathbf{M}_3$ is the composition of $\mathbf{M}_2$ with one more simple machine.

The standard base of $\mathbf{M}_3$ is taken to be

$$P_0Q_0P_1Q_1R_1Q_2R_2Q_3P_4Q_4$$

As in the construction of $\mathbf{M}_2$, any part of this standard base given the same name as a part of the standard base of $\mathbf{M}_2$ contains more letters than its predecessor. The makeup of these parts is clear from the definition of the rules below.

The tape alphabet of the P_0Q_0- and Q_0P_1-sectors are copies of Y_1, while all other tape alphabets naturally correspond to those of $\mathbf{M}_2$. The P_0Q_0-sector is the input sector of the machine.

The idea of the function of $\mathbf{M}_3$ is the following. Given an input configuration, an accepting computation first moves all the letters to the right into the Q_0P_1-sector while all other sectors are locked. Then, the subword $Q_0P_1Q_1R_1Q_2R_2Q_3P_4Q_4$ is operated upon as the standard base of $\mathbf{M}_2$ while the P_0Q_0-sector is locked.

To be precise, we view $\mathbf{M}_3$ as the concatenation of $4n$ submachines: The $4n-1$ submachines $\mathbf{M}_3(2), \dots, \mathbf{M}_3(4n)$ corresponding to the submachines of $\mathbf{M}_2$ and the new machine $\mathbf{M}_3(1)$.

The submachines corresponding to those of $\mathbf{M}_2$ differ only in that the newly introduced part of the standard base consists of a single letter and the newly introduced sector remains locked.

For each part of the standard base, the subset corresponding to the submachine $\mathbf{M}_3(1)$ is a singleton. The positive rules of this submachine are in correspondence with $\mathcal{A}$. For $a \in \mathcal{A}$, the corresponding rule has the part $q_0(1) \to a_1^{-1}q_0(1)a_1'$, where $q_0(1) \in Q_0$ and a_1 (respectively a_1') is the copy of a in the tape alphabet of the P_0Q_0-sector (respectively the Q_0P_1-sector). All other sectors of the standard base are locked by this rule.

We also introduce more transition rules, defined in the same way as for previous machines to force the natural order of the concatenation of these submachines. The transition rules $\theta(i,i+1)^{\pm 1}$ for $2 \leq i \leq 4n-1$ correspond to the rules of the same name in $\mathbf{M}_2$, operating on the natural copy of the standard base of $\mathbf{M}_2$ in the same way and locking all other sectors. Meanwhile, the transition rules $\theta(12)^{\pm 1}$ connect $\mathbf{M}_3(1)$ and $\mathbf{M}_3(2)$, locking all sectors of the standard base of $\mathbf{M}_3$ except for the Q_0P_1-sector.

5.7 Standard computations of $\mathbf{M}_3$

The step history of a reduced computation of $\mathbf{M}_3$ is defined in a manner similar to how it was for reduced computations of $\mathbf{M}_2$. The only new letters corresponding to this machine are (1), (12), and (21), which correspond to maximal subcomputations of $\mathbf{M}_3(1)$ and transition rules in the obvious way.

Further, for $3 \leq i \leq 4n-3$ odd, we define the submachines $\mathbf{M}_3(i^-)$ and $\mathbf{M}_3(i^+)$ as the submachines of $\mathbf{M}_3(i)$ in the same way as the corresponding submachines of $\mathbf{M}_2(i)$. Similarly, for $1 \leq j \leq k$, we define the submachines $\mathbf{M}_3((4n-1)_j)$.

As a result, Lemmas 5.16 and 5.17 have obvious analogues in $\mathbf{M}_3$. The following statement is similar in nature to those, dealing with the newly added steps. Its proof is identical to that of Lemma 5.1.

Lemma 5.28. *Let $\mathcal{C}$ be a reduced computation of $\boldsymbol{M}_3$ with base B.*

(a) If B contains a subword B' of the form $(P_0Q_0)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be $(21)(1)(12)$.

(b) If B contains a subword B' of the form $(R_2Q_3)^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be $(12)(2)(21)$.

Much of the same terminology regarding reduced computations is carried over from $\mathbf{M}_2$.

For example, the history H of a reduced computation $\mathcal{C}$ of $\mathbf{M}_3$ is controlled if $\mathcal{C}$ is the natural copy of a reduced computation of $\mathbf{M}_2$ whose history is controlled.

However, a configuration of $\mathbf{M}_3$ is *tame* if, in addition to its P_1Q_1-, Q_1R_1-, Q_2R_2-, and P_4Q_4-sectors being empty, its P_0Q_0-sector is also empty.

Lemma 5.29. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation such that W_0 is an end configuration. If the history H of $\mathcal{C}$ contains a transition rule, then W_t is not an end configuration.*

Proof. Assuming toward contradiction, Lemma 5.28(a) implies that the step history of $\mathcal{C}$ has no occurrence of (1), (12), or (21).

But then $\mathcal{C}$ can be viewed as a reduced computation of $\mathbf{M}_2$, so that it contradicts Lemma 5.20.

□

Lemma 5.30. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_3$ in the standard base. Suppose W_0 is an input configuration and the step history of $\mathcal{C}$ does not contain the letter $(4n)$. Then $|W_0|_a \leq 9n|W_t|_a$.*

Proof. If $\mathcal{C}$ is a one-step computation with step (1), then $|W_0|_a \leq |W_t|_a$ by a projection argument.

Otherwise, let $W_0 \to \cdots \to W_r$ be the subcomputation with step history $(1)(12)$. Then as above $|W_0|_a \leq |W_r|_a$.

By Lemmas 5.16, 5.17, and 5.28, the subcomputation $W_r \to \cdots \to W_t$ can be identified with a reduced computation of $\mathbf{M}_2$. But then Lemma 5.21 implies $|W_r|_a \leq 9n|W_t|_a$.

□

An analogous proof immediately implies the following statement.

Lemma 5.31. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_3$ in the standard base. Suppose the first letter of the step history of $\mathcal{C}$ is $(4n, 4n-1)$. Then $|W_0|_a \leq 12n|W_t|_a$.*

Lemma 5.32. *(a) Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_3$. Suppose W_0 is a start configuration and W_t is an end configuration. Then there exists $u \in F(\mathcal{A})$ such that the projection of W_0 onto $F(\mathcal{A})$ is $u^n \in \mathcal{L}$.*

(b) *For all* $u \in F(\mathcal{A})$*, there exists a unique reduced computation* $\mathcal{D}_3(u) : W_0 \to \cdots \to W_t$ *of* $\boldsymbol{M}_3$ *in the standard base with step history* $(12)(2)(3)\dots(4n-1)(4n-1,4n)$ *and such that the projection of* W_0 *onto* $F(\mathcal{A})$ *is* u^n.

(c) *Let* t *be the length of* $\mathcal{D}_3(u)$ *and* ℓ *be the length of the subcomputation with step history* $(4n-2,4n-1)(4n-1)(4n-1,4n)$*. Then* $\ell = 2k\|u\|+2k+1$ *and* $t-\ell \leq (c_0+1)(\|u\|+1)$.

Proof. (a) As an application of a rule of step history (1) does not change the projection of a configuration onto $F(\mathcal{A})$, the statement follows from a projection argument and Lemma 5.23(a).

Statement (b) follows immediately from Lemmas 5.23(b),(c) and 4.3(a).

(c) Let $\mathcal{E}$ be the maximal subcomputation of $\mathcal{D}_3(u)$ with step history $(12)(2)(3)\dots(4n-2)$ and $\mathcal{E}'$ be the maximal subcomputation with step history $(12)(2)$. Then the length of $\mathcal{E}$ is $t-\ell$ and, for ℓ' the length of $\mathcal{E}'$, Lemma 5.23(c) implies $t-\ell-\ell' \leq c_0(\|u\|+1)$.

But Lemma 4.3 implies that $\ell' = \|u\|+1$, so that the statement follows.

□

Lemma 5.33. *The language of accepted inputs of* $\boldsymbol{M}_3$ *is* $\mathcal{L}$*. Moreover, for any* $u^n \in \mathcal{L}$*, there exists a unique accepting computation* $\mathcal{C}_3(u)$*.*

Proof. Lemma 5.32(a) implies that any accepted input must be an element of $\mathcal{L}$.

Conversely, for $u^n \in \mathcal{L}$, let H_1 be the natural copy of u^n read right to left in the rules of step history (1) and H_2 be the history of $\mathcal{C}_2(u)$. Then, for $H \equiv H_1\theta(12)H_2$, the input configuration with input u^n is H-admissible with $W \cdot H$ the accept configuration.

The uniqueness of this computation follows immediately from Lemmas 5.32(b) and 4.3(a).

□

Lemma 5.34. *Let* $\mathcal{C} : W_0 \to \cdots \to W_t$ *be a reduced computation of* $\boldsymbol{M}_3$*. Suppose* W_0 *is an input configuration and* W_t *is either an input or the accept configuration. Then the sum of the lengths of the subcomputations of* $\mathcal{C}$ *whose step histories are of the form* $(4n-2,4n-1)(4n-1)(4n-1,4n)$ *or* $(4n,4n-1)(4n-1)(4n-1,4n-2)$ *is at least* $\left(1-\frac{1}{c_0}\right)t$.

Proof. By Lemmas 4.3 and 5.28, the step history of $\mathcal{C}$ has prefix $(1)(2)\dots(4n-1)(4n-1,4n)$. Let $W_0 \to \dots \to W_s$ be the subcomputation with this step history. Lemma 5.32(a) then implies that there exists $u \in F(\mathcal{A})$ such that the input of W_0 is u^n.

Let $W_0 \to \dots \to W_r$ be the maximal subcomputation with step history (1). Then Lemma 4.3 implies $r = \|u^n\|$.

By Lemma 5.32(b), the subcomputation $W_r \to \dots \to W_s$ must be $\mathcal{D}_3(u)$. Letting ℓ_1 be the length of the subcomputation with step history $(4n-2,4n-1)(4n-1)(4n-1,4n)$, Lemma 5.32(c) then implies $s-\ell_1 \leq \|u^n\| + (c_0+1)(\|u\|+1)$ and $\ell_1 = 2k\|u\|+2k+1$.

If W_t is the accept configuration, then Lemma 4.3 implies $t-s=\|u\|$, so that the parameter choice $c_0 >> n$ yields

$$t-\ell_1 \leq \|u^n\| + (c_0+2)(\|u\|+1) \leq 2c_0(\|u\|+1)$$

while $\ell_1 \geq 2k(\|u\|+1)$. So, $\ell_1 \geq \frac{k}{c_0}(t-\ell_1)$, implying $\ell_1 \geq \frac{k}{k+c_0} = 1-\frac{c_0}{k+c_0}$. The parameter choice $k \geq c_0^2$ then implies the statement.

Now suppose W_t is an input configuration. Then we may apply the same arguments to the inverse subcomputation, so that:

- the input of W_t is v^n for some $v \in F(\mathcal{A})$,
- there exists a maximal subcomputation $\mathcal{E} : W_x \to \dots \to W_t$ of $\mathcal{C}$ whose step history is $(4n,4n-1)(4n-1)\dots(2)(1)$
- for ℓ_2 the length of the subcomputation with step history $(4n,4n-1)(4n-1)(4n-1,4n-2)$, $t-x-\ell_2 \leq \|v^n\| + (c_0+n)(\|v\|+1)$ and $\ell_2 = 2k\|v\|+2k+1$.

By Lemma 5.29, the subcomputation $W_s \to \dots \to W_x$ has step history $(4n)$. So, Lemma 4.3 implies $x-s \leq \|u\|+\|v\|$. Combining these inequalities and taking $c_0 >> n$ then yields

$$t-(\ell_1+\ell_2) \leq \|u^n\|+\|v^n\|+(c_0+1)\left(\|u\|+\|v\|+2\right)+\|u\|+\|v\| \leq 2c_0(\|u\|+\|v\|+2)$$

while $\ell_1+\ell_2 \geq 2k(\|u\|+\|v\|+2)$. So, $\ell \geq \frac{k}{c_0}(t-\ell)$ for $\ell=\ell_1+\ell_2$, so that the statement follows as above.

□

Lemma 5.35. *For any reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ of $\boldsymbol{M}_3$ in the standard base, $t \leq 2c_1 \max(\|W_0\|, \|W_t\|)$.*

Proof. If $\mathcal{C}$ is a one-step computation with step (1), then the statement follows from Lemma 4.3.

So, by Lemma 5.28, we may assume there exists a maximal subcomputation $\mathcal{C}_2 : W_r \to \cdots \to W_s$ of $\mathcal{C}$ whose step history has no occurrence of the letters (1), (12), or (21).

Then, $\mathcal{C}_2$ can be viewed as a computation of $\mathbf{M}_2$, so that we have $s-r \leq c_1 \max(\|W_r\|, \|W_s\|)$ by Lemma 5.26.

If the subcomputation $W_s \to \cdots \to W_t$ is nonempty, then it must be a one-step computation with step (1). But then this implies $t-s \leq |W_t|_a + 1$ and $|W_s|_a \leq |W_t|_a$.

The symmetric argument implies $r \leq |W_0|_a + 1$ and $|W_r|_a \leq |W_0|_a$.

Hence, $t \leq (c_1+2)\max(\|W_0\|, \|W_t\|)$, so that the statement follows by $c_1 \geq 2$.

□

A two-letter subword of the standard base of $\mathbf{M}_3$ is defined to be left-active (or right-active) in $\mathbf{M}_3(i)$ in the same way as subwords of the standard base of $\mathbf{M}_2$.

For example, the subwords P_0Q_0 and Q_0P_1 are right-active and left-active, respectively, in $\mathbf{M}_3(1)$.

Lemma 5.36. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of $\boldsymbol{M}_3(i)$ in the standard base for some i. Assume that for some index j, $|W_j|_a > 4|W_0|_a$. Then there are subwords $U_\ell V_\ell$ and $U_r V_r$ of the standard base such that $U_\ell V_\ell$ is left-active in $\boldsymbol{M}_3(i)$, $U_r V_r$ is right-active in $\boldsymbol{M}_3(i)$, and for $W_0' \to \cdots \to W_t'$ the restriction of $\mathcal{C}$ to either sector, $|W_j'|_a < |W_{j+1}'|_a < \cdots < |W_t'|_a$.*

Proof. By Lemma 5.27, it suffices to assume that $i=1$. But then the statement follows immediately from Lemma 4.3.

□

5.8 The machine $\mathbf{M}_4$

The machine $\mathbf{M}_4$ is the 'circular' analogue of a simple tweak to the machine $\mathbf{M}_3$.

The standard base of $\mathbf{M}_4$ adds just one part to that of $\mathbf{M}_3$. In particular, setting B_3 as the standard base of $\mathbf{M}_3$, the standard base of $\mathbf{M}_4$ is $\{t\}B_3$, where $\{t\}$ consists of a single letter (which, clearly, acts as both the start and end letter of its part). The tape alphabet of the new sector in the standard base, i.e the $\{t\}P_0$-sector, is empty. All other tape alphabets are carried over from $\mathbf{M}_3$.

A major difference between $\mathbf{M}_4$ and the machines constructed in previous sections is that a tape alphabet is assigned to the space after the final letter Q_4 of B_3, which corresponds to the $Q_4\{t\}$-sector. As such, it is possible for an admissible word of $\mathbf{M}_4$ to have base

$$Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}P_4Q_4\{t\}P_0Q_0$$

i.e it essentially 'wraps around' the standard base. An S-machine with this property is called a *cyclic machine*, as one can think of the standard base as being written on a circle.

In this machine, the tape alphabet assigned to the $Q_4\{t\}$-sector is empty. The positive rules of $\mathbf{M}_4$ correspond to those of $\mathbf{M}_3$, operating on the copy of the hardware of $\mathbf{M}_3$ in the same way and locking the new sectors.

As such, we define the submachines $\mathbf{M}_4(i)$ as in $\mathbf{M}_3$ and define the step history of a reduced computation in the natural way.

The input sector of $\mathbf{M}_4$ is the same as that of $\mathbf{M}_3$, i.e the P_0Q_0-sector.

There are obvious analogues of the statements from previous sections. Instead of reformulating them here, we reference the previous statements even when in reference to this machine.

The base of an admissible word of $\mathbf{M}_4$ (or any cyclic S-machine) is called *revolving* if:

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

(b) none of its proper subwords satisfy (a).

An unreduced revolving base is called *faulty*.

Suppose W is an admissible word of a cyclic S-machine **S** whose base $B \equiv xvx$ is revolving. If v has the form v_1yv_2, for some letter y, then there exists a naturally formed admissible word W' with revolving base $B' \equiv yv_2xv_1y$ and satisfying $|W'|_a = |W|_a$. In this case, B' (respectively W') is called a *cyclic permutation* of B (respectively of W).

Note that for any reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ of **S** with base B and history H, there exists a reduced computation $\mathcal{C}' : W_0' \to \cdots \to W_t'$ with base B' and history H and so that $|W_j|_a = |W_j'|_a$ for all $0 \leq j \leq t$.

Lemma 5.37. *For every reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ of $\boldsymbol{M}_4$ with faulty base B, $|W_j|_a \leq c_0 \max(|W_0|_a, |W_t|_a)$ for all $0 \leq j \leq t$.*

Proof. Note that we may assume that $t > 1$ and $|W_j|_a > \max(|W_0|_a, |W_t|_a)$ for all $0 < j < t$, as otherwise the statement follows from an obvious inductive argument. In particular, since a transition rule (resp χ-rule, connecting rule) does not alter the a-length of an admissible word, we may assume that neither the first nor the last letter of the history H of $\mathcal{C}$ is a transition rule (resp χ-rule, connecting rule.).

1. Suppose $\mathcal{C}$ is a computation of $\mathbf{M}_4(i)$ for some $i \in \{1, 2, 4, \ldots, 4n\}$. Then the restriction $\mathcal{C}' : W_0' \to \cdots \to W_t'$ of $\mathcal{C}$ to any two-letter subword of B has fixed a-length, satisfies the hypotheses of Lemma 4.3, or satisfies the hypotheses of Lemma 4.4. In each case, $|W_j'|_a \leq \max(|W_0'|_a, |W_t'|_a)$. So, $|W_j|_a = \sum |W_j'|_a \leq \sum \max(|W_0'|_a, |W_t'|_a) \leq 2\max(|W_0|_a, |W_t|_a)$ for all $0 \leq j \leq t$.

2. Suppose $\mathcal{C}$ is a computation of $\mathbf{M}_4(i^-)$ for some $i \in \{3, 5, \ldots, 4n-3\}$.

If H contains no connecting rule, then an identical argument to the one used in Step 1 applies. So, assume that H contains such a connecting rule, locking the Q_0P_1-sector of the standard base.

If B has no occurrence of the letters $P_1^{\pm 1}$, then no rule of H changes the a-length of an admissible word with base B. So, assuming that B contains such a letter, Lemma 4.2 and the definition of faulty imply that B has a subword UV of the form $(Q_0P_1)^{\pm 1}$.

If more than one connecting rule occurs in H, then H must have a subword $\zeta H' \zeta^{-1}$, where ζ is a connecting rule and H' contains no connecting rule. Letting $\mathcal{C}'$ be the subcomputation with history H', the restriction of $\mathcal{C}'$ to the UV-sector satisfies the hypotheses of Lemma 4.3. But then H' must be empty, yielding a contradiction.

So, H contains exactly one connecting rule. Let $W_{r-1} \to W_r$ be the subcomputation corresponding to this connecting rule.

By the definition of faulty, any subword of B of the form $(P_1Q_1)^{\pm 1}$ is contained in a subword of a cyclic permutation of B of the form $(Q_0P_1Q_1)^{\pm 1}$. Letting $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be the restriction of (a cyclic permutation of) $\mathcal{C}$ to a subword of the form $(Q_0P_1Q_1)^{\pm 1}$, Lemma 5.13 implies $|W_r'|_a \leq \cdots \leq |W_t'|_a$.

Further, any subword of B of the form $P_1P_1^{-1}$ is contained in a subword of a cyclic permutation of B of the form $Q_0P_1P_1^{-1}Q_0^{-1}$. Letting $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be the restriction of (a cyclic permutation of) $\mathcal{C}$ to a subword of the form $Q_0P_1P_1^{-1}Q_0^{-1}$, Lemma 5.14 implies the inequalities $|W_r'|_a \leq \cdots \leq |W_t'|_a$.

The tape word of any other sector is fixed throughout $\mathcal{C}$, so that $|W_r|_a \leq |W_t|_a$. By assumption, we must then have $r = t$. But then the final letter of H is a connecting rule, contradicting our assumption.

Analogous arguments yield the same inequalities if $\mathcal{C}$ is a computation of $\mathbf{M}_4(i^+)$ for some i.

3. Suppose $\mathcal{C}$ is a computation of $\mathbf{M}_4(i)$ for some $i \in \{3, 5, \dots, 4n-3\}$.

Suppose H has a suffix χH^- where χ is a χ-rule and H^- is the history of a maximal subcomputation of $\mathbf{M}_4(i^-)$. Let $\mathcal{C}^- : W_r \to \cdots \to W_t$ be the subcomputation with history H^-.

Note that if i is of the form $4\ell - 1$ (resp $4\ell + 1$), then the only sectors of the standard base that χ does not lock are the Q_0P_1-sector and the R_2Q_3-sector (resp R_1Q_2-sector).

So, by the definition of faulty, any subword of B of the form $(Q_0P_1)^{\pm 1}$ is contained in a subword of a cyclic permutation of B of the form $(Q_0P_1Q_1)^{\pm 1}$. The restriction of (a cyclic permutation of) $\mathcal{C}^-$ to this subword then satisfies the hypotheses of Lemma 5.13.

Further, any subword of B of the form $P_1^{-1}P_1$ is contained in a subword of a cyclic permutation of B of the form $Q_1^{-1}P_1^{-1}P_1Q_1$. The restriction of (a cyclic permutation of) $\mathcal{C}^-$ to this subword then satisfies the hypotheses of Lemma 5.14.

All other sectors have fixed a-length throughout $\mathcal{C}^-$, so that $|W_r|_a \leq |W_t|_a$. But this contradicts our assumption.

So, H has no suffix of the form χH^-. By Step 2, H must contain a χ-rule, so that it must have a suffix of the form χH^+ where χ is a χ-rule and H^+ is the history of a maximal subcomputation of $\mathbf{M}_4(i^+)$.

But then an analogous argument yields a similar contradiction.

4. Suppose $\mathcal{C}$ is a computation of $\mathbf{M}_4((4n-1)_j)$ for some $1 \leq j \leq k$.

As in Step 2, H must contain a connecting rule, as otherwise we may apply the argument used in Step 1.

If B has no occurrence of the letters $R_2^{\pm 1}$ or $P_4^{\pm 1}$, then no rule of H changes the a-length of an admissible word with base B. So, we assume that B has a subword of the form $(R_2Q_3)^{\pm 1}$ or $(Q_3P_4)^{\pm 1}$.

As in Step 2, this implies that H contains exactly one connecting rule. Let $W_{r-1} \to W_r$ be the subcomputation corresponding to this connecting rule.

Note the following consequences of the definition of faulty and Lemma 4.2:

- any subword of B of the form $(Q_2R_2)^{\pm 1}$ is contained in a cyclic permutation of B of the form $(Q_2R_2Q_3)^{\pm 1}$
- any subword of B of the form $(P_4Q_4)^{\pm 1}$ is contained in a cyclic permutation of B of the form $(Q_3P_4Q_4)^{\pm 1}$
- any subword of B of the form $R_2^{-1}R_2$ is contained in a cyclic permutation of B of the form $Q_3^{-1}R_2^{-1}R_2Q_3$
- any subword of B of the form $P_4P_4^{-1}$ is contained in a cyclic permutation of B of the form $Q_3P_4P_4^{-1}Q_4^{-1}$

Then, the restriction of the subcomputation $W_r \to \cdots \to W_t$ to any of the subwords above satisfies the hypotheses of Lemma 5.13(5) or Lemma 5.14. As any other sector has fixed tape word, this implies $|W_r|_a \leq \cdots \leq |W_t|_a$. But this leads to a contradiction in the same way as in Step 2.

5. Suppose $\mathcal{C}$ is a computation of $\mathbf{M}_4(4n-1)$.

By Step 4, H must then have suffix χH_j where χ is a χ-rule and H_j is the history of a maximal subcomputation of $\mathbf{M}_4((4n-1)_j)$. Let $\mathcal{C}_j : W_r \to \cdots \to W_t$ be the subcomputation with history H_j.

Note that χ locks every sector of the standard base except for the R_2Q_3- and Q_3P_4-sectors, while these sectors are locked by any connecting rule. So, H_j must contain no connecting rule.

Then, any unreduced two-letter subword of B must be of the form $R_2R_2^{-1}$, $Q_3^{-1}Q_3$, $Q_3Q_3^{-1}$, or $P_4^{-1}P_4$. We then have the following consequences of the definition of faulty and Lemma 4.2:

- any subword of B of the form $(R_2Q_3)^{\pm 1}$ is contained in a cyclic permutation of B of the form $(Q_2R_2Q_3)^{\pm 1}$
- any subword of B of the form $(Q_3P_4)^{\pm 1}$ is contained in a cyclic permutation of B of the form $(Q_3P_4Q_4)^{\pm 1}$
- any subword of B of the form $R_2R_2^{-1}$ is contained in a cyclic permutation of B of the form $Q_2R_2R_2^{-1}Q_2^{-1}$
- any subword of B of the form $P_4^{-1}P_4$ is contained in a cyclic permutation of B of the form $Q_4^{-1}P_4^{-1}P_4Q_4$

As in Step 4, Lemmas 5.13(5) and 5.14 imply $|W_r|_a \leq |W_t|_a$, yielding a contradiction.

Hence, it suffices to assume that H contains some transition rule.

6. Suppose H has a suffix of the form $\theta(i-1,i)H_i$ where H_i is the history of a maximal subcomputation with step history (i) for $3 \leq i \leq 4n-3$ of the form $4\ell-1$.

Let $\mathcal{C}_i : W_r \to \cdots \to W_t$ be the subcomputation with history H_i.

Then $\mathcal{C}_i$ is a reduced computation of $\mathbf{M}_4(i)$, so that there exists a maximal subcomputation $W_r \to \cdots \to W_s$ of $\mathcal{C}_i$ which operates as $\mathbf{M}_4(i^-)$.

As $\theta(i-1,i)$ locks every sector of the standard base except for the Q_0P_1-sector and the R_2Q_3-sector, any unreduced two-letter subword of B must be of the form $Q_0Q_0^{-1}$, $P_1^{-1}P_1$, $R_2R_2^{-1}$, or $Q_3^{-1}Q_3$.

By the definition of faulty, any subword of B of the form $(Q_0P_1)^{\pm 1}$ is contained in a subword of a cyclic permutation of B of the form $(Q_0P_1Q_1)^{\pm 1}$. Similarly, any subword of B of the form $P_1P_1^{-1}$ is contained in a subword of a cyclic permutation of B of the form $Q_1^{-1}P_1^{-1}P_1Q_1$. As in previous steps, Lemmas 5.13 and 5.14 then imply $|W_r|_a \leq \dots \leq |W_s|_a$.

As a result, we may assume that $t > s$, so that H_i contains the letter χ_i. As H_i must also contain the connecting rule of $\mathbf{M}_4(i^-)$, every unreduced two-letter subword of B must be of the form $R_2R_2^{-1}$ or $Q_3^{-1}Q_3$. In particular, B must be a cyclic permutation of

$$Q_2R_2R_2^{-1}Q_2^{-1}R_1^{-1}Q_1^{-1}P_1^{-1}Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}Q_3P_4Q_4\{t\}P_0Q_0P_1Q_1R_1Q_2$$

Let $\mathcal{C}_i^+ : W_{s+1} \to \dots \to W_x$ be the maximal subcomputation of $\mathcal{C}_i$ which operates as $\mathbf{M}_4(i^+)$. As the connecting rule of $\mathbf{M}_4(i^+)$ locks the R_2Q_3-sector, the restriction of $\mathcal{C}_i^+$ to the Q_2R_2-sector satisfies the hypotheses of Lemma 4.3. So, since W_{s+1} is χ_i^{-1}-admissible, W_x cannot be. In particular, $x = t$.

Then, the restriction of $\mathcal{C}_i^+$ to the subword $Q_2R_2R_2^{-1}Q_2^{-1}$ satisfies the hypotheses of Lemma 5.14. Since the tape word of any other sector is fixed throughout $\mathcal{C}_i^+$, this implies $|W_s|_a \leq |W_t|_a$, yielding a contradiction.

7. Suppose H has a suffix of the form $\theta(i-1,i)H_i$ where H_i is the history of a maximal subcomputation with step history (i) for $3 \leq i \leq 4n-3$ of the form form $4\ell+1$.

Let $\mathcal{C}_i : W_r \to \dots \to W_t$ be the subcomputation with history H_i.

As in Step 6, $\mathcal{C}_i$ must have a maximal subcomputation $W_r \to \dots \to W_s$ which operates as $\mathbf{M}_4(i^-)$ such that $s < t$. So, since H_i must contain the connecting rule of $\mathbf{M}_4(i^-)$, every unreduced two-letter subword of B must be of the form $R_1R_1^{-1}$ or $Q_2^{-1}Q_2$. As a result, B must be a cyclic permutation of

$$Q_1R_1R_1^{-1}Q_1^{-1}P_1^{-1}Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}Q_2R_2Q_3P_4Q_4\{t\}P_0Q_0P_1Q_1$$

Let $\mathcal{C}_i^+ : W_{s+1} \to \cdots \to W_x$ be the maximal subcomputation of $\mathcal{C}_i$ which operates as $\mathbf{M}_4(i^+)$. As the connecting rule of $\mathbf{M}_4(i^+)$ locks the R_1Q_2-sector, we again have $x = t$. Applying Lemma 5.14 to the restriction of $\mathcal{C}_i^+$ to the subword $Q_1R_1R_1^{-1}Q_1^{-1}$ then implies $|W_s|_a \leq |W_t|_a$.

8. Suppose H has a suffix of the form $\theta(i+1, i)H_i$ where H_i is the history of a maximal subcomputation with step history (i) for $3 \leq i \leq 4n-3$ odd.

Let $\mathcal{C}_i : W_r \to \cdots \to W_t$ be the subcomputation with history H_i.

Letting $W_r \to \cdots \to W_s$ be the maximal subcomputation with step history (i^+), as in Steps 4 and 5 we have $|W_r|_a \leq \cdots \leq |W_s|_a$. As a result, it suffices to assume that $s < t$.

So, any unreduced two-letter subword of B must be of the form $Q_0Q_0^{-1}$ or $P_1^{-1}P_1$. In particular, B must be a cyclic permutation of

$$Q_1^{-1}P_1^{-1}P_1Q_1R_1Q_2R_2Q_3P_4Q_4\{t\}P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}R_1^{-1}Q_1^{-1}$$

Let $\mathcal{C}_i^- : W_{s+1} \to \cdots \to W_x$ be the maximal subcomputation of $\mathcal{C}_i$ which operates as $\mathbf{M}_4(i^-)$. As in Steps 6 and 7, the presence of the subword P_0Q_0 implies that we must have $x = t$. But then an application of Lemma 5.14 to the restriction of $\mathcal{C}_i^-$ to the subword $Q_1^{-1}P_1^{-1}P_1Q_1$ implies $|W_s|_a \leq |W_t|_a$.

Hence, we may assume that the last letter of the step history of $\mathcal{C}$ is not of the form (i) for some $i \in \{3, 5, \dots, 4n-3\}$. Moreover, the symmetric argument allows the same assumption to be made about the first letter of the step history.

9. Suppose H has a subword θH_{4n-1} where θ is a transition rule and H_{4n-1} is the history of a maximal subcomputation with step history $(4n-1)$.

Note that both $\theta(4n-2, 4n-1)$ and $\theta(4n, 4n-1)$ lock every sector of the standard base except for the R_2Q_3- and Q_3P_4-sectors. Meanwhile, for any j, the connecting rule of $\mathbf{M}_4((4n-1)_j)$ locks these two sectors. So, by Lemma 4.2, no connecting rule can appear in H_{4n-1}.

As a result, H_{4n-1} is the history of a computation of $\mathbf{M}_4((4n-1)_1)$ if $\theta = \theta(4n-2, 4n-1)$ or $\mathbf{M}_4((4n-1)_k)$ if $\theta = \theta(4n, 4n-1)$.

Note the following consequences of the definition of faulty and Lemma 4.2:

- any subword of B of the form $(R_2Q_3)^{\pm 1}$ is contained in a cyclic permutation of B of the form $(Q_2R_2Q_3)^{\pm 1}$
- any subword of B of the form $(Q_3P_4)^{\pm 1}$ is contained in a cyclic permutation of B of the form $(Q_3P_4Q_4)^{\pm 1}$
- any subword of B of the form $R_2R_2^{-1}$ is contained in a cyclic permutation of B of the form $Q_2R_2R_2^{-1}Q_2^{-1}$
- any subword of B of the form $P_4^{-1}P_4$ is contained in a cyclic permutation of B of the form $Q_4^{-1}P_4^{-1}P_4Q_4$

Let $\mathcal{C}' : W_r \to \cdots \to W_s$ be the subcomputation with history H_{4n-1}. Then, as in previous steps, Lemmas 5.13 and 5.14 imply $|W_r|_a \leq \cdots \leq |W_s|_a$. So, we must have $s < t$.

As no connecting rule can occur in H_{4n-1}, the subsequent rule of H must be θ^{-1}, i.e H has a subword $\theta H_{4n-1}\theta^{-1}$.

If B contains a subword of the form $(Q_2R_2)^{\pm 1}$ or $(P_4Q_4)^{\pm 1}$, then the restriction of $\mathcal{C}'$ to this sector satisfies the hypotheses of Lemma 4.3. But then H_{4n-1} must be empty, yielding a contradiction.

So, B cannot contain such a subword. By the definition of faulty, it follows that B cannot contain the letters $R_2^{\pm 1}$ or $P_4^{\pm 1}$. In particular, B must be a cyclic permutation of $Q_3Q_3^{-1}Q_3$.

However, no rule alters the a-length of an admissible with such a base, so that $|W_j|_a = |W_0|_a$ for all j.

Hence, H has no such subword. What's more, by the symmetric argument, H has no subword of the form $H_{4n-1}\theta$. So, we may assume that the step history of $\mathcal{C}$ has no occurrence of $(4n-1)$, $(4n)$, $(4n-1, 4n)$, $(4n, 4n-1)$, $(4n-2, 4n-1)$, or $(4n-1, 4n-2)$.

10. Suppose the step history of $\mathcal{C}$ contains the letter (12).

Then H has a subword $H_1\theta(12)H_2$ where H_1 and H_2 are histories of maximal subcomputations of step history (1) and (2), respectively.

As $\theta(12)$ locks all sectors of the standard base except for the Q_0P_1-sector, B must be a cyclic permutation of

$$P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}R_1^{-1}Q_1^{-1}P_1^{-1}P_1Q_1R_1Q_2R_2Q_3P_4Q_4\{t\}P_0$$

Let $\mathcal{C}_1 : W_r \to \cdots \to W_s$ be the subcomputation with history H_1. Then the restriction of $\mathcal{C}_1$ to the P_0Q_0-sector satisfies the hypotheses of Lemma 4.3. So, since W_s is $\theta(12)$-admissible, W_r cannot be. In particular, $r = 0$.

Let $\mathcal{C}_1' : W_0' \to \cdots \to W_s'$ be the restriction of $\mathcal{C}_1$ to the subword $P_0Q_0Q_0^{-1}P_0^{-1}$. Then we may apply Lemma 5.14 to $\mathcal{C}_1'$, so that $|W_s'|_a \leq |W_0'|_a$.

But every other sector must have fixed a-length throughout $\mathcal{C}_1$, so that $|W_s|_a \leq |W_0|_a$.

Hence, we may assume that the step history of $\mathcal{C}$ has no occurrence of the letter (12) and, by the symmetric argument, no occurrence of the letter (21). In particular, $\mathcal{C}$ has no subcomputation with step history (1) and every rule of H locks the P_0Q_0-sector.

11. Suppose H contains a connecting rule ζ of $\mathbf{M}_4(i^-)$ for some $3 \leq i \leq 4n-3$ of the form $4\ell - 1$.

Note that every sector of the standard base is locked by at least one of ζ, χ_i, or the connecting rule of $\mathbf{M}_4(i^+)$. So, by Steps 6-8, H must contain a subword $\theta(i-1,i)H_i\theta(i,i-1)$ where H_i contains ζ and is the history of a subcomputation $\mathcal{C}_i$ of $\mathbf{M}_4(i)$.

The only sector of the standard base not locked by at least one of $\theta(i-1,i)$ or ζ is the R_2Q_3-sector. So, B must be a cyclic permutation of

$$Q_2R_2R_2^{-1}Q_2^{-1}R_1^{-1}Q_1^{-1}P_1^{-1}Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}Q_3P_4Q_4\{t\}P_0Q_0P_1Q_1R_1Q_2$$

Suppose H_i contains the letter χ_i. As the connecting rule of $\mathbf{M}_4(i^+)$ locks the R_2Q_3-sector, it cannot occur in H_i. So, H_i must have a subword $\chi_i H_i' \chi_i^{-1}$, where H_i' is the history of a subcomputation $\mathcal{C}_i'$ of $\mathbf{M}_4(i^+)$ not containing a connecting rule. But then applying Lemma 4.3 to the restriction of $\mathcal{C}_i'$ to the Q_2R_2-sector implies that H_i' is empty.

So, $\mathcal{C}_i$ must be a computation of $\mathbf{M}_4(i^-)$. But then applying Lemma 5.13(4) to the restriction of $\mathcal{C}_i$ to the subword $Q_0P_1Q_1$ implies that H_i is empty, yielding a contradiction.

12. Suppose H contains a connecting rule ζ of $\mathbf{M}_4(i^-)$ for some $3 \leq i \leq 4n-3$ of the form $4\ell+1$.

As in Step 11, H must contain a subword $\theta(i-1,i)H_i\theta(i,i-1)$ where H_i contains ζ and is the history of a computation $\mathcal{C}_i$ of $\mathbf{M}_4(i)$.

The only sector of the standard base not locked by at least one of $\theta(i-1,i)$ or ζ is the R_1Q_2-sector. So, B must be a cyclic permutation of

$$Q_1R_1R_1^{-1}Q_1^{-1}P_1^{-1}Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}Q_2R_2Q_3P_4Q_4\{t\}P_0Q_0P_1Q_1$$

Suppose H_i contains the letter χ_i. As in Step 11, H_i must then have a subword $\chi_i H_i' \chi_i^{-1}$, where H_i' is the history of a subcomputation $\mathcal{C}_i'$ of $\mathbf{M}_4(i^+)$ not containing a connecting rule. But then applying Lemma 4.3 to the restriction of $\mathcal{C}_i'$ to the Q_1R_1-sector implies that H_i' is empty.

So, $\mathcal{C}_i$ must be a computation of $\mathbf{M}_4(i^-)$. But then applying Lemma 5.13(4) to the restriction of $\mathcal{C}_i$ to the subword $Q_0P_1Q_1$ implies that H_i is empty, yielding a contradiction.

13. Suppose H contains a connecting rule ζ of $\mathbf{M}_4(i^+)$ for some $3 \leq i \leq 4n-3$ odd.

Similar to the arguments in Steps 11 and 12, H must contain a subword $\theta(i+1,i)H_i\theta(i,i+1)$, where H_i contains ζ and is the history of a computation $\mathcal{C}_i$ of $\mathbf{M}_4(i)$.

The only sector of the standard base not locked by at least one of $\theta(i+1,i)$ or ζ is the Q_0P_1-sector. So, B must be a cyclic permutation of

$$\{t\}P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}R_1^{-1}Q_1^{-1}P_1^{-1}P_1Q_1R_1Q_2R_2Q_3P_4Q_4\{t\}$$

Suppose H_i contains the letter χ_i^{-1}. As the connecting rule of $\mathbf{M}_4(i^-)$ locks the Q_0P_1-sector, it cannot occur in H_i. So, H_i must have a subword $\chi_i^{-1}H_i'\chi_i$ where H_i' is the history of a subcomputation $\mathcal{C}_i'$ of $\mathbf{M}_4(i^-)$ not containing a connecting rule. Applying Lemma 4.3 to the restriction of $\mathcal{C}_i'$ to the P_1Q_1-sector then implies that H_i' is empty, yielding a contradiction.

So, $\mathcal{C}_i$ must be a computation of $\mathbf{M}_4(i^+)$. But then we may apply Lemma 5.13(4) to the restriction of $\mathcal{C}_i$ to the subword $Q_2R_2Q_3$ (respectively $Q_1R_1Q_2$) if i is of the from $4\ell-1$ (respectively $4\ell+1$), so that that H_i must be empty.

Hence, by Steps 11-13, we may assume that H contains no connecting rule.

14. Suppose the step history of $\mathcal{C}$ contains the letter $(4n-3,4n-2)$.

Then H has a subword $H_{4n-3}\theta(4n-3,4n-2)H_{4n-2}$ where each H_i is the history of a maximal subcomputation of step history (i).

Note that the rule $\theta(4n-3,4n-2)$ locks every sector of the standard base except for the Q_0P_1- and R_1Q_2-sectors. So, every unreduced two-letter subword of B must be of the form $Q_0Q_0^{-1}$, $P_1^{-1}P_1$, $R_1R_1^{-1}$, or $Q_2^{-1}Q_2$.

By Steps 6-8, H_{4n-3} cannot be a prefix of H. So, since H_{4n-3} contains no connecting rule, $\theta(4n-2,4n-3)H_{4n-3}\theta(4n-3,4n-2)$ must be a subword of H.

Further, the subcomputation $\mathcal{C}_{4n-3}$ of history H_{4n-3} must be a computation of $\mathbf{M}_4((4n-3)^+)$. If B contains a subword of the form $(Q_1R_1)^{\pm1}$, then the restriction of $\mathcal{C}_{4n-3}$ to this sector satisfies the hypotheses of Lemma 4.3. But then H_{4n-3} must be empty, so that H is not reduced.

So, B cannot contain the letters $R_1^{\pm1}$. This implies that B must be a cyclic permutation of

$$P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}Q_2R_2Q_3P_4Q_4\{t\}P_0$$

Let $\mathcal{C}_{4n-2}:W_r\to\cdots\to W_s$ be the subcomputation with history H_{4n-2}. The restriction of $\mathcal{C}_{4n-2}$ to the R_2Q_3-sector satisfies the hypotheses of Lemma 4.3. So, since W_r is $\theta(4n-2,4n-3)$-admissible, W_s cannot be. By Step 9, this implies that $s=t$ and H_{4n-2} is a suffix of H.

Let $\mathcal{C}_{4n-2}':W_r'\to\cdots\to W_t'$ be the restriction of $\mathcal{C}_{4n-2}$ to the subword $Q_3^{-1}R_2^{-1}Q_2^{-1}Q_2R_2Q_3$. As every rule with step history $(4n-2)$ locks the Q_2R_2-sector, we may view the subwords with base $(Q_2R_2)^{\pm 1}$ as a single state letter. With this view, $\mathcal{C}'_{4n-2}$ satisfies the hypotheses of Lemma 5.14, so that $|W'_r|_a \leq \cdots \leq |W'_t|_a$.

For $\mathcal{C}''_{4n-2} : W''_r \to \cdots \to W''_t$ the restriction of $\mathcal{C}_{4n-2}$ to a subword $(Q_3P_4)^{\pm 1}$, Lemma 4.3 implies that $|W''_r|_a \leq \cdots \leq |W''_t|_a$.

Any other sector must have fixed tape word throughout $\mathcal{C}_{4n-2}$. But then $|W_r|_a \leq |W_t|_a$, contradicting our assumption.

Hence, we may assume that the step history of $\mathcal{C}$ has no occurrence of $(4n-2)$, so that the Q_3P_4-sector is locked by every rule of H.

15. Suppose the step history of $\mathcal{C}$ contains the letter $(2i, 2i+1)$ for some $1 \leq i \leq 2n-2$ odd.

Note that $\theta(2i, 2i+1)$ locks all sectors of the standard base except for the Q_0P_1-sector and the R_2Q_3-sector. So, any unreduced two-letter subword of B must be of the form $Q_0Q_0^{-1}$, $P_1^{-1}P_1$, $R_2R_2^{-1}$, or $Q_3^{-1}Q_3$.

By Steps 6-8, the step history of $\mathcal{C}$ must then have a subword $(2i, 2i+1)(2i+1)(2i+1, 2i)$. Let $\mathcal{C}_{2i+1}$ be the maximal subcomputation with step history $(2i+1)$ in this subword. By Steps 11-13, the history H_{2i+1} of $\mathcal{C}_{2i+1}$ contains no connecting rule. So, $\mathcal{C}_{2i+1}$ is a computation of $\mathbf{M}_4((2i+1)^-)$.

If B contains a subword of the form $(P_1Q_1)^{\pm 1}$, then the restriction of $\mathcal{C}_{2i+1}$ to this subword satisfies the hypotheses of Lemma 4.3, so that H_{2i+1} must be empty. So, B cannot contain such a subword and, by the definition of faulty, cannot contain the letters $P_1^{\pm 1}$. In particular, B must be a cyclic permutation of

$$P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}Q_3P_4Q_4\{t\}P_0$$

But then Steps 9, 10, and 14 imply that the application of any rule of H does not alter the tape word of an admissible word with such a base, so that $|W_j|_a = |W_0|_a$ for all j.

16. Suppose the step history of $\mathcal{C}$ contains the letter $(2i, 2i+1)$ for some $1 \leq i \leq 2n-2$ even.

Then H has a subword $H_{2i}\theta(2i, 2i+1)H_{2i+1}$ where H_{2i} and H_{2i+1} are histories of maximal subcomputations of step history $(2i)$ and $(2i+1)$, respectively.

Note that $\theta(2i, 2i+1)$ locks all sectors of the standard base except for the Q_0P_1- and R_1Q_2-sectors. So, any unreduced two-letter subword of B must be of the form $Q_0Q_0^{-1}$, $P_1^{-1}P_1$, $R_1R_1^{-1}$, or $Q_2^{-1}Q_2$.

As in Step 15, we then have that H also contains a subword $\theta(2i, 2i+1)H_{2i+1}\theta(2i+1, 2i)$. Again, this implies that B cannot contain the letters $P_1^{\pm 1}$. So, B must be a cyclic permutation of

$$P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}Q_2R_2Q_3P_4Q_4\{t\}P_0$$

Let $\mathcal{C}_{2i} : W_r \to \cdots \to W_s$ be the subcomputation with history H_{2i}. The restriction of $\mathcal{C}_{2i}$ to the R_2Q_3-sector satisfies the hypotheses of Lemma 4.3, so that $r = 0$.

Let $\mathcal{C}'_{2i} : W'_0 \to \cdots \to W'_s$ be the restriction of $\mathcal{C}_{2i}$ to the subword $Q_3^{-1}R_2^{-1}Q_2^{-1}Q_2R_2Q_3$. As in Step 14, Lemma 5.14 then implies that $|W'_s|_a \leq |W'_0|_a$.

But all other sectors have fixed tape word throughout $\mathcal{C}_{2i}$, so that $|W_s|_a \leq |W_0|_a$.

Hence, Steps 15 and 16 imply that the step history $\mathcal{C}$ contains no letter of the form $(2i, 2i+1)$. The symmetric argument further implies that it contains no letter of the form $(2i+1, 2i)$.

17. Suppose the step history of $\mathcal{C}$ contains the letter $(2i, 2i-1)$ for some $2 \leq i \leq 2n-2$ even.

As $\theta(2i, 2i-1)$ locks every sector of the standard base except for the Q_0P_1- and R_2Q_3-sectors, any unreduced two-letter subword of B must be of the form $Q_0Q_0^{-1}$, $P_1^{-1}P_1$, $R_2R_2^{-1}$, or $Q_3^{-1}Q_3$.

By Steps 6-8, the step history of $\mathcal{C}$ must contain the subword $(2i, 2i-1)(2i-1)(2i-1, 2i)$. Let $\mathcal{C}_{2i-1}$ be the maximal subcomputation with step history $(2i-1)$ in this subword. By Steps 11-13, the history H_{2i-1} of $\mathcal{C}_{2i-1}$ contains no connecting rule. So, $\mathcal{C}_{2i-1}$ is a computation of $\mathbf{M}_4((2i-1)^+)$.

If B contains a subword of the form $(Q_2R_2)^{\pm 1}$, then the restriction of $\mathcal{C}_{2i-1}$ to this subword satisfies the hypotheses of Lemma 4.3, so that H_{2i-1} must be empty. So, B cannot contain the letters $R_2^{\pm 1}$. In particular, B must be a cyclic permutation of

$$P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}Q_3P_4Q_4\{t\}P_0$$

As in Step 15, this implies $|W_j|_a = |W_0|_a$ for all j.

The same argument implies that the step history of $\mathcal{C}$ cannot contain a letter $(2i-1,2i)$ for $1\leq i\leq 2n-2$ even.

So, every transition rule of H must be of the form $\theta(2i-1,2i)^{\pm 1}$ for $2\leq i\leq 2n-2$ odd.

18. Finally, assume that $\mathcal{C}$ contains the letter $(2i,2i-1)$ for some $2\leq i\leq 2n-2$ odd.

Then H must contain a subword $H_{2i}\theta(2i,2i-1)H_{2i-1}$ where H_{2i-1} and H_{2i} are histories of maximal subcomputations of step history $(2i-1)$ and $(2i)$, respectively.

Since $\theta(2i,2i-1)$ locks every sector of the standard base except for the Q_0P_1- and R_1Q_2-sectors, any unreduced two-letter subword of B must be of the form $Q_0Q_0^{-1}$, $P_1^{-1}P_1$, $R_1R_1^{-1}$, or $Q_2^{-1}Q_2$.

As in Step 17, the step history of $\mathcal{C}$ must contain the subword $(2i,2i-1)(2i-1)(2i-1,2i)$. Let $\mathcal{C}_{2i-1}$ be the maximal subcomputation with step history $(2i-1)$ in this subword. By Steps 11-13, the history H_{2i-1} of $\mathcal{C}_{2i-1}$ contains no connecting rule, so that $\mathcal{C}_{2i-1}$ is a computation of $\mathbf{M}_4((2i-1)^+)$.

If B contains a subword of the form $(Q_1R_1)^{\pm 1}$, then the restriction of $\mathcal{C}_{2i-1}$ to this subword satisfies the hypotheses of Lemma 4.3, so that H_{2i-1} must be empty. So, B cannot contain $R_1^{\pm 1}$, and so is a cyclic permutation of

$$P_0Q_0Q_0^{-1}P_0^{-1}\{t\}^{-1}Q_4^{-1}P_4^{-1}Q_3^{-1}R_2^{-1}Q_2^{-1}Q_2R_2Q_3P_4Q_4\{t\}P_0$$

As in Step 16, letting $\mathcal{C}_{2i}:W_r\to\cdots\to W_s$ be the subcomputation with history H_{2i}, we must then have $r=0$ and $|W_s|_a\leq |W_0|_a$. Thus, as we can apply the symmetric argument, we reach a final contradiction.

□

5.9 The machines $\mathbf{M}_{5,1}$ and $\mathbf{M}_{5,2}$

The cyclic machine $\mathbf{M}_{5,1}$ functions as the 'parallel' composition of the machine $\mathbf{M}_4$ with itself a large number of times.

Letting $\{t(i)\}B_3(i)$ be a copy of the standard base of $\mathbf{M}_4$ for $i \in \{1, \dots, L\}$, the standard base of $\mathbf{M}_{5,1}$ is

$$\{t(1)\}B_3(1)\{t(2)\}B_3(2)\dots\{t(L)\}B_3(L)$$

For any letter of $\{t(i)\}B_3(i)$ (or its inverse), the index i is called its *coordinate*.

The tape alphabet of any sector formed by a one-letter part $\{t(i)\}$ of the standard base (including the $Q_4(L)\{t(1)\}$-sector) is defined to be empty. The tape alphabets of all other sectors arise from $\mathbf{M}_4$ in the natural way.

The rules of $\mathbf{M}_{5,1}$ are in correspondence with those of $\mathbf{M}_4$, with each rule operating in parallel on each of the copies of the standard base of $\mathbf{M}_4$ in the same way as its corresponding rule.

The copies of the input sector are taken as the input sectors of the machine.

Naturally, there arise submachines $\mathbf{M}_{5,1}(i)$ corresponding to the submachines $\mathbf{M}_4$. As such, the definition of step history and controlled history extend to reduced computations of $\mathbf{M}_{5,1}$.

The statements of Chapter 5.7 have natural analogues in $\mathbf{M}_{5,1}$. For example, letting $I_5(w)$ be the input configuration with the natural copy of w in each $P_0(i)Q_0(i)$-sector, then the following is the analogue of Lemma 5.33.

Lemma 5.38. *An input configuration W is accepted by $\mathbf{M}_{5,1}$ if and only if $W \equiv I_5(u^n)$ for some $u^n \in \mathcal{L}$. Moreover, for any $u^n \in \mathcal{L}$, there exists a unique accepting computation $\mathcal{C}_{5,1}(u)$ of the input configuration $I_5(u^n)$.*

The cyclic machine $\mathbf{M}_{5,2}$ is constructed in much the same way as $\mathbf{M}_{5,1}$, but with one fundamental difference: Each rule locks the first input sector, i.e the $P_0(1)Q_0(1)$-sector.

The definitions of $\mathbf{M}_{5,1}$ extend in an obvious way to $\mathbf{M}_{5,2}$, and many of the statements of Chapter 5.7 again have natural analogues. For example, letting $J_5(w)$ be the input configuration that is obtained from emptying the $P_0(1)Q_0(1)$-sector of the natural copy of $I_5(w)$, the following is the analogue of Lemma 5.33 (and Lemma 5.38).

Lemma 5.39. *An input configuration W is accepted by $\mathbf{M}_{5,2}$ if and only if $W \equiv J_5(u^n)$ for some $u^n \in \mathcal{L}$. Moreover, for any $u^n \in \mathcal{L}$, there exists a unique accepting computation $\mathcal{C}_{5,2}(u)$ of the* *input configuration* $J_5(u^n)$.

Chapter 6

The machine $\mathbf{M}$

6.1 Definition of the machine

The final step of our construction is to combinine the machines $\mathbf{M}_{5,1}$ and $\mathbf{M}_{5,2}$ to create the cyclic machine $\mathbf{M}$ that is sufficient for the proof of Theorem A.

Similar to $\mathbf{M}_{5,1}$ and $\mathbf{M}_{5,2}$, the standard base of $\mathbf{M}$ is of the form $\{t(1)\}B_3(1)\dots\{t(L)\}B_3(L)$, with the sectors of the form $P_0(i)Q_0(i)$ taken to be the input sectors. However, each of the parts making up $B_3(i)$ consists of more state letters than its counterparts in $\mathbf{M}_{5,1}$ and $\mathbf{M}_{5,2}$.

To be precise, any part of the standard base that is not a one-letter part $\{t(i)\}$ consists of a copy of the corresponding part of the standard base of $\mathbf{M}_{5,1}$, a (disjoint) copy of the corresponding part of the standard base of $\mathbf{M}_{5,2}$, and two new letters which function as the part's start and end letters. The accept configuration of $\mathbf{M}$ is denoted W_{ac}.

The set of rules Θ of $\mathbf{M}$ is partitioned into two symmetric sets, Θ_1 and Θ_2. The positive rules of each consist of a set of 'working' rules and two more transition rules. Unlike in previous constructions, though, these two sets are not concatenated in order to force them to run sequentially, rather in order to force them to operate 'one or the other'.

The rules of Θ_1^+ are defined as follows:

- The transition rule $\theta(s)_1$ locks all sectors other than the input sectors. It switches the state letters from the start state of $\mathbf{M}$ to the copy of the start state of $\mathbf{M}_{5,1}$.
- The positive 'working' rules of Θ_1^+ are copies of the positive rules of the machine $\mathbf{M}_{5,1}$.
- The transition rule $\theta(a)_1$ locks all sectors and switches the state letters from the copies of the end letters of $\mathbf{M}_{5,1}$ to the end letters of $\mathbf{M}$.

The rules of Θ_2^+ are defined as follows:

- The transition rule $\theta(s)_2$ locks each of the sectors locked by $\theta(s)_1$, but also locks the $P_0(1)Q_0(1)$-sector. It switches the state letters from the start state of **M** to the copy of the start state of $\mathbf{M}_{5,2}$.
- The positive 'working' rules of Θ_2^+ are copies of the positive rules of the machine $\mathbf{M}_{5,2}$.
- The transition rule $\theta(a)_2$ locks all sectors and switches the state letters from the copies of the end letters of $\mathbf{M}_{5,2}$ to the end letters of **M**.

By the definition of the rules, one might infer that the first input sector $P_0(1)Q_0(1)$ is of particular significance. Hence, it is referred to as the *'special' input sector*.

For $w \in F(\mathcal{A})$, the natural copy of $I_5(w)$ (respectively $J_5(w)$) in the hardware of this machine is $\theta(s)_1^{-1}$-admissible (respectively $\theta(s)_2^{-1}$-admissible). We denote $I(w)$ (respectively $J(w)$) as the input configuration satisfying $I(w) \equiv I_5(w) \cdot \theta(s)_1^{-1}$ (respectively $J(w) \equiv J_5(w) \cdot \theta(s)_2^{-1}$). Note that both $I(w)$ and $J(w)$ are $\theta(s)_1$-admissible, while $I(w)$ is not $\theta(s)_2$-admissible if $w \neq 1$.

6.2 Standard computations of **M**

Next, we adapt the definition of step history to computations of **M**. To this end, let the letters $(s)_j^{\pm 1}$ and $(a)_j^{\pm 1}$ represent the transition rules $\theta(s)_j^{\pm 1}$ and $\theta(a)_j^{\pm 1}$ of Θ_j, respectively, and add the subscript j to each letter of the step history of a maximal subcomputation whose history consists of working rules of Θ_j.

So, an example of a step history of a reduced computation of **M** is $(s)_1(1)_1(12)_1(2)_1$, while a general step history is some concatenation of the letters

$$\left\{\begin{array}{l} (1)_j,\ (2)_j,\ \dots,\ (4n)_j,\ (12)_j,\ (23)_j,\ \dots,\ (4n-1,4n)_j, \\ (21)_j,\ (32)_j,\ \dots,\ (4n,4n-1)_j,\ (s)_j^{\pm 1},\ (a)_j^{\pm 1};\ j=1,2 \end{array}\right\}$$

A one-step computation of **M** is defined similar to how it was defined in previous machines. For example, reduced computations with step history $(s)_2^{-1}(s)_1(1)_1(12)_1$ or $(4n)_2(a)_2(a)_1^{-1}$ are one-step computations of **M**.

A reduced computation is called a *one-machine computation* if every letter of its step history has the same index. If this index is i, then the computation is called a *one-machine computation of the i-th machine*.

For example, a reduced computation with step history $(s)_1(1)_1(12)_1(2)_1$ is a one-machine computation of the first machine, while a reduced computation with step history $(1)_1(s)_1^{-1}(s)_2(1)_2$ is not a one-machine computation, i.e it is a *multi-machine* computation.

As with previous machines, some subwords clearly cannot appear in the step history of a reduced computation, while other impossibilities are less obvious. However, there are clear analogues of Lemmas 5.16, 5.17, and 5.28(b) (after adding the same index to each letter of the step histories), as **M** operates on the standard base as parallel copies of $\mathbf{M}_4$ in any one-machine computation whose step history does not contain $(s)_i^{\pm 1}$, $(a)_i^{\pm 1}$, or $(1)_2$.

The following is the analogue of Lemma 5.28(a) and is proved in exactly the same way.

Lemma 6.1. *Suppose the base B of a reduced computation $\mathcal{C}$ of **M** contains a subword UV of the form $(P_0(i)Q_0(i))^{\pm 1}$. Then the step history of $\mathcal{C}$ cannot be $(21)_1(1)_1(12)_1$. Moreover, if $i \neq 1$, then the step history of $\mathcal{C}$ cannot be $(21)_2(1)_2(12)_2$.*

Lemma 6.2. *Let $\mathcal{C}$ be a reduced computation of **M** with base B.*

(a) If B contains a subword UV of the form $(Q_0(i)P_1(i))^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be $(s)_j(1)_j(s)_j^{-1}$ for $j = 1, 2$.

(b) If B contains a subword UV of the form $(R_2(i)Q_3(i))^{\pm 1}$ or $(Q_3(i)P_4(i))^{\pm 1}$, then the step history of $\mathcal{C}$ cannot be $(a)_j^{-1}(4n)_j(a)_j$ for $j = 1, 2$.

Proof. Both statements follow from an application of Lemma 4.3(a) to the restriction of $\mathcal{C}$ to the UV-sector.

□

Lemma 6.3. *Let $\mathcal{C}$ be a reduced computation with base $\{t(i)\}B_4(i)$ for some $i \in \{2, \dots, L\}$. Suppose $\mathcal{C}$ contains at least $8n$ distinct maximal one-step computations. Then $\mathcal{C}$ contains a subword of the form $(4n-2, 4n-1)_j(4n-1)_j(4n-1, 4n)_j$ or $(4n, 4n-1)_j(4n-1)_j(4n-1, 4n-2)_j$.*

Proof. Assuming the step history has no such subword, Lemmas 5.16, 5.17, 5.28, 6.1, and 6.2 imply that the step history is a subword of

- $(4n-1)_1(4n-2)_1\dots(1)_1(s)_1^{-1}(s)_2(1)_2\dots(4n-2)_2(4n-1)_2$, or
- $(4n-1)_1(4n)_1(a)_1(a)_2^{-1}(4n)_2(4n-1)_2$

But then $\mathcal{C}$ has at most $8n-2$ distinct maximal one-step computations.

□

Lemma 6.4. *Let $\mathcal{C} : W_0 \to W_1 \to W_2$ be a reduced computation with step history $((s)_1^{-1}(s)_2)^{\pm 1}$ and base $(P_0(1)Q_0(1))^{\pm 1}$. Then $|W_i|_a = 0$ for $0 \leq i \leq 2$.*

Lemma 6.5. *Let $\mathcal{C} : W_0 \to \dots \to W_t$ of **M** be a one-machine computation of the i-th machine in the standard base. Suppose the step history of $\mathcal{C}$ is of the form $(s)_i h_i (s)_i^{-1}$. Then there exist $u, v \in F(\mathcal{A})$ with $u \neq v$ such that*

- *$W_0 \equiv I(u^n)$ and $W_t \equiv I(v^n)$ if $i = 1$ or*
- *$W_0 \equiv J(u^n)$ and $W_t \equiv J(v^n)$ if $i = 2$.*

Proof. Since W_0 is $\theta(s)_i$-admissible, it is an input configuration.

Further, as h_i cannot be empty, Lemmas 5.16, 5.17, 5.28(b), and 6.2(a) imply that it has prefix

$$(1)_i(2)_i\dots(4n-1)_i(4n-1,4n)_i$$

Let $W_1 \to \dots \to W_s$ be the subcomputation with this step history and $W_1 \to \dots \to W_r$ be the subcomputation with step history $(1)_i$. Then W_r is $\theta(12)_i$-admissible, so that its input sectors are empty.

If $i = 1$, then Lemma 4.3(a) implies that W_0 must have a copy of the same word $w \in F(\mathcal{A})$ written in each input sector. So, $W_0 \equiv I(w)$.

If $i = 2$, then the only difference is that the 'special' input sector must be empty, i.e $W_0 \equiv J(w)$.

The restriction of $W_1 \to \cdots \to W_s$ to $B_3(2)$ can be identified with a reduced computation of $\mathbf{M}_3$. Lemma 5.32(a) then implies that there exists $u \in F(\mathcal{A})$ such that $w = u^n$.

The same argument applied to the inverse computation implies that there exists $v \in F(\mathcal{A})$ such that $W_t \equiv I(v^n)$ if $i = 1$ and $W_t \equiv J(v^n)$ if $i = 2$. Let $W_x \to \cdots \to W_t$ be the subcomputation with step history $(4n, 4n-1)_i(4n-1)_i \dots (2)_i(1)_i(s)_i^{-1}$. Then applying Lemma 5.29 to the restriction of $W_s \to \cdots \to W_x$ to $B_3(2)$ implies that its step history is $(4n)_i$. As this subcomputation cannot be empty, Lemma 4.3 implies that $W_s \neq W_x$. Hence, by Lemma 5.32(b), $u \neq v$.

□

Lemma 6.6. *An input configuration W is accepted by the machine **M** if and only if $W \equiv I(u^n)$ or $W \equiv J(u^n)$ for some $u^n \in \mathcal{L}$. Moreover, for any $u \in F(\mathcal{A})$, there exists a unique one-machine computation of the first (respectively second) machine accepting $I(u^n)$ (respectively $J(u^n)$).*

Proof. Let $\mathcal{C}$ be an accepting computation for W and $\mathcal{C}'$ be the maximal one-machine computation serving as a prefix of $\mathcal{C}$.

Then the step history of $\mathcal{C}'$ must either be of the form $(s)_i h_i (s)_i^{-1}$ or $(s)_i h_i (a)_i$. By Lemma 6.5, it suffices to suppose the step history is of the form $(s)_i h_i (a)_i$.

The maximal subcomputation with step history h_i must then be an accepting computation of the natural copy of $W \cdot \theta(s)_i$ in $\mathbf{M}_{5,i}$. So, Lemma 5.38 implies that $W \equiv I(u^n)$ if $i = 1$, while Lemma 5.39 implies $W \equiv J(u^n)$ if $i = 2$.

The existence and uniqueness of an accepting one-machine computation similarly follow from Lemmas 5.38 and 5.39.

□

6.3 Components of a configuration

For a configuration W and $1 \leq i \leq L$, the *i-th component* of W, $W(i)$, is defined to be the admissible subword of W with base $\{t(i)\}B_3(i)$. So, since the tape alphabet of the $Q_4(i)\{t(i+1)\}$-sector is empty for each i, $W \equiv W(1) \dots W(L)$ for any configuration W. It is useful to note that if a rule θ is applicable to some configuration W, then θ operates on each $W(j)$ in parallel for $j \geq 2$ (but may not operate on $W(1)$ in the analogous way).

Particularly, for $1 \leq i \leq L$, we denote the components $A(i) \equiv W_{ac}(i)$, $I(w,i) \equiv (I(w))(i)$, and $J(w,i) \equiv (J(w))(i)$ for all $w \in F(\mathcal{A})$.

The history H of a reduced computation $\mathcal{C}$ of $\mathbf{M}$ is called *controlled* if $\mathcal{C}$ is a one-machine computation and H corresponds to a controlled computation of $\mathbf{M}_4$. As such, the next statement follows immediately from Lemma 5.18.

Lemma 6.7. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of* ***M*** *with controlled history H. Then the base of the computation is a reduced word and all configurations are uniquely defined by the history H and the base of $\mathcal{C}$.*
Moreover, if $\mathcal{C}$ is a computation in the standard base, then $|W_j|_a = |W_0|_a$ for all $0 \leq j \leq t$, $\|H\| = |W_0(i)|_a + 3$ for all $1 \leq i \leq L$, and W_0 is accepted.

Let V be an admissible word with base B and suppose there exists $i \in \{1, \dots, L\}$ such that every letter of B has coordinate i. Then, a *coordinate shift* of V is an admissible word V' obtained by changing each of the state letters' coordinates from i to j for some $j \in \{1, \dots, L\}$ and taking the natural copies of the tape words. For example, if W is an accepted configuration, then $W(i)$ and $W(j)$ are coordinate shifts of one another for $i, j \geq 2$, while $J(w,1)$ is not a coordinate shift of $J(w,2)$ if $w \neq 1$.

Lemma 6.8. *For $i \in \{2, \dots, L\}$, let $\mathcal{C} : A(i) \to \cdots \to A(i)$ be a nonempty reduced computation of* ***M****. Then $\mathcal{C}$ is not a one-machine computation.*

Proof. Assume toward contradiction that $\mathcal{C}$ is a one-machine computation of the j-th machine. Then H can be factored as $H \equiv \theta(a)_j^{-1} H' \theta(a)_j$ where H' has no letters of the form $\theta(s)_j^{\pm 1}$ or $\theta(a)_j^{\pm 1}$.

Let $\mathcal{C}'$ be the subcomputation with history H'. Then, we can identify $\mathcal{C}'$ with a reduced computation of $\mathbf{M}_4$. This computation starts and ends with the accept configuration of $\mathbf{M}_4$, so that Lemma 5.29 implies that it cannot contain a transition rule.

But then $\mathcal{C}$ has step history $(a)_j^{-1}(4n)_j(a)_j$, so that it contradicts Lemma 6.2(b).

□

Lemma 6.9. *Let $V_0 \to \cdots \to V_t$ be a one-machine computation of the j-th machine with history H and base $\{t(i)\}B_3(i)$ for some $i \in \{2, \ldots, L\}$. Then there exists a one-machine computation $W_0 \to \cdots \to W_t$ in the standard base with history H such that $W_\ell(i) \equiv V_\ell$ for all $0 \leq \ell \leq t$.*

Proof. For each $\ell \in \{0, \ldots, t\}$ and $x \in \{2, \ldots, L\}$, define $V_\ell(x)$ as the coordinate shift of V_ℓ with base $\{t(x)\}B_3(x)$.

If $j = 1$, then similarly define $V_\ell(1)$ as the coordinate shift of V_ℓ with base $\{t(1)\}B_3(1)$. Conversely, if $j = 2$, then define $V_\ell(1)$ as the admissible word obtained from emptying the 'special' input sector of the coordinate shift of V_ℓ.

Now define $W_\ell \equiv V_\ell(1) \ldots V_\ell(L)$ for each $0 \leq \ell \leq t$. Clearly, $W_\ell(i) \equiv V_\ell$ for all ℓ.

Letting $H \equiv \theta_1 \ldots \theta_t$, it follows from construction that $W_{\ell-1}$ is θ_ℓ-admissible and $W_{\ell-1} \cdot \theta_\ell \equiv W_\ell$ for all $1 \leq \ell \leq t$.

□

Using Lemma 6.9, the following statement is an immediate consequence of Lemma 6.5.

Lemma 6.10. *Let $\mathcal{C} : V_0 \to \cdots \to V_t$ be a one-machine computation of the j-th machine with base $\{t(i)\}B_3(i)$ for some $i \in \{2, \ldots, L\}$. Suppose the step history of $\mathcal{C}$ is of the form $(s)_j h_j (s)_j^{-1}$. Then there exist $u, v \in F(\mathcal{A})$ with $u \neq v$ such that $V_0 \equiv I(u^n, i)$ and $V_t \equiv I(v^n, i)$.*

Moreover, for H the history of $\mathcal{C}$, $I(u^n) \cdot H \equiv I(v^n)$ if $j = 1$ and $J(u^n) \cdot H \equiv J(v^n)$ if $j = 2$.

Similarly, the following is an immediate consequence of Lemmas 6.6 and 6.9.

Lemma 6.11. *If W_0 is an admissible subword of a start configuration with base $\{t(i)\}B_3(i)$ for some $i \in \{2, \ldots, L\}$, then there exists a one-machine computation $W_0 \to \cdots \to A(i)$ of the first machine (respectively of the second machine) if and only if $W_0 \equiv I(u^n, i)$ (respectively $W_0 \equiv J(u^n, i)$) for some $u^n \in \mathcal{L}$.*

Finally, the next statement is an immediate consequence of Lemmas 6.8, 6.10, and 6.11.

Lemma 6.12. *For $i \in \{2, \dots, L\}$, suppose $\mathcal{C} : A(i) \to \cdots \to A(i)$ is a reduced computation of* ***M*** *with history H. Let $H_1 \dots H_\ell$ be the factorization of H such that for all $j \in \{1, \dots, \ell\}$, H_j is the history of a maximal one-machine subcomputation $\mathcal{C}_j : U_j \to \cdots \to V_j$ of $\mathcal{C}$. Then $\ell \geq 2$ and for all j, either:*

(a) $V_j \equiv A(i)$, or

(b) $V_j \equiv I(w_j, i)$ or $J(w_j, i)$ for some $w_j \in \mathcal{L}$.

In case (a), set $W_j^{(1)} \equiv W_j^{(2)} \equiv W_{ac}$; in case (b), set $W_j^{(1)} \equiv I(w_j)$ and $W_j^{(2)} \equiv J(w_j)$. Further, set $W_0^{(1)} \equiv W_0^{(2)} \equiv W_{ac}$.

If $\mathcal{C}_j$ is a one-machine computation of the z_j-th machine, then for each $1 \leq j \leq \ell$ there exists a reduced computation in the standard base $\mathcal{C}_j' : W_{j-1}^{(z_j)} \to \cdots \to W_j^{(z_j)}$ with history H_j.

In other words, Lemma 6.12 says that except for the insertion/deletion of elements of $\mathcal{L}$ in the 'special' input sector between its maximal one-machine subcomputations, the computation $\mathcal{C}$ can be 'almost-extended' to a reduced computation $\mathcal{C}' : W_{ac} \to \cdots \to W_{ac}$ (though such a computation need not exist).

Lemma 6.13. *Let W be an accepted configuration and $\theta \in \Theta$. For $i \in \{2, \dots, L\}$, suppose $W(i)$ is θ-admissible while W is not. Then either:*

(1) $\theta = \theta(s)_2$ and $W \equiv I(u^n)$ for some $u^n \in \mathcal{L} \setminus \{1\}$, or

(2) $\theta = \theta(12)_1$ and W has u^n written in the 'special' input sector for some $u^n \in \mathcal{L} \setminus \{1\}$.

In particular, the configuration obtained from W by emptying the 'special' input sector is θ-admissible.

Proof. The symmetry of the rules implies that $W(j)$ is θ-admissible for each $j \geq 2$. So, $W(1)$ must not be θ-admissible. By the definition of the rules, θ must lock the 'special' input sector while that sector is not empty in W.

1. Suppose $\theta \in \Theta_2$.

As W is accepted, it must be θ'-admissible for some $\theta' \in \Theta$. But each rule of Θ_2 locks the 'special' input sector, so that $\theta' \in \Theta_1$. As a result, $W(i)$ is admissible for rules from both Θ_1 and Θ_2, which implies that W must either be a start or an end configuration.

But the only accepted end configuration is W_{ac}, which has empty 'special' input sector. So, W must be an accepted start configuration, $\theta' = \theta(s)_1$, and $\theta = \theta(s)_2$. Finally, since W has nonempty 'special' input sector, Lemma 6.6 yields $W \equiv I(u^n)$ for $u \neq 1$.

2. Suppose $\theta \in \Theta_1$.

Let $\mathcal{C}' : W_{ac} \equiv V_0 \to \cdots \to V_t \equiv W$ be the inverse of an accepting computation of W.

As the rules of Θ_1 operate in parallel as $\mathbf{M}_4$, $W(1)$ cannot be a coordinate shift of $W(i)$. So, $\mathcal{C}'$ cannot be a one-machine computation of the first machine.

As a result, there exists a maximal one-machine subcomputation $\mathcal{D}' : V_r \to \cdots \to V_s$ of the second machine such that the subsequent subcomputation $\mathcal{E}' : V_s \to \cdots \to V_t \equiv W$ is a (perhaps empty) one-machine computation of the first machine. The parallel nature of the rules of Θ_1 imply $V_s(1)$ is not a coordinate shift of $V_s(i)$.

Note that if $\mathcal{E}'$ is empty, then $V_s \equiv W$, so that $V_s(i)$ is θ-admissible. Otherwise, $V_s(i)$ is θ'-admissible for $\theta' \in \Theta_1$ the first rule in the history of $\mathcal{E}'$.

So, since $V_s(i)$ is $(\theta'')^{-1}$-admissible for $\theta'' \in \Theta_2$ the final rule in the history of $\mathcal{D}'$, it follows that V_s is either an accepted start or end configuration. Lemma 6.6 then implies that $V_s \equiv J(w)$ for some $w \in \mathcal{L} \setminus \{1\}$.

As no rule of a computation with step history $(s)_1(1)_1$ locks the special input sector, θ cannot be such a rule. So, the step history of $\mathcal{E}'$ has prefix $(s)_1(1)_1$.

Lemma 6.2(a) then implies that any subsequent letter of the step history must be $(12)_1$. But since V_s has empty 'special' input sector while every other input sector is nonempty, V_t cannot be $\theta(12)_1$-admissible. So, the entire step history of $\mathcal{C}'_1$ is $(s)_1(1)_1$ and $\theta = \theta(12)_1$.

Since $W(i)$ is θ-admissible, Lemma 4.3 implies that the 'special' input sector of W contains the natural copy of the word $w^{-1} \in \mathcal{L} \setminus \{1\}$. Note that removing w^{-1} from the 'special' input sector of W yields a configuration that is θ-admissible.

□

For W an accepted configuration of **M** such that $W \neq W_{ac}$, let $A(W)$ be the set of accepting computations of W. For $\mathcal{C} \in A(W)$, define $\ell(\mathcal{C})$ as the number of maximal one-machine subcomputations of $\mathcal{C}$. Then, define $\ell(W) = \min\{\ell(\mathcal{C}) \mid \mathcal{C} \in A(W)\}$.

For simplicity, further define $\ell(W_{ac}) = 0$.

Lemma 6.14. *For any accepted configuration W of **M**, $\ell(W) \leq 2$.*

Proof. Suppose $\ell = \ell(W) \geq 3$ and set $\mathcal{C} \in A(W)$ such that $\ell(\mathcal{C}) = \ell$.

Let H be the history of the inverse computation of $\mathcal{C}$. Then, factor $H \equiv H_1 \dots H_\ell$ such that each H_i is the history of a maximal one-machine subcomputation.

For $j \in \{1, \dots, \ell{-}1\}$, let $V_j \equiv W_{ac}{\cdot}(H_1 \dots H_j)$. If $V_j \equiv W_{ac}$ for some j, then $(H_{j+1} \dots H_\ell)^{-1}$ is the history of an accepting computation $\mathcal{C}'$ of W with $\ell(\mathcal{C}') < \ell$, contradicting the definition of ℓ.

Lemma 6.6 then implies that for all $1 \leq j \leq \ell - 1$, there exists $w_j \in \mathcal{L}$ such that $V_j \equiv I(w_j)$ or $J(w_j)$. Lemma 6.6 then provides a one-machine computation $\mathcal{D}$ accepting V_2.

Let H' be the history of $\mathcal{D}$. Then $(H_3 \dots H_\ell)^{-1}H'$ is the history of an accepting computation of W whose number of maximal one-machine computations is less than ℓ, again contradicting the definition of ℓ.

□

Lemma 6.15. *Let W be an accepted configuration with $\ell(W) = 2$ and set $\mathcal{C} \in A(W)$ such that $\ell(\mathcal{C}) = 2$. Factor the history H of $\mathcal{C}$ as $H \equiv H_1H_2$, where each H_i is the history of a one-machine computation. Then H_j is the history of a one-machine computation of the j-th machine and $W \cdot H_1 \equiv W_{ac} \cdot H_2^{-1} \equiv J(w)$ for some $w \in \mathcal{L} \setminus \{1\}$.*

Proof. As in the proof of Lemma 6.14, $W \cdot H_1 \equiv W_{ac} \cdot H_2^{-1}$ must be an accepted input configuration.

Note that the final rule of H_1 is $\theta(s)_i^{-1}$ and the first rule of H_2 is $\theta(s)_j$ for $i \neq j$. Lemma 6.4 then implies that $W \cdot H_1$ has empty 'special' input sector, so that $W_{ac} \cdot H_2^{-1} \equiv J(w)$ for some $w \in \mathcal{L}$ by Lemma 6.6.

Suppose H_2 is the history of a one-machine computation of the first machine. Then, since every rule of the first machine operates in parallel on the input sectors, $W_{ac} \cdot H_2^{-1} \equiv I(w)$. This implies $I(w) \equiv J(w)$, so that $w = 1$. Lemma 6.6 provides a one-machine computation $\mathcal{D}$ of the second machine accepting $J(1)$. Let H' be the history of $\mathcal{D}$. Then, the reduced form of H_1H' is the history of a one-machine computation accepting W, contradicting the hypothesis.

Hence, H_j is the history of a one-machine computation of the j-th machine.

If $w = 1$, then there exists a one-machine computation $\mathcal{E}$ of the first machine accepting $I(w)$ by Lemma 6.6. But then for H'' the history of $\mathcal{E}$, the reduced form of H_1H'' is the history of a one-machine computation accepting W. Thus, $w \in \mathcal{L} \setminus \{1\}$.

□

Lemma 6.16. *Let W be an accepted configuration. Then $|W(1)|_a \leq 2|W(j)|_a$ for all $2 \leq j \leq L$.*

Proof. The symmetry of the rules implies that $|W(j)|_a$ is constant for $j \geq 2$.

Let $\mathcal{C}$ be an accepting computation of W with $\ell(\mathcal{C}) = \ell(W)$. As the statement is obvious for $W \equiv W_{ac}$, we may assume that $\ell = \ell(\mathcal{C}) \geq 1$.

If $\mathcal{C}$ is a one-machine computation of the first machine, then $W(1)$ and $W(j)$ are coordinate shifts of one another, so that $|W(1)|_a = |W(j)|_a$.

If it is a one-machine computation of the second machine, then the 'special' input sector is empty while any other sector of $W(1)$ is a coordinate shift of the corresponding admissible subword of $W(j)$. So, in this case $|W(1)|_a \leq |W(j)|_a$.

Hence, we may assume $\ell = 2$. Let $H \equiv H_1H_2$ be the factorization of the history of $\mathcal{C}$ provided by Lemma 6.15, so that $W \cdot H_1 \equiv J(w)$ for some $w \in \mathcal{L} \setminus \{1\}$.

In particular, the 'special' input sector of $W \cdot H_1$ is empty while each of its other input sectors is not. So, since each rule of the first machine operates in parallel on the input sectors, H_1 cannot contain the letter $\theta(12)_1^{\pm 1}$.

Let $\mathcal{C}_1$ be the subcomputation with history H_1. As the application of a transition rule does not alter the a-length of any sector, we may assume that the step history of $\mathcal{C}_1$ is $(1)_1(s)_1^{-1}$.

Factor $H_1 \equiv H_1'\theta(s)_1^{-1}$ and let v be the natural copy of H_1' read right to left in $F(\mathcal{A})$. Letting $\mathcal{C}_1'$ be the subcomputation with history H_1', we may apply Lemma 4.3 to the restriction of $\mathcal{C}_1'$ to the subwords $P_0(i)Q_0(i)$ and $Q_0(i)P_1(i)$. It then follows that W has (the natural copy of):

- v written in the 'special' input sector,
- wv written in every other input sector, and
- v^{-1} written in every $Q_0(i)P_1(i)$-sector.

As all other sectors of W are empty, $|W(1)|_a = \|v\| + \|v^{-1}\| = 2\|v\| \leq 2(\|v^{-1}\| + \|wv\|) = 2|W(j)|_a$.

□

Lemma 6.17. *Let W be an accepted configuration of **M** and $\mathcal{C} : W \equiv W_0 \to \cdots \to W_t \equiv W_{ac}$ be an accepting computation with $\ell(\mathcal{C}) = \ell(W)$. Then $t \leq c_2\|W(i)\|$ for all $i \in \{2, \ldots, L\}$.*

Proof. The statement is clear for $W \equiv W_{ac}$, so we may assume $\ell(W) \geq 1$.

Suppose $\ell(W) = 1$, so that $\mathcal{C}$ is a one-machine computation of the j-th machine. The history of $\mathcal{C}$ can then be factored as $H''H'\theta(a)_j$, where:

- H'' is either empty or $\theta(s)_j$, and
- H' does not contain the letters $\theta(a)_j^{\pm 1}$ or $\theta(s)_j^{\pm 1}$.

Let $\mathcal{C}' : W_0' \to \cdots \to W_s'$ be the subcomputation with history H' and let $\mathcal{C}_i'$ be its restriction to the base $\{t(i)\}B_3(i)$ for some $i \geq 2$. We can then identify $\mathcal{C}_i'$ with a reduced computation of $\mathbf{M}_4$ in the standard base.

Lemma 5.35 implies that $s \leq 2c_1 \max(\|W_0'(i)\|, \|W_s'(i)\|)$. As W_s' is $\theta(a)_j$-admissible, $|W_s'|_a = 0$, so that $\|W_0'(i)\| \geq \|W_s'(i)\|$. Further, since transition rules do not change the tape word of any sector, $\|W_0'(i)\| = \|W(i)\|$.

Hence, $t \leq 2c_1\|W(i)\| + 2 \leq 3c_1\|W(i)\|$. The statement then follows from the parameter choice $c_2 >> c_1$.

So, we may assume that $\ell(W) = 2$. Then, factor the history $H \equiv H_1H_2$ of $\mathcal{C}$ as in Lemma 6.15. Let $\mathcal{C}_1 : W_0 \to \cdots \to W_r$ be the subcomputation with history H_1. Then, $W_r \equiv J(w)$ for $w \in \mathcal{L} \setminus \{1\}$.

As $w \neq 1$ and the rules with step history $(1)_1$ operate in parallel on all input sectors, the step history of $\mathcal{C}_1$ must be $(1)_1(s)_1^{-1}$ (with perhaps empty maximal subcomputation with step history $(1)_1$). For $i \geq 2$, applying Lemma 4.3 to the restriction of $\mathcal{C}_1$ to the $Q_0(i)P_1(i)$-sector then implies that $r - 1 \leq |W(i)|_a$. Further, a projection argument yields $|W(i)|_a \geq |W_r(i)|_a$.

The subcomputation $W_r \to \cdots \to W_t$ is a one-machine computation accepting W_r, so that as above $t - r \leq 3c_1\|W_r(i)\|$ for all $i \geq 2$.

Thus, $t \leq 3c_1\|W(i)\| + \|W(i)\| \leq 4c_1\|W(i)\|$, so that the statement again follows from the parameter choice $c_2 >> c_1$.

□

Lemma 6.18. *For any reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ of **M** in the standard base, $\|W_i\| \leq c_2 \max(\|W_0\|, \|W_t\|)$ for all $0 \leq i \leq t$.*

Proof. As the application of a transition rule does not change the length of a configuration, we may assume without loss of generality that neither the first nor the last rule of the history H of $\mathcal{C}$ is a transition rule.

1. Suppose $\mathcal{C}$ is a one-machine computation of the first machine. Then H cannot contain the letters $\theta(s)_j^{\pm 1}$ or $\theta(a)_j^{\pm 1}$.

So, for each $1 \leq j \leq L$, the restriction $\mathcal{C}(j) : W_0(j) \to \cdots \to W_t(j)$ to $\{t(j)\}B_3(j)$ can be identified with a reduced computation of $\mathbf{M}_4$ in the standard base. Lemma 5.35 then implies $t \leq 2c_1 \max(\|W_0(j)\|, \|W_t(j)\|)$.

Note that the application of any rule alters the a-length of any component by at most four. So, applying the same argument as used in the proof of Lemma 5.11, we get

$$\|W_i(j)\| \leq 5c_1 \max(\|W_0(j)\|, \|W_t(j)\|)$$

for all $0 \leq i \leq t$. So, for all i,

$$\|W_i\| = \sum_{j=1}^{L} \|W_i(j)\| \leq 5c_1 \sum_{j=1}^{L} \max(\|W_0(j)\|, \|W_t(j)\|) \leq 5c_1 \left(\sum_{j=1}^{L} \|W_0(j)\| + \sum_{j=1}^{L} \|W_t(j)\| \right)$$
$$= 5c_1(\|W_0\| + \|W_t\|) \leq 10c_1 \max(\|W_0\|, \|W_t\|)$$

2. Suppose $\mathcal{C}$ is a one-machine computation of the second machine. If the step history of $\mathcal{C}$ does not contain the letter $(1)_2$, then the same argument as above implies $\|W_i\| \leq 10c_1 \max(\|W_0\|, \|W_t\|)$ for all i.

If the step history of $\mathcal{C}$ is $(1)_2$, then the restriction of $\mathcal{C}$ to any unlocked sector satisfies the hypotheses of Lemma 4.3. As a result, it follows that $|W_i|_a \leq 2\max(|W_0|_a, |W_t|_a)$ for all i.

So, we assume that the step history contains $(1)_2$ as a proper subword. Lemma 6.1 then implies that any occurrence of $(1)_2$ is as the first or last letter of the step history. Let $W_r \to \cdots \to W_s$ be the maximal subcomputation of $\mathcal{C}$ such that its step history has no occurrence of $(1)_2$. Then $\|W_i\| \leq 10c_1 \max(\|W_r\|, \|W_s\|)$ for all $r \leq i \leq s$.

If $\mathcal{C}_1 : W_s \to \cdots \to W_t$ is nonempty, then it has step history $(1)_2$ and W_s is $\theta(12)_2$-admissible. Lemma 5.13 then implies that $|W_s(j)|_a \leq \cdots \leq |W_t(j)|_a$ for $j \geq 2$. Further, Lemma 4.3 applies to the restriction of $\mathcal{C}_1$ to the $P_0(j)Q_0(j)$-sector for any $j \geq 2$, implying $t - s \leq |W_t(j)|_a$. Then, for all $s \leq i \leq t$, Lemma 4.3 implies that $|W_i(1)|_a \leq |W_t(1)|_a + t - i \leq |W_t(1)|_a + |W_t(j)|_a$ for any $j \geq 2$. So, $|W_i|_a \leq 2|W_t|_a$ for all $s \leq i \leq t$.

Similarly, if $W_0 \to \cdots \to W_r$ is nonempty, then $|W_i|_a \leq 2|W_0|_a$ for all $0 \leq i \leq r$. Combining these inequalities yields $\|W_i\| \leq 20c_1 \max(\|W_0\|, \|W_t\|)$ for all $0 \leq i \leq t$.

Hence, by the parameter choice $c_2 >> c_1$, we may assume that $\mathcal{C}$ is a multi-machine computation.

Moreover, we may induct on the number of maximal one-machine subcomputations of $\mathcal{C}$.

3. Suppose W_0 is not an accepted configuration.

As $\mathcal{C}$ is multi-machine, there exists $0 < s < t$ such that W_s is either a start or an end configuration.

If W_s is an end configuration, then it must be $\theta(a)_j^{-1}$-admissible, so that $W_s \equiv W_{ac}$. But then W_0 must be an accepted configuration, contradicting our assumption.

So, W_s must be an input configuration. Lemma 6.4 then implies that it has empty 'special' input sector. If all input sectors are empty, then $W_s \equiv I(1)$, so that W_0 is an accepted configuration. As a result, W_s must have a nonempty input sector.

Perhaps taking the inverse computation, we may assume without loss of generality that there exists a maximal one-machine computation $\mathcal{C}_1 : W_r \to \cdots \to W_s$ of the first machine. Since W_s has empty 'special' input sector, Lemmas 4.3 and 6.2(a) imply that the step history of $\mathcal{C}_1$ is $(1)_1(s)_1^{-1}$, so that $r = 0$. Lemmas 4.3 and 5.13 then imply that $|W_i|_a \leq |W_0|_a$ for all $i \leq s$.

As $W_s \to \cdots \to W_t$ consists of one less maximal one-machine subcomputation than $\mathcal{C}$, the inductive hypothesis implies $\|W_i\| \leq c_2 \max(\|W_s\|, \|W_t\|)$ for all $s \leq i \leq t$. This yields $\|W_i\| \leq c_2 \max(\|W_0\|, \|W_t\|)$ for all i.

Hence, we may assume that W_0 is an accepted configuration.

4. Suppose there exists $w \in \mathcal{L} \setminus \{1\}$ such that $W_s \equiv I(w)$ or $J(w)$ for some $s \in \{0, \dots, t\}$.

If $s = 0$ (resp $s = t$), then the first (resp last) rule of H must be a transition rule of the form $\theta(s)_j^{\pm 1}$. But this contradicts our assumption. Lemma 6.4 then implies that the 'special' input sector of W_s must be empty, so that $W_s \equiv J(w)$.

As a result, we may assume without loss of generality that there exists a maximal one-machine subcomputation $\mathcal{C}_1 : W_r \to \cdots \to W_s$ of the first machine. As in Step 3, Lemmas 4.3 and 6.2(a) then imply that $r = 0$ and $|W_i|_a \leq |W_0|_a$ for all $0 \leq i \leq s$.

But then the inductive hypothesis again yields $\|W_i\| \leq c_2 \max(\|W_0\|, \|W_t\|)$ for all i.

5. Finally, suppose that for any $s \in \{0, \dots, t\}$ such that W_s is a start (resp end) configuration, $W_s \equiv I(1)$ (resp $W_s \equiv W_{ac}$).

As $\mathcal{C}$ is not a one-machine computation, there exists such an s. Further, by the same reasoning as used in Step 4, we may assume that $s \in \{1, \dots, t-1\}$.

Then, $W_0 \to \dots \to W_s$ and $W_s \to \dots \to W_t$ each consist of less maximal one-machine subcomputations than does $\mathcal{C}$, so that the inductive hypothesis implies $\|W_i\| \leq c_2 \max(\|W_0\|, \|W_s\|)$ for all $0 \leq i \leq s$ and $\|W_i\| \leq c_2 \max(\|W_s\|, \|W_t\|)$ for all $s \leq i \leq t$.

But $|W_s|_a = 0$, so that $\max(\|W_0\|, \|W_s\|) = \|W_0\|$ and $\max(\|W_s\|, \|W_t\|) = \|W_t\|$. Thus, the statement is proved.

□

6.4 Computations of **M** with long history

Lemma 6.19. *Let $\mathcal{C} : V_0 \to \dots \to V_t$ be a reduced computation of **M** with base $\{t(i)\}B_3(i)$ for some $i \in \{2, \dots, L\}$. Suppose $t > c_3 \max(\|V_0\|, \|V_t\|)$. Then:*

(a) There exist accepted configurations W_0 and W_t such that $W_0(i) \equiv V_0$ and $W_t(i) \equiv V_t$,

(b) Let $\mathcal{C}_0$ and $\mathcal{C}_t$ be accepting computations of W_0 and W_t, respectively, with $\ell(\mathcal{C}_j) = \ell(W_j)$. For H_j the history $\mathcal{C}_j$, $\|H_0\| + \|H_t\| \leq t/500$

(c) The sum of the lengths of all subcomputations of $\mathcal{C}$ whose step histories are of the form $(4n-2, 4n-1)_j(4n-1)_j(4n-1, 4n)_j$ or $(4n, 4n-1)_j(4n-1)_j(4n-1, 4n-2)_j$ is at least $0.98t$.

Proof. For H' the history of $\mathcal{C}$, factor $H' \equiv H'_1 \dots H'_m$ for $m \geq 1$ so that each H'_j is the history of a maximal one-machine subcomputation of $\mathcal{C}$.

Suppose $m = 1$, i.e $\mathcal{C}$ is a one-machine computation. Then the letters $\theta(s)_j^{\pm 1}$ or $\theta(a)_j^{\pm 1}$ can only occur in H' as the first or last letter. So, other than perhaps these two rules, $\mathcal{C}$ can be viewed as a reduced computation of $\mathbf{M}_4$ in the standard base. But then Lemma 5.35 implies that $t - 2 \leq 2c_1 \max(\|V_0\|, \|V_t\|)$, so that the parameter choice $c_3 >> c_1$ provides a contradiction. So, we may assume that $m \geq 2$, i.e $\mathcal{C}$ is a multi-machine computation.

For each $1 \leq j \leq m$, let $\mathcal{C}_j : V_{y(j)} \to \cdots \to V_{z(j)}$ be the subcomputation with history H'_j. By Lemma 6.9, there exists a one-machine computation $\mathcal{C}'_j : W'_{y(j)} \to \cdots \to W'_{z(j)}$ in the standard base with history H'_j and such that $W'_x(i) \equiv V_x$ for $y(j) \leq x \leq z(j)$. Note that for $1 \leq j \leq m-1$, $W'_{z(j)}$ and $W'_{y(j+1)}$ are not necessarily the same configuration; to differentiate them, they are represented with these indices.

Suppose $W'_{z(1)}$ is not an accepted configuration. Since $V_{z(1)} \equiv V_{y(2)}$ is admissible for the last rule of H'_1 and the first rule of H'_2, $W'_{z(1)}$ must be a start or an end configuration.

Note that the only end configuration that is admissible for any rule is W_{ac}, which is accepted. So, since $W'_{z(1)}$ is admissible for the last rule of H'_1, it must be a start configuration which is admissible for $\theta(s)_j$ for some j.

By construction, there exists $w \in F(\mathcal{A})$ such that $W'_{z(1)} \equiv I(w)$ (or $W'_{z(1)} \equiv J(w)$) if $\mathcal{C}_1$ is a one-machine computation of the first (or second) machine. By Lemma 6.6, we must then have $w \notin \mathcal{L}$. Lemmas 6.10 and 6.11 then imply that $m = 2$.

Suppose $\mathcal{C}_2$ is a one-machine computation of the j-th machine. Then as above, we can view this as a reduced computation of $\mathbf{M}_4$ after disregarding the first rule. Lemmas 5.32 and 5.28 then imply that the step history is a subword of $(s)_j(1)_j \dots (4n-1)_j$. So, the length of $\mathcal{C}_2$ must be at most $c_2\|V_t\|$ by Lemmas 5.35 and 5.30.

Similarly, the length of $\mathcal{C}_1$ must be at most $c_2\|V_0\|$.

So, $t \leq c_2(\|V_0\| + \|V_t\|)$, so that the parameter choice $c_3 >> c_2$ yields a contradiction.

Hence, $W'_{z(1)}$ must be an accepted configuration. Consequently, $W_0 \equiv W'_{y(1)}$ is an accepted configuration with $W_0(i) \equiv V_0$.

By the construction of the one-machine computations in the standard base outlined in the proof of Lemma 6.9, that $W'_{z(1)}$ is accepted implies that $W'_{y(2)}$ is accepted. As a result, $W'_{z(2)}$ is accepted. Continuing, we have that $W_t \equiv W'_{z(m)}$ is an accepted configuration with $W_t(i) \equiv V_t$.

Thus, (a) is satisfied.

By Lemma 6.17, we then have $\|H_0\| + \|H_t\| \leq 2c_2 \max(\|V_0\|, \|V_t\|) \leq t/500$ as $c_3 >> c_2$.

For $2 \leq j \leq m-1$, let ℓ_j be the sum of the lengths of the subcomputations of $\mathcal{C}_j$ whose step histories are of the form given in (c).

As above, for such j, by neglecting the first and last rules, we may view $\mathcal{C}_j$ as a reduced computation of $\mathbf{M}_4$ in the standard base. So, Lemma 5.34 implies that $\ell_j \geq \left(1-\frac{1}{c_0}\right)(\|H_j'\|-2)$. As $\|H_j'\| \geq k$ by Lemma 5.32(c), taking c_0 sufficiently large then yields $\ell_j \geq 0.99\|H_j'\|$.

Let $y(m)=y$, so that $\mathcal{C}_m : V_y \to \cdots \to V_t$ is the subcomputation of $\mathcal{C}$ with history H_m'. Then, as $\mathcal{C}_m$ is a one-machine computation, Lemma 5.35 implies $\|H_m'\| = t-y \leq c_2 \max(\|V_y\|, \|V_t\|)$.

If $\|V_y\| \leq c_2\|V_t\|$, then $\|H_m'\| \leq c_2^2\|V_t\| \leq t/200$ by the parameter choice $c_3 >> c_2$.

Otherwise, $\|V_y\| > c_2\|V_t\|$. If V_y is the i-th component of an end configuration, then it must be $\theta(a)_j^{-1}$-admissible. But then $V_y \equiv A(i)$, so that $\|V_y\| \leq \|V_t\|$.

So, V_y must be the i-th component of a start configuration. As above, we may choose this start configuration to be accepted, so that $V_y \equiv I(u^n, i)$ for some $u^n \in \mathcal{L}$.

Identifying $\mathcal{C}_m$ with a reduced computation of $\mathbf{M}_4$, Lemma 5.30 then implies that its step history must contain the letter $(4n)_j$. Let $\mathcal{C}_m'' : V_y \to \cdots \to V_z$ be the maximal subcomputation whose step history is $(s)_j(1)_j \dots (4n)_j$ and $V_x \to \cdots \to V_z$ be the maximal subcomputation with step history $(4n)_j$. So, by Lemmas 5.30 and 5.32, $|V_y|_a \leq 9n|V_x|_a$ and $|V_x|_a = 2\|u\|$.

If $z \neq t$, then the first letter of the step history of $V_z \to \cdots \to V_t$ is either $(a)_j$ or $(4n, 4n-1)_j$. Lemma 5.31 then implies that $|V_z|_a \leq 12n|V_t|_a$, so that $|V_y|_a > c_1|V_z|_a$ and $t-z \leq t/200$. So, taking c_1 sufficiently large, $|V_z|_a \leq |V_x|_a$. Hence, Lemma 4.3 implies $z-x \leq |V_x|_a = 2\|u\|$.

Let ℓ_m be the sum of the lengths of the subcomputations of $\mathcal{C}_m''$ whose step histories are of the form described in the statement. By Lemmas 5.32(c) and 4.3, we then have

$$z-y-\ell_m \leq (c_0+1)(\|u\|+1) + 2\|u\| + \|u^n\| + 1 \leq 2c_0(\|u\|+1)$$

and $\ell_m = 2k\|u\| + 2k + 1 \geq 2k(\|u\|+1)$. As in the proof of Lemma 5.34, this then implies $\ell_m \geq \left(1-\frac{1}{c_0}\right)(z-y) \geq 0.99(z-y)$.

We can then do the same for the subcomputation $\mathcal{C}_1 : V_0 \to \cdots \to V_s$ with history H_1', finding $r \in \{0, \dots, s\}$ with $r \leq t/200$ such that for ℓ_1 the sum of the lengths of the subcomputations of

$V_r \to \cdots \to V_s$ whose step histories are of the form described in (c), $\ell_1 \geq 0.99(s-r)$.

Let ℓ be the sum of the lengths of the subcomputations of $V_r \to \cdots \to V_z$ whose step histories are of the form given in (c). Then $\ell = \sum_{i=1}^m \ell_i \geq 0.99(z-r)$ while $z-r \geq 0.99t$. Thus, $\ell \geq 0.98t$.

□

Lemma 6.20. *Let $\mathcal{C} : V_0 \to \cdots \to V_t$ be a reduced computation of **M** with base $\{t(i)\}B_3(i)$ for some $i \in \{2, \dots, L\}$. If $t > c_3 \max(\|V_0\|, \|V_t\|)$, then the history of any subcomputation $\mathcal{D} : V_r \to \cdots \to V_s$ of $\mathcal{C}$ (or the inverse of $\mathcal{D}$) of length at least $0.4t$ contains a controlled subword.*

Proof. By Lemma 6.19(c), the sum of the lengths of all subcomputations of $\mathcal{C}$ with step histories of the form $(4n-2, 4n-1)_j(4n-1)_j(4n-1, 4n)_j$ or $(4n, 4n-1)_j(4n-1)_j(4n-1, 4n-2)_j$ is at least $0.98t$. So, there exists such a subcomputation $\mathcal{C}'$ such that $\mathcal{D}$ contains a subcomputation $\mathcal{D}'$ which is also a subcomputation of $\mathcal{C}'$. Moreover, for H' and K' the histories of $\mathcal{C}'$ and $\mathcal{D}'$, we may assume $\|K'\| \geq 0.3\|H'\|$.

But $\mathcal{C}'$ repeats k copies of a controlled history (with an overlap of one rule), so that taking k sufficiently large implies that K' must contain a controlled subword.

□

A two-letter subword UV of the standard base of **M** is defined to be left-active (resp right-active) with respect to the step $(i)_j$ if any rule of step history $(i)_j$ that alters the tape word of an admissible word with base UV inserts/deletes one letter on the left (resp right) of the tape word.

Note that if $(i)_j$ is not $(1)_2$ or the subword does not correspond to the 'special' input sector, then UV is left-active (resp right-active) with respect to $(i)_j$ if and only if the corresponding subword of the standard base of $\mathbf{M}_4$ is left-active (resp right-active) for $\mathbf{M}_4(i)$. Hence, the following is an immediate consequence of Lemma 5.36.

Lemma 6.21. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of **M** with step history $(i)_m$ and base $B \equiv \{t(x)\}B_3(x)$ for some $2 \leq x \leq L$. Assume that for some index j, $|W_j|_a > 4|W_0|_a$. Then there are subwords $U_\ell V_\ell$ and $U_r V_r$ of B such that $U_\ell V_\ell$ is left-active with respect to $(i)_m$, $U_r V_r$ is* *right-active with respect to the step* $(i)_m$*, and for* $W_0' \to \cdots \to W_t'$ *the restriction of* $\mathcal{C}$ *to either sector,* $|W_j'|_a < |W_{j+1}'|_a < \cdots < |W_t'|_a$.

6.5 Reverted Bases

Let B be the base of an admissible word W of **M**. Then the *reversion* of B, denoted $\pi(B)$, is the word obtained from B by 'forgetting' the coordinates of its letters. In this case, $\pi(B)$ is called the *reverted base* of W.

For example, the reverted base of any configuration is the concatenation of L copies of the standard base of $\mathbf{M}_4$. Similarly, if $B = Q_4(1)\{t(2)\}P_0(2)Q_0(2)P_1(2)Q_1(2)Q_1(2)^{-1}P_1(2)^{-1}$, then

$$\pi(B) = Q_4\{t\}P_0Q_0P_1Q_1Q_1^{-1}P_1^{-1}$$

Lemma 6.22. *Let* B *be the base of an admissible word* W *of* **M***. Then there exists an admissible word* W' *of* $\mathbf{M}_4$ *with base* $\pi(B)$ *and such that* $|W'|_a = |W|_a$*. Moreover, if none of the state letters of* W *are start or end letters (or their inverses), then* W' *can be chosen to be the natural copy of* W *in the hardware of* $\mathbf{M}_4$.

Proof. Let $W \equiv q_0w_1q_1\dots w_rq_r$, $B \equiv V_0\dots V_r$, and $\pi(B) \equiv U_0\dots U_r$.

Suppose $V_0 \equiv V_1^{-1}$, so that $U_0 \equiv U_1^{-1}$. Then the tape alphabet corresponding to the V_0V_1-sector is a copy of that corresponding to the U_0U_1-sector, so that there exists a natural copy w_1' of w_1 in this alphabet.

Further, if neither q_0 nor q_1 corresponds to a start or end letter in V_0 or V_1, then we can choose q_0' and q_1' as the natural copies of q_0 and q_1 in U_0 and U_1, respectively. Otherwsise, either q_0 and q_1 both correspond to start letters or both correspond to end letters. In this case, we can choose q_0' as any state letter of U_0 and q_1' as its inverse.

Now suppose $V_0 \neq V_1^{-1}$. Then since the tape alphabet of the $Q_s\{t\}$-sector of $\mathbf{M}_4$ is empty, we again have that the tape alphabet corresponding to the V_0V_1-sector is a copy of that corresponding to the U_0U_1-sector. This allows us to construct w_1' as a copy of w_1 as above.

Further, if neither q_0 nor q_1 corresponds to a start or end letter in V_0 or V_1, then we can again choose q_0' and q_1' as their natural copies in U_0 and U_1. Meanwhile, if q_0 (or q_1) corresponds to a start or end letter, then we can choose q_0' (or q_1') as any letter from U_0 (respectively U_1).

With q_1' now chosen, we can apply the same construction to obtain q_2' and w_2'.

Iterating, we construct the admissible word $W' \equiv q_0' w_1' q_1' \dots w_r' q_r'$ satisfying the statement.

□

Lemma 6.23. *Suppose $\mathcal{C} : W_0 \to W_1$ is a one-rule computation of $\boldsymbol{M}$ with history $\theta \in \Theta$, where $\theta \notin \{\theta(s)_i, \theta(a)_i : i = 1, 2\}^{\pm 1}$. Further, suppose that either:*

(a) the step history of $\mathcal{C}$ is not $(1)_2$, or

(b) the base of $\mathcal{C}$ does not contain a subword of the form $(P_0(1)Q_0(1))^{\pm 1}$.

Then there exists a one-rule computation of $\boldsymbol{M}_4$ $\mathcal{C}' : W_0' \to W_1'$ with history θ', where θ' is the natural copy of θ in $\Theta(\boldsymbol{M}_4)$ and W_0' and W_1' are the natural copies of W_0 and W_1, respectively, in the hardware of $\boldsymbol{M}_4$.

Proof. As $\theta \notin \{\theta(s)_i, \theta(a)_i : i = 1, 2\}^{\pm 1}$, none of the state letters of W_0 or W_1 are start or end letters (or their inverses). So, applying Lemma 6.22, we can find admissible words W_0' and W_1' which are the natural copies of W_0 and W_1, respectively, in the hardware of $\mathbf{M}_4$.

Let θ' be the natural copy of θ in $\Theta(\mathbf{M}_4)$. If (a) holds, then θ operates on each sector of the standard base of $\mathbf{M}$ in the same way as θ' operates on the copy of the corresponding sector of the standard base of $\mathbf{M}_4$.

Conversely, if θ is a rule of step history $(1)_2$, then all sectors of the standard base of $\mathbf{M}$ other than the 'special' input sector are again operated on by θ in the same way as θ' operates on their copy. As θ locks the 'special' input sector, Lemma 4.2 implies that this sector is not present in W_0 if the base of $\mathcal{C}$ satisfies (b).

□

The base B of an admissible word of $\mathbf{M}$ is called *hyperfaulty* (or *pararevolving*) if its reversion $\pi(B)$ is faulty (or revolving) as the base of an admissible word of $\mathbf{M}_4$. Note that a base is hyperfaulty if and only if it is pararevolving and unreduced.

A hyperfaulty base is necessarily faulty, while a faulty base need not be hyperfaulty. For example, if $B \equiv Q_0(3)Q_0(3)^{-1}P_0(3)^{-1}\{t(3)\}^{-1}\dots Q_4(1)^{-1}Q_4(1)\dots\{t(3)\}P_0(3)Q_0(3)$, where gaps correspond to strings of letters that follow the order of the standard base of $\mathbf{M}$ or its inverse, then B is faulty but not hyperfaulty.

Conversely, a pararevolving base that is not hyperfaulty (for example, $\{t(1)\}\dots\{t(2)\}$) is not revolving, while a revolving base that is not faulty (for example, $\{t(1)\}\dots\{t(1)\}$) is not pararevolving.

As the standard base of $\mathbf{M}_4$ has length 11, a pararevolving base has length at most 23 while a revolving base of $\mathbf{M}$ has length at most $22L+1$.

Lemma 6.24. *Suppose $\mathcal{C}: W_0 \to \dots \to W_t$ is a one-machine computation of* ***M*** *with hyperfaulty base B. Then for all $0 \le j \le t$, $|W_j|_a \le c_0 \max(|W_0|_a, |W_t|_a)$.*

Proof. As in the proof of Lemma 5.37, we may assume that $|W_r|_a > \max(|W_0|_a, |W_t|_a)$ for all $0 < r < t$. In particular, we assume that neither the first nor the last letter of the history H of $\mathcal{C}$ is a transition rule.

As $\mathcal{C}$ is a one-machine computation, any occurrence of a letter of the form $\theta(s)_i^{\pm1}$ or $\theta(a)_i^{\pm1}$ in H would have to be either the first or the last letter. So, no such letter occurs in H.

Suppose the hypotheses of Lemma 6.23 are satisfied by each rule of $\mathcal{C}$. Then, we obtain the reduced computation $\mathcal{C}': W_0' \to \dots \to W_t'$ of $\mathbf{M}_4$ with base $\pi(B)$ such that $|W_j'|_a = |W_j|_a$ for all $0 \le j \le t$. But $\pi(B)$ is faulty, so that Lemma 5.37 implies that $|W_j|_a \le c_0 \max(|W_0|_a, |W_t|_a)$ for all $0 \le j \le t$.

So, it suffices to assume that $\mathcal{C}$ is a one-machine computation of the second machine, that the step history of $\mathcal{C}$ contains the letter $(1)_2$, and that B contains a subword of the form $(P_0(1)Q_0(1))^{\pm1}$.

Suppose that the step history of $\mathcal{C}$ is $(1)_2$. Then the restriction of $\mathcal{C}$ to any two-letter subword of B has fixed tape word, satisfies the hypotheses of Lemma 4.3, or satisfies the hypotheses of

Lemma 4.4. As in Step 1 of the proof of Lemma 5.37, this implies the statement for $c_0 \geq 2$.

So, H (or its inverse) must contain a subword of the form $H_1\theta(12)_2H_2$, where H_i is the history of a maximal subcomputation with step history $(i)_2$. As $\theta(12)_2$ locks every sector of the standard base other than those of the form $(Q_0(i)P_1(i))^{\pm 1}$, any unreduced two-letter subword of B must be of the form $Q_0(i)Q_0(i)^{-1}$ or $P_1(i)^{-1}P_1(i)$. As a result, B must be a cyclic permutation of

$$P_0(1)Q_0(1)Q_0(1)^{-1}P_0(1)^{-1}\{t(1)\}^{-1}Q_4(L)^{-1}\dots Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)\dots\{t(1)\}P_0(1)$$

As B contains the subword $R_2(L)Q_3(L)$, Lemma 5.28(b) implies that the step history of $\mathcal{C}$ cannot contain a subword of the form $(12)_2(2)_2(21)_2$.

Let $\mathcal{C}_2 : W_r \to \dots \to W_s$ be the subcomputation with history H_2 and let $\mathcal{C}_2' : W_r' \to \dots \to W_s'$ be its restriction to the subword

$$Q_3(L)^{-1}R_2(L)^{-1}Q_2(L)^{-1}R_1(L)^{-1}Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)R_1(L)Q_2(L)R_2(L)Q_3(L)$$

Note that every rule with step history $(2)_2$ locks each sector of an admissible subword with base $(P_1(L)Q_1(L)R_1(L)Q_2(L)R_2(L))^{\pm 1}$. So, we may view these subwords as a single state letter. With this view, we may apply Lemma 5.14 to $\mathcal{C}_2'$, so that $|W_r'|_a \leq \dots \leq |W_s'|_a$.

As any other sector has fixed tape word throughout $\mathcal{C}_2$, this yields $|W_r|_a \leq |W_s|_a$. So, H must contain a subword $H_2\theta(23)_2H_3$, where H_3 is the history of a maximal subcomputation with step history $(3)_2$.

Note that the connecting rule of $\mathbf{M}_4(3^-)$ locks the Q_0P_1-sector, so that it cannot occur in H_3. So, letting $\mathcal{C}_3 : W_{s+1} \to \dots \to W_x$ be the subcomputation with history H_3, the restriction of $\mathcal{C}_3$ to the $P_1(L)Q_1(L)$-sector satisfies the hypotheses of Lemma 4.3. As a result, W_x cannot be $\theta(32)_2$-admissible. Hence, $x = t$ and $\mathcal{C}_3$ is the copy of computation of $\mathbf{M}_4(3^-)$.

Letting $\mathcal{C}_3' : W_{s+1}' \to \dots \to W_t'$ be the restriction of $\mathcal{C}_3$ to $Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)$, Lemma 5.14 implies $|W_{s+1}'|_a \leq \dots \leq |W_t'|_a$. But the tape word in each other sector remains unchanged throughout $\mathcal{C}_3$, so that $|W_s|_a \leq |W_t|_a$.

□

Lemma 6.25. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation of **M** with hyperfaulty base B. Then for all $0 \leq j \leq t$, $|W_j|_a \leq c_0 \max(|W_0|_a, |W_t|_a)$.*

Proof. As in the proofs of Lemmas 5.37 and 6.24, we assume that $|W_r|_a > \max(|W_0|_a, |W_t|_a)$ for all $0 < r < t$, so that neither the first nor the last letter of the history H of $\mathcal{C}$ is a transition rule.

By Lemma 6.24, it suffices to assume that $\mathcal{C}$ is a multi-machine computation. Further, as B must be unreduced and the rules $\theta(a)_i^{\pm 1}$ lock each sector of the standard base, the step history of $\mathcal{C}$ must contain a subword of the form $((s)_1^{-1}(s)_2)^{\pm 1}$ by Lemma 4.2.

The only sectors of the standard base not locked by $\theta(s)_1^{\pm 1}$ or $\theta(s)_2^{\pm 1}$ are those of the form $P_0(i)Q_0(i)$ for $i \geq 2$. As a result, any unreduced two-letter subword of B must be of the form $P_0(i)P_0(i)^{-1}$ or $Q_0(i)^{-1}Q_0(i)$ for $i \geq 2$.

As both $\theta(12)_1$ and $\theta(12)_2$ lock all $P_0(i)Q_0(i)$-sectors, the step history of $\mathcal{C}$ cannot contain the letters $(12)_j$ and $(21)_j$.

So, if B does not contain the letters $Q_0(i)^{\pm 1}$, then no rule of $\mathcal{C}$ alters the a-length of an admissible word with base B. As a result, we may assume that B contains such a letter.

By the definition of hyperfaulty, B must contain a subword of the form $(Q_0(i)P_1(i))^{\pm 1}$. Lemma 6.2(a) then implies that the step history of $\mathcal{C}$ cannot contain a subword of the form $(s)_j(1)_j(s)_j^{-1}$.

Hence, the step history of $\mathcal{C}$ (or its inverse) is $(1)_2(s)_2^{-1}(s)_1(1)_1$. Let $\mathcal{C}_1 : W_s \to \cdots \to W_t$ be the maximal subcomputation with step history $(1)_1$.

Note that any subword of B of the form $(P_0(i)Q_0(i))^{\pm 1}$ is contained in a subword of a cyclic permutation of B of the form $(P_0(i)Q_0(i)P_1(i))^{\pm 1}$.

Further, any subword of B of the form $Q_0(i)^{-1}Q_0(i)$ is contained in a subword of a cyclic permutation of B of the form $P_1(i)^{-1}Q_0(i)^{-1}Q_0(i)P_1(i)$.

So, since W_s is $\theta(s)_1^{-1}$-admissible, Lemmas 5.13 and 5.14 imply $|W_s|_a \leq |W_t|_s$, contradicting our assumption.

□

Lemma 6.26. *Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be a reduced computation with revolving base B. Then $\|W_j\| \leq c_4 \max(\|W_0\|, \|W_t\|)$ for all $0 \leq j \leq t$.*

Proof. If B is reduced, then the statement follows from Lemma 6.18 and the parameter choice $c_4 >> c_2$. Similarly, if B is hyperfaulty, then the statement follows from Lemma 6.25 and the parameter choice $c_4 >> c_0$.

So, it suffices to assume that B is faulty but not hyperfaulty. As a result, B contains a reduced pararevolving subword B'. Fix i such that B' contains a subword $(P_0(i)Q_0(i))^{\pm 1}$. Since all non-input sectors are operated on in parallel across coordinates, we may assume that B' is of the form $\{t(i)\}B_3(i)\{t(i+1)\}$ (where we take $L+1$ to be 1).

Let $\mathcal{C}' : W_0' \to \cdots \to W_t'$ be the restriction to B' and suppose $i \neq 1$. Since B is unreduced, Lemma 6.7 implies that the history H of $\mathcal{C}$ cannot contain a controlled subword. So, Lemma 6.19 yields $t \leq c_3 \max(\|W_0'\|, \|W_t'\|) \leq c_3 \max(\|W_0\|, \|W_t\|)$. Hence, the statement follows from the parameter choice $c_4 >> c_3$.

Further, if $i = 1$ and the step history of $\mathcal{C}$ does not contain the letter $(1)_2$, then we may construct a coordinate shift of $\mathcal{C}'$, implying the statement in the same way.

So, we may assume that the step history of $\mathcal{C}$ contains the letter $(1)_2$ and that any reduced pararevolving subword of B contains a subword $(P_0(1)Q_0(1))^{\pm 1}$.

If the step history of $\mathcal{C}$ is $(1)_2$, then the restriction of $\mathcal{C}$ to any two-letter subword has fixed tape word, satisfies the hypotheses of Lemma 4.3, or satisfies the hypotheses of Lemma 4.4. As in Step 1 of the proof of Lemma 5.37, this implies $\|W_j\| \leq 2\max(\|W_0\|, \|W_t\|)$.

So, H must contain a transition rule.

As in previous proofs, we may assume that $\|W_r\| > \max(\|W_0\|, \|W_t\|)$ for all $0 < r < t$, so that neither the first nor the last letter of H is a transition rule (or χ-rule or connecting rule).

Suppose the step history of $\mathcal{C}$ contains the letter $(12)_2$. Then, H must have a subword of the form $H_1\theta(12)_2H_2$, where each H_1 (respectively H_2) is the history of a maximal subcomputation with step history $(1)_2$ (respectively $(2)_2$).

As $\theta(12)_2$ locks every sector of the standard base except for those of the form $Q_0(i)P_1(i)$, any unreduced two-letter subword of B must be of the form $Q_0(i)Q_0(i)^{-1}$ or $P_1(i)^{-1}P_1(i)$. As a result, B must be a cyclic permutation of

$$Q_0(2)Q_0(2)^{-1}\dots\{t(1)\}^{-1}Q_4(L)^{-1}\dots Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)\dots Q_4(L)\{t(1)\}\dots Q_0(2)$$

Let $\mathcal{C}_2 : W_r \to \dots \to W_s$ be the subcomputation of $\mathcal{C}$ with history H_2 and let $\mathcal{C}_2' : W_r' \to \dots \to W_s'$ be its restriction to the subword

$$Q_3(L)^{-1}R_2(L)^{-1}Q_2(L)^{-1}R_1(L)^{-1}Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)R_1(L)Q_2(L)R_2(L)Q_3(L)$$

As in the proof of Lemma 6.24, we may view the admissible subwords whose bases are of the form $(P_1(L)Q_1(L)R_1(L)Q_2(L)R_2(L))^{\pm 1}$ as a single state letter. With this view, Lemma 5.14 implies $|W_r'|_a \leq \dots \leq |W_s'|_a$.

Let $\mathcal{C}_2'' : W_r'' \to \dots \to W_s''$ be the restriction of $\mathcal{C}_2$ to a subword of the form $(\{t(1)\}B_3(1))^{\pm 1}$. Then a projection argument implies $|W_r''|_a \leq \dots \leq |W_s''|_a$.

As any other sector has fixed tape word throughout $\mathcal{C}_2$, this yields $|W_r|_a \leq |W_s|_a$. So, we may assume $t > s$.

Since B contains the subword $R_2(1)Q_3(1)$, Lemma 5.28(b) implies that H must have a subword $H_2\theta(23)_2H_3$, where H_3 is the history of a maximal subcomputation with step history $(3)_2$.

The connecting rule of $\mathbf{M}_4(3^-)$ locks the Q_0P_1-sector, so that H_3 cannot contain an occurrence of a copy of this rule. So, since B contains the subword $P_1(1)Q_1(1)$, Lemma 4.3 implies that H_3 must be a suffix of H.

Let $\mathcal{C}_3 : W_{s+1} \to \dots \to W_t$ be the subcomputation with history H_3 and $\mathcal{C}_3' : W_{s+1}' \to \dots \to W_t'$ be its restriction to a subword of the form $(Q_0(1)P_1(1)Q_1(1))^{\pm 1}$. Then, Lemma 5.13 implies $|W_{s+1}'|_a \leq \dots \leq |W_t'|_a$.

Letting $\mathcal{C}_3'' : W_{s+1}'' \to \dots \to W_t''$ be the restriction of $\mathcal{C}_3$ to $Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)$, Lemma 5.14 implies $|W_{s+1}''|_a \leq \dots \leq |W_t''|_a$.

As any other sector has fixed tape word throughout $\mathcal{C}_3$, it follows that $|W_s|_a \leq |W_t|_a$, contradicting our assumption.

Hence, we may assume that the step history of $\mathcal{C}$ does not contain the letters $(12)_2$ or $(21)_2$. So, since the step history contains the letter $(1)_2$ and a letter corresponding to a transition rule, it must contain a letter of the form $(s)_2^{\pm 1}$.

As $\theta(s)_2$ locks every sector of the standard base except for those of the form $P_0(i)Q_0(i)$ for $i \geq 2$, B must be a cyclic permutation of

$$P_0(2)P_0(2)^{-1}\dots\{t(1)\}^{-1}Q_4(L)^{-1}\dots P_1(L)^{-1}Q_0(L)^{-1}Q_0(L)P_1(L)\dots Q_4(L)\{t(1)\}\dots P_0(2)$$

As B contains the subword $Q_0(L)P_1(L)$, Lemma 6.2(a) implies that the step history of $\mathcal{C}$ cannot contain a subword of the form $(s)_j(1)_j(s)_j^{-1}$. So, the step history of $\mathcal{C}$ (or its inverse) must be $(1)_2(s)_2^{-1}(s)_1(1)_1$.

Let $\mathcal{C}_1 : W_s \to \cdots \to W_t$ be the maximal subcomputation of $\mathcal{C}$ with step history $(1)_1$ and $\mathcal{C}_1' : W_s' \to \cdots \to W_t'$ be its restriction to the subword $P_1(L)^{-1}Q_0(L)^{-1}Q_0(L)P_1(L)$. Then, Lemma 5.14 implies $|W_s'|_a \leq \cdots \leq |W_t'|_a$.

Letting $\mathcal{C}_1'' : W_s'' \to \cdots \to W_t''$ be the restriciton of $\mathcal{C}_1$ to a subword $(P_0(1)Q_0(1)P_1(1))^{\pm 1}$, Lemma 5.13 (or a projection argument) implies $|W_s''|_a \leq \cdots \leq |W_t''|_a$.

As any other sector has fixed tape word throughout $\mathcal{C}_1$, $|W_s|_a \leq |W_t|_a$. Thus, we reach a final contradiction.

□

Chapter 7

Groups Associated to an S-machine and their Diagrams

7.1 The groups

As in previous literature (for example [18], [22], [25]), we now associate two finitely presented groups to a cyclic S-machine **S**. These groups are denoted $M(\mathbf{S})$ and $G(\mathbf{S})$ and 'simulate' the work of **S** in the precise sense described in Chapter 7.3.

Let **S** be a cyclic recognizing S-machine with hardware (Y, Q), where $Q = \sqcup_{i=0}^{s} Q_i$ and $Y = \sqcup_{i=1}^{s+1} Y_i$, and software the set of rules $\Theta = \Theta^+ \sqcup \Theta^-$. For notational purposes, set $Q_0 = Q_{s+1}$ and denote the accept word of **S** by W_{ac}.

For $\theta \in \Theta^+$, Lemma 4.1 allows us to assume that θ takes the form

$$\theta = [q_0 \to v_{s+1} q_0' u_1,\ q_1 \to v_1 q_1' u_2,\ \dots,\ q_{s-1} \to v_{s-1} q_{s-1}' u_s,\ q_s \to v_s q_s' u_{s+1}]$$

where $q_i, q_i' \in Q_i$, u_i and v_i are either empty or single letters in $Y_i^{\pm 1}$, and some of the arrows may take the form $\xrightarrow{\ell}$. Note that if θ locks the i-th sector, then both u_i and v_i are necessarily empty.

Define $R = \{\theta_i : \theta \in \Theta^+, 0 \leq i \leq s\}$. For notational convenience, set $\theta_{s+1} = \theta_0$ for all $\theta \in \Theta^+$.

The group $M(\mathbf{S})$ is then defined by taking the (finite) generating set $\mathcal{X} = Q \cup Y \cup R$ and subjecting it to the (finite number of) relations:

- $q_i \theta_{i+1} = \theta_i v_i q_i' u_{i+1}$ for all $\theta \in \Theta^+$ and $0 \leq i \leq s$,
- $\theta_i a = a \theta_i$ for all $0 \leq i \leq s$ and $a \in Y_i(\theta)$.

As in the language of computations of S-machines, letters from $Q^{\pm 1}$ are called *q-letters* and those from $Y^{\pm 1}$ are called *a-letters*. Additionally, those from $R^{\pm 1}$ are called *θ-letters*. The relations of the form $q_i \theta_{i+1} = \theta_i v_i q_i' u_{i+1}$ are called *(θ, q)-relations*, while those of the form $\theta_i a = a \theta_i$ are called (θ, a)-*relations*.

Note that the number of a-letters in any part of θ, and so in any defining relation of $M(\mathbf{S})$, is at most two.

To simplify these relations, it is convenient to omit reference to the indices of the letters of R. This notational quirk may make it appear as though θ commutes with the letters of $Y_i(\theta)$ and conjugates q_i to $v_i q_i' u_{i+1}$ for each i; it should be noted that these statements are not strictly true. Further, it is useful to note that if θ locks the i-th sector, then $Y_i(\theta) = \emptyset$ so that θ has no relation with the elements of Y_i.

However, this group evidently lacks any reference to the accept configuration. To amend this, the group $G(\mathbf{S})$ is constructed by adding one more relation to the presentation of $M(\mathbf{S})$, namely the *hub-relation* $W_{ac} = 1$. In other words, $G(\mathbf{S}) \cong M(\mathbf{S})/\langle\langle W_{ac}\rangle\rangle$.

Moreover, it is useful for the purposes of our construction to consider extra relations, called a-*relations*, within the language of tape letters. If Ω is the set of relators defining these a-relations, then we denote the groups arising from the addition of a-relations by $M_\Omega(\mathbf{S})$ and $G_\Omega(\mathbf{S})$. Note that $M_\Omega(\mathbf{S}) \cong M(\mathbf{S})/\langle\langle\Omega\rangle\rangle$ and $G_\Omega(\mathbf{S}) \cong G(\mathbf{S})/\langle\langle\Omega\rangle\rangle$.

It is henceforth taken as an assumption that any a-relation adjoined to the groups associated to the machine **M** correspond to words over the alphabet of the 'special' input sector.

For the purposes of Chapter 12 and the proof of Theorem A, the set of a-relators $\mathcal{S}$ is taken to be exactly the words that represent the trivial element in $B(\mathcal{A}, n)$, where the tape alphabet of the sector is identified with $\mathcal{A}$. However, in the proof of Theorem C presented in Chapter 14, the a-relators are taken to be a larger set of words. So, for the sake of generality, until Section 11, the set of a-relators Ω is taken to be some set of words over $\mathcal{A}$ containing $\mathcal{S}$ as a subset.

Note that though they remain finitely generated, $M_\Omega(\mathbf{S})$ and $G_\Omega(\mathbf{S})$ may no longer be finitely presented. In fact, in all situations encountered in what follows, $M_\Omega(\mathbf{M})$ and $G_\Omega(\mathbf{M})$ are not finitely presented.

7.2 Bands and annuli

Many of the arguments presented in the forthcoming chapters rely on van Kampen diagrams (see Chapter 2) over the presentations of the groups introduced in Section 7.1. To present these arguments efficiently, we first differentiate between the types of edges and cells that arise in such diagrams in a way similar to that employed in [18] and [25].

For simplicity, we will often disregard the presence of 0-cells in these diagrams. For example, we do not differentiate between adjacent edges, so that any edge not on the boundary of a diagram is on the boundary of two $\mathcal{R}$-cells (for $\mathcal{R}$ the defining relators of the corresponding group). Additionally, we will adopt the convention that the contour of any diagram, subdiagram, or cell is traced in the counterclockwise direction.

An edge labelled by a q-letter is called a *q-edge*. Similarly, an edge labelled by an a-letter is called an *a-edge* and one labelled by a θ-letter is a *θ-edge*.

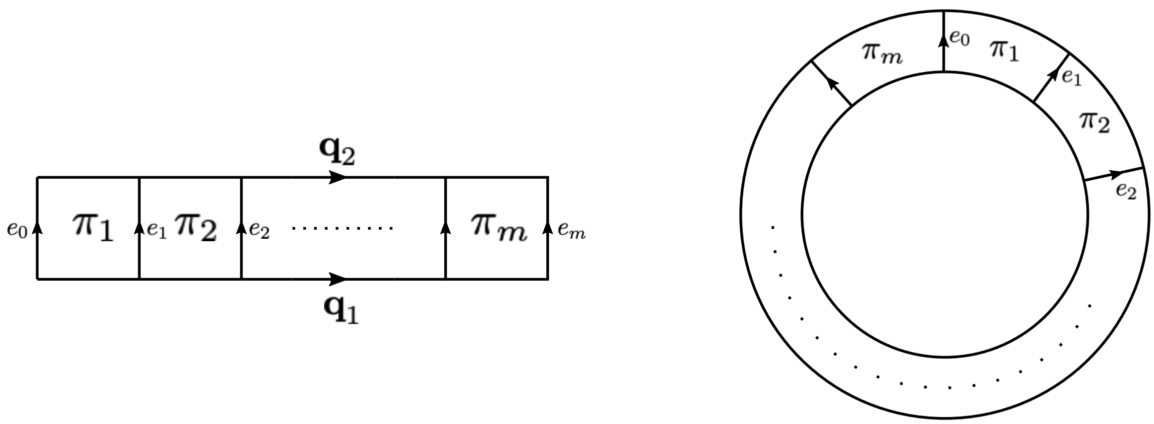
For a path $\mathbf{p}$ in Δ, the (combinatorial) length of $\mathbf{p}$ is denoted $\|\mathbf{p}\|$. Further, the path's *a-length* $|\mathbf{p}|_a$ is the number of a-edges in the path. The path's *θ-length* and *q-length*, denoted $|\mathbf{p}|_\theta$ and $|\mathbf{p}|_q$, respectively, are defined similarly.

A cell whose contour label corresponds to a (θ, q)-relation is called a *(θ, q)-cell*. Similarly, there are *(θ, a)-cells*, *a-cells*, and *hubs*.

In the general setting of a reduced diagram Δ over a presentation $\langle X \mid \mathcal{R}\rangle$, let $\mathcal{Z} \subseteq X$. For $m \geq 1$, a sequence of (distinct) cells $\mathcal{B} = (\Pi_1, \dots, \Pi_m)$ in Δ is called a *$\mathcal{Z}$-band* of length m if:

- every two consecutive cells Π_i and Π_{i+1} have a common boundary edge $\mathbf{e}_i$ labeled by a letter from $\mathcal{Z}^{\pm 1}$ and
- for every i, $\partial\Pi_i$ has exactly two edges labelled by a letter from $\mathcal{Z}^{\pm 1}$, $\mathbf{e}_{i-1}^{-1}$ and $\mathbf{e}_i$, and $\text{Lab}(\mathbf{e}_{i-1})$ and $\text{Lab}(\mathbf{e}_i)$ are either both positive or both negative.

For convenience, we extend this definition by saying that any edge labelled by a letter of $\mathcal{Z}^{\pm 1}$ is a $\mathcal{Z}$-band of length zero.



(a) Non-annular $\mathcal{Z}$-band of length m

(b) Annular $\mathcal{Z}$-band of length m

Figure 7.1:

A $\mathcal{Z}$-band is *maximal* if it is not contained in any other $\mathcal{Z}$-band. Note that every edge labelled by a letter of $\mathcal{Z}^{\pm 1}$ is contained in a maximal $\mathcal{Z}$-band.

In a $\mathcal{Z}$-band $\mathcal{B}$ of length $m \geq 1$ made up of the cells $(\Pi_1, \dots, \Pi_m)$, using only edges from the contours of $\Pi_1, \dots, \Pi_m$, there exists a closed path $\mathbf{e}_0^{-1}\mathbf{q}_1\mathbf{e}_m\mathbf{q}_2^{-1}$ such that $\mathbf{q}_1$ and $\mathbf{q}_2$ are simple (perhaps closed) paths. In this case, $\mathbf{q}_1$ is called the bottom of $\mathcal{B}$, denoted $\mathbf{bot}(\mathcal{B})$, while $\mathbf{q}_2$ is called the *top* of $\mathcal{B}$ and denoted $\mathbf{top}(\mathcal{B})$. When $\mathbf{q}_1$ and $\mathbf{q}_2$ need not be distinguished, they are called the *sides* of the band.

If $\mathbf{e}_0 = \mathbf{e}_m$ in a $\mathcal{Z}$-band $\mathcal{B}$ of length $m \geq 1$, then $\mathcal{B}$ is called a *$\mathcal{Z}$-annulus*. If $\mathcal{B}$ is a non-annular $\mathcal{Z}$-band of length $m \geq 1$, then $\mathbf{e}_0^{-1}\mathbf{q}_1\mathbf{e}_m\mathbf{q}_2^{-1}$ is called the *standard factorization* of the contour of $\mathcal{B}$. If either $(\mathbf{e}_0^{-1}\mathbf{q}_1\mathbf{e}_m)^{\pm 1}$ or $(\mathbf{e}_m\mathbf{q}_2^{-1}\mathbf{e}_0^{-1})^{\pm 1}$ is a subpath of $\partial\Delta$, then $\mathcal{B}$ is called a *rim $\mathcal{Z}$-band*.

A $\mathcal{Z}_1$-band and a $\mathcal{Z}_2$-band *cross* if they have a common cell and $\mathcal{Z}_1 \cap \mathcal{Z}_2 = \emptyset$.

In particular, in a reduced diagram over the canonical presentations of the groups of interest, there exist *q-bands* corresponding to bands arising from $\mathcal{Z} = Q_i^{\pm 1}$ for some i, where every cell is a (θ, q)-cell. Similarly, there exist *θ-bands* for $\theta \in \Theta^+$ and *a-bands* for $a \in Y$. However, it is useful to restrict the definition of an a-band so that they consist only of (θ, a)-cells.

Note that by definition, distinct maximal q-bands (θ-bands, a-bands) cannot intersect.

Given an a-band $\mathcal{B}$, the makeup of the groups' relations dictates that the defining a-edges $\mathbf{e}_0, \dots, \mathbf{e}_m$ are labelled identically. Similarly, the θ-edges of a θ-band correspond to the same rule;

however, the (suppressed) index of two such θ-edges may differ.

If a maximal a-band contains a cell with an a-edge that is also on the contour of a (θ, q)-cell, then the a-band is said to *end* (or *start*) on that (θ, q)-cell and the corresponding a-edge is said to be the *end* (or *start*) of the band. This definition extends similarly, so that:

- a maximal a-band can end on a (θ, q)-cell, on an a-cell, or on the diagram's contour,
- a maximal θ-band can end only on the diagram's contour, and
- a maximal q-band can end on a hub or on the diagram's contour.

Note that if a maximal θ-band (a-band, q-band) ends as above in one part of the diagram, then it must also end in another part of the diagram as it cannot be a θ-annulus (a-annulus, q-annulus).

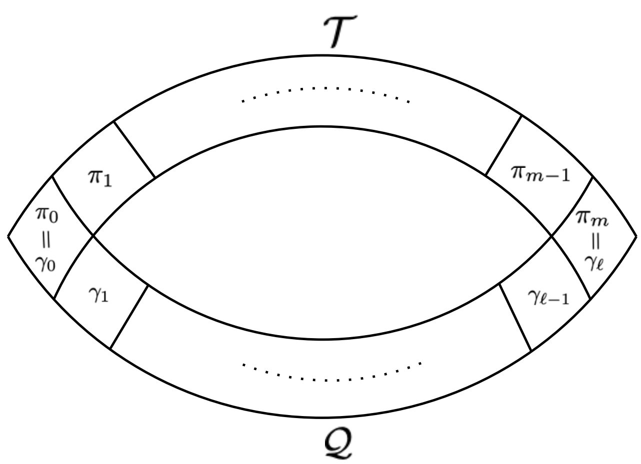


Figure 7.2: (θ, q)-annulus with defining θ-band $\mathcal{T}$ and q-band $\mathcal{Q}$

The natural projection of the label of the top (or bottom) of a q-band onto $F(\Theta^+)$ is called the *history* of the band; the *step history* of the band is then defined in the obvious way. The natural projection (without reduction) of the top (or bottom) of a θ-band onto the alphabet $\{Q_0, \dots, Q_s\}$ is called the *base* of the band.

Let $\mathcal{T}$ be a maximal θ-band in a reduced diagram Δ over $G_\Omega(\mathbf{M})$ with two ends on $\partial\Delta$. Suppose that any cell between one side of $\mathcal{T}$ and $\partial\Delta$ is an a-cell. Then $\mathcal{T}$ is called a *quasi-rim θ-band*. Note that a rim θ-band is a quasi-rim θ-band.

Suppose the sequence of cells $(\pi_0, \pi_1, \dots, \pi_m)$ comprises a θ-band and $(\gamma_0, \gamma_1, \dots, \gamma_\ell)$ a q-band such that $\pi_0 = \gamma_0$, $\pi_m = \gamma_\ell$, and no other cells are shared. Suppose further that $\partial\pi_0$ and $\partial\pi_m$ both contain edges on the outer countour of the annulus bounded by the two bands. Then the union of these two bands is called a (θ, q)-*annulus* and π_0 and π_m are called its *corner* cells. A (θ, a)-*annulus* is defined similarly.

The following statement is proved in a more general setting in [17]:

Lemma 7.1. *(Lemma 6.1 of [17]) A reduced diagram over $M(\mathbf{S})$ contains no:*

(1) (θ, q)-annuli

(2) (θ, a)-annuli

(3) a-annuli

(4) q-annuli

(5) θ-annuli

As a result, in a reduced diagram Δ over $M(\mathbf{S})$, if a maximal θ-band and a maximal q-band (respectively a-band) cross, then their intersection is exactly one (θ, q)-cell (respectively (θ, a)-cell). Further, every maximal θ-band and maximal q-band ends on $\partial\Delta$ in two places.

7.3 Trapezia

Let Δ be a reduced diagram over the canonical presentation of $M(\mathbf{S})$ whose contour is 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 $\mathbf{q}_1$ and $\mathbf{q}_2$ are maximal parts of the sides of θ-bands whose labels start and end with q-letters. Then Δ is called a *trapezium*.

In this case, $\mathbf{q}_1$ and $\mathbf{q}_2$ are called the *bottom* and *top* of the trapezium, respectively, while $\mathbf{p}_1$ and $\mathbf{p}_2$ are the *left* and *right* sides. Further, $\mathbf{p}_1^{-1}\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2^{-1}$ is called the *standard factorization* of the contour.

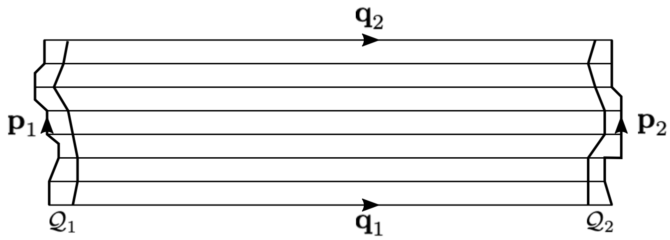


Figure 7.3: Trapezium with side q-bands $\mathcal{Q}_1$ and $\mathcal{Q}_2$

The *(step) history* of the trapezium is the (step) history of the rim q-band with $\mathbf{p}_2$ as one of its sides and the length of this history is the trapezium's *height*. The base of Lab($\mathbf{q}_1$) is called the *base* of the trapezium.

It's easy to see from this definition that a θ-band $\mathcal{T}$ whose first and last cells are (θ, q)-cells can be viewed as a trapezium of height 1 as long as its top and bottom start and end with q-edges. We extend this to all such θ-bands by merely disregarding any a-edges of the top and bottom that precede the first q-edge or follow the final q-edge. The paths formed by disregarding these edges are called the *trimmed* top and bottom of the band and are denoted **ttop**($\mathcal{T}$) and **tbot**($\mathcal{T}$).

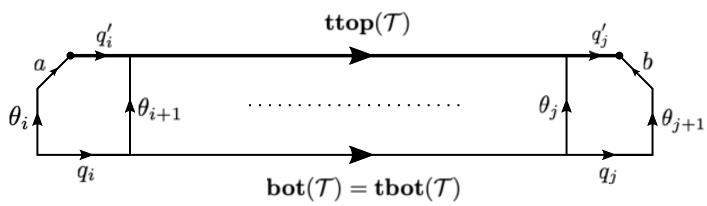


Figure 7.4: θ-band $\mathcal{T}$ with trimmed top

Lemma 7.2. *Let $\mathcal{T}$ be a θ-band in a reduced diagram Δ over the canonical presentation of $M(\boldsymbol{S})$ whose first and last cells are (θ, q)-cells. Then Lab(**tbot**($\mathcal{T}$)) and Lab(**ttop**($\mathcal{T}$)) are admissible words. Moreover, for θ the rule corresponding to the band $\mathcal{T}$, Lab(**tbot**($\mathcal{T}$)) is θ-admissible and Lab(**tbot**($\mathcal{T}$)) $\cdot\, \theta \equiv$ Lab(**ttop**($\mathcal{T}$)).*

Proof. Suppose $\theta \in \Theta^+$.

Further, suppose $\mathcal{T}$ consists of one (θ, q)-cell π. Then **top**($\mathcal{T}$) and **bot**($\mathcal{T}$) contain just one q-edge, which is a part of $\partial\pi$ (or its inverse). So, **ttop**($\mathcal{T}$) and **tbot**($\mathcal{T}$) each consist of this one q-edge. It follows from the definition of (θ, q)-relations that Lab(**tbot**($\mathcal{T}$)) $\cdot\, \theta \equiv$ Lab(**ttop**($\mathcal{T}$)).

Now suppose $\mathcal{T}$ contains at least two (θ, q)-cells and let $\mathbf{e}_1, \mathbf{e}_2$ be the first two q-edges of $\mathbf{bot}(\mathcal{T})$ with $q_1 = \text{Lab}(\mathbf{e}_1)$ and $q_2 = \text{Lab}(\mathbf{e}_2)$. So, $\text{Lab}(\mathbf{tbot}(\mathcal{T}))$ has prefix $q_1 w q_2$ for some $w \in F(Y)$. For $j = 1, 2$, let π_j be the (θ, q)-cell of $\mathcal{T}$ so that $\mathbf{e}_j$ is an edge of $\partial\pi_j$.

For $0 \leq i \leq s$, suppose $q_1 \in Q_i$. Then the i-th part of θ must be $q_1 \to u_i q_1' v_{i+1}$ for some $q_1' \in Q_i$, $u_i \in F(Y_i(\theta))$, and $v_{i+1} \in F(Y_{i+1}(\theta))$ with $\|u_i\|, \|v_{i+1}\| \leq 1$. So, we have $\text{Lab}(\partial\pi_1) \equiv \theta_i^{-1} q_1 \theta_{i+1} v_{i+1}^{-1} (q_1')^{-1} u_i^{-1}$. If there exists any cell of $\mathcal{T}$ between π_1 and π_2, it must be a (θ, a)-cell with an edge labelled by θ_{i+1} on its contour. Hence, $w \in F(Y_{i+1}(\theta))$.

What's more, the label of $\partial\pi_2$ must have a subword $\theta_{i+1}^{-1} q_2$. By the definition of the (θ, q)-relations, this means one of two things:

(a) $q_2 \in Q_{i+1}$ and the $(i+1)$-th part of θ is $q_2 \to u_{i+1} q_2' v_{i+2}$ for some $q_2' \in Q_{i+1}$, $u_{i+1} \in F(Y_{i+1}(\theta))$, and $v_{i+2} \in F(Y_{i+2}(\theta))$ with $\|u_{i+1}\|, \|v_{i+2}\| \leq 1$; or

(b) $q_2 = q_1^{-1}$

In case (a), the subword $q_1 w q_2$ of $\text{Lab}(\mathbf{tbot}(\mathcal{T}))$ satisfies condition (1) in the requirements for subwords of admissible words (see Section 4.1).

In case (b), it satisfies condition (2) as long as there is some (θ, a)-cell between them; but this is required in the band, as otherwise π_1 and π_2 would be a pair of cancellable cells.

Let $\mathcal{T}_1 = (\pi_1, \dots, \pi_2)$ be the corresponding subband of $\mathcal{T}$. Then $\text{Lab}(\mathbf{tbot}(\mathcal{T}_1)) \equiv q_1 w q_2$. The above arguments make it clear that $q_1 w q_2$ is θ-admissible. Further, it is easy to see that $\text{Lab}(\mathbf{ttop}(\mathcal{T}_1)) \equiv (q_1 w q_2) \cdot \theta$.

If $q_1 \in Q_i^{-1}$, then an analogous argument yields the same conclusion.

If $\mathbf{tbot}(\mathcal{T})$ has more than two q-edges, then the argument above can be iterated to apply to the whole band, implying the statement.

Conversely, if $\theta \in \Theta^-$, then the analogous arguments apply to $\mathbf{ttop}(\mathcal{T})$ to show that $\text{Lab}(\mathbf{ttop}(\mathcal{T}))$ is θ^{-1}-admissible with $\text{Lab}(\mathbf{ttop}(\mathcal{T})) \cdot \theta^{-1} \equiv \text{Lab}(\mathbf{tbot}(\mathcal{T}))$.

But then $\text{Lab}(\mathbf{tbot}(\mathcal{T}))$ is θ-admissible with $\text{Lab}(\mathbf{tbot}(\mathcal{T})) \cdot \theta \equiv \text{Lab}(\mathbf{ttop}(\mathcal{T}))$.

$\square$

Lemma 7.3. *Let $U \to V$ be a computation of* **S** *with history H of length 1, so that $H = \theta \in \Theta$. Then there exists a θ-band $\mathcal{T}$ corresponding to the rule θ whose first and last cells are (θ, q)-cells such that Lab(**tbot**($\mathcal{T}$)) $\equiv U$ and Lab(**ttop**($\mathcal{T}$)) $\equiv V$.*

Proof. Suppose $\theta \in \Theta^+$ and set $U \equiv q_0^{\varepsilon_0} w_1 q_1^{\varepsilon_1} \dots w_\ell q_\ell^{\varepsilon_\ell}$ so that for each $0 \leq i \leq \ell$, $q_i \in Q_{j(i)}$ for some $0 \leq j(i) \leq s$ and $\varepsilon_i \in \{\pm 1\}$.

Then $q_i \in Q(\theta)$ for each $0 \leq i \leq \ell$, so that the $j(i)$-th part of θ takes the form $q_i \to u_{j(i)} q_i' v_{j(i)+1}$ for some $q_i' \in Q_{j(i)}$, $u_{j(i)} \in F(Y_{j(i)}(\theta))$, and $v_{j(i)+1} \in F(Y_{j(i)+1}(\theta))$. So, there are relations of $M(\mathbf{S})$ of the form $R_i = \theta_{j(i)}^{-1} q_i \theta_{j(i)+1} (u_{j(i)} q_i' v_{j(i)+1})^{-1}$ for all i.

If $\varepsilon_i = 1$, then each letter of w_{i+1} is an element of $Y_{j(i)+1}(\theta)$ since U is θ-admissible, so that there are (θ, a)-relations of the form $\theta_{j(i)+1}^{-1} a \theta_{j(i)+1} a^{-1}$ for each letter a of w_{i+1}. Gluing along the edges labelled by $\theta_{j(i)+1}^{\pm 1}$, one can construct a θ-band $\mathcal{T}_{i+1}$ with contour label $\theta_{j(i)+1}^{-1} w_{i+1} \theta_{j(i)+1} w_{i+1}^{-1}$.

If $\varepsilon_i = -1$, then each letter of of w_{i+1} is in $Y_{j(i)}(\theta)$ since U is θ-admissible. So, there are relations of $M(\mathbf{S})$ of the form $\theta_{j(i)}^{-1} a \theta_{j(i)} a^{-1}$ for each letter a of w_{i+1}. Then, gluing along the edges labelled by $\theta_{j(i)}^{\pm 1}$, one can construct a θ-band $\mathcal{T}_{i+1}$ with contour label $\theta_{j(i)}^{-1} w_{i+1} \theta_{j(i)} w_{i+1}^{-1}$.

Now, let π_i be a cell with boundary labelled by $R_i^{\varepsilon_i}$. For either possibility of ε_i, one can glue $\mathcal{T}_i$ and $\mathcal{T}_{i+1}$ to the left and right of π_i, respectively.

After 0-refinement (or gluing) to cancel any adjacent edges with mutually inverse labels, this process produces a θ-band $\mathcal{T}$ corresponding to the rule θ with Lab(**tbot**($\mathcal{T}$)) $\equiv U$. By the makeup of the band, it is easy to see that Lab(**ttop**($\mathcal{T}$)) $\equiv V$.

If $\theta \in \Theta^-$, then the same construction forms a θ-band $\mathcal{T}$ corresponding to the rule θ^{-1} with Lab(**tbot**($\mathcal{T}$)) $\equiv V$ and Lab(**ttop**($\mathcal{T}$)) $\equiv U$. Taking the 'inverse' of this band (i.e inverting the label of each cell) produces a θ-band corresponding to θ as in the statement.

□

By Lemma 7.1, any trapezium Δ of height $h \geq 1$ can be decomposed into θ-bands $\mathcal{T}_1, \dots, \mathcal{T}_h$ connecting the left and right sides of the trapezium, with **bot**($\mathcal{T}_1$) and **top**($\mathcal{T}_h$) making up the bottom and top of Δ, respectively. Moreover, the first and last cells of each $\mathcal{T}_i$ are (θ, q)-cells and **ttop**($\mathcal{T}_i$) = **tbot**($\mathcal{T}_{i+1}$) for all $1 \leq i \leq h-1$.

The following two statements are clear from the previous two and exemplify how the group $M(\mathbf{S})$ simulates the work of the S-machine:

Lemma 7.4. *Let Δ be a trapezium with history $H \equiv \theta_1 \dots \theta_h$ for $h \geq 1$ and maximal θ-bands $\mathcal{T}_1, \dots, \mathcal{T}_h$ enumerated from bottom to top. If $U_j \equiv$ Lab($\textbf{tbot}(\mathcal{T}_j)$) and $V_j \equiv$ Lab($\textbf{ttop}(\mathcal{T}_j)$) for all j, then H is a reduced word, U_j and V_j are admissible words, and $V_j \equiv U_j \cdot \theta_j$ for all j.*

Lemma 7.5. *For any reduced computation $U \to \cdots \to U \cdot H \equiv V$ of the S-machine **S** with $\|H\| \geq 1$, there exists a trapezium Δ with trimmed bottom label U, trimmed top label V, and history H.*

Chapter 8

Modified length and area functions

8.1 Modified length function

To assist with the proofs to come, we now modify the length function on words over the groups associated to an S-machine and paths in diagrams over their presentations. This is done in the same way as in [18] and [25]. The standard length of a word/path will henceforth be referred to as its *combinatorial length* and the modified length simply as its *length*.

Define a word consisting of no q-letters, one θ-letter, and one a-letter as a (θ, a)-*syllable*. Then, define the length of:

- any q-letter as 1
- any θ-letter as 1
- any a-letter as the parameter δ (as indicated in Chapter 4.3, this should be thought of as a very small positive number)
- any (θ, a)-syllable as 1

For a word w over the generators of the canonical presentation of $G_\Omega(\mathbf{S})$ (or any group associated to $\mathbf{S}$), define a *decomposition* of w as a factorization of w into a product of letters and (θ, a)-syllables. The length of a decomposition of w is then defined to be the sum of the lengths of the factors.

Finally, the length of w, denoted $|w|$, is defined to be the minimum of the lengths of its decompositions.

The length of a path in a diagram over the presentations of the groups associated to $\mathbf{S}$ is defined to be the length of its label.

The following gives some basic properties of the length function. Its proof is a consequence of Lemma 4.1.

Lemma 8.1. *(Lemma 6.2 of [25]) Let* $\mathbf{s}$ *be a path in a diagram* Δ *over the canonical presentation of* $G_\Omega(\mathbf{S})$ *(or any of the groups associated to* $\mathbf{S}$*) consisting of* c θ*-edges and* d a*-edges. Then:*

(a) $|\mathbf{s}| \geq \max(c, c + (d - c)\delta)$

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

(c) For any product $\mathbf{s} = \mathbf{s}_1\mathbf{s}_2$ *of two paths in a diagram,* $|\mathbf{s}_1| + |\mathbf{s}_2| - \delta \leq |\mathbf{s}| \leq |\mathbf{s}_1| + |\mathbf{s}_2|$*.*

(d) Let $\mathcal{T}$ *be a* θ*-band with base of length* l_b*. If* ***top***$(\mathcal{T})$ *(or* ***bot***$(\mathcal{T})$*) has* l_a a*-edges, then the number of cells in* $\mathcal{T}$ *is between* $l_a - l_b$ *and* $l_a + 3l_b$*.*

8.2 Disks and weights

Next, we add extra relations to the groups $G(\mathbf{S})$ and $G_\Omega(\mathbf{S})$ that will aid with later estimates. This is done in the same way as in [18] and [25] (though no group $G_\Omega(\mathbf{S})$ was present in those sources).

These relations, called *disk relations*, are of the form $W = 1$ for any configuration W accepted by the machine $\mathbf{S}$.

Lemma 8.2. *If the configuration* W *is accepted by the machine* $\mathbf{S}$ *and* $Y_{s+1} = \emptyset$*, then the word* W *is trivial over the groups* $G(\mathbf{S})$ *and* $G_\Omega(\mathbf{S})$*.*

Proof. Let $\mathcal{C}$ be an accepting computation of W and H be its history. By Lemma 7.5, there exists a trapezium Δ corresponding to $\mathcal{C}$ with trimmed bottom label W and trimmed top label W_{ac}.

As this is a computation of the standard base and every rule locks the Q_sQ_0-sector, one can further assume that no trimming was necessary in Δ, i.e the labels of the bottom and top of Δ are W and W_{ac}, respectively. Finally, it follows that the sides of the trapezium are labelled identically; specifically, they are labelled by the copy of H obtained by adding the index 0 to each letter.

So, W and W_{ac} are conjugate in $M(\mathbf{S})$. Taking into account the hub relation in both $G(\mathbf{S})$ and $G_\Omega(\mathbf{S})$ then implies the relation $W = 1$.

□

As a result of Lemma 8.2, the presentation obtained by adding the disk relations to the group $G(\mathbf{S})$ (respectively $G_\Omega(\mathbf{S})$) defines a group isomorphic to $G(\mathbf{S})$ (respectively $G_\Omega(\mathbf{S})$). The presentation containing disk relations will be referred to in what follows as the *disk presentation* of the group $G(\mathbf{S})$ (respectively $G_\Omega(\mathbf{S})$). A cell of a diagram over the disk presentation corresponding to a disk relation (or its inverse) is referred to simply as a *disk*.

One should note the following when considering diagrams over a disk presentation rather than diagrams over a canonical presentation:

- The disk presentation of $G(\mathbf{S})$ or of $G_\Omega(\mathbf{S})$ need not be finite. In particular, there may be infinitely many disk relations in this presentation. In particular, the disk presentations of $G(\mathbf{M})$ and of $G_\Omega(\mathbf{M})$ are not finitely presented.

- For a word $w \in F(\mathcal{X})$ that represents the trivial element of $G(\mathbf{S})$, the minimal area of diagrams over the disk presentation with contour label w can be drastically different than that of diagrams over the canonical presentation of $G(\mathbf{S})$.

- As in Section 7.2, we insist that an a-band in a diagram over the disk presentation of $G_\Omega(\mathbf{S})$ consist only of (θ, a)-cells. As a consequence, a maximal a-band may end on a disk in addition to the other possibilities outlined in Section 7.2.

Similar to how we modified the length function in Section 8.1, we now alter the definition of the area of a diagram over the disk presentations of $G(\mathbf{S})$ and $G_\Omega(\mathbf{S})$.

We do this first by introducing a weight function on the cells of such diagrams, wt, defined by:

- $\text{wt}(\Pi) = 1$ if Π is a (θ, q)-cell or a (θ, a)-cell
- $\text{wt}(\Pi) = C_1|\partial\Pi|^2$ if Π is a disk
- $\text{wt}(\Pi) = C_1\|\partial\Pi\|^2$ if Π is an a-cell

Naturally, we extend this to define the weight of a reduced diagram Δ, $\text{wt}(\Delta)$, as the sum of the weights of its cells.

8.3 Mixtures

We now recall an invariant of reduced diagrams over the relevant presentations, first introduced in [24], that will prove useful in future numerical estimates.

Let O be a circle containing a finite two-colored set of points, with the two colors taken to be black and white. The circle O is called a *necklace* while the corresponding points are called *white beads* and *black beads*.

Let P_j be the set of ordered pairs of distinct white beads, (o_1, o_2), such that the counterclockwise simple arc on O from o_1 to o_2 contains at least j black beads.

Define $\mu_J(O) = \sum_{j=1}^{J} \#P_j$ as the J-*mixture* of O, where J is the parameter specified in Section 4.3. The following statement then gives some basic properties of the J-mixture. The proof is straightforward.

Lemma 8.3. *(Lemma 6.1 of [24]) Let O be a necklace with x white beads and y black beads.*

(a) $\mu_J(O) \leq J(x^2 - x)$

(b) If O' is a necklace obtained from O through the removal of one white bead, then for every j, $\#P_j - 2x < \#P'_j \leq \#P_j$, and so $\mu_J(O) - 2Jx < \mu_J(O') \leq \mu_J(O)$

(c) If O' is a necklace obtained from O through the removal of one black bead, then for every j, $\#P'_j \leq \#P_j$, and so $\mu_J(O') \leq \mu_J(O)$

(d) Suppose v_1, v_2, v_3 are three black beads on O such that the counterclockwise arc from v_1 to v_3, $v_1 - v_3$, has at most J black beads (excluding v_1 and v_3). Let y_1 and y_2 be the number of white beads on the counterclockwise arcs $v_1 - v_2$ and $v_2 - v_3$, respectively. If O' is the necklace obtained from O through the removal of v_2, then $\mu_J(O') \leq \mu_J(O) - y_1 y_2$.

Let Δ be a reduced diagram over a group associated to an S-machine **S**. Let O be a circle partitioned by subarcs labeled by the edges of $\partial\Delta$. At the midpoint of a subarc labeled by a θ-edge (respectively a q-edge), place a white bead (respectively a black bead). Then, define the *mixture on* Δ $\mu(\Delta)$ as the J-mixture of the corresponding necklace, i.e $\mu(\Delta) = \mu_J(O)$.

Chapter 9

Diagrams without disks

In this chapter, we study diagrams over $M_\Omega(\mathbf{M})$, with the ultimate goal of bounding the 'size' of such a diagram in terms of its perimeter. To do this, we first define a special class of diagrams for which this bound will hold.

9.1 M-minimal diagrams

A reduced diagram Δ over the canonical presentation of $M_\Omega(\mathbf{M})$ is called M*-minimal* if:

(MM1) for any a-cell π and any θ-band $\mathcal{T}$, at most half of the edges of $\partial\pi$ mark the start of an a-band that crosses $\mathcal{T}$, and

(MM2) no maximal a-band ends on two different a-cells.

It follows immediately from this definition that a subdiagram of an M-minimal diagram is M-minimal.

9.2 Annuli

Our first step is to rule out the possible existence of certain types of subdiagrams in an M-minimal diagram.

Lemma 9.1. *A reduced diagram Δ over $G_\Omega(\boldsymbol{M})$ contains no:*

(1) (θ, q)-annuli

(2) (θ, a)-annuli

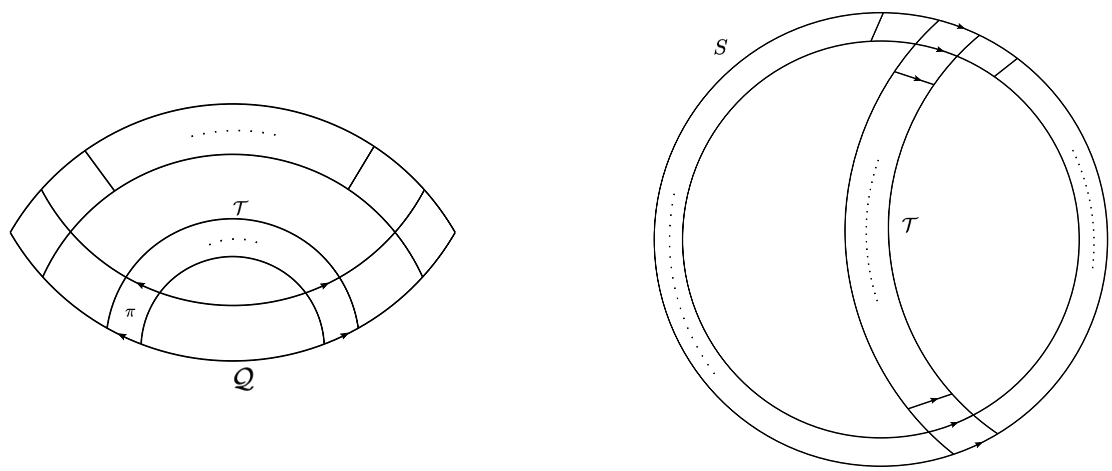
(3) a-annuli

(4) q-annuli

Proof. (1) Suppose Δ contains a (θ, q)-annulus S. Let Δ_S be the subdiagram bounded by the outer component of the contour of S and $\mathcal{Q}$ be the defining q-band (see Figure 9.1(a)).

By the definition of the annulus, the history H of $\mathcal{Q}$ must be of the form $\theta w \theta^{-1}$ for some rule $\theta \in \Theta$ and some word $w \in F(\Theta^+)$. Since H must be reduced, w cannot be trivial, and so $\mathcal{Q}$ must contain a (θ, q)-cell π with neither q-edge on $\partial\Delta_S$.

Note that each cell of $\mathcal{Q}$ has a θ-edge on its boundary that is shared with $\partial\Delta_S$. Indeed, all θ-edges of $\partial\Delta_S$ arise in this way.



(a) Δ_S for S a (θ, q)-annulus

(b) Δ_S for S an a-annulus

Figure 9.1:

Let $\mathcal{T}$ be the maximal θ-band in Δ_S containing π, so that $\mathcal{T}$ starts at the θ-edge of $\partial\pi$ shared with $\partial\Delta_S$. Then $\mathcal{T}$ must also end on a θ-edge of $\partial\Delta_S$, and so defines a (θ, q)-annulus S' with some subband of $\mathcal{Q}$. Note that the history of the q-band defining S' is a proper subword of w.

Iterating, there exists a θ-band that starts and ends on the boundary of adjacent cells of $\mathcal{Q}$. But then these two cells are cancellable, contradicting the assumption that Δ is reduced.

(2) is proved by an identical argument to (1).

(3) Suppose Δ contains an a-annulus S and let Δ_S be the subdiagram bounded by the outer component of the contour of S (see Figure 9.1(b)).

Recall that each cell comprising S is a (θ, a)-cell. By the definition of (θ, a)-relations, each of these cells must have two θ-edges on its boundary, one of which is shared with $\partial\Delta_S$. The maximal

θ-band $\mathcal{T}$ starting at such an edge must also end on $\partial\Delta_S$, i.e at a θ-edge on the boundary of another cell of S.

But then $\mathcal{T}$ and a subband of S form a (θ, a)-annulus, contradicting (2).

(4) As each cell comprising a q-annulus must be a (θ, q)-cell, an identical argument to (3) produces a (θ, q)-annulus which contradicts (1).

□

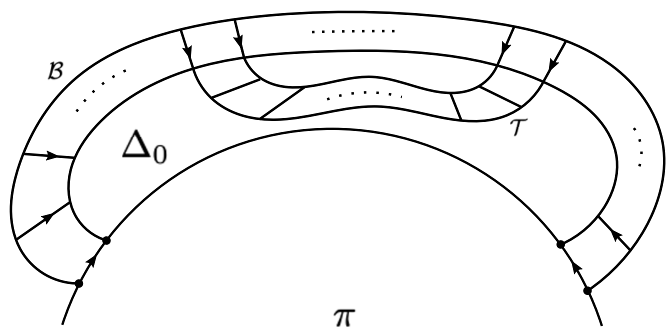


Figure 9.2: a-band ending twice on an a-cell

Lemma 9.2. *For any a-cell π in a reduced diagram Δ over $G_\Omega(\boldsymbol{M})$, no a-band can have two ends on π.*

Proof. Suppose $\mathcal{B}$ is an a-band ending twice on π. As $\text{Lab}(\partial\pi)$ is a reduced word, $\mathcal{B}$ must have nonzero length.

Consider the subdiagram Δ_0 of Δ bounded by a side of $\mathcal{B}$ and the subpath of $\partial\pi$ whose initial and terminal edges correspond to the two ends of $\mathcal{B}$ (see Figure 9.2).

Since each cell of $\mathcal{B}$ is a (θ, a)-cell, the portion of $\partial\Delta_0$ coinciding with a side of $\mathcal{B}$ is comprised entirely of θ-edges. Moreover, as $\partial\pi$ is comprised entirely of a-edges, any θ-edge of $\partial\Delta_0$ must lie on this side of $\mathcal{B}$.

So, a maximal θ-band $\mathcal{T}$ of Δ_0 starting on the side of $\mathcal{B}$ must also end on this side. But then $\mathcal{T}$ and a subband of $\mathcal{B}$ form a (θ, a)-annulus in Δ, contradicting Lemma 9.1(2).

□

Lemma 9.3. *Let Δ be a reduced diagram over $M_\Omega(\boldsymbol{M})$.*

(1) *Suppose Δ contains a θ-annulus S and let Δ_S be the subdiagram of Δ bounded by the outer component of the contour of S. Then Δ_S contains no (θ, q)-cells and Lab$(\partial\Delta_S)$ is a word over the tape alphabet of the 'special' input sector.*

(2) *If Δ is M-minimal, then it contains no θ-annuli.*

Proof. (1) Suppose Δ_S contains a (θ, q)-cell and let $\mathcal{Q}$ be the maximal q-band containing this cell. Lemma 9.1(4) then implies that $\mathcal{Q}$ must have two ends on $\partial\Delta_S$. But then $\mathcal{Q}$ and a subband of S define a (θ, q)-annulus in Δ, contradicting Lemma 9.1(1). Hence, every cell of S is a (θ, a)-cell, so that $\partial\Delta_S$ consists entirely of a-edges.

Similarly, if a maximal a-band $\mathcal{B}$ of Δ_S has both ends on $\partial\Delta_S$, then $\mathcal{B}$ and a subband of S define a (θ, a)-annulus in Δ, contradicting Lemma 9.1(2). So, every edge of $\partial\Delta_S$ is an a-edge marking the start of a maximal a-band in Δ_S which must end on an a-cell.

Thus, as the boundary of an a-cell is labelled by tape letters from the 'special' input sector and the a-edges of an a-band are labelled identically, the proof is complete.

(2) Suppose Δ contains a θ-annulus. As θ-bands cannot cross, the θ-annuli of Δ are partially ordered by the relation:

$$S' \leq S \text{ if } S' \text{ is contained in } \Delta_S$$

where S and S' are θ-annuli in Δ and Δ_S is the subdiagram of Δ bounded by the outer contour of S. Since Δ is finite, it is clear that there exists a minimal θ-annulus T with respect to this partial order.

Let Δ_T be the subdiagram of Δ bounded by the outer contour of T. If $\Delta_T \setminus T$ is empty, then T must contain a pair of cancellable cells, contradicting the assumption that Δ is reduced.

Suppose $\Delta_T \setminus T$ contains a (θ, a)-cell π. Then, letting T' be the maximal θ-band of Δ_T containing π, T' cannot cross T and so must be a θ-annulus. But then $T' < T$, contradicting the minimality of T.

Hence, by (1), any cell π of $\Delta_T \setminus T$ must be an a-cell. Then, (MM2) and Lemma 9.2 imply that every maximal a-band starting at an edge of $\partial\pi$ must either end on a cell of T or cross T. But since (1) implies that T consists entirely of (θ, a)-cells, π and T form a counterexample to (MM1), contradicting the assumption that Δ is M-minimal.

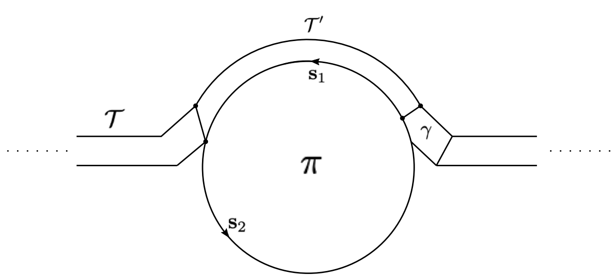
□

9.3 Transpositions of a θ-band with an a-cell

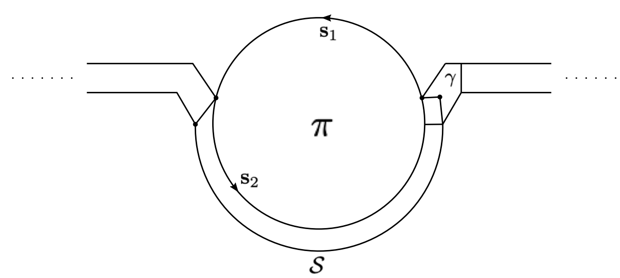
Let Δ be a reduced diagram over $G_\Omega(\mathbf{M})$ containing an a-cell π and a θ-band $\mathcal{T}$ subsequently crossing some of the a-bands starting at π. As the cells shared by these bands and $\mathcal{T}$ are (θ, a)-cells, the domain of the rule θ corresponding to $\mathcal{T}$ must be nonempty in the 'special' input sector. So, by the definition of the rules of $\mathbf{M}$, the domain of θ in this sector is the entire alphabet.

Suppose there are no other cells between π and the bottom of $\mathcal{T}$, i.e there is a subdiagram formed by π and $\mathcal{T}$.

Let $\mathbf{s}_1$ be the maximal subpath of $\partial\pi$ so that each edge is on the boundary of a (θ, a)-cell of $\mathcal{T}$. Further, let $\mathbf{s}_2$ be the complement of $\mathbf{s}_1$ in $\partial\pi$ so that $\partial\pi = \mathbf{s}_1\mathbf{s}_2$ and let $\mathcal{T}'$ be the subband of $\mathcal{T}$ satisfying $\mathbf{bot}(\mathcal{T}') = \mathbf{s}_1$.



(a) The subdiagram Γ



(b) The resulting subdiagram Γ'

Figure 9.3: The transposition of a θ-band with an a-cell, γ a (θ, q)-cell

Let $V_1 \equiv \text{Lab}(\mathbf{s}_1)$ and $V_2 \equiv \text{Lab}(\mathbf{s}_2)$. Finally, let Γ be the subdiagram formed by π and $\mathcal{T}'$.

Then, we can construct the θ-band $\mathcal{S}$ corresponding to θ consisting only of (θ, a)-cells and with top label V_2. Let Γ' be the subdiagram obtained by gluing a copy of π to $\mathbf{top}(\mathcal{S})$ in the clear way.

As $\text{Lab}(\mathbf{top}(\mathcal{T}')) \equiv V_1^{-1}$, $\text{Lab}(\partial\Gamma) \equiv \text{Lab}(\partial\Gamma')$. So, we may replace the Γ with Γ', attaching the first and last cells of $\mathcal{S}$ to the complement of $\mathcal{T}'$ in $\mathcal{T}$ and making any necessary cancellations in the resulting band.

This process is called the *transposition* of the θ-band with the a-cell.

Note that the diagram $\tilde{\Delta}$ resulting from the transposition has the same contour label as Δ. Further, if a maximal a-band of Δ has one end on the a-cell π, then the other end is not changed by the transposition.

Hence, if Δ is M-minimal, then $\tilde{\Delta}$ satisfies (MM2). However, $\tilde{\Delta}$ may not be M-minimal, as the transposed θ-band may cross the maximal a-bands emanating from more than half of the a-edges on the boundary of the transposed a-cell.

Further, since the number of (θ, a)-cells is altered by the transposition, the weight of the diagrams Δ and $\tilde{\Delta}$ may differ considerably.

Despite these disadvantages, this process will prove valuable in forthcoming arguments.

9.4 a-trapezia

We now generalize the concept of trapezium defined in Section 7.3 to the setting of M-minimal diagrams, allowing the existence of a-cells within the diagram.

To be specific, an *a-trapezium* Δ is an M-minimal diagram with contour of the form $\mathbf{p}_1^{-1}\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2^{-1}$, where each $\mathbf{p}_i$ is the side of a q-band and each $\mathbf{q}_i$ is the maximal subpath of the side of a θ-band that starts and ends with q-edges. As with trapezia, the factorization $\mathbf{p}_1^{-1}\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2^{-1}$ of the boundary is called the *standard factorization* of $\partial\Delta$.

The *history*, *step history*, *height*, and *base* of an a-trapezium are defined in the same way they are defined for a trapezium.

Note that the history of an a-trapezium must be reduced. Further, by Lemma 7.2, the base of an a-trapezium must be the base of an admissible word. So, in an a-trapezium Δ, the subdiagram Γ bounded by two consecutive q-bands is an a-trapezium with base UV corresponding to these q-bands' makeups. In this case, Γ is called a *UV-sector* in Δ. As with admissible words, an a-trapezium may contain sectors of the same name.

Lemma 9.4. *Suppose Δ is an a-trapezium containing an a-cell π. Then π is contained in a $(P_0(1)Q_0(1))^{\pm 1}$-, $P_0(1)P_0(1)^{-1}$-, or $Q_0(1)^{-1}Q_0(1)$-sector. Moreover, the step history of Δ must contain the letter $(1)_1$.*

Proof. Let Δ' be the sector of Δ containing π. Enumerate the maximal θ-bands of Δ' as $\mathcal{T}_1', \dots, \mathcal{T}_h'$. Then, there exists $j \in \{1, \dots, h-1\}$ such that π sits between $\mathcal{T}_j'$ and $\mathcal{T}_{j+1}'$.

By (MM2) and Lemma 9.2, each edge of $\partial\pi$ marks the start of a maximal a-band that must end on $\partial\Delta$ or on a (θ, q)-cell of one of the q-bands bounding Δ'. So, such a band must cross $\mathcal{T}_j'$, cross $\mathcal{T}_{j+1}'$, or end on a (θ, q)-cell in one of these bands.

Suppose an a-band ends on a (θ, q)-cell of $\mathcal{T}_j'$. Then by Lemma 7.2, the rule corresponding to $\mathcal{T}_j'$ must be of step history $(1)_1$ and the tape alphabet corresponding to Δ' must be the same as that of the 'special' input sector. So, the base of Δ' must be of the form $(P_0(1)Q_0(1))^{\pm 1}$ or $Q_0(1)^{-1}Q_0(1)$.

If an a-band ends on a (θ, q)-cell of $\mathcal{T}_{j+1}'$, then the same conclusion may be reached.

So, we may assume that all maximal a-bands with one end on π must cross either $\mathcal{T}_j'$ or $\mathcal{T}_{j+1}'$. Taking $n \geq 3$, property (MM1) implies that there must exist such a-bands crossing each of these θ-bands.

Lemma 7.2 then implies that the base of Δ' is of one of the forms in the statement. Moreover, the rules corresponding to $\mathcal{T}_j'$ and $\mathcal{T}_{j+1}'$ cannot lock the 'special' input sector, so that the step history corresponding to each rule is either of the form $(s)_1^{\pm 1}$ or $(1)_1$. As the rules cannot be mutually inverse, at least one contributes to an occurrence of $(1)_1$ in the step history of Δ'.

□

As a result of Lemmas 9.4, 7.2, and 4.2, if Δ is an a-trapezium with base B containing at least one a-cell, then every unreduced two-letter subword of B must be of the form:

(a) $P_0(i)P_0(i)^{-1}$ or $Q_0(i)^{-1}Q_0(i)$, or

(b) $Q_0(i)Q_0(i)^{-1}$ or $P_1(i)^{-1}P_1(i)$.

Note the following immediate consequences of Lemmas 7.2 and 4.2:

(i) if B contains a subword of form (a), then H cannot contain $\theta(12)_j^{\pm 1}$ for $j = 1, 2$

(ii) if B contains a subword of form (b), then H cannot contain $\theta(s)_1^{\pm 1}$ or a copy of the connecting rule of $\mathbf{M}_4(3^-)$

An a-trapezium is called *standard* if its base is pararevolving and its history contains a controlled subword. Note that the subdiagram of a standard a-trapezium bounded by the θ-bands corresponding to the controlled subword of the history is a trapezium.

Further, an a-trapezium is called *big* if its base is revolving, it contains a-cells, and it contains a subdiagram that is a standard trapezium. Note that the base of a big a-trapezium is necessarily reduced.

An a-trapezium is called *exceptional* if it contains a-cells and its base B is a cyclic shift of either:

- $Q_0(1)Q_0(1)^{-1}Q_0(1)$, or
- $P_0(1)Q_0(1)Q_0(1)^{-1}P_0(1)^{-1}\{t(1)\}^{-1}Q_4(L)^{-1}\dots P_1(L)^{-1}P_1(L)\dots Q_4(L)\{t(1)\}P_0(1)$

where gaps correspond to strings of letters that follow the order of the standard base (or its inverse).

Note that the base of an exceptional a-trapezium is hyperfaulty.

A *partition* of an a-trapezium Δ is a (finite) collection of subdiagrams $\{\Delta_i\}_{i=1}^m$ such that each Δ_i consists of a number of sectors of Δ, $\Delta_i \cap \Delta_j$ is either empty or a q-band for $i \neq j$, and each sector is a subdiagram of some Δ_i. Note that $\|\mathbf{tbot}(\Delta)\| = \sum_i \|\mathbf{tbot}(\Delta_i)\| - m$ and, similarly, $\|\mathbf{ttop}(\Delta)\| = \sum_i \|\mathbf{ttop}(\Delta_i)\| - m$. Moreover, as $\mathbf{tbot}(\Delta)$ and $\mathbf{ttop}(\Delta)$ each have at least m q-edges, $\sum_i \|\mathbf{tbot}(\Delta_i)\| \leq 2\|\mathbf{tbot}(\Delta)\|$ and $\sum_i \|\mathbf{ttop}(\Delta_i)\| \leq 2\|\mathbf{ttop}(\Delta)\|$.

Clearly, given a partition $\{\Delta_i\}$ of an a-trapezium Δ, $\text{wt}(\Delta) \leq \sum_i \text{wt}(\Delta_i)$.

Let Δ be an a-trapezium with revolving base B and let B' be a cyclic permutation of B. Then, there exists an a-trapezium Δ' with revolving base B' such that $\text{wt}(\Delta') = \text{wt}(\Delta)$. This diagram is constructed by cutting along a maximal q-band $\mathcal{Q}$ of Δ, pasting together the left and right q-bands of Δ, and pasting a copy of $\mathcal{Q}$ onto the side of the diagram. As with reduced computations, Δ' is called a cyclic permutation of Δ.

Note that by Lemma 8.1(d), for any maximal θ-band $\mathcal{T}$ in an a-trapezium Δ, the length of $\mathcal{T}$ is at most $|\mathbf{tbot}(\mathcal{T})|_q + 3|\mathbf{tbot}(\mathcal{T})|_a \leq 3\|\mathbf{tbot}(\mathcal{T})\|$.

Lemma 9.5. *Let Δ be a trapezium with history H and revolving base B. Then for $h = \|H\|$,*

$$wt(\Delta) \leq 3c_4 h \max(\|\boldsymbol{tbot}(\Delta)\|, \|\boldsymbol{ttop}(\Delta)\|)$$

Proof. Enumerate the maximal θ-bands of Δ as $\mathcal{T}_1, \dots, \mathcal{T}_h$. Then, letting ℓ_i be the length of $\mathcal{T}_i$, the definition of weight implies $\text{wt}(\Delta) = \sum_i \ell_i \leq \sum_i 3\|\mathbf{tbot}(\mathcal{T}_i)\|$.

By Lemmas 7.5 and 6.26, $\|\mathbf{tbot}(\mathcal{T}_i)\| \leq c_4 \max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|)$ for all i.

Hence, the statement follows.

□

Lemma 9.6. *Let Δ be an a-trapezium with history H and hyperfaulty base B. Then either Δ is exceptional or for $h = \|H\|$,*

$$wt(\Delta) \leq 3c_0 h \max(\|\boldsymbol{tbot}(\Delta)\|, \|\boldsymbol{ttop}(\Delta)\|) + C_1(|\boldsymbol{tbot}(\Delta)|_a + |\boldsymbol{ttop}(\Delta)|_a)^2$$

Proof. We proceed by cases:

1. Suppose Δ contains no a-cells.

Then Δ is a trapezium, so that Lemma 7.4 yields a corresponding reduced computation with a hyperfaulty base. For any maximal θ-band $\mathcal{T}$ of Δ, Lemma 6.25 then implies $|\mathbf{tbot}(\mathcal{T})|_a \leq c_0 \max(|\mathbf{ttop}(\Delta)|_a, |\mathbf{tbot}(\Delta)|_a)$. Hence,

$$\text{wt}(\Delta) \leq 3c_0 h \max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|)$$

So, it suffices to assume that Δ contains at least one a-cell.

2. Suppose B contains a subword $B' \equiv P_0(1)P_0(1)^{-1}$.

The definition of hyperfaulty and Lemma 9.4 then imply that the B'-sector Δ' is the only sector containing a-cells.

By (i), the step history of Δ cannot contain $(12)_1^{\pm}$. Further, Lemmas 7.2 and 4.2 imply that the step history cannot contain the letter $(s)_2^{\pm 1}$. Hence, the step history must be a subword of $(s)_1(1)_1(s)_1^{-1}$.

Let Δ'' be any sector of Δ other than Δ' and $\mathcal{T}''$ be a maximal θ-band in Δ''. Lemma 7.4 yields a reduced computation $\mathcal{C}''$ corresponding to Δ'' with the same history as Δ. The maximal computation of $\mathcal{C}''$ with step history $(1)_1$ has fixed tape word, satisfies the hypotheses of Lemma 4.3 or satisfies the hypotheses of Lemma 4.4. In each case, $\|\textbf{tbot}(\mathcal{T}'')\| \leq \max(\|\textbf{tbot}(\Delta'')\|, \|\textbf{ttop}(\Delta'')\|)$, so that $\text{wt}(\Delta'') \leq 3h \max(\|\textbf{tbot}(\Delta'')\|, \|\textbf{ttop}(\Delta'')\|)$.

By Lemma 9.2 and (MM2), every maximal a-band starting on the boundary of an a-cell must have another end on $\textbf{tbot}(\Delta')$ or $\textbf{ttop}(\Delta')$. So, the sum of the combinatorial perimeters of all a-cells in Δ' is at most $|\textbf{tbot}(\Delta')|_a + |\textbf{ttop}(\Delta')|_a \leq |\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a$.

Further, for $\mathcal{T}'$ a maximal θ-band in Δ', every a-edge of $\textbf{bot}(\mathcal{T}')$ must be part of a maximal a-band which has at least one end on $\textbf{tbot}(\Delta')$ or $\textbf{ttop}(\Delta')$. So, $\|\textbf{tbot}(\mathcal{T}')\| \leq \|\textbf{tbot}(\Delta')\| + \|\textbf{ttop}(\Delta')\|$.

Combining these gives $\text{wt}(\Delta') \leq 3h(\|\textbf{tbot}(\Delta')\| + \|\textbf{ttop}(\Delta')\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2$.

The set of sectors $\{\Delta_i\}$ then form a partition of Δ, so that

$$\begin{aligned}\text{wt}(\Delta) &\leq 3h \sum(\|\textbf{tbot}(\Delta_i)\| + \|\textbf{ttop}(\Delta_i)\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2 \\ &\leq 6h(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2 \\ &\leq 12h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2\end{aligned}$$

The statement is then satisfied by the parameter choice $c_0 \geq 4$.

3. Suppose B contains a subword $B' \equiv Q_0(1)^{-1}Q_0(1)$.

Again, the B'-sector Δ' must be the only sector containing a-cells and the step history of Δ must be a subword of $(s)_1(1)_1(s)_1^{-1}$.

As Δ is not exceptional, B must be a cyclic permutation of

$$P_1(1)^{-1}Q_0(1)^{-1}Q_0(1)P_1(1)\dots Q_4(1)\{t(2)\}P_0(2)P_0(2)^{-1}\{t(2)\}^{-1}Q_4(1)^{-1}\dots P_1(1)^{-1}$$

As in Step 2, we then have that $\text{wt}(\Delta'') \leq 3h\max(\|\textbf{tbot}(\Delta'')\|, \|\textbf{ttop}(\Delta'')\|)$ for any sector Δ'' other than Δ'.

However, a maximal a-band in Δ' need not have one end on $\textbf{tbot}(\Delta')$ or $\textbf{ttop}(\Delta')$, as it may end on a (θ, q)-cell of one of the bounding q-bands. Note that the rule corresponding to such a (θ, q)-cell must be of step history $(1)_1$. Let H_1 be the maximal subword of H comprising all letters with step history $(1)_1$ and $h_1 = \|H_1\|$.

Then, Lemma 9.2 and (MM2) imply that the sum of the combinatorial perimeters of all a-cells in Δ' is at most $|\textbf{tbot}(\Delta')|_a + |\textbf{ttop}(\Delta')|_a + 2h_1$. Further, for $\mathcal{T}'$ a maximal θ-band in Δ', $\|\textbf{tbot}(\mathcal{T}')\| \leq \|\textbf{tbot}(\Delta')\| + \|\textbf{ttop}(\Delta')\| + 2h_1$.

Let Δ_1'' be the $Q_0(1)P_1(1)$-sector of Δ. By Lemma 7.4, there exists a reduced computation $\mathcal{C}_1''$ corresponding to Δ_1''. The maximal subcomputation of $\mathcal{C}_1''$ with step history $(1)_1$ satisfies the hypotheses of Lemma 4.3, so that $h_1 \leq |\textbf{tbot}(\Delta_1'')|_a + |\textbf{ttop}(\Delta_1'')|_a$.

Similarly, letting Δ_2'' be the $P_1(1)^{-1}Q_0(1)^{-1}$-sector of Δ, $h_1 \leq |\textbf{tbot}(\Delta_2'')|_a + |\textbf{ttop}(\Delta_2'')|_a$

So, the sum of the combinatorial perimeters of the a-cells in Δ' is at most

$$|\textbf{tbot}(\Delta')|_a + |\textbf{ttop}(\Delta')|_a + \sum |\textbf{tbot}(\Delta_i'')|_a + |\textbf{ttop}(\Delta_i'')|_a \leq |\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a$$

This implies that $\text{wt}(\Delta') \leq 3h(\sum \|\textbf{tbot}(\mathcal{T}')\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2$, where this sum is taken over all maximal θ-bands $\mathcal{T}'$ in Δ'.

Meanwhile, for $\mathcal{T}'$ a maximal θ-band in Δ',

$$\|\textbf{tbot}(\mathcal{T}')\| \leq \|\textbf{tbot}(\Delta')\| + \|\textbf{ttop}(\Delta')\| + \sum_{i=1}^{2} \|\textbf{tbot}(\Delta_i'')\| + \|\textbf{ttop}(\Delta_i'')\|$$

So, taking the set of sectors $\{\Delta_i\}$ as a partition of Δ, we have

$$\begin{aligned}\mathrm{wt}(\Delta) &\leq 6h\sum(\|\mathbf{tbot}(\Delta_i)\| + \|\mathbf{ttop}(\Delta_i)\|) + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2 \\ &\leq 12h(\|\mathbf{tbot}(\Delta)\| + \|\mathbf{ttop}(\Delta)\|) + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2 \\ &\leq 24h\max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|) + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2\end{aligned}$$

so that the statement follows for $c_0 \geq 8$.

4. By Steps 2 and 3, B must contain a subword of the form $(P_0(1)Q_0(1))^{\pm 1}$. So, B must be a cyclic permutation of

$$P_0(1)Q_0(1)Q_0(1)^{-1}P_0(1)^{-1}\{t(1)\}^{-1}Q_s(L)^{-1}\dots P_1(L)^{-1}P_1(L)\dots Q_s(L)\{t(1)\}P_0(1)$$

But we may assume that Δ contains an a-cell by Step 1, so that Δ is exceptional.

□

Lemma 9.7. *Let Δ be an a-trapezium with history H and revolving base B. If Δ is neither big nor exceptional, then for $h = \|H\|$,*

$$wt(\Delta) \leq 3C_1h\max(\|\textbf{\textit{tbot}}(\Delta)\|, \|\textbf{\textit{ttop}}(\Delta)\|) + C_2(\|\textbf{\textit{tbot}}(\Delta)\| + \|\textbf{\textit{ttop}}(\Delta)\|)^2$$

Proof. By Lemma 9.6, we may assume that B is not hyperfaulty. As a result, B must contain a reduced pararevolving subword.

Further, Lemma 9.5 allows us to assume that Δ contains an a-cell.

1. Suppose B contains a reduced pararevolving subword B' such that B' has a subword of the form $(P_0(i)Q_0(i))^{\pm 1}$ for some $i \geq 2$.

Let Δ' be the maximal subdiagram of Δ which is an a-trapezium with base B'. By Lemma 9.4, Δ' is a trapezium.

Then, Lemma 7.4 yields a reduced computation $\mathcal{C}' : V_0' \to \dots \to V_h'$ with base B' corresponding to Δ'. By the parallel nature of the rules, we may assume that the base of this computation is

$\{t(i)\}B_3(i)\{t(i+1)\}$.

If $h > c_3 \max(\|V_0'\|, \|V_h'\|)$, then we may apply Lemma 6.19 to $\mathcal{C}'$, so that its history must contain a controlled subword. But then Lemma 6.7 implies that Δ is a big a-trapezium.

So, $h \leq c_3 \max(\|V_0'\|, \|V_h'\|) \leq c_3 \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|)$.

Let $\mathcal{T}$ be a maximal θ-band in Δ. By Lemma 9.2 and (MM2), any a-edge of $\textbf{tbot}(\mathcal{T})$ is part of a maximal a-band which must have at least one end on $\textbf{tbot}(\Delta)$, on $\textbf{ttop}(\Delta)$, or on a maximal q-band of Δ.

By the definition of revolving, there are at most $22L+1$ maximal q-bands in Δ, each of which consists of h (θ, q)-cells. Further, each such (θ, q)-cell has at most two a-edges on its contour. So, $|\textbf{tbot}(\mathcal{T})|_a \leq |\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a + (44L+2)h$.

As C_1 is chosen after L and c_3, this implies $\|\textbf{tbot}(\mathcal{T})\| \leq C_1 \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|)$.

Similarly, any maximal a-band of Δ with one end on an a-cell has its other end on $\textbf{tbot}(\Delta)$, on $\textbf{ttop}(\Delta)$, or on a maximal q-band corresponding to a letter $Q_0(1)^{\pm 1}$. As there are at most two such q-bands bounding a 'special' input sector, the sum of the combinatorial perimeters of a-cells must be at most $|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a + 2h \leq (2c_3+1)(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|)$.

Hence, $\text{wt}(\Delta) \leq 3C_1 h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|) + C_1(2c_3+1)^2(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|)^2$. The parameter choices $C_2 >> C_1 >> c_3$ then imply the statement.

Thus, we may assume that B is faulty but not hyperfaulty and that every reduced pararevolving subword of B contains a subword of the form $(P_0(1)Q_0(1))^{\pm 1}$. As a result, B has exactly two sectors corresponding to the 'special' input sector, each of which is of this form.

As Δ must contain a-cells, Lemma 9.4 implies that its step history must contain the letter $(1)_1$.

2. Suppose the step history of Δ is $(1)_1$.

For any sector Δ'' not containing a-cells, the corresponding computation must have fixed tape word, satisfy the hypotheses of Lemma 4.3, or satisfy the hypotheses of Lemma 4.4. So, for any maximal θ-band $\mathcal{T}''$ of Δ'', $\|\textbf{tbot}(\mathcal{T}'')\| \leq \max(\|\textbf{tbot}(\Delta'')\|, \|\textbf{ttop}(\Delta'')\|)$. Consequently, $\text{wt}(\Delta'') \leq 3h \max(\|\textbf{tbot}(\Delta'')\|, \|\textbf{ttop}(\Delta'')\|)$.

Let Δ' be a sector containing a-cells and $\mathcal{T}'$ be a maximal θ-band of Δ'.

Then any a-edge of **tbot**$(\mathcal{T}')$ is part of a maximal a-band which must end on **tbot**(Δ'), on **ttop**(Δ'), or on the maximal q-band corresponding to $Q_0(1)^{\pm 1}$. This then implies that $|\mathbf{tbot}(\mathcal{T}')|_a \leq |\mathbf{tbot}(\Delta')|_a + |\mathbf{ttop}(\Delta')|_a + h$. Similarly, the sum of the combinatorial perimeters of the a-cells in Δ' is at most $|\mathbf{tbot}(\Delta')|_a + |\mathbf{ttop}(\Delta')|_a + h$.

So, $\mathrm{wt}(\Delta') \leq 6h \max(\|\mathbf{tbot}(\Delta')\|, \|\mathbf{ttop}(\Delta')\|) + 3h^2 + C_1(|\mathbf{tbot}(\Delta')|_a + |\mathbf{ttop}(\Delta')|_a + h)^2$.

As B must contain a reduced pararevolving subword, it must have a subword B_0'' of the form $(Q_0(i)P_1(i))^{\pm 1}$. Let Δ_0'' be the B_0''-sector of Δ. Then the corresponding computaton satisfies the hypotheses of Lemma 4.3, so that $h \leq |\mathbf{tbot}(\Delta_0'')|_a + |\mathbf{ttop}(\Delta_0'')|_a$.

Thus, letting $\{\Delta_i\}$ be the partition of Δ given by its sectors, we have

$$\begin{aligned}
\mathrm{wt}(\Delta) &\leq 6h \sum (\|\mathbf{tbot}(\Delta_i)\| + \|\mathbf{ttop}(\Delta_i)\|) + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2 \\
&\leq 12h(\|\mathbf{tbot}(\Delta)\| + \|\mathbf{ttop}(\Delta)\|) + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2 \\
&\leq 24h \max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|) + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2
\end{aligned}$$

3. Suppose the step history of Δ contains a letter $(12)_j$.

By (i), any unreduced two-letter subword of B must be of form (b).

Suppose B contains a subword of the form $(Q_0(1)P_1(1))^{\pm 1}$. Then, by the definition of faulty, it must contain a reduced pararevolving subword of the form $(Q_0(1)\dots Q_0(2))^{\pm 1}$. But then B satisfies the hypothesis of Step 1, so that the statement follows.

Suppose B contains a subword of the form $(Q_0(L)P_1(L))^{\pm 1}$. Then by the definition of faulty, it must also contain a sector of the form $(P_0(L)Q_0(L))^{\pm 1}$, so that it contains a reduced pararevolving subword of the form $(P_0(L)\dots P_0(1))^{\pm 1}$. Again, B then satisfies the hypothesis of Step 1.

But assuming these to be false, since B must contain a subword of the form $(P_0(1)Q_0(1))^{\pm 1}$, B must be hyperfaulty and Δ exceptional.

So, the step history of Δ contains no letter of the form $(12)_j$ or, by symmetry, $(21)_j$.

4. By Steps 2 and 3, we assume that the step history of Δ contains a letter of the form $(s)_j^{\pm 1}$.

By (ii), every unreduced two-letter subword must be of form (a).

So, B must be a cyclic permutation of

$$P_0(2)P_0(2)^{-1}\dots P_1(L)^{-1}Q_0(L)^{-1}Q_0(L)P_1(L)\dots P_0(2)$$

where gaps correspond to strings of letters following the order of the standard base or its inverse written on a circle.

As B contains a subword of the form $(Q_0(1)P_1(1))^{\pm 1}$, then applying Lemma 6.2(a) to the corresponding reduced computation with such a base allows us to assume that the step history is a subword of $(1)_1(s)_1^{-1}(s)_2(1)_2$.

Let Δ'' be the maximal subtrapezium of Δ with base $P_1(L)^{-1}Q_0(L)^{-1}Q_0(L)P_1(L)$. Further, let $W_0''\to\dots\to W_h''$ be the corresponding reduced computation and $W_0''\to\dots\to W_r''$ be the maximal subcomputation with step history $(1)_1$. Then, Lemma 5.14 applies to the maximal subcomputations with step history $(1)_j$, so that $|W_r''|_a\leq\dots\leq|W_0''|_a$ and $|W_r''|_a\leq\dots\leq|W_h''|_a$. So, for any maximal θ-band $\mathcal{T}''$ of Δ'', $\|\textbf{tbot}(\mathcal{T}'')\|\leq\max(\|\textbf{tbot}(\Delta'')\|,\|\textbf{ttop}(\Delta'')\|)$, yielding $\text{wt}(\Delta'')\leq 3h\max(\|\textbf{tbot}(\Delta'')\|,\|\textbf{ttop}(\Delta'')\|)$.

Let Δ''' be a sector of Δ with base of the form $(Q_0(1)P_1(1))^{\pm 1}$. Then as above, Lemma 4.3 implies $\text{wt}(\Delta''')\leq 3h\max(\|\textbf{tbot}(\Delta''')\|,\|\textbf{ttop}(\Delta''')\|)$ and $r=|\textbf{tbot}(\Delta''')|_a$.

Let Δ' be a sector of Δ with base of the form $(P_0(1)Q_0(1))^{\pm 1}$. Then for any maximal θ-band $\mathcal{T}'$, an a-edge of $\textbf{tbot}(\mathcal{T}')$ is part of a maximal a-band with one end on $\textbf{tbot}(\Delta')$, on $\textbf{ttop}(\Delta')$, or on a (θ,q)-cell corresponding to the base letter $Q_0(1)^{\pm 1}$ and a rule of step history $(1)_1$. So, $\|\textbf{tbot}(\mathcal{T}')\|\leq\|\textbf{tbot}(\Delta')\|+\|\textbf{ttop}(\Delta')\|+r$. Similarly, the sum of the combinatorial perimeters of the a-cells of Δ' is at most $|\textbf{tbot}(\Delta')|_a+|\textbf{ttop}(\Delta')|_a+r$. Hence,

$$\text{wt}(\Delta')\leq 3h(\|\textbf{tbot}(\Delta')\|+\|\textbf{ttop}(\Delta')\|+r)+C_1(|\textbf{tbot}(\Delta')|_a+|\textbf{ttop}(\Delta')|_a+r)^2$$

For Δ_0' any other sector of Δ not already accounted for, its tape word is fixed, so that $\text{wt}(\Delta_0')\leq 3h\max(\|\textbf{tbot}(\Delta_0')\|,\|\textbf{ttop}(\Delta_0')\|)$.

Let $\{\Delta_i\}$ be the partition of Δ given by the subdiagrams detailed above. Then,

$$\begin{aligned}\mathrm{wt}(\Delta) &\leq 6h\sum(\|\textbf{tbot}(\Delta_i)\| + \|\textbf{ttop}(\Delta_i)\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2 \\ &\leq 12h(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2 \\ &\leq 24h\max(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|) + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2\end{aligned}$$

Thus, the statement follows from the parameter choices $C_2 >> C_1 \geq 8$.

□

9.5 Combs and Subcombs

Let Γ be an M-minimal diagram containing a maximal q-band $\mathcal{Q}$ such that $\textbf{bot}(\mathcal{Q})$ is a subpath of $\partial\Delta$ and every maximal θ-band of Δ ends at an edge of $\textbf{bot}(\mathcal{Q})$. Then Γ is called a *comb* and $\mathcal{Q}$ its *handle*.

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

Note that every a-trapezium (or trapezium) may be viewed as a comb with either maximal side q-band its handle.

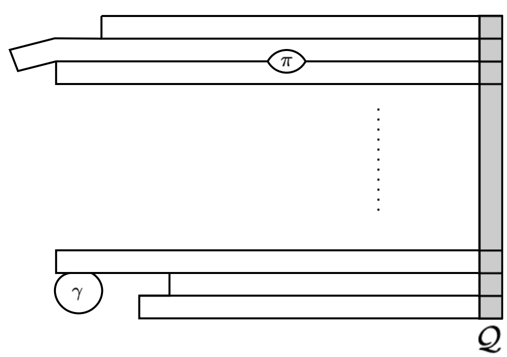


Figure 9.4: Comb with handle $\mathcal{Q}$ containing a-cells π and γ

Lemma 9.8. *Let Γ be a comb with height h, basic width b, and $|\partial\Gamma|_a = \alpha$. Let $\mathcal{T}_1, \dots, \mathcal{T}_h$ be the consecutive maximal θ-bands of Γ enumerated from bottom to top. Factor $\partial\Gamma =$ **yxz**, where **z** is the bottom of the handle of Γ and **x** is the maximal subpath below $\mathcal{T}_1$. Then:*

(1) $wt(\Gamma) \leq c_0 bh^2 + 2\alpha h + C_1(bh + \alpha)^2$

(2) $|\textbf{\textit{bot}}(\mathcal{T}_1)|_a \leq |\textbf{\textit{y}}|_a + 4bh$

Proof. (1) Let n_i be the length of $\mathcal{T}_i$.

By Lemma 9.2 and (MM2), every a-band starting on an a-cell must either end on a (θ, q)-cell or on $\partial\Gamma$. Since every (θ, q)-relation contains at most one a-letter from the ‘special’ input sector, the sum of the combinatorial perimeters of all a-cells in Γ is at most $bh + \alpha$.

So, $\text{wt}(\Gamma) \leq C_1(bh + \alpha)^2 + \sum_{i=1}^{h} n_i$.

Let α_i (respectively α_i') be the number of (unoriented) a-edges of $\textbf{bot}(\mathcal{T}_i)$ (respectively $\textbf{top}(\mathcal{T}_i)$) that are shared with $\partial\Gamma$. Similarly, let β_i (respectively β_i') be the number of a-edges of $\textbf{bot}(\mathcal{T}_i)$ (respectively $\textbf{top}(\mathcal{T}_i)$) that are on the boundary of an a-cell. Note that $\sum_{i=1}^h(\alpha_i + \alpha_i') \leq \alpha$ and $\sum_{i=1}^h(\beta_i + \beta_i') \leq bh + \alpha$.

By the definition of a comb, any cell of Γ below $\mathcal{T}_1$ must be an a-cell. So, any a-edge of $\textbf{bot}(\mathcal{T}_1)$ that is not shared with $\partial\Gamma$ is on the boundary of an a-cell below $\mathcal{T}_1$. Hence, $|\textbf{bot}(\mathcal{T}_1)|_a = \alpha_1 + \beta_1$.

Similarly, $|\textbf{top}(\mathcal{T}_h)|_a = \alpha_h' + \beta_h'$.

Lemma 8.1(d) implies $\alpha_1 + \beta_1 - b \leq n_1 \leq \alpha_1 + \beta_1 + 3b$ and $\alpha_h' + \beta_h' - b \leq n_h \leq \alpha_h' + \beta_h' + 3b$.

Suppose an a-edge of $\textbf{top}(\mathcal{T}_i)$ is not shared with $\textbf{bot}(\mathcal{T}_{i+1})$. Then either this edge is counted in α_i' or is on the boundary of an a-cell between $\mathcal{T}_i$ and $\mathcal{T}_{i+1}$, so that it is counted in β_i'. Similarly, an a-edge of $\textbf{bot}(\mathcal{T}_{i+1})$ not shared with $\textbf{top}(\mathcal{T}_i)$ is either counted in α_{i+1} or in β_{i+1}.

So, the difference in the number of a-edges of $\textbf{top}(\mathcal{T}_i)$ and $\textbf{bot}(\mathcal{T}_{i+1})$ is at most $\alpha_i' + \alpha_{i+1} + \beta_i' + \beta_{i+1}$. Lemma 8.1(d) then implies that $|n_{i+1} - n_i| \leq 4b + \alpha_i' + \alpha_{i+1} + \beta_i' + \beta_{i+1}$ for $1 \leq i \leq h-1$.

Hence, for all $1 \leq i \leq h$, we have:

$$\begin{aligned} n_i &\leq 3b + 4b(i-1) + \alpha_1 + \alpha_1' + \dots + \alpha_{i-1} + \alpha_{i-1}' + \alpha_i + \beta_1 + \beta_1' + \dots + \beta_{i-1} + \beta_{i-1}' + \beta_i \\ &\leq 5bh + 2\alpha \end{aligned}$$

Thus, $\text{wt}(\Gamma) \leq C_1(bh + \alpha)^2 + 5bh^2 + 2\alpha h$.

(2) For $i = 1, 2, 3$, define $\mathbf{A}_i$ as the subset of the set of (unoriented) a-edges of $\mathbf{y}$ as follows:

- $\mathbf{A}_1$ is the subset of edges that are on the boundary of an a-cell
- $\mathbf{A}_2$ is the subset of edges marking the start of a maximal a-band which ends on an edge of $\partial\Gamma$ shared with $\mathbf{bot}(\mathcal{T}_1)$
- $\mathbf{A}_3$ is the subset of edges that mark the start of a maximal a-band of nonzero length which ends on an a-cell

Note that these three sets are disjoint, so that $\gamma_1 + \gamma_2 + \gamma_3 \leq |\mathbf{y}|_a$ for $\gamma_i = \#\mathbf{A}_i$.

Let π be an a-cell in Γ such that some edge of $\partial\pi$ contributes to β_i'. Let β_π' be the number of edges of $\partial\pi$ contributing to β_i' and β_π'' be the number of such edges on the contour of a (θ, q)-cell of $\mathcal{T}_i$. Property (MM1) implies $\beta_\pi' \leq \frac{1}{2}\|\partial\pi\| + \beta_\pi''$. So,

$$\sum_{i=1}^{h} \beta_i' \leq \sum_{\pi} \beta_\pi' \leq \sum_{\pi} \left(\frac{1}{2}\|\partial\pi\| + \beta_\pi''\right)$$

Note that any edge of $\partial\pi$ not contributing to β_π' is either part of $\partial\Gamma$ or on $\mathbf{bot}(\mathcal{T}_{i+1})$, and so contributes to γ_1 or β_{i+1}, respectively. So, since there are at least $\frac{1}{2}\|\partial\pi\| - \beta_\pi''$ such edges,

$$\gamma_1 + \sum_{i=2}^{h} \beta_i \geq \sum_{\pi} \left(\frac{1}{2}\|\partial\pi\| - \beta_\pi''\right)$$

As the contour of any (θ, q)-cell contains at most one a-edge corresponding to the 'special' input sector, we then have

$$\sum_{i=1}^{h} \beta_i' \leq \gamma_1 + \sum_{i=2}^{h} \beta_i + 2\sum_{\pi} \beta_\pi'' \leq \gamma_1 + \sum_{i=2}^{h} \beta_i + 2bh$$

Next, let $\mathbf{e}$ be an a-edge of $\mathbf{bot}(\mathcal{T}_1)$ contributing to α_1. Then the maximal a-band starting at $\mathbf{e}$ ends on $\mathbf{y}$, ends on a (θ, q)-cell, or ends on an a-cell. Those that end on $\mathbf{y}$ correspond to edges of $\mathbf{A}_2$ while those that end on an a-cell correspond to a-edges contributing to β_i' for some i. So,

$$\alpha_1 \leq \gamma_2 + bh + \sum_{i=1}^{h} \beta_i' \leq \gamma_1 + \gamma_2 + \sum_{i=2}^{h} \beta_i + 3bh$$

Finally, let $\mathbf{e}$ be an a-edge of $\mathbf{bot}(\mathcal{T}_i)$ contributing to β_i. By Lemma 9.2 and (MM2), the maximal a-band starting on $\mathbf{e}$ must end on $\mathbf{y}$ or on a (θ, q)-cell. As those that end on $\mathbf{y}$ correspond to edges of $\mathbf{A}_3$, we have $\sum_{i=1}^h \beta_i \leq \gamma_3 + bh$. Thus,

$$|\mathbf{bot}(\mathcal{T}_1)|_a = \alpha_1 + \beta_1 \leq \gamma_1 + \gamma_2 + \sum_{i=1}^h \beta_i + 3bh \leq \gamma_1 + \gamma_2 + \gamma_3 + 4bh \leq |\mathbf{y}|_a + 4bh$$

□

A base word B is *tight* if it is of the form $uxvx$ for some letter x and words u and v, where:

(1) xvx is revolving, and

(2) no letter from u occurs in xvx.

Note that any tight base has length at most $K_0 = 22L + 1$, while any base with length at least K_0 must have a tight prefix.

A comb Δ is called *tight* if:

(C1) one of its maximal θ-bands $\mathcal{T}$ has a tight base when read toward the handle, and

(C2) all maximal θ-bands have tight bases or bases without tight prefixes

If Δ is an M-minimal diagram over $M_\Omega(\mathbf{M})$, then a subdiagram Γ is a *subcomb* of Δ if Γ is a comb and its handle divides Δ into two parts, one of which is Γ.

Let Γ be a comb with handle $\mathcal{C}$ and $\mathcal{B}$ be another maximal q-band in Γ. Then $\mathcal{B}$ cuts Γ into two parts, where the part not containing $\mathcal{C}$ is a subcomb Γ' with handle $\mathcal{B}$. Note that each maximal θ-band $\mathcal{T}$ of Γ crossing $\mathcal{B}$ has a subband $\mathcal{T}_0$ connecting $\mathcal{B}$ with $\mathcal{C}$. If $\mathcal{T}_0$ has no (θ, q)-cells, then Γ' is called a *derivative subcomb* of Γ.

Note that no maximal θ-band of a comb can cross the handles of more than one derivative subcomb.

Lemma 9.9. *(Compare with Lemma 6.6 of [18] and Lemma 6.10 of [25]) Let Δ be an M-minimal diagram such that $|\partial\Delta|_\theta > 0$ and every quasi-rim θ-band has base of length at least K. Then Δ contains a tight subcomb.*

Proof. As maximal θ-bands cannot cross, there exists a quasi-rim θ-band $\mathcal{T}_0$ in Δ. Taking $K > 2K_0$, the base of $\mathcal{T}_0$ has disjoint prefix and suffix, B_1 and B_2, of lengths K_0. As a result, B_1 has a prefix B_1' which is tight, while B_2 has a suffix B_2' such that $(B_2')^{-1}$ is tight.

Let π be the (θ, q)-cell of $\mathcal{T}_0$ corresponding to the last base letter of B_1' and $\mathcal{Q}'$ be the maximal q-band of Δ containing π. Let Γ' be the subdiagram of Δ bounded by $\mathcal{Q}'$ containing the subband of $\mathcal{T}_0$ with base B_1'.

Note that we may do the same with B_2' to construct a subdiagram Γ''.

Hence, there exists a maximal q-band $\mathcal{Q}$ such that for one of the subdiagrams Γ of Δ bounded by $\mathcal{Q}$, there exists a maximal θ-band $\mathcal{T}$ whose base is tight when read toward $\mathcal{Q}$. Choose such a $\mathcal{Q}$ and Γ such that wt(Γ) is minimal.

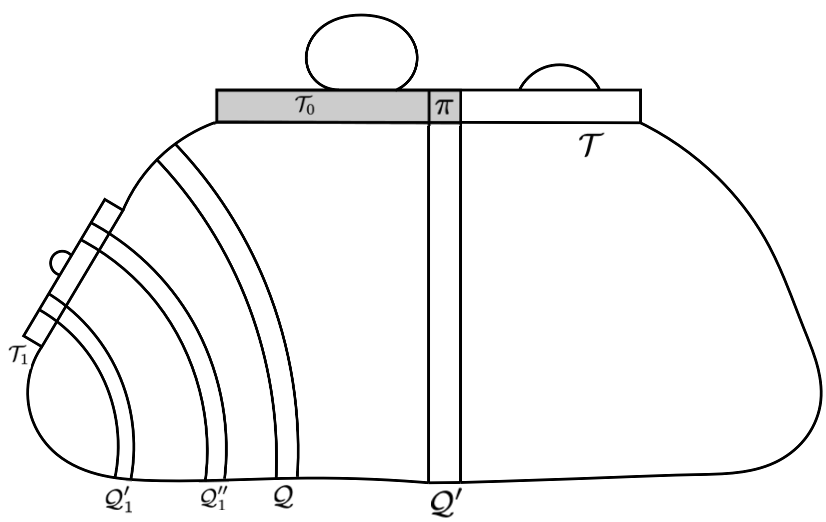


Figure 9.5: Lemma 9.9

Suppose there exists a θ-band in Γ which does not cross $\mathcal{Q}$. Then, there exists a quasi-rim θ-band $\mathcal{T}_1$ not crossing $\mathcal{Q}$. As the base of $\mathcal{T}_1$ has length at least K, we may repeat the argument above. This produces disjoint subdiagrams Γ_1' and Γ_1'' of Γ bounded by the maximal q-bands $\mathcal{Q}_1'$ and $\mathcal{Q}_1''$, respectively, such that the subband of $\mathcal{T}_1$ which is a maximal θ-band of Γ_1' (resp Γ_1'') has tight base when read toward $\mathcal{Q}_1'$ (resp $\mathcal{Q}_1''$). One of these subdiagrams, say Γ_1', does not contain $\mathcal{Q}$, and so is a subcomb of Δ. But then $\text{wt}(\Gamma_1') < \text{wt}(\Gamma)$, so that $\mathcal{Q}_1'$ and Γ_1' contradict the choice of $\mathcal{Q}$ and Γ.

Thus, Γ is a comb with handle $\mathcal{Q}$ satisfying condition (C1).

Now suppose there exists a maximal θ-band $\mathcal{T}'$ in Γ with a tight proper prefix B'. Let π' be the (θ, q)-cell of $\mathcal{T}'$ corresponding to the last letter of B' and $\mathcal{Q}'$ be the maximal q-band in Γ containing π'. Then for Γ' the subcomb of Γ with handle $\mathcal{Q}'$, $\mathcal{Q}'$ and Γ' contradict the choice of $\mathcal{Q}$ and Γ. Hence, Γ must be a tight subcomb.

□

9.6 G-weight

The goal of this section is to bound the size of an M-minimal diagram over $M_\Omega(\mathbf{M})$ in terms of its perimeter. However, this bound will not be given in terms of the area or weight of the diagram. Instead, we give the bound in terms of the artifical concept of *G-weight* (adapted from the concept of G-area in [18] and [25]).

Let Γ be an a-trapezium with base B and history H. Suppose B is of the form $(P_0(1)Q_0(1))^{\pm 1}$ or $Q_0(1)^{-1}Q_0(1)$, the step history of Γ is $(1)_1$, and H has a factorization $H_1 H_2^\ell H_3$ for some $\ell \geq 0$. Then Γ is called an *impeding* a-trapezium.

In this case, let $\eta = \|H_1\| + n\|H_2\| + \|H_3\|$ and $h = \|H\|$. Then we define the G-weight of Γ, denoted $\text{wt}_G(\Gamma)$, to be the minimum of half its weight and:

$$3h \max(\|\mathbf{tbot}(\Gamma)\|, \|\mathbf{ttop}(\Gamma)\|) + 3C_1 h\eta + C_1(|\mathbf{tbot}(\Gamma)|_a + |\mathbf{ttop}(\Gamma)|_a + 2\eta)^2$$

Similarly, if Γ is a big a-trapezium with height h then its G-weight is defined to be the minimum of half its weight and:

$$c_5 \max(\|\mathbf{ttop}(\Gamma)\|, \|\mathbf{tbot}(\Gamma)\|) + 4C_1(\|\mathbf{tbot}(\Gamma)\| + \|\mathbf{ttop}(\Gamma)\|)^2$$

Finally, any single cell in Γ is assigned G-weight equal to its weight.

For a reduced diagram Δ over $G_\Omega(\mathbf{M})$, consider a family of subdiagrams $\mathbf{P}$ such that:

- if $P \in \mathbf{P}$, then P is a single cell, a big a-trapezium, or an impeding trapezium,
- every cell of Δ belongs to an element of $\mathbf{P}$, and
- if there exist $P_1, P_2 \in \mathbf{P}$ with nonempty intersection, then both P_1 and P_2 are a-trapezia and this intersection is a q-band.

In this case, $\mathbf{P}$ is called a *covering* of Δ. The G-weight of $\mathbf{P}$, $\text{wt}_G(\mathbf{P})$, is defined to be the sum of the G-weights of its elements.

Note that any reduced diagram over $G_\Omega(\mathbf{M})$ has a covering, namely the one given by its cells. So, we may define the G-weight of Δ, $\text{wt}_G(\Delta)$, as the minimum of the G-weights of its coverings.

Further, since the G-weight of a big or impeding a-trapezium does not exceed half of its weight and any cell belongs to at most two elements of a covering, the inequality $\text{wt}_G(\Delta) \leq \text{wt}(\Delta)$ holds for all Δ.

Lemma 9.10. *Let Δ be a reduced diagram over $G_\Omega(\boldsymbol{M})$ and suppose every cell π of Δ belongs in one of the subdiagrams $\Delta_1, \dots, \Delta_m$, where any nonempty intersection $\Delta_i \cap \Delta_j$ is a q-band. Then $wt_G(\Delta) \leq \sum_{i=1}^m wt_G(\Delta_i)$.*

Proof. Let $\mathbf{P}_1, \dots, \mathbf{P}_m$ be coverings of $\Delta_1, \dots, \Delta_m$, respectively, so that the G-weight of $\mathbf{P}_i$ is equal to that of Δ_i. Then $\mathbf{P} = \mathbf{P}_1 \cup \dots \cup \mathbf{P}_m$ is a covering of Δ with $\text{wt}_G(\mathbf{P}) \leq \sum_{i=1}^m \text{wt}_G(\mathbf{P}_i)$, implying the statement.

□

In particular, note that Lemma 9.10 implies that if $\{\Delta_i\}$ is a partition of the a-trapezium Δ, then $\text{wt}_G(\Delta) \leq \sum \text{wt}_G(\Delta_i)$.

Lemma 9.11. *Suppose Δ is an a-trapezium with revolving base B and history H. Then for $h = \|H\|$,*

$$wt_G(\Delta) \leq C_2 h \max(\|\boldsymbol{tbot}(\Delta)\|, \|\boldsymbol{ttop}(\Delta)\|) + C_2(\|\boldsymbol{tbot}(\Delta)\| + \|\boldsymbol{ttop}(\Delta)\|)^2$$

Proof. By Lemma 9.7 and the assignment of G-weight to big trapezia, we may assume that Δ is exceptional.

1. Suppose the step history of Δ is $(1)_1$.

Let Δ' be the $Q_0(1)Q_0(1)^{-1}$-sector of Δ and let $\mathcal{C}' : W_0' \to \cdots \to W_h'$ be the reduced computation corresponding to Δ'. Lemma 4.4 implies that $|W_j'|_a \leq \max(|W_0'|_a, |W_h'|_a)$ for all $0 \leq j \leq h$, so that $\text{wt}_G(\Delta') \leq \text{wt}(\Delta') \leq 3h \max(\|\textbf{tbot}(\Delta')\|, \|\textbf{ttop}(\Delta')\|)$.

What's more, Lemma 4.4 implies that the history H of $\mathcal{C}'$ can be factored as $H_1 H_2^\ell H_3$ for some $\ell \geq 0$ with $\|H_1\| \leq \frac{1}{2}|W_0'|_a$, $\|H_3\| \leq \frac{1}{2}|W_h'|_a$, and $\|H_2\| \leq \min(|W_0'|_a, |W_h'|_a)$.

So, for Δ'' any sector of Δ corresponding to the 'special' input sector, Δ'' is an impeding a-trapezium with $\eta \leq \frac{n+1}{2}(|W_0'|_a + |W_h'|_a) \leq c_0(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)$.

Setting $S = |\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a$, this implies:

$$\text{wt}_G(\Delta'') \leq 3h \max(\|\textbf{tbot}(\Delta'')\|, \|\textbf{ttop}(\Delta'')\|) + 3c_0 C_1 h S + C_1(2c_0 + 1)^2 S^2$$

Every sector Δ''' not of these forms is a trapezium whose corresponding computation has fixed a-length, so that $\text{wt}(\Delta''') \leq 3h \max(\|\textbf{tbot}(\Delta''')\|, \|\textbf{ttop}(\Delta''')\|)$.

Let $\{\Delta_i\}$ be the partition of Δ given by its sectors. Then $\text{wt}_G(\Delta) \leq \sum \text{wt}_G(\Delta_i)$ by Lemma 9.10. Recall that $\sum \max(\|\textbf{tbot}(\Delta_i)\|, \|\textbf{ttop}(\Delta_i)\|) \leq 4 \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|)$.

Since at most two sectors correspond to the 'special' input sector, the parameter choices $C_2 >> C_1 >> c_0$ yield:

$$\begin{aligned}
\text{wt}_G(\Delta) &\leq 12h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|) + 6c_0 C_1 h S + 2C_1(3c_0)^2 S^2 \\
&\leq 12(c_0 C_1 + 1)h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|) + 18c_0^2 C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2 \\
&\leq C_2 h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|) + C_2(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2
\end{aligned}$$

2. Thus, we may assume that the step history of Δ is not $(1)_1$.

By (i) and (ii), we may then assume that the base B of Δ is a cyclic permutation of

$$P_0(1)Q_0(1)Q_0(1)^{-1}P_0(1)^{-1}\{t(1)\}^{-1}Q_4(L)^{-1}\dots P_1(L)^{-1}P_1(L)\dots Q_4(L)\{t(1)\}P_0(1)$$

As a result, H cannot contain a letter of the form $\theta(s)_j^{\pm 1}$ or the copy of a connecting rule of $\mathbf{M}_4(3^-)$. So, the step history of Δ must contain the letter $(12)_1$ or $(21)_1$.

Let Δ' be the $R_2(L)Q_3(L)$-sector of Δ. Lemma 9.4 then implies that Δ' is a trapezium, so that Lemma 7.4 gives a corresponding reduced computation. As a result, Lemma 5.28(b) implies that the step history of Δ has no subword of the form $(12)_1(2)_1(21)_1$.

Similarly, as B contains a subword $P_1(L)Q_1(L)$, Lemma 4.3 implies that the step history of Δ has no subword of the form $(23)_1(3)_1(32)_1$.

Hence, the step history of Δ is a subword of

$$(3)_1(32)_1(2)_1(21)_1(1)_1(12)_1(2)_1(23)_1(3)_1$$

containing the letter $(1)_1$.

Let Δ_1 be the maximal subdiagram of Δ which is an a-trapezium with step history $(1)_1$.

Suppose **top**(Δ_1) does not coincide with **top**(Δ).

Let $\mathcal{T}_1$ be the maximal θ-band of Δ_1 such that **top**$(\mathcal{T}_1)$ $=$ **top**(Δ_1). Then, there exists a maximal θ-band $\mathcal{T}_1'$ of Δ corresponding to the rule $\theta(12)_1$ and such that some edges of **bot**$(\mathcal{T}_1')$ coincide with those of **top**$(\mathcal{T}_1)$.

Suppose there is an a-cell π in Δ between $\mathcal{T}_1$ and $\mathcal{T}_1'$. As $\theta(12)_1$ locks the 'special' input sector, no a-band starting on $\partial\pi$ can cross $\mathcal{T}_1'$ or end on a (θ,q)-cell of $\mathcal{T}_1'$. Lemma 9.4 implies that π must belong to a $(P_0(1)Q_0(1))^{\pm 1}$ sector of Δ, so that at most one a-band starting on $\partial\pi$ can end on a (θ,q)-cell of $\mathcal{T}_1$. Taking $n\geq 3$, more than half of the a-bands starting on $\partial\pi$ must cross $\mathcal{T}_1$. But then π and $\mathcal{T}_1$ contradict property (MM1). So, **bot**$(\mathcal{T}_1')=$ **top**$(\mathcal{T}_1)$.

Let Δ_1' be the maximal subdiagram of Δ which is an a-trapezium with **bot**(Δ_1') $=$ **bot**$(\mathcal{T}_1')$. Lemma 9.4 then implies that Δ_1' is a trapezium, so that Lemma 7.4 yields a corresponding reduced computation $\mathcal{C} : W_0 \to \cdots \to W_t$ with base B and step history a prefix of $(12)_1(2)_1(23)_1(3)_1$.

We now prove that $|W_0|_a \leq \cdots \leq |W_t|_a$. Assuming toward contradiction, the step history of $\mathcal{C}$ cannot be $(12)_1$. Let $\mathcal{C}_2 : W_0 \to \cdots \to W_r$ be the maximal subcomputation with step history $(12)_1(2)_1$ and $\mathcal{C}_2' : W_0' \to \cdots \to W_r'$ be the restriction to the subword

$$Q_3(L)^{-1}R_2(L)^{-1}\dots Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)\dots R_2(L)Q_3(L)$$

of B. As every rule of $\mathcal{C}_2'$ locks the $Q_1(L)R_1(L)$-, $R_1(L)Q_2(L)$-, and $Q_2(L)R_2(L)$-sectors, we may view the subwords of the form $(Q_1(L)R_1(L)Q_2(L)R_2(L))^{\pm 1}$ as a single state letter. With this view, we may apply Lemma 5.14 to $\mathcal{C}_2'$, so that $|W_0'|_a \leq \cdots \leq |W_r'|_a$. As all other sectors have fixed tape word throughout $\mathcal{C}_2$, this implies $|W_0|_a \leq \cdots \leq |W_r|_a$.

So, we may assume that $t > r$. As a result, there exists a subcomputation $\mathcal{C}_3 : W_r \to \cdots \to W_t$ with step history $(23)_1(3)_1$. Letting $\mathcal{C}_3' : W_r' \to \cdots \to W_t'$ be the restriction of $\mathcal{C}_3$ to the subword $Q_1(L)^{-1}P_1(L)^{-1}P_1(L)Q_1(L)$, Lemma 5.14 implies $|W_r'|_a \leq \cdots \leq |W_t'|_a$. As above, all other sectors have fixed tape word throughout $\mathcal{C}_3$, so that $|W_r|_a \leq \cdots \leq |W_t|_a$.

As a result, for any maximal θ-band $\mathcal{T}$ of Δ_1', $\|\mathbf{tbot}(\mathcal{T})\| \leq \|W_t\| = \|\mathbf{ttop}(\Delta_1')\| = \|\mathbf{ttop}(\Delta)\|$. In particular, $\|\mathbf{ttop}(\Delta_1)\| \leq \|\mathbf{ttop}(\Delta)\|$. Hence, for h_1' the height of Δ_1', $\mathrm{wt}(\Delta_1') \leq 3h_1'\|\mathbf{ttop}(\Delta)\|$.

Similarly, if $\mathbf{bot}(\Delta_1)$ does not coincide with $\mathbf{bot}(\Delta)$, then there exists a subdiagram Δ_1'' of Δ which is a trapezium satisfying $\mathbf{top}(\Delta_1'') = \mathbf{bot}(\Delta_1)$ and $\mathbf{bot}(\Delta_1'') = \mathbf{bot}(\Delta)$. By analogous arguments, $\|\mathbf{tbot}(\Delta_1)\| \leq \|\mathbf{tbot}(\Delta)\|$ and $\mathrm{wt}(\Delta_1'') \leq 3h_1''\|\mathbf{tbot}(\Delta)\|$ for h_1'' the height of Δ_1''.

By Step 1, letting h_1 be the height of Δ_1, we have

$$\begin{aligned}\mathrm{wt}_G(\Delta_1) &\leq C_2h_1\max(\|\mathbf{tbot}(\Delta_1)\|, \|\mathbf{ttop}(\Delta_1)\|) + C_2(|\mathbf{tbot}(\Delta_1)|_a + |\mathbf{ttop}(\Delta_1)|_a)^2 \\ &\leq C_2h_1\max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|) + C_2(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2\end{aligned}$$

Thus, Lemma 9.10 yields

$$\mathrm{wt}_G(\Delta) \leq (C_2h_1 + 3h_1' + 3h_1'')\max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|) + C_2(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2$$

$$\leq C_2 h \max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|) + C_2(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2$$

□

9.7 Quadratic upper bound

Our goal throughout the rest of this section is to prove that for any M-minimal diagram Δ,

$$\mathrm{wt}_G(\Delta) \leq N_2|\partial\Delta|^2 + N_1\mu(\Delta) \tag{9.1}$$

for the parameters N_1 and N_2.

We do this by arguing toward contradiction, considering a 'minimal counterexample' diagram Δ. In other words, Δ is an M-minimal diagram over $M_\Omega(\mathbf{M})$ satisfying the inequality $\mathrm{wt}_G(\Delta) > N_2|\partial\Delta|^2 + N_1\mu(\Delta)$, while (9.1) holds for all M-minimal diagrams Γ over $M_\Omega(\mathbf{M})$ satisfying $|\partial\Gamma| < |\partial\Delta|$.

Lemma 9.12. *If Γ is an M-minimal diagram over $M_\Omega(\boldsymbol{M})$, with no q-edges on its boundary, then $wt_G(\Gamma) \leq C_2|\partial\Gamma|^2$.*

Proof. Since any q-edge in Γ would give rise to a maximal q-band which, by Lemma 9.1, can only end on the boundary of the diagram, Γ cannot have any q-edges. So, Γ is comprised entirely of (θ, a)-cells and a-cells.

In particular, Γ contains no a-trapezia (or trapezia), so that the only covering of Γ is by single cells. Hence, $\mathrm{wt}_G(\Gamma) = \mathrm{wt}(\Gamma)$.

Lemma 9.2 and (MM2) then imply that any maximal a-band with one end on an a-cell must have its other end on the boundary, so that the sum of the (combinatorial) perimeters of the a-cells is at most $\|\partial\Gamma\|$. It follows that the sum of the weights of the a-cells is at most $C_1\|\partial\Gamma\|^2$.

Further, Lemma 9.3 implies that any maximal θ-band must start and end on $\partial\Gamma$, so that there are at most $\frac{1}{2}\|\partial\Gamma\|$ maximal θ-bands in Γ. As Lemma 9.1 implies that each maximal a-band must have at least one end on $\partial\Gamma$ and each θ-band intersects each a-band in at most one cell, the length of each θ-band is at most $\|\partial\Gamma\|$. So, the sum of the lengths of all maximal θ-bands, and so the number of (θ, a)-cells, is at most $\frac{1}{2}\|\partial\Gamma\|^2$.

Taking into account the modified definition of perimeter, the statement follows from an appropriate choice of C_2 in terms of C_1 and δ.

□

The parameter choice $N_2 >> C_2$ and Lemma 9.12 allow us to assume that $\partial\Delta$ consists of at least two q-edges, i.e $|\partial\Delta| \geq 2$.

Lemma 9.13. *Let π be an a-cell contained in Δ. Suppose $\partial\pi$ has a subpath **s** shared with $\partial\Delta$. Then $\|\mathbf{s}\| \leq \frac{2}{3}\|\partial\pi\|$.*

Proof. Let $\partial\pi = \mathbf{st}$ and $\partial\Delta = \mathbf{ss}_0$.

Assuming toward contradiction that $\|\mathbf{s}\| > \frac{2}{3}\|\partial\pi\|$, we have $\|\mathbf{s}\| > 2\|\mathbf{t}\|$ and $\|\mathbf{s}\| > \frac{2}{3}n \geq 8$ by a parameter choice.

Let Δ_0 be the subdiagram of Δ obtained by removing π. So, $\partial\Delta_0 = \mathbf{t}^{-1}\mathbf{s}_0$.

By Lemma 8.1(c), $|\partial\Delta_0| \leq |\mathbf{s}_0| + |\mathbf{t}| = |\mathbf{s}_0| + \delta\|\mathbf{t}\|$ and $|\partial\Delta| \geq |\mathbf{s}| + |\mathbf{s}_0| - 2\delta \geq |\mathbf{s}_0| + \delta(\|\mathbf{s}\| - 2)$.

So, $|\partial\Delta| - |\partial\Delta_0| \geq \delta(\|\mathbf{s}\| - \|\mathbf{t}\| - 2) \geq \frac{1}{2}\delta(\|\mathbf{s}\| - 4) \geq \frac{1}{4}\delta\|\mathbf{s}\| > 0$.

The inductive hypothesis then applies to Δ_0, yielding

$$\mathrm{wt}_G(\Delta_0) \leq N_2|\partial\Delta_0|^2 + N_1\mu(\Delta_0) \leq N_2(|\partial\Delta| - \delta\|\mathbf{s}\|/4)^2 + N_1\mu(\Delta_0)$$

As $\delta\|s\|/4 \leq |\partial\Delta|$, $(|\partial\Delta| - \delta\|\mathbf{s}\|/4)^2 \leq |\partial\Delta|^2 - \frac{1}{4}\delta\|\mathbf{s}\||\partial\Delta|$.

By Lemma 9.10, we have $\mathrm{wt}_G(\Delta) \leq \mathrm{wt}_G(\Delta_0) + \mathrm{wt}(\pi)$. Further, the necklaces corresponding to $\partial\Delta$ and $\partial\Delta_0$ are identical, so that $\mu(\Delta) = \mu(\Delta_0)$. So, since the combinatorial perimeter of π is $\|\mathbf{s}\| + \|\mathbf{t}\|$, Lemma 9.10 then implies:

$$\mathrm{wt}_G(\Delta) \leq N_2|\partial\Delta|^2 - \frac{1}{4}N_2\delta\|\mathbf{s}\||\partial\Delta| + N_1\mu(\Delta) + C_1(\|\mathbf{s}\| + \|\mathbf{t}\|)^2$$

So, we reach the contradiction $\text{wt}_G(\Delta) \leq N_2|\partial\Delta|^2 + N_1\mu(\Delta)$ if

$$\frac{1}{4}N_2\delta\|\mathbf{s}\||\partial\Delta| \geq C_1(\|\mathbf{s}\| + \|\mathbf{t}\|)^2$$

As $\|\mathbf{t}\| < \frac{1}{2}\|\mathbf{s}\|$, we have $C_1(\|\mathbf{s}\| + \|\mathbf{t}\|)^2 \leq \frac{9}{4}C_1\|\mathbf{s}\|^2$. So, since $|\partial\Delta| \geq \frac{1}{4}\delta\|\mathbf{s}\|$, it suffices to show $N_2\delta^2 \geq 36C_1$. But this follows from the parameter choices $N_2 >> C_1 >> \delta^{-1}$.

□

The following is the direct analogue of Lemma 6.12 of [18] and Lemma 6.16 of [25]. The method of proof is identical to the ones presented in those sources, though many of the estimates differ.

Lemma 9.14. *(1) 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 $\boldsymbol{x}$ of $\partial\Delta$ with $|\boldsymbol{x}|_q \leq c_0$.*

(2) If Γ is a subcomb of Δ with basic width $s \leq K$, $|\partial\Gamma|_q = 2s$.

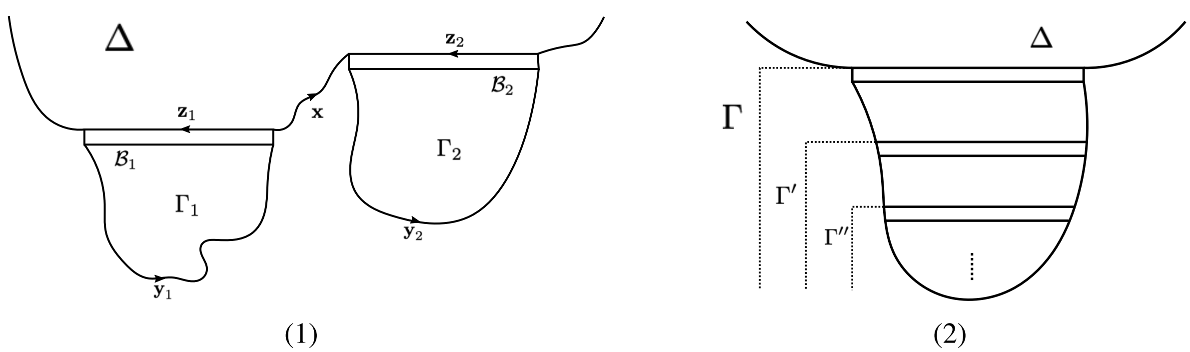


Figure 9.6: Lemma 9.14

Proof. We prove (1) and (2) simultaneously, inducting on $W = \text{wt}(\Gamma_1) + \text{wt}(\Gamma_2)$ for (1) and $W = \text{wt}(\Gamma)$ for (2). In other words, we consider a counterexample to one of these two with minimal value of W.

Suppose the minimal counterexample is of the form (1).

As $\mathrm{wt}(\Gamma_i) < W$ for $i = 1, 2$, the inductive hypothesis implies that (2) holds for each. So, $\partial\Gamma_i$ has at most $2K$ q-edges.

Let h_i be the height of Γ_i and assume without loss of generality that $h_1 \leq h_2$. For $i = 1, 2$, let $\partial\Gamma_i = \mathbf{y}_i\mathbf{z}_i$ where $\mathbf{y}_i$ is a subpath of $\partial\Delta$ and $\mathbf{z}_i = \mathbf{bot}(\mathcal{B}_i)$. Without loss of generality, assume $\mathbf{y}_1\mathbf{x}\mathbf{y}_2$ is a subpath of $\partial\Delta$.

Then each θ-edge of $\mathbf{y}_1$ is separated in $\partial\Delta$ from each θ-edge of $\mathbf{y}_2$ by at most $4K + c_0$ q-edges, and so by at most J q-edges by the choice of parameters. Hence, each (correctly ordered) pair of such edges (or the white beads corresponding to these edges) makes a contribution to $\mu(\Delta)$.

Let Δ' be the diagram obtained by removing the subdiagram Γ_1 from Δ. When passing from $\partial\Delta$ to $\partial\Delta'$, one replaces each θ-edge of $\mathbf{y}_1$ with the corresponding θ-edge of $\mathbf{z}_1$ belonging to the same θ-band. But since $\mathcal{B}_1$ is removed, there is at least one less q-edge separating any of the h_1h_2 (correctly ordered) pairs of θ-edges described above. So, $\mu(\Delta) - \mu(\Delta') \geq h_1h_2$ by Lemma 8.3(d).

Letting $|\partial\Gamma_1|_a = \alpha$, Lemma 9.8 yields $\mathrm{wt}_G(\Gamma_1) \leq \mathrm{wt}(\Gamma_1) \leq c_0Kh_1^2 + 2\alpha h_1 + C_1(Kh_1 + \alpha)^2$.

By Lemma 8.1(b), we have $|\mathbf{z}_1| = h_1$. Moreover, each of the h_1 (θ, q)-cells of $\mathcal{B}_1$ contributes at most one a-edge to $\mathbf{z}_1$.

So, $\mathbf{y}_1$ consists of h_1 θ-edges, at least two q-edges, and at least $\max(0, \alpha - h_1)$ a-edges. Lemma 8.1(a) then implies $|\mathbf{y}_1| \geq \max(h_1 + 2, h_1 + 2 + (\alpha - 2h_1)\delta)$.

Letting $\mathbf{s}$ be the complement of $\mathbf{y}_1$ in $\partial\Delta$, $\mathbf{s}$ is also the complement of $\mathbf{z}_1^{-1}$ in $\partial\Delta'$. So, Lemma 8.1(c) implies that $|\partial\Delta'| \leq |\mathbf{z}_1| + |\mathbf{s}| = h_1 + |\mathbf{s}|$ and

$$|\partial\Delta| \geq |\mathbf{y}_1| + |\mathbf{s}| - 2\delta \geq h_1 + |\mathbf{s}| + 2 - 2\delta + \max(0, (\alpha - 2h_1)\delta)$$

Hence, taking $\delta^{-1} > 2$, we have

$$|\partial\Delta| - |\partial\Delta'| \geq \gamma = \max(1, (\alpha - 2h_1)\delta) \tag{9.2}$$

In particular, $|\partial\Delta'| < |\partial\Delta|$, so that the inductive hypothesis implies

$$\mathrm{wt}_G(\Delta') \leq N_2|\partial\Delta'| + N_1\mu(\Delta') \leq N_2(|\partial\Delta| - \gamma)^2 + N_1(\mu(\Delta) - h_1h_2)$$

Noting that $\gamma \leq |\partial\Delta|$, we have $(|\partial\Delta| - \gamma)^2 \leq |\partial\Delta|^2 - \gamma|\partial\Delta|$, so that

$$\mathrm{wt}_G(\Delta') \leq N_2|\partial\Delta|^2 - N_2\gamma|\partial\Delta| + N_1\mu(\Delta) - N_1h_1h_2$$

Combining this with the G-weight of Γ_1, Lemma 9.10 then implies:

$$\mathrm{wt}_G(\Delta) \leq N_2|\partial\Delta|^2 - N_2\gamma|\partial\Delta| + N_1\mu(\Delta) - N_1h_1h_2 + c_0Kh_1^2 + 2\alpha h_1 + C_1(Kh_1 + \alpha)^2$$

So, in order to reach the contradiction $\mathrm{wt}_G(\Delta) \leq N_2|\partial\Delta|^2 + N_1\mu(\Delta)$, it suffices to show:

$$-N_2\gamma|\partial\Delta| - N_1h_1h_2 + c_0Kh_1^2 + 2\alpha h_1 + C_1(Kh_1 + \alpha)^2 \leq 0$$

As $h_1 \leq h_2$, this amounts to proving:

$$(c_0K + C_1K^2)h_1^2 + 2(C_1K + 1)\alpha h_1 + C_1\alpha^2 \leq N_2\gamma|\partial\Delta| + N_1h_1^2 \tag{9.3}$$

If $\alpha \leq 4h_1$, then the inequality (9.3) follows from the parameter choice of N_1, as it is chosen after c_0, K, and C_1.

Otherwise, we have $\alpha > 4h_1$. The parameter choice $N_1 \geq c_0K + C_1K^2$ means that it suffices only to show that:

$$\left(\frac{C_1K + 1}{2} + C_1\right)\alpha^2 \leq N_2\gamma|\partial\Delta| \tag{9.4}$$

But then $\alpha - 2h_1 \geq \alpha/2$, so that $\gamma \geq \frac{1}{2}\delta\alpha$. Hence, $N_2\gamma|\partial\Delta| \geq \frac{1}{4}N_2\delta^2\alpha^2$, so that (9.4) follows from the parameter choices $N_2 >> C_1 >> \delta^{-1} >> K$.

Now suppose we have a minimal counterexample of the form (2).

As each derivative subcomb of Γ is connected with the handle $\mathcal{B}$ of Γ by θ-bands, they can be ordered in a natural way.

Consider two neighbor derivative subcombs, Γ_1 and Γ_2. The handle of Γ_i is intersected by two disjoint collections of θ-bands which connect them with $\mathcal{B}$. If there is any θ-band between these two collections, then it cannot intersect any q-bands except for $\mathcal{B}$, as otherwise it intersects a derivative subcomb between Γ_1 and Γ_2. So, the subpath **x** of $\partial\Delta$ between the handles of Γ_1 and Γ_2 satisfies $|\mathbf{x}|_q = 0$.

Hence, Γ_1 and Γ_2 form a contradiction to (1). However, $\text{wt}_G(\Gamma_1) + \text{wt}_G(\Gamma_2) < \text{wt}_G(\Gamma) = W$ since they contain no cells of $\mathcal{B}$, contradicting the minimality of the counterexample.

Thus, Γ contains at most one derivative subcomb Γ'. In turn, Γ' contains at most one derivative subcomb Γ'', and so on. Thus, there are s maximal q-bands in Γ, so that Lemma 9.1 implies that $|\partial\Gamma|_q = 2s$.

□

Similarly, the next statement is a direct analogue of Lemma 6.14 in [18] and Lemma 6.17 in [25] with altered estimates.

Lemma 9.15. *Suppose Γ is a subcomb of Δ whose basic width is at most K_0 and whose handle $\mathcal{B}$ has length ℓ. If Γ' is a subcomb of Γ with handle $\mathcal{B}'$ of length ℓ', then $\ell' > \ell/2$.*

Proof. Assume toward contradiction that Γ' is a subcomb of Γ whose handle $\mathcal{B}'$ has length $\ell' \leq \ell/2$. Then, we can choose Γ' so that ℓ' is minimal for all subcombs in Γ and so that Γ' has no proper subcombs, i.e the basic width of Γ' is 1. Then, letting $\alpha = |\partial\Gamma'|_a$, Lemma 9.8 implies

$$\text{wt}_G(\Gamma') \leq \text{wt}(\Gamma') \leq c_0(\ell')^2 + 2\alpha\ell' + C_1(\ell' + \alpha)^2$$

Let Δ' be the diagram obtained from Δ by removing Γ'. Then the following inequality arises as the analogue of (9.2):

$$|\partial\Delta| - |\partial\Delta'| \geq \gamma = \max(1, (\alpha - 2\ell')\delta) \tag{9.5}$$

In particular, $|\partial\Delta'| < |\partial\Delta|$, so that

$$\text{wt}_G(\Delta') \leq N_2|\partial\Delta'|^2 + N_1\mu(\Delta') \leq N_2(|\partial\Delta| - \gamma)^2 + N_1\mu(\Delta') \tag{9.6}$$

Every maximal θ-band of Γ passing through $\mathcal{B}'$ connects it to $\mathcal{B}$. The cells of $\mathcal{B}$ that such bands end on form a subband $\mathcal{C}$ of $\mathcal{B}$ with length ℓ'.

Then, the maximal θ-bands of Γ starting from $\mathcal{C}$ bound a comb with handle $\mathcal{C}$. So, there exists a maximal subdiagram Γ'' of Γ which is a comb with handle $\mathcal{C}$. Note that Γ'' contains Γ'.

The components of $\mathcal{B} \setminus \mathcal{C}$ are handles of combs E_1 and E_2, respectively, which comprise the complement of Γ'' in Γ. Letting ℓ_i be the height of E_i, we then have $\ell_1 + \ell_2 = \ell - \ell' \geq \ell'$.

Let $\partial\Gamma = \mathbf{yz}$ be the factorization such that $\mathbf{z} = \mathbf{bot}(\mathcal{B})$ and $\mathbf{y}$ is a subpath of $\partial\Delta$. So, there are ℓ_i θ-edges on the common subpath $\mathbf{x}_i$ of $\mathbf{y}$ and ∂E_i and ℓ' θ-edges on the common subpath $\mathbf{x}$ of $\mathbf{y}$ and $\partial\Gamma''$. Further, as the basic width of Γ is at most K_0, Lemma 9.14(2) implies that $\mathbf{y}$ contains at most K q-edges.

So, for any edge from $\mathbf{x}$ and any edge from $\mathbf{x}_i$, there are at most K q-edges between the pair in $\mathbf{y}$, and so at most J such edges by the choice of parameters. Hence, each of these $\ell'(\ell_1 + \ell_2)$ (appropriately ordered) pairs of edges contributes to $\mu(\Delta)$.

When passing from Δ to Δ', the θ-edges of $\mathbf{y}$ are replaced with the corresponding edges of $\mathbf{bot}(\mathcal{B}')$. However, the q-edges of $\mathcal{B}'$ are removed, so that there is at least one less q-edge between a θ-edge corresponding to an edge of $\mathbf{x}$ and one corresponding to an edge of $\mathbf{x}_i$. So, Lemma 8.3(d) implies $\mu(\Delta) - \mu(\Delta') \geq \ell'(\ell_1 + \ell_2)$. Substituting this into (9.6) then implies

$$\text{wt}_G(\Delta') \leq N_2(|\partial\Delta| - \gamma)^2 + N_1\mu(\Delta) - N_1\ell'(\ell_1 + \ell_2) \tag{9.7}$$

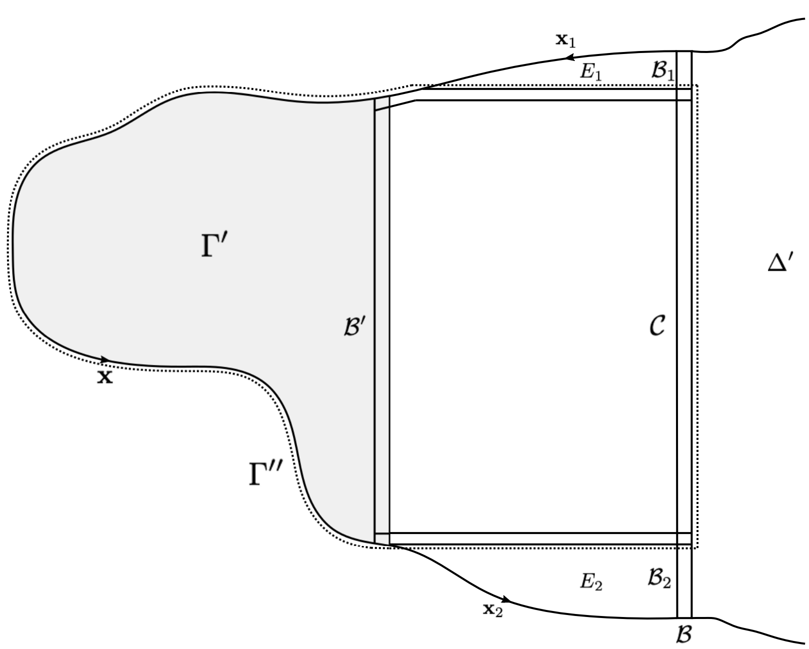


Figure 9.7: Lemma 9.15

Note that $|\partial\Delta| \geq \gamma$, so that $(|\partial\Delta| - \gamma)^2 \leq |\partial\Delta|^2 - \gamma|\partial\Delta|$. Factoring in Γ' and applying Lemma 9.10 then yields

$$\text{wt}_G(\Delta) \leq N_2|\partial\Delta|^2 - N_2\gamma|\partial\Delta| + N_1\mu(\Delta) - N_1\ell'(\ell_1 + \ell_2) + c_0(\ell')^2 + 2\alpha\ell' + C_1(\ell' + \alpha)^2$$

So, it suffices to show

$$-N_2\gamma|\partial\Delta| - N_1\ell'(\ell_1 + \ell_2) + c_0(\ell')^2 + 2\alpha\ell' + C_1(\ell' + \alpha)^2 \leq 0 \tag{9.8}$$

Suppose $\alpha \leq 4\ell'$. Then $c_0(\ell')^2 + 2\alpha\ell' + C_1(\ell' + \alpha)^2 \leq c_0(\ell')^2 + 8(\ell')^2 + C_1(5\ell')^2$. As $\ell_1 + \ell_2 \geq \ell'$, (9.8) then follows from the parameter choices $N_1 >> C_1 >> c_0$.

Otherwise, $\alpha > 4\ell'$, so that $\gamma \geq \frac{1}{2}\delta\alpha$. Hence, $|\partial\Delta| \geq \gamma$ and (9.8) imply that it suffices to show

$$c_0(\ell')^2 + 2\alpha\ell' + C_1(\ell' + \alpha)^2 \leq \frac{1}{4}N_2\delta^2\alpha^2 + N_1(\ell')^2 \tag{9.9}$$

Note that $2\alpha\ell' + C_1(\ell' + \alpha)^2 \leq \left(\frac{25C_1+8}{16}\right)\alpha^2$, so that (9.9) follows by the parameter choices

$N_1 >> c_0$ and $N_2 >> C_1 >> \delta^{-1}$.

□

Lemma 9.16. *If $\mathcal{T}$ is a quasi-rim θ-band in Δ, then the base of $\mathcal{T}$ has length $s > K$.*

Proof. Suppose $\mathcal{T}$ is a quasi-rim θ-band in Δ with base of length $s \leq K$. Without loss of generality, say that any cell between **top**$(\mathcal{T})$ and $\partial\Delta$ is an a-cell. Let $\mathbf{P}_1$ be the set of such a-cells.

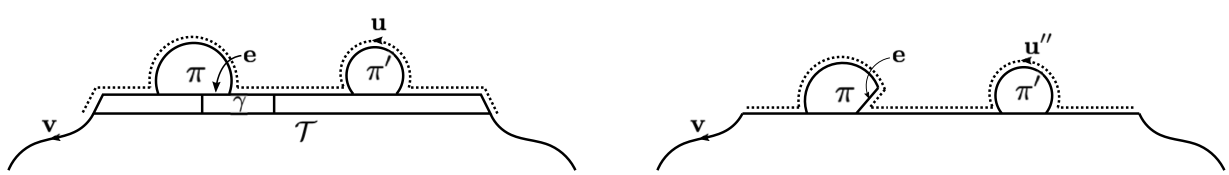
Let **u** be the subpath of $\partial\Delta$ bounded by the two end θ-edges of $\mathcal{T}$ and **v** be its complement in $\partial\Delta$. For $\pi \in \mathbf{P}_1$, factor $\partial\pi = \mathbf{p}_\pi \mathbf{p}'_\pi$ where $\mathbf{p}_\pi$ is a subpath of **u** and $\mathbf{p}'_\pi$ is a subpath of **top**$(\mathcal{T})$.

Let b_π be the number of edges of $\mathbf{p}'_\pi$ that are on the boundary of a (θ, q)-cell of $\mathcal{T}$.

By Lemma 9.13, $\|\mathbf{p}'_\pi\| \geq \frac{1}{3}\|\partial\pi\|$. Further, by Lemma 3.6, $\|\partial\pi\| \geq (1-\beta)n \geq n/2$ by the parameter choice for β (see Section 3.7). As a result, $\|\mathbf{p}'_\pi\| \geq 3$ and $b_\pi \leq 2$, so that $\mathbf{p}'_\pi$ has a maximal subpath $\mathbf{p}''_\pi$ consisting of edges on the boundary of (θ, a)-cells of $\mathcal{T}$.

Consider the diagram Δ' obtained from Δ by cutting along **bot**$(\mathcal{T})$, removing $\mathcal{T}$ and the a-cells of $\mathbf{P}_1$. For $\pi \in \mathbf{P}_1$, the subpath $\mathbf{p}''_\pi$ can be identified with a subpath of **bot**$(\mathcal{T})$, so that we may paste π to Δ' along this subpath.

Let Δ'' be the diagram obtained by pasting all cells of $\mathbf{P}_1$ to Δ'. Note that **v** can be identified with a subpath of $\partial\Delta''$. Let $\mathbf{u}''$ be the complement of **v** in $\partial\Delta''$.



(a) Δ, γ a (θ, q)-cell (b) Δ''

Figure 9.8: Lemma 9.16

For any $\pi \in \mathbf{P}_1$, the edges of $\partial\pi$ contributing to b_π belong to $\mathbf{u}''$ after this pasting. So, at least $\|\mathbf{p}_\pi\| + b_\pi \geq \frac{1}{2}\|\partial\pi\|$ edges of $\partial\pi$ are shared with $\partial\Delta''$. It is thus clear from construction that Δ'' is M-minimal.

Meanwhile, by Lemma 4.1, each (θ, q)-cell of $\mathcal{T}$ contributes at most two a-edges to **bot**$(\mathcal{T})$. Any other edge of $\mathbf{u}''$ corresponds to an edge of **u**.

As each a-edge contributing to b_π for some $\pi \in \mathbf{P}$ is labelled by a letter from the alphabet of the 'special' input sector and each (θ, q)-relation has at most one such letter, $\sum b_\pi \leq s$. So, since two θ-edges are removed from $\mathbf{u}$, Lemma 8.1 implies

$$|\mathbf{u}| - |\mathbf{u}''| \geq 2 - (2s+2)\delta - \delta \sum b_\pi \geq 2 - (3s+2)\delta \geq 2 - (3K+2)\delta \geq 1$$

The parameter choice $\delta^{-1} >> K$ and Lemma 8.1 then imply

$$|\partial\Delta| - |\partial\Delta''| \geq (|\mathbf{u}| + |\mathbf{v}| - \delta) - (|\mathbf{u}''| + |\mathbf{v}|) \geq 2 - (3K+3)\delta \geq 1$$

Hence, the inductive hypothesis may be applied to Δ'', so that

$$\text{wt}_G(\Delta'') \leq N_2|\partial\Delta''|^2 + N_1\mu(\Delta'') \leq N_2(|\partial\Delta| - 1)^2 + N_1\mu(\Delta'')$$

Note that the necklace corresponding to Δ'' is obtained from that corresponding to Δ by the removal of two white beads. Lemma 8.3(a) then yields $\mu(\Delta'') \leq \mu(\Delta)$.

Let $\mathbf{P}''$ be a minimal covering of Δ''. As each a-cell of $\mathbf{P}_1$ has a boundary edge shared with $\partial\Delta''$, it cannot be contained in a trapezium in Δ''. So, $\mathbf{P}_1 \subset \mathbf{P}''$.

Let $\mathbf{P}$ be the covering of Δ given by $\mathbf{P}''$ and the cells of $\mathcal{T}$. Then for ℓ the length of $\mathcal{T}$,

$$\text{wt}_G(\Delta) \leq \text{wt}_G(\mathbf{P}) = \mathbf{wt}_G(\mathbf{P}'') + \ell \leq N_2|\partial\Delta|^2 - N_2|\partial\Delta| + N_1\mu(\Delta) + \ell$$

Hence, it suffices to show that $N_2|\partial\Delta| \geq \ell$.

For $\pi \in \mathbf{P}_1$, (MM1) implies $\|\mathbf{p}'_\pi\| \leq \|\mathbf{p}_\pi\| + 2b_\pi$. So, since $\sum b_\pi \leq s$, $|\mathbf{top}(\mathcal{T})|_a \leq |\mathbf{u}|_a + 2s$.

By Lemma 8.1(d), $\ell \leq |\mathbf{top}(\mathcal{T})|_a + 3|\mathbf{top}(\mathcal{T})|_q \leq |\mathbf{u}|_a + 5s$. As each a-edge of $\mathbf{u}$ contributes at least δ to $|\Delta|$ and there are s q-edges of $\mathbf{u}$, $\ell \leq \delta^{-1}(|\partial\Delta| - s) + 5s \leq \delta^{-1}|\partial\Delta|$.

But then the statement follows from the parameter choice $N_2 >> \delta^{-1}$.

□

Thus, Lemmas 9.9 and 9.16 imply that there exists a tight subcomb Γ in Δ. By the definition of tight combs, the basic width of Γ is at most K_0 (see Figure 9.9).

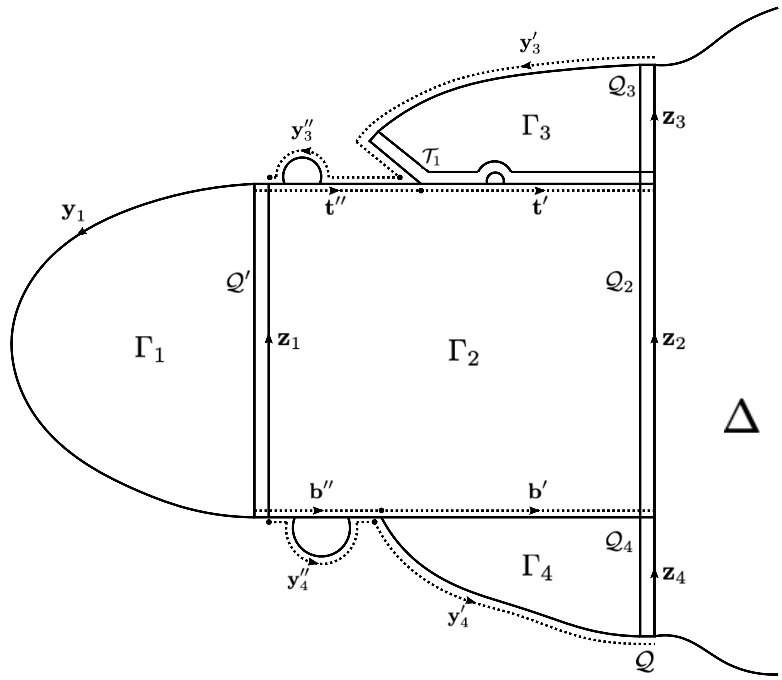


Figure 9.9: Tight subcomb Γ

Let $\mathcal{T}$ be a maximal θ-band in Γ with tight base B. Then B has the form $uxvx$, where x does not occur in u or v and the final letter corresponds to the handle $\mathcal{Q}$ of Γ. Let $\mathcal{Q}'$ be the q-band corresponding to the first occurrence of x in B.

Every maximal θ-band in Γ crossing $\mathcal{Q}'$ has a subband connecting $\mathcal{Q}$ and $\mathcal{Q}'$. The (θ, q)-cells of $\mathcal{Q}$ on which these θ-bands end form a subband $\mathcal{Q}_2$ of $\mathcal{Q}$ with length ℓ'.

Let Γ_2 be the a-trapezium with side q-bands $\mathcal{Q}'$ and $\mathcal{Q}_2$ and bounded by the θ-bands connecting these two. By the definition of tight, the base of Γ_2 is revolving.

Cutting along $\mathbf{bot}(\mathcal{Q}')$ separates Δ into two components, one of which and is a subcomb Γ_1 with handle $\mathcal{Q}'$. Further, $\Gamma' = \Gamma_1 \cup \Gamma_2$ is a comb contained in Γ with handle $\mathcal{Q}_2$.

Let $\mathcal{Q}_3$ and $\mathcal{Q}_4$ be the components of $\mathcal{Q} \setminus \mathcal{Q}_2$. Then there exist maximal subdiagrams Γ_3 and Γ_4 of Γ that are combs with handles $\mathcal{Q}_3$ and $\mathcal{Q}_4$, respectively.

Let ℓ, ℓ', ℓ_3, and ℓ_4 be the heights of Γ, Γ_1, Γ_3, and Γ_4, respectively. By Lemma 9.15, $\ell' > \ell/2$.

Let $\partial\Gamma = \mathbf{y}\mathbf{z}$ be the factorization given by $\mathbf{z} = \mathbf{bot}(\mathcal{Q})$. Similarly, let $\partial\Gamma_1 = \mathbf{y}_1\mathbf{z}_1$ be the factorization given by $\mathbf{z}_1 = \mathbf{bot}(\mathcal{Q}')$.

Note that $\mathbf{t} := \mathbf{ttop}(\Gamma_2)$ can be factored as $\mathbf{t} = \mathbf{t}''\mathbf{t}'$ such that $\mathbf{t}'$ is a maximal subpath shared with $\partial\Gamma_3$. Similarly, $\mathbf{b} := \mathbf{tbot}(\Gamma_2)$ has a factorization $\mathbf{b} = \mathbf{b}''\mathbf{b}'$ such that $(\mathbf{b}')^{-1}$ is a maximal subpath shared with $\partial\Gamma_4$. Note that $\Gamma \setminus (\Gamma' \sqcup \Gamma_3 \sqcup \Gamma_4)$ consists of a-cells that are attached to $\mathbf{t}''$ or $\mathbf{b}''$.

Factor $\mathbf{y} = \mathbf{y}_3\mathbf{y}_1\mathbf{y}_4$. Then, $\mathbf{y}_3$ can be factored as $\mathbf{y}_3'\mathbf{y}_3''$ where $\mathbf{y}_3'$ is a maximal subpath shared with $\partial\Gamma_3$. Note that every edge of $\mathbf{y}_3''$ is either shared with $(\mathbf{t}'')^{-1}$ or is on the boundary of an a-cell attached to $\mathbf{t}''$. Similarly, we may factor $\mathbf{y}_4 = \mathbf{y}_4''\mathbf{y}_4'$.

Finally, factor $\mathbf{z} = \mathbf{z}_4\mathbf{z}_2\mathbf{z}_3$ where $\mathbf{z}_i = \mathbf{bot}(\mathcal{Q}_i)$.

Let $\mathcal{T}_1$ be the bottom θ-band of Γ_3. Then, let Γ_3' be the comb contained in Γ_3 obtained by removing any a-cells below $\mathcal{T}_1$. So, $\mathbf{bot}(\mathcal{T}_1)$ is a subpath of $\partial\Gamma_3'$.

Similarly, define Γ_4' by removing any a-cells above the top θ-band of Γ_4.

Lemma 9.17. *(1)* $|\boldsymbol{t}''|_a \leq |\boldsymbol{y}_3''|_a + 4$ *and* $|\boldsymbol{b}''|_a \leq |\boldsymbol{y}_4''|_a + 4$

(2) $|\boldsymbol{y}_3|_a \geq |\boldsymbol{t}|_a - 2\ell_3 K - 4$ *and* $|\boldsymbol{y}_4|_a \geq |\boldsymbol{b}|_a - 2\ell_4 K - 4$.

Proof. Let $\mathbf{E}$ be the set of (unoriented) a-edges of $\mathbf{t}$. Then $\mathbf{E} = \sqcup_{i=1}^4 \mathbf{E}_i$, where:

- $\mathbf{E}_1$ is the set of such edges shared with $\partial\Delta$,
- $\mathbf{E}_2$ is the set of such edges shared with the boundary of an a-cell not contained in Γ_3,
- $\mathbf{E}_3$ is the set of such edges shared with the boundary of an a-cell contained in Γ_3, and
- $\mathbf{E}_4$ is the set of such edges shared with $\mathbf{bot}(\mathcal{T}_1)$.

Further, let $\mathbf{E}_2'$ (resp $\mathbf{E}_3'$) be the subset of $\mathbf{E}_2$ (resp $\mathbf{E}_3$) consisting of the edges which are on the boundary of a (θ, q)-cell in Γ_2.

Let $\alpha_i = \#\mathbf{E}_i$ for all $1 \leq i \leq 4$. Similarly, let $\alpha_i' = \#\mathbf{E}_i'$ for $i = 2, 3$.

Note that $\alpha_1 + \alpha_2 = |\mathbf{t}''|_a$ and $\mathbf{E}_1$ is a subset of the a-edges of $\mathbf{y}_3''$.

Further, every edge of $\mathbf{E}'_2 \sqcup \mathbf{E}'_3$ is labelled by a letter from the 'special' input sector and is on the boundary of a (θ, q)-cell of the same θ-band (the top θ-band of Γ_2). Any (θ, q)-cell with such an edge on its boundary must correspond to the base letter $Q_0(1)^{\pm 1}$, in which case it has exactly one such letter on its boundary. Hence, by the definition of revolving, $\alpha'_2 + \alpha'_3 \leq 2$.

For any a-edge $\mathbf{e}$ of $\mathbf{t}''$, either $\mathbf{e} \in \mathbf{E}_1$ or $\mathbf{e}$ is on the boundary of some a-cell π. In the latter case, $\partial\pi$ can be factored as $\mathbf{s}_1\mathbf{s}_2$, where $\mathbf{s}_1$ is a subpath of $\mathbf{y}''_3$ and $\mathbf{s}_2$ is a subpath of $\mathbf{t}''$. Let b be the number of edges of $\mathbf{s}_2$ that are on the boundary of a (θ, q)-cell contained in Γ_2. Then by condition (MM1), $\|\mathbf{s}_2\| - b \leq \frac{1}{2}\|\partial\pi\|$. So, $\|\mathbf{s}_1\| \geq \|\mathbf{s}_2\| - 2b$.

Applying the same reasoning to all such a-cells, we have $\alpha_1 + \max(0, \alpha_2 - 2\alpha'_2) \leq |\mathbf{y}''_3|_a$.

Hence, $|\mathbf{t}''|_a = \alpha_1 + \alpha_2 \leq |\mathbf{y}''_3|_a + 2\alpha'_2 \leq |\mathbf{y}''_3|_a + 4$.

Next, let $\mathbf{F}_3$ be the edges of $\mathbf{y}'_3$ on the boundary of an a-cell below $\mathbf{bot}(\mathcal{T}_1)$ and set $\beta = \#\mathbf{F}_3$.

Let π be an a-cell with a boundary edge contributing to α_3. Then as above, for b the number of edges of $\partial\pi$ on the boundary of a (θ, q)-cell of Γ_2, at most $\frac{1}{2}\|\partial\pi\| + b$ of the edges of $\partial\pi$ are shared with $\mathbf{t}'$. Note that the other edges of $\partial\pi$ are either part of $\mathbf{bot}(\mathcal{T}_1)$ or contribute to β.

So, $\alpha_4 + \max(0, \alpha_3 - 2\alpha'_3) \leq \beta + |\mathbf{bot}(\mathcal{T}_1)|_a$.

Hence, $|\mathbf{t}|_a = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 \leq |\mathbf{y}''_3|_a + \beta + |\mathbf{bot}(\mathcal{T}_1)|_a + 2(\alpha'_2 + \alpha'_3) \leq |\mathbf{y}''_3|_a + \beta + |\mathbf{bot}(\mathcal{T}_1)|_a + 4$.

Let $\partial\Gamma'_3 = \mathbf{x}'\mathbf{x}\mathbf{z}_3$ be the factorization given by $\mathbf{x} = \mathbf{bot}(\mathcal{T}_1)$. Note that $\mathbf{x}'$ is a subpath of $\mathbf{y}'_3$ not containing any edges of $\mathbf{F}_3$, so that $\beta + |\mathbf{x}'|_a \leq |\mathbf{y}'_3|_a$.

Then, applying Lemma 9.8 to Γ'_3, we have $|\mathbf{x}|_a \leq |\mathbf{x}'|_a + 4K_0\ell_3$.

Thus, $|\mathbf{t}|_a \leq |\mathbf{y}''_3|_a + \beta + |\mathbf{x}'|_a + 4K_0\ell_3 + 4 \leq |\mathbf{y}_3|_a + 2\ell_3 K + 4$.

Applying the analogous argument to Γ_4 yields the inequalities $|\mathbf{b}''|_a \leq |\mathbf{y}''_3|_a + 4$ and $|\mathbf{b}|_a \leq |\mathbf{y}_4|_a + 2\ell_4 K + 4$.

□

Lemma 9.18. *Set* $M = \max(|\boldsymbol{b}|_a, |\boldsymbol{t}|_a)$. *Then* $2K\ell > M$.

Proof. As $|\mathbf{y}_i|_a \leq |\mathbf{y}|_a$ and $\ell_i \leq \ell/2$ for $i = 3, 4$, Lemma 9.17(2) implies

$$M \leq |\mathbf{y}|_a + K\ell + 4 \leq |\mathbf{y}|_a + 3K\ell/2$$

Assuming that $2K\ell \leq M$, we then have $|\mathbf{y}|_a \geq \frac{1}{2}K\ell \geq K_0\ell$.

By Lemma 9.8, we have

$$\begin{aligned}\mathrm{wt}(\Gamma) &\leq c_0 K_0 \ell^2 + 2(|\mathbf{y}|_a + \ell)\ell + C_1(K_0\ell + |\mathbf{y}|_a + \ell)^2 \\ &\leq (c_0K_0 + 2 + C_1(K_0+1)^2)\ell^2 + (2 + 2C_1(K_0+1))|\mathbf{y}|_a\ell + C_1|\mathbf{y}|_a^2\end{aligned}$$

So, as C_2 is chosen after C_1, K_0, and c_0, we have:

$$\mathrm{wt}(\Gamma) \leq C_2|\mathbf{y}|_a^2$$

Since $|\mathbf{y}|_\theta = \ell$ and $|\mathbf{y}|_q \geq 2$, Lemma 8.1(a) implies that $|\mathbf{y}| \geq 2 + \ell + (|\mathbf{y}|_a - \ell)\delta \geq \ell + 2 + \frac{1}{2}\delta|\mathbf{y}|_a$.

Let Δ' be the M-minimal diagram formed from Δ by removing Γ. Then in $\partial\Delta'$, **y** is replaced with **z**. Lemma 8.1(b) implies that $|\mathbf{z}| = \ell$.

Letting **s** be the complement of **z** in $\partial\Delta'$, Lemma 8.1(c) implies $|\partial\Delta'| \leq |\mathbf{s}| + \ell$ and $|\partial\Delta| \geq |\mathbf{s}| + |\mathbf{y}| - 2\delta$. So, $|\partial\Delta| - |\partial\Delta'| \geq \gamma = \max(1, \frac{1}{2}\delta|\mathbf{y}|_a)$.

Hence, we may apply the inductive hypothesis to Δ', yielding

$$\mathrm{wt}_G(\Delta') \leq N_2|\partial\Delta'|^2 + N_1\mu(\Delta') \leq N_2(|\partial\Delta| - \gamma)^2 + N_1\mu(\Delta')$$

As $\gamma \leq |\partial\Delta|$, $(|\partial\Delta| - \gamma)^2 \leq |\partial\Delta|^2 - \gamma|\partial\Delta|$. Lemma 8.3 further implies that $\mu(\Delta') \leq \mu(\Delta)$. So, adding in the weight of Γ, Lemma 9.10 implies:

$$\mathrm{wt}_G(\Delta) \leq N_2|\partial\Delta|^2 - N_2\gamma|\partial\Delta| + N_1\mu(\Delta) + C_2|\mathbf{y}|_a^2$$

So, it suffices to show that $N_2\gamma|\partial\Delta| \geq C_2|\mathbf{y}|_a^2$.

But $\frac{1}{2}\delta|\mathbf{y}|_a \leq \gamma \leq |\partial\Delta|$, so that $N_2\gamma|\partial\Delta| \geq \frac{1}{4}N_2\delta^2|\mathbf{y}|_a^2$. So, the desired inequality follows from the parameter choices $N_2 >> C_2 >> \delta^{-1}$.

□

Lemma 9.19. *The counterexample diagram Δ does not exist.*

Proof. Let Δ_1 be the diagram obtained from Δ by removing $\Gamma \setminus \mathcal{Q}$.

As the base of Γ_2 is revolving, the bands $\mathcal{Q}'$ and $\mathcal{Q}_2$ are labelled identically. So, we may construct a diagram Δ_0 by pasting Γ_1 to Δ_1 along $\mathcal{Q}'$ and $\mathcal{Q}_2$ (see Figure 9.10, compare with Figure 9.9).

Since an a-band cannot cross a q-band, any counterexample to (MM1) or (MM2) in Δ_0 is contained in one of Δ_1 or Γ_1. But Δ_1 and Γ_1 are M-minimal as subdiagrams of Δ. Hence, Δ_0 must be M-minimal.

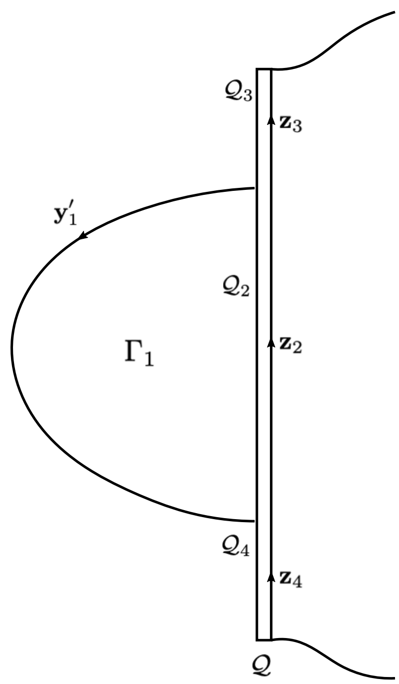


Figure 9.10: The construction of Δ_0

Let **P** be a covering of Δ_0. Suppose there exists $P \in$ **P** that is not completely contained in Γ_1 or Δ_1. Then P is a big a-trapezium containing a maximal q-band $\mathcal{B}$ that is a subband of $\mathcal{Q}_2$ (and $\mathcal{Q}'$) and which is not a side q-band of P. Then, the history of P is a subword of the history of Γ_2, so that Γ_2 is itself a big a-trapezium by Lemma 6.7.

Let $\mathcal{Q}''$ be the maximal q-band of Γ corresponding to the first letter of the base of P. Further, let $\mathcal{T}$ be a θ-band in Γ connecting $\mathcal{Q}$ to $\mathcal{Q}''$. Then the base of $\mathcal{T}$ (read toward $\mathcal{Q}$) is B_1B_2, where

B_1 is a prefix of the base of P and B_2 is the base of Γ_2. By Lemma 6.7, both the base of P and B_2 are reduced. Moreover, since the first letter of B_2 appears in the base of P, B_1B_2 must be reduced.

As B_2 is revolving, the first letter of B_1 appears in B_2. But then the base of the maximal θ-band of Γ containing $\mathcal{T}$ has a tight prefix, contradicting the assumption that Γ is a tight comb.

So, for any covering of Δ_0, each element is either contained completely in Γ_1 or completely in Δ_1. Hence, given a minimal covering $\mathbf{P}$ of Δ_0, we may construct coverings $\mathbf{P}'$ and $\mathbf{P}''$ of Γ_1 and Δ_1, respectively, by including only the elements belonging to these subdiagrams and perhaps adding in the cells of $\mathcal{Q}'$ or $\mathcal{Q}_2$. As at most the ℓ' cells of $\mathcal{Q}_2$ are counted twice in these coverings, we have $\text{wt}_G(\Delta_0) \geq \text{wt}_G(\Gamma_1) + \text{wt}_G(\Delta_1) - \ell'$.

Lemma 9.10 then implies:

$$\text{wt}_G(\Delta) \leq \text{wt}_G(\Delta_1) + \text{wt}_G(\Gamma) \leq \text{wt}_G(\Delta_0) + \text{wt}_G(\Gamma_2) + \text{wt}_G(\Gamma_3) + \text{wt}_G(\Gamma_4) + A + \ell' \quad (9.10)$$

where A is the sum of the weights of the a-cells attached to $\mathbf{t}''$ or $\mathbf{b}''$.

For $i = 3, 4$, note that the subpath $\mathbf{y}_i''$ has no θ-edges, while $\mathbf{y}_i'$ consists of ℓ_i θ-edges and at least one q-edge. So, Lemma 8.1(a) implies $|\mathbf{y}_i| \geq 1 + \ell_i + \delta \max(0, |\mathbf{y}_i|_a - \ell_i, |\mathbf{y}_i''|_a - 1)$.

Letting $\mathbf{s}$ be the complement of $\mathbf{y}$ in $\partial\Delta$, Lemma 8.1(c) then yields

$$|\partial\Delta| \geq |\mathbf{s}| + |\mathbf{y}| - 2\delta \geq |\mathbf{s}| + |\mathbf{y}_3| + |\mathbf{y}_1| + |\mathbf{y}_4| - 4\delta$$

Next, let $\mathbf{y}_1'$ be the subpath of $\mathbf{y}_1$ not containing the first or last edge. Note that both of these edges are q-edges corresponding to $\mathcal{Q}'$, so that $|\mathbf{y}_1| = |\mathbf{y}_1'| + 2$.

Then, Lemma 8.1(c) implies $|\partial\Delta_0| \leq |\mathbf{s}| + 1 + |\mathbf{top}(\mathcal{Q}_3)| + |\mathbf{y}_1'| + |\mathbf{top}(\mathcal{Q}_4)| + 1$. Further, Lemma 8.1(b) implies $|\mathbf{top}(\mathcal{Q}_i)| = \ell_i$ for $i = 3, 4$. So, $|\partial\Delta_0| \leq |\mathbf{s}| + |\mathbf{y}_1| + \ell_3 + \ell_4$.

Hence,

$$\begin{aligned} |\partial\Delta| - |\partial\Delta_0| \geq \gamma &= (|\mathbf{y}_3| - \ell_3) + (|\mathbf{y}_4| - \ell_4) - 4\delta \\ &\geq 2 - 4\delta + \delta \max(0, |\mathbf{y}_3|_a - \ell_3, |\mathbf{y}_3''|_a - 1) + \delta \max(0, |\mathbf{y}_4|_a - \ell_4, |\mathbf{y}_4''|_a - 1) \end{aligned}$$

So, taking $\delta^{-1} \geq 4$, $|\partial\Delta| - |\partial\Delta_0| \geq \gamma \geq 2 - 4\delta \geq 1$.

Hence, we may apply the inductive hypothesis to Δ_0, so that

$$\mathrm{wt}_G(\Delta_0) \leq N_2|\partial\Delta_0|^2 + N_1\mu(\Delta_0) \leq N_2(|\partial\Delta| - \gamma)^2 + N_1\mu(\Delta_0)$$

In $\mathbf{y}$, any θ-edge of $\mathbf{y}_1$ is separated from a θ-edge of $\mathbf{y}_3$ or $\mathbf{y}_4$ by a q-edge at the end of $\mathcal{Q}'$. Moreover, since the basic width of Γ is at most K_0, the parameter choice $J >> K_0$ implies that each of these (correctly ordered) pairs contribute to $\mu(\Delta)$. But the black beads corresponding to $\mathcal{Q}'$ are removed in the formation of the necklace for Δ_0, so that Lemma 8.3(d) implies

$$\mu(\Delta) - \mu(\Delta_0) \geq \ell'(\ell_3 + \ell_4)$$

Noting that $\gamma \leq |\partial\Delta|$, we then have:

$$\mathrm{wt}_G(\Delta_0) \leq N_2|\partial\Delta|^2 - N_2\gamma|\partial\Delta| + N_1\mu(\Delta) - N_1\ell'(\ell_3 + \ell_4)$$

Hence, by (9.10), it suffices to show that:

$$N_2\gamma|\partial\Delta| + N_1\ell'(\ell_3 + \ell_4) \geq \mathrm{wt}_G(\Gamma_2) + \mathrm{wt}_G(\Gamma_3) + \mathrm{wt}_G(\Gamma_4) + A + \ell' \tag{9.11}$$

Setting $\nu_i = |\partial\Gamma_i|_a$ for $i = 3, 4$, Lemma 9.8 implies:

$$\mathrm{wt}_G(\Gamma_i) \leq \mathrm{wt}(\Gamma_i) \leq c_0K_0\ell_i^2 + 2\nu_i\ell_i + C_1(K_0\ell_i + \nu_i)^2 \leq C_2(\ell_i + \nu_i)^2$$

By Lemmas 9.11,

$$\begin{aligned}\mathrm{wt}_G(\Gamma_2) &\leq C_2\ell'\max(\|\mathbf{t}\|, \|\mathbf{b}\|) + C_2(\|\mathbf{t}\| + \|\mathbf{b}\|)^2 \\ &\leq C_2\ell'(M + K_0) + 4C_2(M + K_0)^2\end{aligned}$$

For any a-cell π whose weight contributes to A, π is attached to either $\mathbf{t}''$ or $\mathbf{b}''$. Lemma 9.13 then implies that at least a third of the edges of $\partial\pi$ are shared with $\mathbf{t}''$ or $\mathbf{b}''$.

So, $A \leq C_1(3|\mathbf{t}''|_a + 3|\mathbf{b}''|_a)^2 \leq C_1(6M)^2 \leq 36C_1M^2$.

By Lemma 9.18 and 9.15, $M \leq 2K\ell \leq 4K\ell'$. So, since $C_2 >> C_1 >> K$, $A \leq C_2\ell'M$.

Hence, the parameter choices $C_3 >> C_2 >> K >> K_0$ imply

$$\begin{aligned}\mathrm{wt}_G(\Gamma_2) + A + \ell' &\leq C_2\ell'M + C_2K_0\ell' + 4C_2M^2 + 8C_2K_0M + 4C_2K_0^2 + C_2\ell'M + \ell' \\ &\leq 2C_2\ell'M + 16C_2K\ell'M + C_2K_0\ell' + 32C_2K_0K\ell' + \ell' + 4C_2K_0^2 \\ &\leq C_3\ell'M + C_3\ell' + C_3\end{aligned}$$

So, by (9.11), it suffices to show that:

$$N_2\gamma|\partial\Delta| + N_1\ell'(\ell_3 + \ell_4) \geq C_3\ell'M + C_3\ell' + C_3 + C_2(\ell_3 + \nu_3)^2 + C_2(\ell_4 + \nu_4)^2 \tag{9.12}$$

Without loss of generality, assume $\nu_4 \leq \nu_3$.

Note that

$$M = \max(|\mathbf{t}|_a, |\mathbf{b}|_a) = \max(|\mathbf{t}''|_a + |\mathbf{t}'|_a, |\mathbf{b}''|_a + |\mathbf{b}'|_a) \leq \max(|\mathbf{t}''|_a, |\mathbf{b}''|_a) + \max(|\mathbf{t}'|_a, |\mathbf{b}'|_a)$$

Since $\mathbf{t}'$ and $\mathbf{b}'$ are subpaths of $\partial\Gamma_3$ and $\partial\Gamma_4$, respectively, we then have $M \leq \max(|\mathbf{t}''|_a, |\mathbf{b}''|_a) + \nu_3$.

Lemma 9.17(1) then yields $M \leq \max(|\mathbf{y}_3''|_a, |\mathbf{y}_4''|_a) + \nu_3 + 4$.

So, $C_3\ell'M \leq C_3\ell' \max(|\mathbf{y}_3''|_a, |\mathbf{y}_4''|_a) + C_3\ell'\nu_3 + 4C_3\ell'$.

As $\gamma \geq 1$, $|\partial\Delta| \geq \ell'$, and $|\partial\Delta| \geq 2$, the parameter choice $N_2 >> C_3$ allows us to assume that

$$\frac{1}{3}N_2\gamma|\partial\Delta| \geq 5C_3\ell' + C_3$$

Hence, it suffices to show that:

$$\frac{2}{3}N_2\gamma|\partial\Delta|+N_1\ell'(\ell_3+\ell_4) \geq C_3\ell' \max(|\mathbf{y}_3''|_a, |\mathbf{y}_4''|_a)+C_3\ell'\nu_3+C_2(\ell_3+\nu_3)^2+C_2(\ell_4+\nu_4)^2 \quad (9.13)$$

Suppose $\max(|\mathbf{y}_3''|_a, |\mathbf{y}_4''|_a) \leq 1$. Then since $\gamma \geq 1$, $|\partial\Delta| \geq \ell'$, and $N_2 >> C_3$, we may take

$$\frac{1}{3}N_2\gamma|\partial\Delta| \geq C_3\ell' \max(|\mathbf{y}_3''|_a, |\mathbf{y}_4''|_a) \quad (9.14)$$

Otherwise, recall that

$$\gamma \geq \delta(|\mathbf{y}_3''|_a - 1) + \delta(|\mathbf{y}_4''|_a - 1) \geq \delta(\max(|\mathbf{y}_3''|_a, |\mathbf{y}_4''|_a) - 1) \geq \frac{1}{2}\delta \max(|\mathbf{y}_3''|_a, |\mathbf{y}_4''|_a)$$

So, since $|\partial\Delta| \geq \ell'$, the parameter choices $N_2 >> C_3 >> \delta^{-1}$ allow us to again assume (9.14) holds.

Thus, by (9.13), it suffices to show that

$$\frac{1}{3}N_2\gamma|\partial\Delta| + N_1\ell'(\ell_3+\ell_4) \geq C_3\ell'\nu_3 + C_2(\ell_3+\nu_3)^2 + C_2(\ell_4+\nu_4)^2 \quad (9.15)$$

1. Suppose $\nu_3 \leq 3J(\ell_3+\ell_4)$.

Then, for $i = 3, 4$, $\ell_i + \nu_i \leq 4J(\ell_3+\ell_4)$. As Lemma 9.15 implies $\ell_3 + \ell_4 \leq \ell'$, this implies $(\ell_i+\nu_i)^2 \leq 16J^2\ell'(\ell_3+\ell_4)$.

So, the parameter choices $C_3 >> C_2 >> J$ imply $C_2(\ell_3+\nu_3)^2+C_2(\ell_4+\nu_4)^2 \leq C_3\ell'(\ell_3+\ell_4)$.

Hence, as $N_1 >> C_3 >> J$, we have

$$N_1\ell'(\ell_3+\ell_4) \geq C_3\ell'\nu_3 + C_2(\ell_3+\nu_3)^2 + C_2(\ell_4+\nu_4)^2$$

Thus, we may assume that $\nu_3 > 3J(\ell_3+\ell_4)$.

As a result, for $i \in \{3,4\}$, $\ell_i + \nu_i \leq \ell_3 + \ell_4 + \nu_3 \leq 2\nu_3$. So, $C_2(\ell_i+\nu_i)^2 \leq 4C_2\nu_3^2$.

It then follows from (9.15) that it suffices to show that

$$\frac{1}{3}N_2\gamma|\partial\Delta| + N_1\ell'(\ell_3+\ell_4) \geq C_3\ell'\nu_3 + 8C_2\nu_3^2 \tag{9.16}$$

2. Suppose $\nu_3 \leq 16$.

So, $C_3\ell'\nu_3 + 8C_2\nu_3^2 \leq 8C_3\ell' + C_3$ by the parameter choice $C_3 >> C_2$.

As $|\partial\Delta| \geq \max(2, \ell')$ and $\gamma \geq 1$, the parameter choices $N_2 >> C_3 >> C_2$ allow us to assume that

$$\frac{1}{3}N_2\gamma|\partial\Delta| \geq C_3\ell'\nu_3 + 8C_2\nu_3^2$$

3. Thus, it suffices to assume that $\nu_3 > \max(3J(\ell_3+\ell_4), 16)$ and show that (9.16) holds.

As any a-edge of $\partial\Gamma_3$ is part of $\mathbf{y}_3'$, $\mathbf{t}'$, or $\mathbf{bot}(\mathcal{Q}_3)$, we have $\nu_3 \leq |\mathbf{y}_3|_a + |\mathbf{t}|_a + \ell_3$. By Lemma 9.17(2), this implies $\nu_3 \leq 2|\mathbf{y}_3|_a + 2K\ell_3 + \ell_3 + 4 \leq 2|\mathbf{y}_3|_a + J\ell_3 + 4$.

Note that $J\ell_3 + 4 < \frac{1}{3}\nu_3 + \frac{1}{4}\nu_3 = \frac{5}{12}\nu_3$, so that $\frac{7}{12}\nu_3 \leq 2|\mathbf{y}_3|_a$.

Recall that $\gamma \geq \delta(|\mathbf{y}_3|_a - \ell_3)$. So, since $\ell_3 < \frac{1}{3J}\nu_3 \leq \frac{1}{24}\nu_3$ by taking $J \geq 8$, we have $\gamma \geq \frac{1}{4}\delta\nu_3$.

As $|\partial\Delta| \geq \gamma$, we then have $\gamma|\partial\Delta| \geq \frac{1}{16}\delta^2\nu_3^2$. So, the parameter choices $N_2 >> C_2 >> \delta^{-1}$ allow us to assume

$$\frac{1}{6}N_2\gamma|\partial\Delta| \geq 8C_2\nu_3^2$$

By (9.16), it then suffices to show:

$$\frac{1}{6}N_2\gamma|\partial\Delta| + N_1\ell'(\ell_3+\ell_4) \geq C_3\ell'\nu_3 \tag{9.17}$$

But $|\partial\Delta| \geq \ell'$, so that the parameter choices $N_2 >> C_3 >> \delta^{-1}$ give us:

$$\frac{1}{6}N_2\gamma|\partial\Delta| \geq \frac{1}{24}N_2\delta\ell'\nu_3 \geq C_3\ell'\nu_3$$

Thus, (9.17) is satisfied, and so the statement is proved.

□

Chapter 10

Diagrams with disks

10.1 Diminished, Minimial, and D-minimal diagrams

A q-letter of the form $t(i)$ for $2 \leq i \leq L$ is called a *t-letter*. Accordingly, a (θ, q)-relation corresponding to a t-letter is called a *(θ, t)-relation*. Note that for each rule θ and each t-letter, the corresponding (θ, t)-relation is of the simple form $\theta_j t(i) = t(i)\theta_{j+1}$.

Now, we modify the definition of a reduced diagram over the canonical presentation of $M_\Omega(\mathbf{M})$ or over the disk presentation of $G_\Omega(\mathbf{M})$. To this end, we introduce the *signature* of such a diagram Δ as the four-tuple $s(\Delta) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ where:

- α_1 is the number of disks in Δ (of course, this is zero if Δ is a diagram over $M_\Omega(\mathbf{M})$),
  - α_2 is the number of (θ, t)-cells,
  - α_3 is the total number of (θ, q)-cells, and
  - α_4 is the number of a-cells.

The signatures of reduced diagrams over the disk presentation of $G_\Omega(\mathbf{M})$ are ordered lexicographically. In particular, if Δ and Γ are such diagrams with $s(\Delta) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and $s(\Gamma) = (\beta_1, \beta_2, \beta_3, \beta_4)$, then $s(\Delta) \leq s(\Gamma)$ if:

- $\alpha_1 \leq \beta_1$
  - for $i \in \{2, 3, 4\}$, if $\alpha_j = \beta_j$ for all $j < i$, then $\alpha_i \leq \beta_i$

A reduced diagram Δ over the disk presentation of $G_\Omega(\mathbf{M})$ is called *diminished* if for any reduced diagram Γ with $\text{Lab}(\partial\Delta) \equiv \text{Lab}(\partial\Gamma)$, we have $s(\Delta) \leq s(\Gamma)$.

Given a reduced diagram Δ over the disk presentation of $G_\Omega(\mathbf{M})$ with $s(\Delta) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, the *2-signature* of Δ is the ordered pair $s_2(\Delta) = (\alpha_1, \alpha_2)$. The *1-signature* $s_1(\Delta)$ is defined similarly and can be interpreted simply as the number of disks in Δ with the natural order on the natural numbers.

A reduced diagram Δ over the disk presentation of $G_\Omega(\mathbf{M})$ is called *D-minimal* if for any reduced diagram Γ with $\text{Lab}(\partial\Gamma) \equiv \text{Lab}(\partial\Delta)$, $s_1(\Delta) \leq s_1(\Gamma)$. By the definition of the lexicographic order, a diminished diagram is necessarily D-minimal.

Finally, reduced diagram Δ over the disk presentation of $G_\Omega(\mathbf{M})$ is called *minimal* if:

(M1) for any a-cell π and any θ-band $\mathcal{T}$, at most half of the edges of $\partial\pi$ mark the start of an a-band that crosses $\mathcal{T}$,

(M2) no maximal a-band ends on two different a-cells, and

(M3) for any reduced diagram Γ with $\text{Lab}(\partial\Delta) \equiv \text{Lab}(\partial\Gamma)$, $s_2(\Delta) \leq s_2(\Gamma)$.

Note that conditions (M1) and (M2) are equivalent to the conditions (MM1) and (MM2) in the definition of M-minimal. As a result, a minimal diagram containing no disks is necessarily M-minimal. Further, a diminished diagram necessarily satisfies (M3).

As with M-minimal diagrams, a subdiagram of a diminished (resp minimal, D-minimal) diagram is necessarily diminished (resp minimal, D-minimal).

In what follows, it is taken implicitly that any diminished, minimal, or D-minimal diagram over $G_\Omega(\mathbf{M})$ is formed over its disk presentation (rather than its canonical presentation).

Lemma 10.1. *A word W over $\mathcal{X}$ represents the trivial element of $M_\Omega(\boldsymbol{M})$ if and only if there exists a diminished diagram Δ over $M_\Omega(\boldsymbol{M})$ such that Lab$(\partial\Delta) \equiv W$ and Δ contains no θ-annuli.*

Proof. The reverse direction follows immediately from van Kampen's Lemma (see Chapter 2).

Let $\mathcal{S}_1$ be the set of relators defining the (θ, a)-relations of $M_\Omega(\mathbf{M})$, i.e the words $[\theta_i, a]$ for $\theta \in \Theta^+$ and $a \in Y_i(\theta)$. Similarly, let $\mathcal{S}_2$ be the set of relators defining the a-relations, $\mathcal{S}_3$ be the set of relators defining the (θ, q)-relations for the q-letters that are not t-letters, and $\mathcal{S}_4$ be the set of relators defining the (θ, t)-relations. Note that any cyclic permutation of an element of $\mathcal{S}_i^{\pm 1}$ is not an element of $\mathcal{S}_j$ for $j \neq i$. So, the partition of the relations given by $\mathcal{S}_1 \sqcup \cdots \sqcup \mathcal{S}_4$ defines a grading on the presentation of $M_\Omega(\mathbf{M})$ (see Section 3.4).

A reduced graded diagram with respect to this grading has minimal signature. Hence, for W a word over $\mathcal{X}$ representing the trivial element of $G_\Omega(\mathbf{M})$, the strengthened version of van Kampen's Lemma (Section 3.4) yields a diminished diagram Δ over $M_\Omega(\mathbf{M})$ with Lab$(\partial\Delta) \equiv W$.

Now suppose Δ contains a θ-annulus. As θ-bands cannot cross, the θ-annuli of Δ are partially ordered as in the proof of Lemma 9.3(2). Since Δ is finite, there exists a minimal θ-annulus S with respect to this partial order.

Let Δ_S be the subdiagram bounded by the outer contour of S and suppose $\Delta_S \setminus S$ contains a (θ, a)-cell π. Then, let S' be the maximal θ-band containing π. Since θ-bands cannot cross, S' must be a θ-annulus contained in $\Delta_S \setminus S$. But this contradicts the minimality of S.

So, since Lemma 9.3(1) implies that Δ_S contains no (θ, q)-cells, $\Delta_S \setminus S$ consists entirely of a-cells. Hence, Lab$(\partial\Delta_S) \equiv$ Lab$(\partial(\Delta_S \setminus S))$ is trivial in the group $B(\mathcal{A}, n)$.

As a result, we may form the reduced diagram Γ by excising Δ_S from Δ, pasting a single a-cell in its place, and making any necessary cancellations. Note that Lab$(\partial\Gamma) \equiv$ Lab$(\partial\Delta)$ and $s(\Gamma) \leq s(\Delta)$, so that Γ must again be diminished. However, the number of θ-annuli in Γ is one less than the number in Δ.

Iterating this process, we remove all θ-annuli in Δ, producing a diminished diagram satisfying the statement.

□

Lemma 10.2. *Every diminished diagram satisfies (M2).*

Proof. Arguing toward a contradiction, let π_1 and π_2 be two a-cells in a diminished diagram Δ connected by an a-band. Let Δ_0 be the subdiagram consisting of π_1, π_2, and this a-band (see Figure 10.1).

As an a-band consists only of (θ, a)-cells, the top and bottom of the a-band have equivalent labels visually equal to a word $H \in F(R)$.

So, Lab$(\partial\Delta_0) \equiv uHvH^{-1}$ for some words $u, v \in F(\mathcal{A})$. Note that for any rule θ corresponding to a letter of H, the makeup of the a-band implies the existence of a (θ, a)-relation corresponding to θ and an a-letter from the 'special' input sector. This then implies that the domain of θ in the 'special' input sector is nonempty, which in turn implies that the domain of θ in this sector is the entire alphabet.

As a result, we can build a reduced annular diagram Γ' over the canonical presentation of $M(\mathbf{M})$ with outer label $uHvH^{-1}$, inner label uv, and made up entirely of (θ, a)-cells. Then, since Δ_0 is a diagram over $M_\Omega(\mathbf{M})$, we have $uv = 1$ in $M_\Omega(\mathbf{M})$.

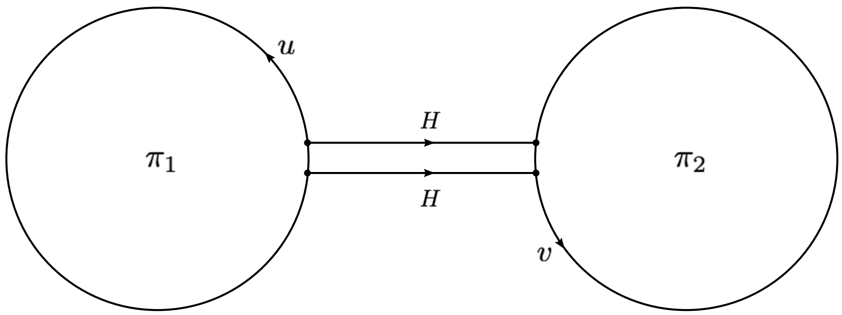


Figure 10.1: The subdiagram Δ_0

Let Ψ be the diminished diagram over $M_\Omega(\mathbf{M})$ with $\text{Lab}(\partial\Psi) \equiv uv$ given by Lemma 10.1. Since $\partial\Psi$ has no θ-edges and Ψ has no θ-annuli, Ψ must consist only of a-cells.

By van Kampen's Lemma, we then have $uv = 1$ over $B(\mathcal{A}, n)$, so that Ψ must consist of exactly one a-cell by the minimality of its signature. Pasting Ψ into the middle of Γ' then yields a reduced diagram Γ_0 over $M_\Omega(\mathbf{M})$ with contour label $uHvH^{-1}$.

Let Γ be the reduced diagram obtained from Δ by excising Δ_0, pasting Γ_0 in its place, and making any necessary cancellations. Then Γ has the same contour label as Δ, one less a-cell, and the same number of disks, (θ, t)-cells, and (θ, q)-cells. Hence, $s(\Gamma) < s(\Delta)$, contradicting the assumption that Δ is diminished.

□

Note that Lemma 10.2 implies that a diminished diagram satisfying (M1) is minimal.

10.2 t-spokes

When considering diminished, minimal, or D-minimal diagrams in what follows, many arguments rely on the q-bands corresponding to t-letters. To distinguish these from bands corresponding to other parts of the base, we adopt the convention of [18] and [25] and refer to them as t-*bands*. Note that the top and the bottom of such a band are each labelled by a copy of the band's history.

In a diminished, minimal, or D-minimal diagram, a maximal q-band with one end on a disk Π is called a *spoke* of Π. A t-*spoke* is then defined in the natural way.

The pairs $\{t(2), t(3)\}, \dots, \{t(L-1), t(L)\}, \{t(L), t(2)\}$ are called *adjacent* t-letters. Two t-spokes of the same disk are called *consecutive* if they correspond to adjacent t-letters.

Lemma 10.3. *For $i \in \{2, \dots, L\}$, let $\mathcal{C} : A(i) \to \cdots \to A(i)$ be a reduced computation of $\boldsymbol{M}$ with history H. Then there exists a reduced diagram Δ over $M_\Omega(\boldsymbol{M})$ with contour label $H(0)^{-1} W_{ac} H(0) W_{ac}^{-1}$, where $H(0)$ is the copy of H in $F(R)$ obtained by adding the subscript 0 to each letter.*

Proof. Consider the factorization $H \equiv H_1 \cdots H_\ell$ for $\ell \geq 2$ given by Lemma 6.12.

Define $H_i(0)$ as the word in $F(R)$ obtained from H_i by adding a subscript 0 to each letter. By Lemma 7.5, for each $1 \leq j \leq \ell$, there exists a trapezium Δ_j with contour label

$$H_j(0)^{-1} W_{j-1}^{(z_j)} H_j(0) (W_j^{(z_j)})^{-1}$$

where $W_j^{(z_j)}$ is defined as in Lemma 6.12.

Recall that for $1 \leq j \leq \ell - 1$, $W_j^{(z_j)}$ differs from $W_j^{(z_{j+1})}$ only by the insertion/deletion of words in $\mathcal{L}$ in the 'special' input sector, while $W_0^{(z_1)} \equiv W_\ell^{(z_\ell)} \equiv W_{ac}$. Note that every word of $\mathcal{L}$ represents the trivial element of $B(\mathcal{A}, n)$, so that $\mathcal{L} \subset \Omega$. For $1 \leq j \leq \ell - 1$, let $\tilde{\Delta}_j$ be the diagram obtained from pasting the a-cell corresponding to this element of $\mathcal{L}$ to the top of Δ_j, so that the 'top' label of $\tilde{\Delta}_j$ is $W_j^{(z_{j+1})}$.

Then, letting $\tilde{\Delta}_\ell = \Delta_\ell$, we may glue the top of $\tilde{\Delta}_j$ to the bottom of $\tilde{\Delta}_{j+1}$. Letting Δ be the reduced diagram that results from these pastings, $\text{Lab}(\partial\Delta) \equiv H(0)^{-1} W_{ac} H(0) W_{ac}^{-1}$.

□

Lemma 10.4. *Let $\mathcal{C} : V_0 \to \cdots \to V_t$ be a reduced computation of $\boldsymbol{M}$ with history H and base $\{t(i)\} B_3(i)$ for some $i \in \{2, \dots, L\}$. Suppose there exists an accepted configuration W_0 such* *that $W_0(i) \equiv V_0$. Then there exists an accepted configuration W_t with $W_t(i) \equiv V_t$ and a reduced diagram Δ over $M_\Omega(\boldsymbol{M})$ with contour label $H(0)^{-1}W_0H(0)W_t^{-1}$, where $H(0)$ is the copy of H in $F(R)$ obtained by adding the subscript 0 to each letter.*

Proof. Let $H \equiv H_1 \dots H_\ell$ be the factorization such that each H_j is the history of a maximal one-machine subcomputation.

For $1 \leq j \leq \ell$, let $\mathcal{C}_j : V_{y(j)} \to \dots \to V_{z(j)}$ be the subcomputation with history H_j. Then, let $\mathcal{C}_j' : W_{y(j)}' \to \dots \to W_{z(j)}'$ be the reduced computation in the standard base given by Lemma 6.9. As in the proof of Lemma 6.19, note that $W_{z(j)}'$ may differ from $W_{y(j+1)}'$.

1. Suppose W_0 is H_1-admissible and $\ell = 1$.

Then, there exists a reduced computation $\mathcal{C}'$ with initial configuration W_0 and history $H \equiv H_1$. So, we may let $W_t \equiv W_0 \cdot H$ and Δ be the trapezium corresponding to $\mathcal{C}'$ from Lemma 7.5.

2. Suppose W_0 is H_1-admissible and $\ell \geq 2$.

Without loss of generality, we may assume that $W_{y(1)}' \equiv W_0$, i.e $\mathcal{C}_1'$ is a reduced computation with initial configuration W_0 and history H_1.

Then $W_{z(1)}'$ is an accepted start or end configuration. By the construction of the computation given in the proof of Lemma 6.9, $W_{y(2)}'$ is also an accepted configuration. Continuing in this way, $W_{z(\ell)}'$ is an accepted configuration with $W_{z(\ell)}'(i) \equiv V_t$, so that we may let $W_t \equiv W_{z(\ell)}'$.

For $1 \leq j \leq \ell$, let Δ_j be the trapezium corresponding to $\mathcal{C}_j'$ given by Lemma 7.5. Then, for $1 \leq j \leq \ell - 1$, Lab(**top**(Δ_j)) $\equiv W_{z(j)}'$ and Lab(**bot**(Δ_{j+1})) $\equiv W_{y(j+1)}'$ differ only by the insertion of a word from $\mathcal{L}$ in the 'special' input sector. So, we may paste Δ_j to Δ_{j+1} along an a-cell corresponding to this difference.

The reduced diagram Δ arising from these pastings then satisfies the statement.

3. Suppose W_0 is not H_1-admissible.

Let $H \equiv H_1'H_1''$ such that H_1' is the maximal (perhaps empty) prefix for which W_0 is H_1'-admissible. Then for θ the first letter of H_1'', $(W_0 \cdot H_1')(i)$ is θ-admissible while $W_0 \cdot H_1'$ is not. Lemma 6.13 then implies the following:

(1) If $\mathcal{C}_1$ is a one-machine computation of the first machine, then there exists $w \in \mathcal{L}$ such that $W_0 \cdot H_1' \equiv J(w) \cdot H'$, where H' is the natural copy of w read right to left in the language of positive rules with step history $(1)_1$

(2) If $\mathcal{C}_1$ is a one-machine computation of the second machine, then there exists $w \in \mathcal{L}$ such that $W_0 \cdot H_1' \equiv I(w)$.

Let $\|H_1'\| = r$. Then by the construction given in Lemma 6.9, $W_r' \equiv I(w) \cdot H'$ in case (1) or $W_r' \equiv J(w)$ in case (2). Either way, Lemma 6.6 implies W_r' is an accepted configuration. As a result, every configuration of in $\mathcal{C}_1'$ is accepted.

As in Step 2, as $W_{z(1)}'$ is accepted, $W_{y(2)}'$ (and so each configuration of $\mathcal{C}_2'$) is accepted. Continuing, this implies $W_{z(\ell)}'$ is accepted with $W_{z(\ell)}'(i) \equiv V_t$, so that we may let $W_t \equiv W_{z(\ell)}'$.

Let Δ_1' be the trapezium corresponding to the reduced computation with initial configuration W_0 and history H_1'. Further, let Δ_1'' be the trapezium corresponding to the reduced computation with initial configuration W_r' and history H_1''.

Then, Lab($\mathbf{top}(\Delta_1')) \equiv W_0 \cdot H_1'$ and Lab($\mathbf{bot}(\Delta_1'')) \equiv W_r'$ differ by the insertion/deleetion of an element of $\mathcal{L}$ in the 'special' input sector. Hence, we may construct a reduced diagram Δ_1 by pasting Δ_1' to Δ_1'' along an a-cell corresponding to this difference.

As in previous steps, we may then construct the reduced diagram Δ satisfying the statement.

□

Lemma 10.5. *Let Δ be a D-minimal diagram over the disk presentation of $G_\Omega(\boldsymbol{M})$. Suppose there exist two disks Π_1 and Π_2 in Δ so that $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are consecutive t-spokes of both. Let Ψ be the subdiagram bounded by the sides of $\mathcal{Q}_i$ and the subpaths of $\partial\Pi_i$ such that Ψ does not contain Π_1 or Π_2. Then Ψ contains a disk.*

Proof. Assume that Π_1 and Π_2 are hubs. Note that if either of these two bands has zero length, then the two hubs are cancellable, contradicting the assumption that Δ is reduced.

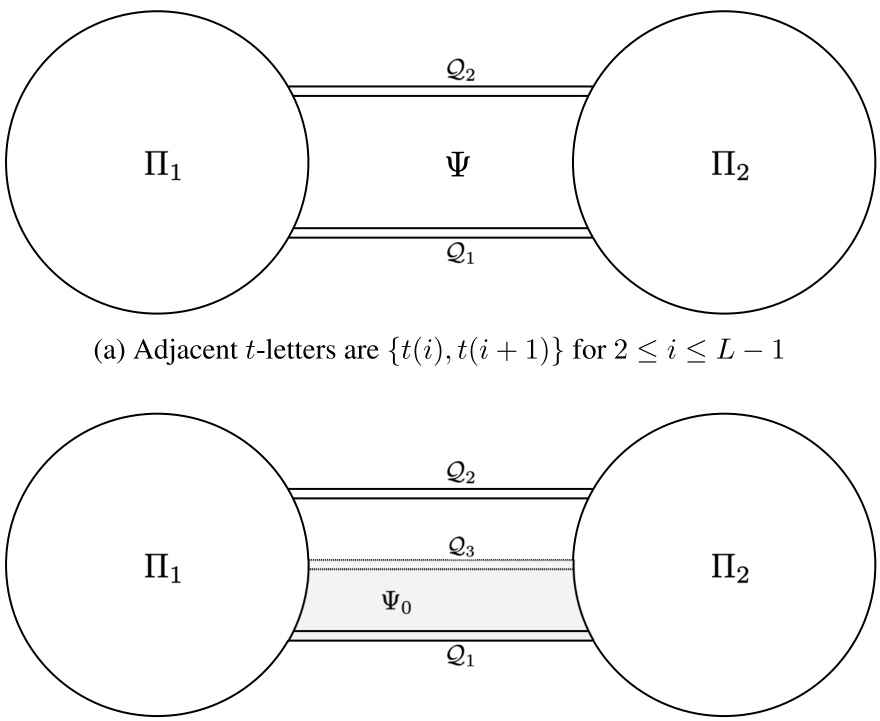
Arguing toward contradiction, suppose Ψ contains no disk.

First, suppose the pair of adjacent t-letters corresponding to $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is $\{t(i), t(i+1)\}$ for some $2 \leq i \leq L-1$ (see Figure 10.2(a)).

By Lemma 10.1, there exists a diminished diagram Λ over $M_\Omega(\mathbf{M})$ with $\text{Lab}(\partial\Lambda) \equiv \text{Lab}(\partial\Psi)$.

Suppose there exists an a-cell π in Λ. Note that no edge of $\partial\Lambda$ is labelled by an a-letter from the 'special' input sector. So, by Lemmas 9.2 and 10.2, any a-band starting on $\partial\pi$ must end on a (θ, q)-cell in Λ. Further, by Lemma 9.1, the maximal q-band $\mathcal{Q}$ containing this (θ, q)-cell must have two ends on $\partial\Lambda$. But the definition of the rules implies that $\mathcal{Q}$ corresponds to $Q_0(1)^{\pm 1}$ while no q-edge of $\partial\Lambda$ corresponds to such a base letter.

Hence, Λ is a reduced diagram over $M(\mathbf{M})$, so that Lemma 7.1 implies that Λ is a trapezium with top and bottom labels $A(i)\{t(i+1)\}$ (up to inversion). By Lemma 7.4, there exists a corresponding computation $\mathcal{C} : A(i) \to \cdots \to A(i)$ with history H. Thus, Lemma 10.3 yields a reduced diagram Γ_1 over $M_\Omega(\mathbf{M})$ with contour label $H(0)^{-1}W_{ac}H(0)W_{ac}^{-1}$.



(a) Adjacent t-letters are $\{t(i), t(i+1)\}$ for $2 \leq i \leq L-1$

(b) Adjacent t-letters are $\{t(L), t(2)\}$

Figure 10.2: Lemma 10.5

The subdiagram Γ_0 of Γ_1 bounded by the two t-bands corresponding to $t(i)$ and $t(i+1)$ has the same contour label as Λ, and so as Ψ. By cyclic permutation, we may assume that Γ_0 can be cut from Γ_1 to produce a reduced diagram Γ over $M_\Omega(\mathbf{M})$.

Let Ψ' be the smallest subdiagram of Δ containing Π_1, Π_2, $\mathcal{Q}_1$, and $\mathcal{Q}_2$. Then $\text{Lab}(\partial\Psi') \equiv \text{Lab}(\partial\Gamma)$. Hence, excising Ψ' from Δ and pasting Γ in its place reduces the number of hubs (and so disks) by two, contradicting the assumption that Δ is D-minimal.

Now suppose the adjacent t-letters corresponding to $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are $t(L)$ and $t(2)$, respectively. Then, the q-band $\mathcal{Q}_3$ starting on Π_1 corresponding to the part $\{t(1)\}$ must end on Π_2. Let Ψ_0 be the subdiagram of Ψ bounded by $\mathcal{Q}_1$ and $\mathcal{Q}_3$ (see Figure 10.2(b)).

Let Λ_0 be a reduced diagram over $M_\Omega(\mathbf{M})$ given by Lemma 10.1, so that Λ_0 satisfies (**M2**) and $\text{Lab}(\partial\Lambda_0) \equiv \text{Lab}(\partial\Psi_0)$. Then, as above, Λ_0 must be a trapezium with top and bottom labels $A(L)\{t(1)\}$. Hence, Lemma 7.4 gives a corresponding computation $\mathcal{C} : A(L) \to \cdots \to A(L)$, so that we may repeat the argument above to reduce the number of hubs.

Finally, suppose Π_1 and Π_2 are not necessarily hubs. Then we can replace these disks with reduced diagrams over $G(\mathbf{M})$ consisting of a hub and a trapezium (as formed in Lemma 8.2). Let $\tilde{\Delta}$ be the resulting reduced diagram, $\tilde{\Pi}_1$ and $\tilde{\Pi}_2$ be the two hubs, $\tilde{\mathcal{Q}}_1$ and $\tilde{\mathcal{Q}}_2$ be the consecutive t-spokes at these hubs, and $\tilde{\Psi}$ be the subdiagram bounded by the sides of $\tilde{\mathcal{Q}}_i$ and the contours of $\tilde{\Pi}_i$.

As $\tilde{\Delta}$ and Δ have the same number of disks and the same contour labels, $\tilde{\Delta}$ is D-minimal. Thus, the same arguments as outlined above can be applied to remove $\tilde{\Pi}_1$ and $\tilde{\Pi}_2$, yielding a contradiction.

□

For each reduced diagram Δ over the disk presentation of $G_\Omega(\mathbf{M})$, there is a corresponding planar graph $\Gamma \equiv \Gamma(\Delta)$ defined by:

(1) $V(\Gamma) = \{v_0, v_1, \dots, v_\ell\}$ where each v_i for $i \geq 1$ corresponds to one of the ℓ disks of Δ and v_0 is one exterior vertex

(2) For $i, j \geq 1$, each shared t-spoke of the disks corresponding to v_i and v_j corresponds to an

edge $(v_i, v_j) \in E(\Gamma)$

(3) For $i \geq 1$, each t-spoke of the disk corresponding to v_i which ends on $\partial\Delta$ corresponds to an edge $(v_0, v_i) \in E(\Gamma)$

Note that the degree of each interior vertex of Γ is $L-1$. The following statement is a consequence of this fact and Lemma 10.5.

Lemma 10.6. *(Lemma 3.2 of [17]) Suppose Δ is a D-minimal diagram containing at least one disk. Then Δ contains a disk Π such that $L-4$ consecutive t-spokes $\mathcal{Q}_1, \dots, \mathcal{Q}_{L-4}$ of Π end on $\partial\Delta$ and such that every subdiagram Γ_i bounded by $\mathcal{Q}_i$, $\mathcal{Q}_{i+1}$, $\partial\Pi$, and $\partial\Delta$ ($i = 1, \dots, L-5$) contains no disks.*

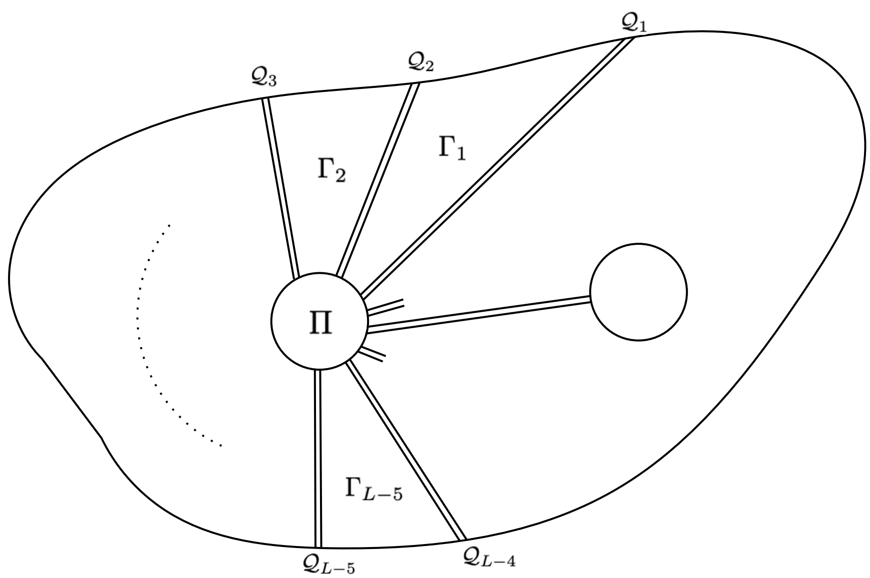


Figure 10.3: Lemma 10.6

10.3 Transposition of a θ-band and a disk

We now describe a procedure, similar to the construction in Section 9.3, for moving a θ-band about a disk.

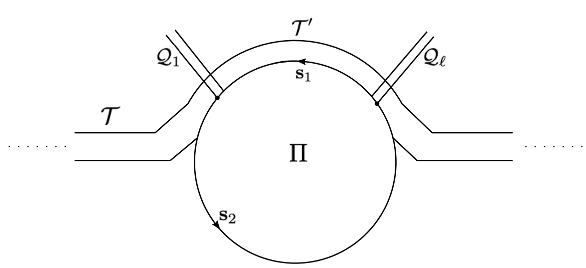
Let Δ be a D-minimal diagram containing a disk Π and a θ-band $\mathcal{T}$ subsequently crossing the t-spokes $\mathcal{Q}_1, \dots, \mathcal{Q}_\ell$ of Π. Assume $\ell \geq 2$ is maximal for Π and $\mathcal{T}$.

First, suppose there are no other cells between Π and the bottom of $\mathcal{T}$, i.e there is a subdiagram formed by Π and $\mathcal{T}$.

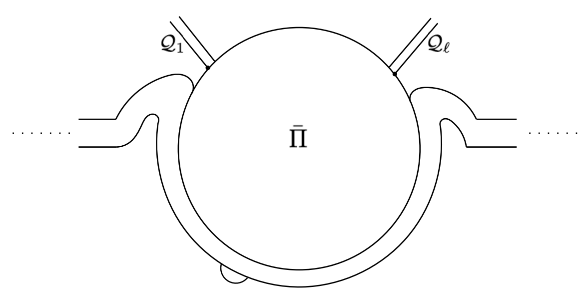
Let $\mathcal{T}'$ be the subband of $\mathcal{T}$ whose bottom path, $\mathbf{s}_1^{-1}$, starts with the t-edge corresponding to the start of $\mathcal{Q}_1$ and ends with that of $\mathcal{Q}_\ell$. Further, let $\mathbf{s}_2$ be the complement of $\mathbf{s}_1$ in $\partial\Pi$ so that $\partial\Pi = \mathbf{s}_1\mathbf{s}_2$. Then as any sector of the standard base containing a t-letter has empty tape alphabet, $\text{Lab}(\mathbf{s}_2)$ is an admissible word.

Let $W \equiv \text{Lab}(\partial\Pi)^{\pm 1}$, $V \equiv \text{Lab}(\mathbf{s}_1)$, and θ be the rule corresponding to $\mathcal{T}$. Further, let Γ be the subdiagram formed by Π and $\mathcal{T}'$. Then, by Lemma 7.2, V^{-1} is θ-admissible with $V^{-1} \cdot \theta \equiv \text{Lab}(\mathbf{ttop}(\mathcal{T}')) = \text{Lab}(\mathbf{top}(\mathcal{T}'))$

Suppose $\text{Lab}(\mathbf{s}_2)$ is θ-admissible. Then W is θ-admissible, so that $W \cdot \theta$ is a disk relator. Let $\bar{\Pi}$ be a disk with contour labelled by $W \cdot \theta$. Let $\mathcal{T}''$ be the auxiliary θ-band corresponding to θ whose top is labelled by $\text{Lab}(\mathbf{s}_2) \cdot \theta$. Then, let $\bar{\Gamma}$ be the diagram obtained from attaching $\mathcal{T}''$ to $\bar{\Pi}$. Finally, let $\bar{\Delta}$ be the reduced diagram obtained from excising Γ from Δ and pasting $\bar{\Gamma}$ in its place, attaching the first and last cells of $\mathcal{T}''$ to the complement of $\mathcal{T}'$ in $\mathcal{T}$ and perhaps making cancellations in the resulting θ-band. Note that $\bar{\Delta}$ has the same contour label as that of Δ.



(a) The subdiagram Γ



(b) The resulting subdiagram $\bar{\Gamma}$

Figure 10.4: The transposition of a θ-band with a disk

Conversely, suppose $\text{Lab}(\mathbf{s}_2)$ is not θ-admissible. Then Lemma 6.13 applies to W, so that $\text{Lab}(\mathbf{s}_2)$ contains the 'special' input sector and would be θ-admissible with the insertion/deletion of some $u^n \in \mathcal{L}$. So, after attaching to Π an a-cell corresponding to u^n, we may construct the disk $\bar{\Pi}$ and the auxiliary θ-band $\mathcal{T}''$ as above. Attaching the mirror a-cell on the other side of $\mathcal{T}''$ then produces a diagram $\bar{\Delta}$ with the same contour label as Δ.

The procedure of excising Γ from Δ to create $\bar{\Delta}$ is called the *transposition* of the disk Π and the θ-band $\mathcal{T}$ in Δ.

Now, consider the situation where there are cells between the θ-band and the disk, each of which is an a-cell.

Suppose the pair of adjacent t-letters corresponding to $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is $\{t(i), t(i+1)\}$ for some $2 \leq i \leq L-1$. Let $\mathcal{T}_1'$ be the subband of $\mathcal{T}'$ between $\mathcal{Q}_1$ and $\mathcal{Q}_2$. Then, let Ψ be the subdiagram of Δ bounded by $\mathcal{T}_1'$ and $\partial\Pi$. By Lemma 10.1, there exists a diminished diagram Λ over $M_\Omega(\mathbf{M})$ with Lab$(\partial\Psi) \equiv$ Lab$(\partial\Lambda)$. Lemmas 9.2 and 10.2 then imply that Λ contains no a-cell, so that Lemma 7.1 implies that Λ consists of a single θ-band. Hence, by Lemma 7.2, $W(i)$ is θ-admissible.

Otherwise, if the pair of adjacent t-letters is $\{t(L), t(2)\}$, then the same argument applies to the subdiagram bounded by the t-band corresponding to $t(L)$, the q-band corresponding to $t(1)$, and $\mathcal{T}'$. As a result, $W(L)$ is θ-admissible.

As above, Lemma 6.13 then implies that, perhaps after attaching an a-cell, we may construct a new disk and auxiliary band that, perhaps after attaching another a-cell, functions as the transposition of Π with $\mathcal{T}$.

The reduced diagram Δ' arising from the transposition has the number of disks and contour label as Δ, and so is D-minimal.

However, the minimality of the 2-signature (and so the signature) need not be preserved by a transposition. This is because many (θ, t)-cells may be added through transposition.

Note that the definition of transposition above differs from that in [18] and [25] only by the presence of a-cells.

Lemma 10.7. *(Compare with Lemma 7.5 of [18] and 7.7 of [25]) Let Δ be a reduced diagram over the disk presentation of $G_\Omega(\boldsymbol{M})$ satisfying (M3).*

(1) Suppose a θ-band $\mathcal{T}$ crosses ℓ t-spokes of a disk Π and there are no disks in the subdiagram bounded by these spokes, $\mathcal{T}$, and $\partial\Pi$. Then $\ell \leq (L-1)/2$.

(2) Suppose $\mathcal{T}$ and $\mathcal{T}'$ are disjoint θ-bands crossing ℓ and ℓ' t-spokes, respectively, of a disk Π.

Suppose further that every cell between the bottom of $\mathcal{T}$ (of $\mathcal{T}'$) and Π is an a-cell. Further, suppose these bands correspond to the same rule θ if the history is read toward the disk. Then $\ell + \ell' \leq (L-1)/2$.

(3) *If S is a θ-annulus in Δ and Δ_S is the subdiagram bounded by the outer contour of S, then Δ_S is a diagram over $M_\Omega(\boldsymbol{M})$.*

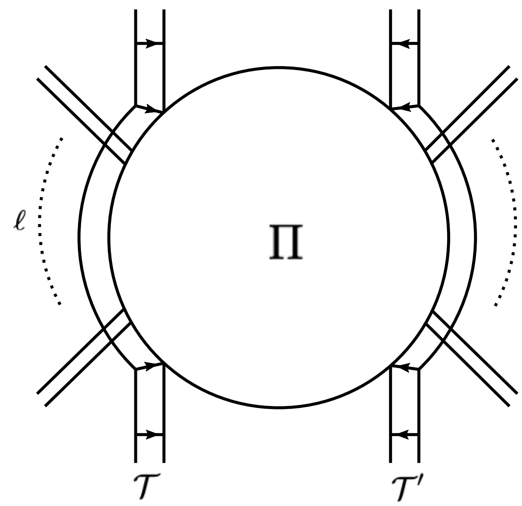


Figure 10.5: Lemma 10.7(2)

Proof. (1) Lemma 9.1 implies that there exists a θ-band $\mathcal{T}_0$ crossing all ℓ spokes such that the only cells between it and Π are a-cells. If $\ell > (L-1)/2$, then the transposition of Π and $\mathcal{T}_0$ in Δ then yields a diagram with the same contour label, the same number of disks, and strictly less (θ, t)-cells. This contradicts the minimality of $s_2(\Delta)$.

(2) The transposition of $\mathcal{T}$ and Π removes ℓ (θ, t)-cells and adds $(L-1) - \ell$ new (θ, t)-cells in the resulting band. However, ℓ' of these cells form cancellable pairs with cells of $\mathcal{T}'$, so that it is possible to cancel $2\ell'$ cells. Hence, the change in the number of (θ, t)-cells is $(L-1) - 2\ell - 2\ell'$, so that the relation $\ell + \ell' > (L-1)/2$ would contradict the minimality of $s_2(\Delta)$.

(3) Suppose Δ_S contains a disk. Then, since Δ is D-minimal, Lemma 10.6 gives a disk Π in Δ_S with $L-4$ consecutive t-spokes that end on $\partial\Delta_S$ and such that the subdiagram of Δ_S bounded by these spokes contains no disks. But then taking $L > 7$, S and Π contradict (1).

□

The following is an immediate consequence of Lemmas 9.3(2) and 10.7(3).

Lemma 10.8. *A minimal diagram Δ contains no θ-annuli.*

The following statement gives a strengthened version of van Kampen's Lemma for $G_\Omega(\mathbf{M})$, specifically for minimal diagrams.

Lemma 10.9. *A word W over $\mathcal{X}$ represents the trivial element of $G_\Omega(\boldsymbol{M})$ if and only if there exists a minimal diagram Δ such that Lab$(\partial\Delta) \equiv W$.*

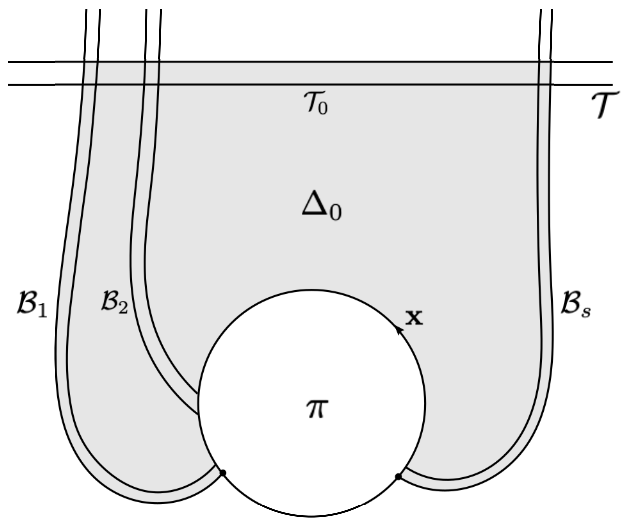


Figure 10.6: Lemma 10.9

Proof. As in the proof of Lemma 10.1, the reverse direction is an immediate consequence of van Kampen's Lemma.

Let $\mathcal{S}_5$ be the set of words defining the disk relations of $G_\Omega(\mathbf{M})$. Then, letting $\mathcal{S}_1, \dots, \mathcal{S}_4$ be as defined in the proof of Lemma 10.1, the partition of the relations $\mathcal{S}_1 \sqcup \dots \sqcup \mathcal{S}_5$ defines a grading on the disk presentation of $G_\Omega(\mathbf{M})$.

By the definition of the grading, a reduced graded diagram has minimal signature. So, the strengthened version of van Kampen's Lemma implies that for any word W representing the trivial element of $G_\Omega(\mathbf{M})$, there exists a diminished diagram Δ with Lab$(\partial\Delta) \equiv W$.

Suppose Δ is not minimal. As Δ is diminished, Lemma 10.2 implies that it satisfies (M2). So, Δ must not satisfy (M1), i.e it contains an a-cell π and a θ-band $\mathcal{T}$ such that for some $s > \frac{1}{2}\|\partial\pi\|$,

s maximal a-bands start on $\partial\pi$ and cross $\mathcal{T}$. Without loss of generality, we assume that π and $\mathcal{T}$ are chosen so that $s/\|\partial\pi\|$ is maximal amongst such pairs.

Enumerate these a-bands $\mathcal{B}_1, \dots, \mathcal{B}_s$ based on where they cross $\mathcal{T}$ and let Δ_0 be the subdiagram containing each of these s bands which is bounded by a side of $\mathcal{B}_1$, a side of $\mathcal{B}_s$, a subpath **x** of $\partial\pi$, and the top of a subband $\mathcal{T}_0$ of $\mathcal{T}$ (see Figure 10.6).

As a-bands consist only of (θ, a)-cells, the sides of $\mathcal{B}_1$ and $\mathcal{B}_s$ consist only of θ-edges. So, any q-edge of $\partial\Delta_0$ must be part of the top of $\mathcal{T}_0$.

Suppose Δ_0 contains a disk. By Lemma 10.6, there exists a disk Π in Δ_0 with at least $L-4$ t-spokes ending on $\textbf{top}(\mathcal{T}_0)$ such that there are no disks in the subdiagram bounded by these spokes. But then the parameter choice $L > 7$ means that $\mathcal{T}_0$ and Π form a counterexample to Lemma 10.7(1).

So, as Δ_0 contains no disks, any q-edge of $\partial\Delta_0$ must mark the start of a maximal q-band which has two ends on $\textbf{top}(\mathcal{T}_1)$. But then this q-band bounds a (θ, q)-annulus with some subband of $\mathcal{T}_0$, contradicting Lemma 9.1(1). So, Lemma 9.1(4) implies Δ_0 contains no (θ, q)-cells.

As Δ_0 satisfies (M2), each edge of **x** is the start of an a-band which ends on the top of $\mathcal{T}_0$, so that this a-band crosses $\mathcal{T}$. Hence, **x** consists entirely of the s edges of $\partial\pi$ marking the start of $\mathcal{B}_1, \dots, \mathcal{B}_s$.

By Lemmas 9.1 and 9.3, any maximal θ-band $\mathcal{T}_0'$ of Δ_0 connects the side of $\mathcal{B}_1$ to the side of $\mathcal{B}_s$, so that all s a-bands must cross this θ-band. Letting $\mathcal{T}'$ be the maximal θ-band of Δ containing $\mathcal{T}_0'$, the maximality of $s/\|\partial\pi\|$ implies that $\mathcal{B}_1, \dots, \mathcal{B}_s$ comprise all maximal a-bands starting on $\partial\pi$ and crossing $\mathcal{T}'$. So, we may pass to $\mathcal{T}'$, assuming that $\mathcal{T}' = \mathcal{T}$ is the θ-band chosen above. As a result, $\mathcal{T}_0$ is the only maximal θ-band of Δ_0.

So, any cell between $\textbf{bot}(\mathcal{T}_0)$ and **x** must be an a-cell. Supposing such an a-cell exists, property (M2) implies that each of the edges on its boundary marks the start of an a-band that must cross $\mathcal{T}_0$, forming another counterexample to (M1). The maximality of $s/\|\partial\pi\|$ then implies $s = \|\partial\pi\|$. Since a-bands cannot cross, we may find a 'minimal' counterexample, i.e an a-cell with no cells between it and $\textbf{bot}(\mathcal{T}_0)$. Passing to this cell, we may assume without loss of generality that $\mathbf{x}^{-1} =$

$\mathbf{bot}(\mathcal{T}_0)$.

As a result, π may be transposed with $\mathcal{T}$ to produce a reduced diagram $\tilde{\Delta}$ in which at most $\|\partial\pi\| - s < \frac{1}{2}\|\partial\pi\|$ maximal a-bands start on $\partial\pi$ and cross the maximal θ-band arising from $\mathcal{T}$. Since the rest of the diagram remains unchanged throughout this process, $\tilde{\Delta}$ is diminished and contains one less counterexample to property (M1). Hence, iterating the process eliminates any a-cell and θ-band violating property (M1), producing the desired minimal diagram.

□

Remark 10.1. *The proof of Lemma 10.9 corresponds to the following sharper statement: A word W over $\mathcal{X}$ represents the trivial element of $G_\Omega(\boldsymbol{M})$ if and only if there exists a diminished diagram Δ satisfying (M1) such that Lab$(\partial\Delta) \equiv W$. However, the statement above suffices for our purposes.*

Lemma 10.9 immedately implies the following strengthened version of van Kampen's Lemma for M-minimal diagrams.

Lemma 10.10. *A word W over $\mathcal{X}$ represents the trivial element of $M_\Omega(\boldsymbol{M})$ if and only if there exists an M-minimal diagram Δ such that Lab$(\partial\Delta) \equiv W$.*

10.4 Quasi-trapezia

Next, the concept of trapezium is generalized to the setting of minimal diagrams over $G_\Omega(\mathbf{M})$.

A *quasi-trapezium* is a minimal diagram defined in much the same way as an a-trapezium (see Section 9.4) except that it is permitted to contain disks. In other words, a quasi-trapezium is a minimal diagram whose boundary can be factored as $\mathbf{p}_1^{-1}\mathbf{q}_1\mathbf{p}_2\mathbf{q}_2^{-1}$, where each $\mathbf{p}_i$ is the side of a q-band and each $\mathbf{q}_i$ is the maximal subpath of the side of a θ-band where the subpath starts and ends with a q-edge.

The *(step) history* of a quasi-trapezium is defined in the same way as for an a-trapezium, as are the *base*, the *height*, and the *standard factorization*.

Note that a quasi-trapezium containing no disks is an a-trapezium, while one without any disks or a-cells is a trapezium.

Indeed, an a-trapezium is necessarily a quasi-trapezium. To see that an a-trapezium satisfies (M3), note that Lemmas 9.1 and 9.3 imply that in any minimal diagram with the same contour label, any maximal θ-band must cross each maximal q-band exactly once.

Lemma 10.11. *Suppose Γ is a reduced diagram over $M_\Omega(\boldsymbol{M})$ with contour $\boldsymbol{p}_1^{-1}\boldsymbol{q}_1\boldsymbol{p}_2\boldsymbol{q}_2^{-1}$ where each $\boldsymbol{p}_j$ is the side of a q-band and each $\boldsymbol{q}_j$ is the maximal subpath of the side of a θ-band that starts and ends with a q-letter. Then there exists a minimal diagram Γ' over $M_\Omega(\boldsymbol{M})$ such that:*

(1) $\partial\Gamma' = (\boldsymbol{p}_1')^{-1}\boldsymbol{q}_1'\boldsymbol{p}_2'(\boldsymbol{q}_2')^{-1}$, where Lab$(\boldsymbol{p}_j') \equiv$ Lab$(\boldsymbol{p}_j)$ and Lab$(\boldsymbol{q}_j') \equiv$ Lab$(\boldsymbol{q}_j)$ for $j = 1, 2$

(2) there exists a simple path $\boldsymbol{s}_1$ (respectively $\boldsymbol{s}_2$) connecting the vertices $(\boldsymbol{p}_1')_-$ and $(\boldsymbol{p}_2')_-$ (respectively the vertices $(\boldsymbol{p}_1')_+$ and $(\boldsymbol{p}_2')_+$) such that

(a) $(\boldsymbol{p}_1')^{-1}\boldsymbol{s}_1\boldsymbol{p}_2'\boldsymbol{s}_2^{-1}$ is the standard factorization of the boundary of an a-trapezium Γ_2 and

(b) any cell above $\boldsymbol{s}_2$ or below $\boldsymbol{s}_1$ is an a-cell.

Proof. By Lemma 10.9, there exists a minimal diagram Γ' with Lab$(\partial\Gamma') \equiv$ Lab$(\partial\Gamma)$. Then $\partial\Gamma'$ can be factored as in (1).

Every q-edge of $\mathbf{q}_j'$ gives rise to a maximal q-band of Γ'. Suppose such a band $\mathcal{Q}$ starts and ends on $\mathbf{q}_j'$ and consider the subdiagram Δ bounded by a side of $\mathcal{Q}$ and $\mathbf{q}_j'$. Since q-bands are comprised entirely of (θ, q)-cells, the side of $\mathcal{Q}$ contains θ-edges that give rise to maximal θ-bands in Δ. Lemma 9.1 then implies that no such θ-band can have both ends on the side of $\mathcal{Q}$, so that it must end on $\mathbf{q}_j'$. But $\mathbf{q}_j'$ contains no θ-edge since Lab$(\mathbf{q}_j') \equiv$ Lab$(\mathbf{q}_j)$.

Hence, by Lemma 9.1, every maximal q-band in Γ' connects an edge of $\mathbf{q}_1'$ with an edge of $\mathbf{q}_2'$.

Now suppose a maximal θ-band of Γ' has two ends on $\mathbf{p}_j'$. Then, as no two θ-bands can cross, there exists a θ-band connecting adjacent θ-letters of $\mathbf{p}_j'$ (with perhaps a-letters between them). Then, the corresponding θ-edges of $\mathbf{p}_j$ in Γ are mutually inverse adjacent θ-edges, so that the corresponding cells of the q-band with side $\mathbf{p}_j$ are cancellable. But this contradicts the assumption that Γ is reduced.

Hence, by Lemma 10.7(3), every maximal θ-band in Γ' connects an edge of $\mathbf{p}_1'$ with an edge of $\mathbf{p}_2'$, and so we can enumerate them from bottom to top $\mathcal{T}_1, \dots, \mathcal{T}_h$ for $h = |\mathbf{p}_j|$.

Let $\mathcal{Q}_1$ be the maximal q-band of Γ such that $\mathbf{p}_1 = \mathbf{top}(\mathcal{Q}_1)$ and let $\mathcal{Q}_1'$ be the maximal q-band of Γ' starting at the first letter of $\mathbf{q}_1'$. Then $\mathcal{Q}_1$ and $\mathcal{Q}_1'$ must correspond to the same base letter. Moreover, since every maximal θ-band must cross $\mathcal{Q}_1'$ exactly once, $\mathcal{Q}_1$ and $\mathcal{Q}_1'$ must have the same history. So, Lab($\mathbf{top}(\mathcal{Q}_1')$) $\equiv$ Lab($\mathbf{p}_1'$).

As Γ' is minimal, any cell between $\mathbf{top}(\mathcal{Q}_1')$ and $\mathbf{p}_1'$ must be a (θ, a)-cell. But removing any such cell from Γ' does not affect the minimality of Γ'. Hence, we may assume that $\mathbf{top}(\mathcal{Q}_1') = \mathbf{p}_1'$.

By an analogous argument, letting $\mathcal{Q}_2'$ be the maximal q-band of Γ' starting at the final letter of $\mathbf{q}_1'$, we may assume that $\mathbf{bot}(\mathcal{Q}_2') = \mathbf{p}_2'$.

Now let $\mathbf{s}_1 = \mathbf{bot}(\mathcal{T}_1)$ and $\mathbf{s}_2 = \mathbf{top}(\mathcal{T}_h)$. By definition, (2a) is satisfied.

Further, as there is no maximal θ-band above $\mathcal{T}_h$ or below $\mathcal{T}_1$, there can be no (θ, q)- or (θ, a)-cells above $\mathbf{s}_2$ or below $\mathbf{s}_1$. Thus, (2b) is satisfied.

$\square$

Lemma 10.12. *Let Γ be a quasi-trapezium with standard factorization of its contour $\boldsymbol{p}_1^{-1}\boldsymbol{q}_1\boldsymbol{p}_2\boldsymbol{q}_2^{-1}$. Then there exists a reduced diagram Γ' such that:*

(1) $\partial\Gamma' = (\boldsymbol{p}_1')^{-1}\boldsymbol{q}_1'\boldsymbol{p}_2'(\boldsymbol{q}_2')^{-1}$, where Lab($\boldsymbol{p}_j'$) $\equiv$ Lab($\boldsymbol{p}_j$) and Lab($\boldsymbol{q}_j'$) $\equiv$ Lab($\boldsymbol{q}_j$) for $j = 1, 2$

(2) the number of disks in Γ' is the same as the number of disks in Γ

(3) there exists a simple path $\boldsymbol{s}_1$ (respectively $\boldsymbol{s}_2$) connecting the vertices $(\boldsymbol{p}_1')_-$ and $(\boldsymbol{p}_2')_-$ (respectively $(\boldsymbol{p}_1')_+$ and $(\boldsymbol{p}_2')_+$) such that

(a) $(\boldsymbol{p}_1')^{-1}\boldsymbol{s}_1\boldsymbol{p}_2'\boldsymbol{s}_2^{-1}$ is the standard factorization of the boundary of an a-trapezium Γ_2 and

(b) any cell above $\boldsymbol{s}_2$ or below $\boldsymbol{s}_1$ is a disk or an a-cell

(4) *there exists $m \in \mathbb{N}$ such that any maximal θ-band of Γ contains m (θ, t)-cells and any maximal θ-band of Γ_2 contains m (θ, t)-cells.*

Proof. By Lemmas 9.1 and 10.7(3), every maximal θ-band of Γ must connect an edge of $\mathbf{p}_1$ with an edge of $\mathbf{p}_2$. So, we can enumerate these bands from bottom to top as $\mathcal{T}_1, \dots, \mathcal{T}_h$ for $h = |\mathbf{p}_1| = |\mathbf{p}_2|$.

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

If Γ contains a disk, then by Lemma 10.6 there exists a disk Π_1 such that at least $L - 4$ of its t-spokes end on $\mathbf{q}_1$ or on $\mathbf{q}_2$. By Lemma 10.7(1), at least $L - 4 - (L-1)/2 \geq 2$ of these spokes must end on $\mathbf{q}_1$ (on $\mathbf{q}_2$). So, for any $j \in \{1, \dots, h\}$, the number of t-spokes of Π_1 crossing $\mathcal{T}_j$ is at least 2.

Fix $j_1 \in \{1, \dots, h-1\}$ such that Π_1 lies between $\mathcal{T}_{j_1}$ and $\mathcal{T}_{j_1+1}$.

If $j_1 \geq i$ (i.e Π_1 lies above $\mathcal{T}_i$), then move Π_1 upwards by transposing it with $\mathcal{T}_{j_1+1}$. Then iterate this process, moving the resulting disk upward until it is transposed with $\mathcal{T}_h$.

If $j_1 < i$, then move Π_1 down in the same way until the corresponding disk is transposed with $\mathcal{T}_1$.

Let Λ_1' be the reduced diagram resulting from this process and Π_1' be the disk arising from Π_1. As Λ_1' is formed by a sequence of transpositions, Λ_1' has the same contour label and number of disks as does Γ, and so must be D-minimal. Factor $\partial\Lambda_1' = \mathbf{u}_1^{-1}\mathbf{b}_1'\mathbf{v}_1(\mathbf{t}_1')^{-1}$, where $\text{Lab}(\mathbf{u}_1) \equiv \text{Lab}(\mathbf{p}_1)$, $\text{Lab}(\mathbf{v}_1) \equiv \text{Lab}(\mathbf{p}_2)$, $\text{Lab}(\mathbf{b}_1') \equiv \text{Lab}(\mathbf{q}_1)$, and $\text{Lab}(\mathbf{t}_1') \equiv \text{Lab}(\mathbf{q}_2)$.

Enumerate the maximal θ-bands of Λ_1' as $\mathcal{S}_1', \dots, \mathcal{S}_h'$ from bottom to top. Then, letting Λ_1 be the subdiagram of Λ_1' given by removing Π_1', we may factor $\partial\Lambda_1 = \mathbf{u}_1^{-1}\mathbf{b}_1\mathbf{v}_1\mathbf{t}_1^{-1}$ such that $\mathbf{b}_1 = \mathbf{tbot}(\mathcal{S}_1')$ and $\mathbf{t}_1 = \mathbf{ttop}(\mathcal{S}_h')$.

If Λ_1 contains a disk, then Lemma 10.6 may be applied to yield a disk Π_2 such that $L - 4$ of its t-spokes end on $\mathbf{b}_1$ or on $\mathbf{t}_1$. Fix $j_2 \in \{1, \dots, h-1\}$ such that Π_2 lies between $\mathcal{S}_{j_2}'$ and $\mathcal{S}_{j_2+1}'$.

Suppose at least two t-spokes of Π_2 end on each of $\mathbf{b}_1$ and $\mathbf{t}_1$. Then we repeat the argument above, moving Π_2 above $\mathcal{S}_h'$ if $j_2 \geq i$ or below $\mathcal{S}_1'$ if $j_2 < i$.

Next, suppose that at most one t-spoke of Π_2 in Λ_1 ends on $\mathbf{b}_1$. Then, there is a set T_2' of at least

$L-5$ t-spokes of Π_2 in Λ_1 such that each ends on $\mathbf{t}_1$. Note that there is a natural bijection between T_2' and a subset T_2 of the t-spokes of Π_2 in Γ. As each t-spoke of T_2' ends on $\mathbf{t}_1$, each t-spoke in T_2 either ends on $\mathbf{q}_2$ or on Π_1. Since Γ satisfies (M3), at most $(L-1)/2$ t-spokes of Π_2 cross $\mathcal{T}_{j_2+1}$, so that at least one of the t-spokes of T_2 does not cross $\mathcal{T}_{j_2+1}$. This spoke must end on Π_1, so that Π_1 must lie below $\mathcal{T}_{j_2+1}$. Hence, $j_2 \geq j_1 \geq i$. In this case, move Π_2 up by transpositions until it is above $\mathcal{S}_h'$.

Finally, if at most one t-spoke of Π_2 in Λ_1 ends on $\mathbf{t}_1$, then the symmetric argument to the one above yields $j_2 \leq j_1 < i$. In this case, move Π_2 down by transpositions until it is below $\mathcal{S}_1'$.

In each case, let Λ_2' be the diagram resulting from applying the corresponding transpositions to Λ_1'. Further, let Π_2' be the disk arising from Π_2. Then Λ_2' must be D-minimal. Factor $\partial\Lambda_2' = \mathbf{u}_2^{-1}\mathbf{b}_2'\mathbf{v}_2(\mathbf{t}_2')^{-1}$ such that $\text{Lab}(\mathbf{u}_2) \equiv \text{Lab}(\mathbf{u}_1)$, $\text{Lab}(\mathbf{v}_2) \equiv \text{Lab}(\mathbf{v}_1)$, $\text{Lab}(\mathbf{b}_2') \equiv \text{Lab}(\mathbf{b}_1')$, and $\text{Lab}(\mathbf{t}_2') \equiv \text{Lab}(\mathbf{t}_1')$. Let Λ_2 be the subdiagram of Λ_2' given by removing Π_2' and enumerate the maximal θ-bands in Λ_2' as $\mathcal{S}_1'', \dots, \mathcal{S}_h''$. Then, we may factor $\partial\Lambda_2 = \mathbf{u}_2^{-1}\mathbf{b}_2\mathbf{v}_2\mathbf{t}_2^{-1}$ such that $\mathbf{b}_2 \equiv \mathbf{tbot}(\mathcal{S}_1'')$ and $\mathbf{t}_2 = \mathbf{ttop}(\mathcal{S}_h'')$.

This process can then be iterated moving every disk above the top θ-band or below the bottom θ-band.

The resulting reduced diagram Γ'' satisfies $\text{Lab}(\partial\Gamma'') \equiv \text{Lab}(\partial\Gamma)$. Enumerating the maximal θ-bands of Γ'' as $\mathcal{T}_1'', \dots, \mathcal{T}_h''$, these θ-bands bound a subdiagram Γ_2'' of Γ'' containing no disks and such that every cell of $\Gamma'' \setminus \Gamma_2''$ is a disk.

Note that the transpositions performed to obtain Γ'' do not alter the side q-bands. So, identifying these q-bands with those in Γ, $\partial\Gamma_2'' = \mathbf{p}_1^{-1}\mathbf{bot}(\mathcal{T}_1'')\mathbf{p}_2\mathbf{top}(\mathcal{T}_h'')^{-1}$. Hence, we may apply Lemma 10.11 to Γ_2'', yielding a minimal diagram Γ_2'.

Factor $\partial\Gamma_2'$ as $(\mathbf{p}_1')^{-1}\mathbf{q}_1'\mathbf{p}_2'(\mathbf{q}_2')^{-1}$ such that $\text{Lab}(\mathbf{p}_j) \equiv \text{Lab}(\mathbf{p}_j')$. Further, let $\mathbf{s}_1$ and $\mathbf{s}_2$ be the simple paths such that the subdiagram Γ_2 of Γ_2' with contour $(\mathbf{p}_1')^{-1}\mathbf{s}_1\mathbf{p}_2'\mathbf{s}_2^{-1}$ is a minimal a-trapezium.

Pasting Γ_2' in place of Γ_2'' in Γ'' and making any necessary cancellations then produces a reduced diagram Γ' satisfying (1) and (3).

In passing from Γ to Γ', no disks are added. So, since Γ is minimal, (2) must be satisfied.

By Lemmas 9.1(1) and 10.8, Γ contains no (θ, q)-annuli or θ-annuli. As no such annulus can be created through a transposition, it follows that every maximal q-band of Γ_2'' crosses every maximal θ-band exactly once.

Since the θ-band $\mathcal{T}_i$ did not participate in any of the transpositions in the construction of Γ'', the resulting maximal θ-band $\mathcal{T}_i''$ in Γ_2'' also contains m (θ, t)-cells. Hence, there are exactly hm (θ, t)-cells in Γ_2''.

The minimality of Γ_2 then implies that it contains at most hm (θ, t)-cells. But Γ is a minimal diagram containing at least hm (θ, t)-cells, so that both Γ and Γ_2 must contain exactly hm (θ, t)-cells.

As Γ contains h maximal θ-bands and each contains at least m (θ, t)-cells, each of these θ-bands must contain exactly m (θ, t)-cells. Conversely, by Lemmas 9.1 and 10.8, each maximal θ-band of Γ_2 must contain the same number of (θ, t)-cells, which again must be m.

□

Remark 10.2. *The concept of D-minimal diagram is introduced in this paper specifically to aid in the iterative step in the proof of Lemma 10.12. It is necessary for this goal as it is both sufficient as a hypothesis for Lemma 10.6 and preserved under transposition (whereas, for example, (M3) satisfies the first condition but not the second).*

10.5 Shafts

We now introduce a concept that, as it was in [18] and [25], will be used to define a valuable measure on minimal diagrams.

Let Π be a disk contained in a minimal diagram and $\mathcal{B}$ be a t-spoke of Π. Suppose there is a subband $\mathcal{C}$ of $\mathcal{B}$ starting on Π whose history H contains a controlled subword. For W the configuration corresponding to Lab$(\partial\Pi)$, suppose $W(i)$ is H-admissible for $i \geq 2$. Then the t-band $\mathcal{C}$ is called a *shaft* of Π.

Note that this definition differs from that used in previous sources (for example, [18] and [25]), where it was required that W be H-admissible. The change here is to allow for 'flexibility' in the ‘special’ input sector, as $W(1)$ need not be H-admissible.

For a disk Π, a shaft $\mathcal{C}$ of Π is called a λ-*shaft* of Π if for every factorization $H \equiv H_1H_2H_3$ satisfying $\|H_1\| + \|H_3\| \leq \lambda\|H\|$, H_2 contains a controlled subword. Note that a shaft is a 0-shaft.

The following is an adaptation of Lemma 7.8 of [18] and Lemma 7.11 of [25] to this setting.

Lemma 10.13. *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 factorization $\mathcal{C} = \mathcal{C}_1\mathcal{C}_2\mathcal{C}_3$ such that*

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

(2) Δ contains a quasi-trapezium Γ such that the bottom (or top) of Γ has L t-edges and $\mathcal{C}_2$ starts on the bottom and ends on the top of Γ.

Proof. Assuming toward contradiction, let H_i be the history of the subband $\mathcal{C}_i$ for $i \in \{1, 2, 3\}$. Then, let $\tilde{\Delta}$ be the reduced diagram obtained from Δ by replacing Γ with the reduced diagram Γ' given in Lemma 10.12 and let Γ_2 be the minimal a-trapezium contained in Γ'.

Lemma 10.12(4) implies that the base of Γ_2 also contains L t-letters. Moreover, as the side labels of Γ_2 and Γ are the same, Γ_2 has history H_2.

By the definition of λ-shaft, H_2 must contain a controlled subword H'. So, Lemmas 6.7 and 7.2 imply that the base of Γ_2 must be reduced. Hence, assuming without loss of generality that the bottom (or top) label of Γ_2 starts and ends with one of its L t-letters, Γ_2 is a big a-trapezium.

Let Λ be the minimal diagram obtained from Γ_2 by removing one of the side t-bands. So, Λ is an a-trapezium whose base B is a cyclic permutation of the standard base (or its inverse). Let Λ' be the subdiagram of Λ that is an a-trapezium with base B and history H'. Then, let Λ_1 and Λ_2 be the two subdiagrams of Λ obtained by cutting along $\mathbf{bot}(\Lambda')$.

Let W be the configuration corresponding to $\partial\Pi$. By the definition of shaft, $W(i)$ is H-admissible for $i \geq 2$. So, by Lemma 10.4, there exists an accepted configuration V and a reduced diagram Ψ over $M_\Omega(\mathbf{M})$ with $\text{Lab}(\partial\Psi) \equiv H(0)^{-1}WH(0)V^{-1}$, where $H(0) \in F(R)$ is the word obtained by adding the subscript 0 to every letter of H.

Recall that Ψ is constructed by pasting together trapezia corresponding to one-machine computations in the standard base along a-cells in the 'special' input sector. So, any subdiagram bounded by two consecutive maximal q-bands not corresponding to the 'special' input sector is a trapezium. As such, we may view Ψ as an a-trapezium (though it may not be M-minimal), referring to its base, history, etc.

Since every rule locks the $Q_4(L)\{t(1)\}$-sector, cutting Ψ along the appropriate q-band and pasting the sides together produces such a diagram so that the base of any maximal θ-band is $B^{\pm 1}$. Perhaps taking the mirror then produces a reduced diagram Ψ_0 with base B.

Let Ψ'' be the subdiagram of Ψ_0 bounded by the maximal θ-bands corresponding to the history H_2. Similarly, let Ψ' be the subdiagram of Ψ'' corresponding to the history H'.

As computations with controlled history are one-machine, Ψ' is a trapezium. Let Ψ_1, Ψ_2 be the two subdiagrams of Ψ'' obtained by cutting along $\mathbf{bot}(\Psi')$, with the 'bottom' of Ψ_1 corresponding to the 'bottom' of Ψ''.

By Lemma 6.7, $\text{Lab}(\mathbf{bot}(\Lambda')) \equiv \text{Lab}(\mathbf{bot}(\Psi'))$, so that $\text{Lab}(\mathbf{top}(\Psi_1)) \equiv \text{Lab}(\mathbf{top}(\Lambda_1))$ and $\text{Lab}(\mathbf{bot}(\Psi_2)) \equiv \text{Lab}(\mathbf{bot}(\Lambda_2))$. Further, as the histories of the side q-bands of Ψ_j are the same as those of the side q-bands of Λ_j, Ψ_j and Λ_j have the same side labels.

Lemma 9.4 then implies that $\text{Lab}(\mathbf{bot}(\Lambda))$ and $\text{Lab}(\mathbf{bot}(\Psi''))$ (or $\text{Lab}(\mathbf{top}(\Lambda))$ and $\text{Lab}(\mathbf{top}(\Psi''))$) differ only by their projection to the 'special' input sector. Gluing together the common contours of Ψ_j and Λ_j, it then follows that these differences correspond to a-relations.

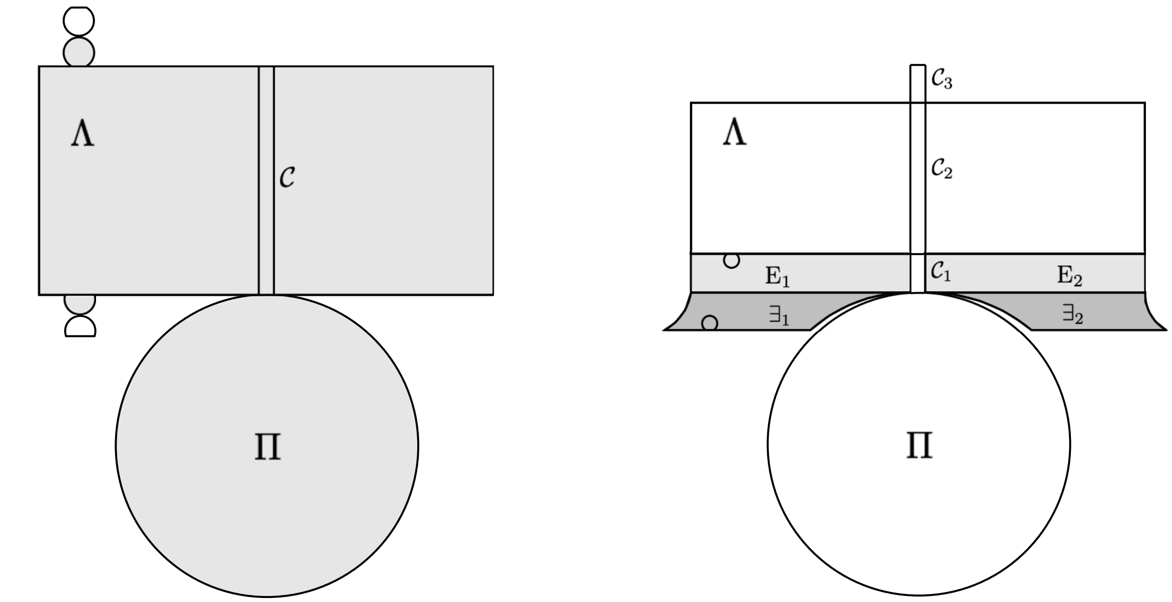
1. Suppose $\|H_1\| + \|H_3\| = 0$.

Then Λ and Π have a common edge, so that they form a subdiagram Δ' of $\tilde{\Delta}$. Perhaps adding two pairs of cancellable a-cells to Δ', we obtain a (perhaps unreduced) diagram Δ_0' with the same contour label as Δ' and containing a subdiagram Δ_0'' such that $\text{Lab}(\partial\Delta_0'') = V^{\pm 1}$ in $F(\mathcal{X})$ and the complement of Δ_0'' in Δ_0' consists of at most two a-cells (see Figure 10.7(a)).

Since V is accepted, there exists a disk relation corresponding to V. So, we can replace Δ_0'' in Δ_0' with one disk, producing the diagram Δ_1'.

Let $\tilde{\Delta}_1$ be the reduced diagram obtained by excising Δ' from $\tilde{\Delta}$, pasting Δ_1' in its place, and making any necessary cancellations. Then the number of disks in $\tilde{\Delta}_1$ is at most the number in $\tilde{\Delta}$, while the number of (θ, t)-cells is strictly smaller.

By Lemma 10.12, $\tilde{\Delta}$ and Δ have the same number of disks and (θ, t)-cells. But then $s_2(\tilde{\Delta}_1) < s_2(\Delta)$, contradicting the minimality of Δ.



(a) The subdiagram Δ_0'' if $\|H_1\| + \|H_3\| = 0$ (b) The construction of $\tilde{\Delta}''$

Figure 10.7: Lemma 10.13

2. Suppose $\|H_1\| + \|H_3\| > 0$.

Let Ψ_1'' be the subdiagram of Ψ_0 with base B and history H_1, so that **top**$(\Psi_1'') =$ **bot**(Ψ_1).

Let E be the diagram obtained by attaching the appropriate a-cell to the top of Ψ_1'' so that the top label is the same as that of **bot**(Λ_1). Further, let $\exists$ be the mirror image of E and $\exists$E be the diagram fomed by gluing $\exists$ to E along the bottom of Ψ_1''. Note that there are at most $\lambda\|H\|(L-1)$ (θ, t)-cells in $\exists$. Then let $\tilde{\Delta}''$ be the (unreduced) diagram obtained from $\tilde{\Delta}$ by gluing the proper components of $\exists$E to the bottom of Λ and along $\mathcal{C}_1$ (see Figure 10.7(b)).

Let Δ' be the subdiagram of $\tilde{\Delta}''$ formed by Π, $\mathcal{C}_1$, Λ, and the components of E. As in the previous case, we may replace Δ' with a diagram made of one disk and perhaps some new a-cells.

After necessary cancellations, the resulting reduced diagram $\tilde{\Delta}_1$ at most as many disks as $\tilde{\Delta}''$, and so the same number as $\tilde{\Delta}$. In passing to this diagram, we added at most $\lambda\|H\|(L-1)$ (θ, t)-

cells from $\exists$, while removing at least $\|H_2\|(L-1) \geq (1-\lambda)\|H\|(L-1)$ (θ, t)-cells of Λ.

Taking $\lambda < 1/2$, it follows that $\tilde{\Delta}_1$ has less (θ, t)-cells than $\tilde{\Delta}$. Thus, as in the previous case, Lemma 10.12 implies that $s_2(\tilde{\Delta}_1) < s_2(\Delta)$, contradicting the minimality of Δ.

□

10.6 Designs on a Disk

In this section, we recall the measure on minimal diagrams, first introduced in [18], that was alluded to in Section 10.5.

Let $\mathcal{D}$ be a disk in the Euclidean plane, **T** be a finite set of disjoint chords, and **Q** be a finite set of disjoint simple curves in $\mathcal{D}$, called *arcs* (as to differentiate them from the chords).

Assume that arcs belong to the open disk $\mathcal{D}^\circ$ and that each chord crosses any arc transversely and at most one, with the intersection not coming at either of the arc's endpoints.

With these assumptions, the pair $(\mathbf{T}, \mathbf{Q})$ is called a *design* on the disk.

The length of an arc $C \in \mathbf{Q}$, denoted $|C|$, is the number of chords crossing it. *Subarcs* are defined in the natural way, so that the inequality $|D| \leq |C|$ is clear for D a subarc of C.

An arc C_1 is *parallel* to an arc C_2, denoted $C_1 \parallel C_2$, if every chord crossing C_1 also crosses C_2. Note that this relation is reflexive and transitive, but not symmetric.

For $\lambda \in (0, 1/2)$ the parameter listed in Section 4.3 and m a positive integer, a design $(\mathbf{T}, \mathbf{Q})$ is said to satisfy property $P(\lambda, m)$ if for any collection of m distinct arcs $C_1, \dots, C_m \in \mathbf{Q}$, there are no subarcs $D_1, \dots, D_m$, respectively, such that $|D_i| > (1-\lambda)|C_i|$ for all i and $D_1 \parallel D_2 \parallel \dots \parallel D_m$.

For a design $(\mathbf{T}, \mathbf{Q})$, define the length of $\mathbf{Q}$, $\ell(\mathbf{Q})$, to be $\ell(\mathbf{Q}) = \sum\limits_{C \in \mathbf{Q}} |C|$.

Lemma 10.14. *(Lemma 8.2 of [18]) There is a constant c dependant on λ and m such that for any design $(\boldsymbol{T}, \boldsymbol{Q})$ satisfying property $P(\lambda, m)$, $\ell(\boldsymbol{Q}) \leq c(\#\boldsymbol{T})$.*

Let Δ be a minimal diagram and $\mathcal{Q}$ be a t-spoke of a disk Π in Δ. Let $\mathcal{Q}_\Pi$ be the subband of $\mathcal{Q}$ which is a λ-shaft at Π of maximal length. Then, define $\sigma_\lambda(\Delta)$ as the sum of the lengths of the λ-shafts $\mathcal{Q}_\Pi$ for all disks Π and t-spokes $\mathcal{Q}$.

If Δ is a minimal diagram, then identify Δ with a disk and construct the design $(\mathbf{T}, \mathbf{Q})$ as follows: Let the middle lines of maximal θ-bands be the chords and the middle lines of maximal λ-shafts be the arcs.

Note that there is a subtle hindrance to this construction: If a maximal t-spoke connects two disks, then it may contain a λ-shaft at each disk, and these λ-shafts may overlap. However, this issue can be remedied simply by 'making room' in the spoke for both arcs to fit and be disjoint.

Note that the length $|C|$ of an arc with respect to this design is the number of cells in the λ-shaft and $\#\mathbf{T} = \frac{1}{2}|\partial\Delta|_\theta \leq \frac{1}{2}|\partial\Delta|$ since every maximal θ-band ends twice on $\partial\Delta$.

Lemma 10.15. *(Lemma 8.5 of [18]) If Δ is a minimal diagram, then $\sigma_\lambda(\Delta) \leq C_1|\partial\Delta|_\theta \leq C_1|\partial\Delta|$.*

Proof. By Lemma 10.14 and the parameter choices $C_1 >> L >> \lambda^{-1}$, it suffices to prove that the design $(\mathbf{T}, \mathbf{Q})$ satisfies Property $P(\lambda, 2L-1)$.

Arguing toward contradiction, there are $2L-1$ maximal λ-shafts $\mathcal{C}_1, \dots, \mathcal{C}_{2L-1}$ such that for some subband $\mathcal{D}$ of $\mathcal{C}_1$, $|\mathcal{D}| > (1-\lambda)|\mathcal{C}_1|$ and every maximal θ-band crossing $\mathcal{D}$ also crosses each of $\mathcal{C}_2, \dots, \mathcal{C}_{2L-1}$. So, since at most two of these λ-shafts correspond to any particular t-spoke, each of the θ-bands crossing $\mathcal{D}$ crosses at least L t-bands.

But then the λ-shaft $\mathcal{C}_1$ crosses a quasi-trapezium of height $|\mathcal{D}| > (1-\lambda)|\mathcal{C}_1|$ whose base has at least L t-letters, contradicting Lemma 10.13.

□

Chapter 11

Upper bound on the weight of minimal diagrams

11.1 Weakly minimal diagrams

The goal in this section is to bound the G-weight of all minimal diagrams Δ in terms of $|\partial\Delta|^2$. In light of Lemma 9.19, it suffices to restrict our attention to minimal diagrams containing disks. However, it proves necessary to consider a larger class of diagrams over the disk presentation of $G_\Omega(\mathbf{M})$, called weakly minimal.

Let Δ be a reduced diagram over the disk presentation of $G_\Omega(\mathbf{M})$ which contains a disk. Then, let $\mathcal{C}$ be a cutting q-band of Δ, i.e $\mathcal{C}$ ends twice on the boundary of Δ. Then $\mathcal{C}$ is called a *stem band* if it is either a rim band of Δ or both components of $\Delta \setminus \mathcal{C}$ contain disks. The unique maximal subdiagram of Δ satisfying the property that every cutting q-band is a stem band is called the *stem* of Δ and denoted Δ^*.

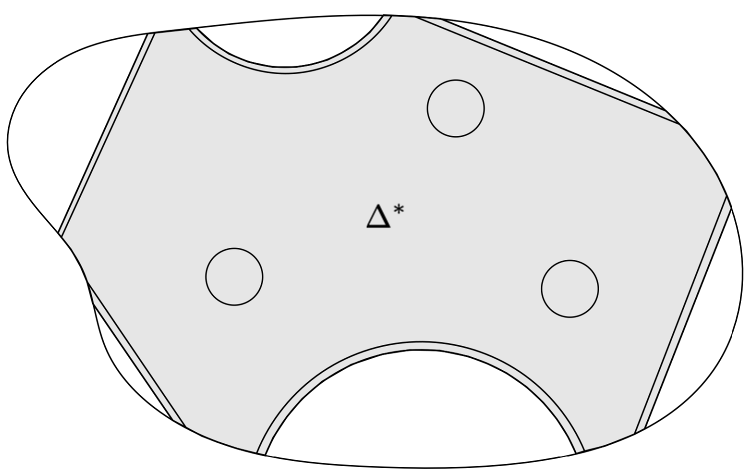


Figure 11.1: The stem of a reduced diagram Δ containing disks

If $\mathcal{C}$ is a cutting q-band that is not a stem band, then exactly one component Γ of $\Delta \setminus \mathcal{C}$ contains no disks. In this situation, the cells of Γ are called *crown* cells. Note that one can construct Δ^* from Δ simply by cutting off all of the crown cells.

Finally, a reduced diagram Δ over the disk presentation of $G_\Omega(\mathbf{M})$ which contains a disk is called *weakly minimal* if:

(WM1) for any a-cell π and any θ-band $\mathcal{T}$, at most half of the edges of $\partial\pi$ mark the start of an a-band that crosses $\mathcal{T}$,

(WM2) no maximal a-band ends on two different a-cells, and

(WM3) its stem Δ^* is a minimal diagram.

Note that conditions (WM1) and (WM2) are identical to conditions (MM1) and (MM2) in the definition of M-minimal (see Section 9.1). As a result, any subdiagram of a weakly minimal diagram which contains no disks is M-minimal.

Conversely, any minimal diagram containing a disk is weakly minimal.

Lemma 11.1. *(Compare to Lemma 9.3 of [18] and Lemma 7.17 of [25])*

(a) If Δ_1 is a subdiagram of a weakly minimal diagram Δ and contains a disk, then Δ_1 is weakly minimal, $\Delta_1^* \subset \Delta^*$, and $\sigma_\lambda(\Delta_1^*) \leq \sigma_\lambda(\Delta^*)$.

(b) For every weakly minimal diagram Δ, $\sigma_\lambda(\Delta^*) \leq C_1|\partial\Delta|$.

(c) A weakly minimal diagram Δ contains no θ-annuli.

(d) Let $\mathcal{C}$ be a cutting q-band of a reduced diagram Δ over the disk presentation of $G_\Omega(\boldsymbol{M})$ and let Δ_1, Δ_2 be the components of $\Delta \setminus \mathcal{C}$. Suppose $\Delta_1 \cup \mathcal{C}$ is M-minimal (over $M_\Omega(\boldsymbol{M})$) and $\Delta_2 \cup \mathcal{C}$ is weakly minimal. Then Δ is weakly minimal.

Proof. (a) Let π be a crown cell of Δ contained in Δ_1. Then there exists a cutting q-band $\mathcal{Q}$ separating π from all disks of Δ. The intersection of $\mathcal{Q}$ with Δ_1 is a cutting q-band separating π from all disks of Δ_1, so that π is a crown cell of Δ_1. Consequently, $\Delta_1^* \subset \Delta^*$, and hence Δ_1^* is minimal being a subdiagram of a minimal diagram.

By definition, every maximal λ-shaft $\mathcal{C}_1$ of Δ_1^* is contained in a maximal λ-shaft $\mathcal{C}$ of Δ^*. The length of $\mathcal{C}_1$ is then at most as large as the length of $\mathcal{C}$, so that $\sigma_\lambda(\Delta_1^*) \leq \sigma_\lambda(\Delta^*)$.

(b) Suppose $\mathcal{C}$ is a cutting q-band that is not a stem band and set Δ_1, Δ_2 as the components of $\Delta \setminus \mathcal{C}$. Then, one of Δ_1 or Δ_2, say Δ_1, is diskless.

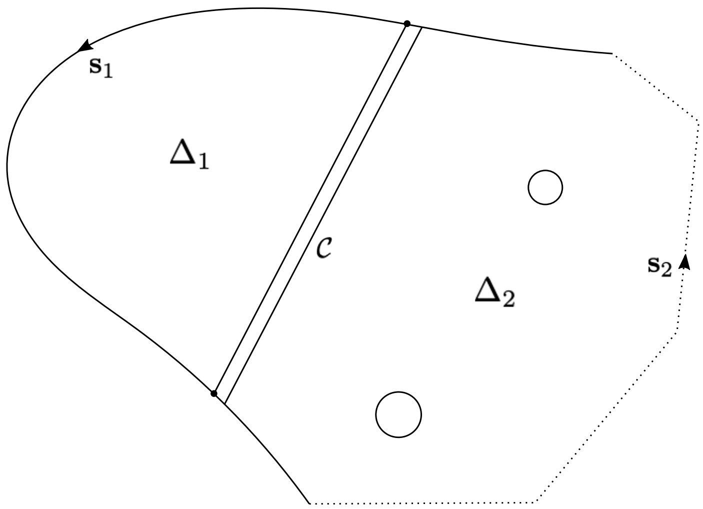


Figure 11.2: Lemma 11.1(b)

Let $\mathbf{s}_1$ be the portion of $\partial\Delta_1$ shared with $\partial\Delta$ and set $\partial\Delta = \mathbf{s}_1\mathbf{s}_2$. By Lemma 9.1, every maximal θ-band of $\Delta_1 \cup \mathcal{C}$ intersects $\mathcal{C}$ at most once. So, for every θ-edge on the side of $\mathcal{C}$, there is a maximal θ-band of $\Delta_1 \cup \mathcal{C}$ with one end on this edge and one end on $\mathbf{s}_1$.

So, letting ℓ be the number of θ-edges of $\mathbf{s}_1$, Lemma 8.1(b) implies $|\mathbf{s}_1| \geq \ell = |\mathbf{bot}(\mathcal{C})| = |\mathbf{top}(\mathcal{C})|$. Hence, for $\Delta' = \Delta_2 \cup \mathcal{C} = \Delta \setminus \Delta_1$, Lemma 8.1(c) implies $|\partial\Delta'| \leq |\mathbf{s}_2| + |\mathbf{top}(\mathcal{C})| \leq |\mathbf{s}_2| + |\mathbf{s}_1|$.

But $\mathbf{s}_2$ starts and ends with q-letters, so that $|\mathbf{s}_2| + |\mathbf{s}_1| = |\partial\Delta|$.

Iterating this process, $|\partial\Delta^*| \leq |\partial\Delta|$.

Thus. the statement is a consequence of Lemma 10.15.

(c) Lemma 9.3(2) implies that no θ-annulus can be contained in a crown of Δ. Since Δ^* is minimal, Lemma 10.8 implies that no θ-annulus can be contained in Δ^*.

Thus, the statement follows from Lemma 9.1, as no θ band can cross a rim q-band of Δ^* twice.

(d) Suppose there exists a counterexample to (WM1) in Δ and fix $j \in \{1, 2\}$ such that the a-cell

π is contained in Δ_j. Let $\mathcal{T}_j$ be the maximal subband of $\mathcal{T}$ contained in Δ_j. Then, since a-bands cannot cross q-bands, π and $\mathcal{T}_j$ form a counterexample to (WM1) in Δ_j. But this contradicts the M-minimality of Δ_1 or the weak minimality of Δ_2.

Similarly, any counterexample to (WM2) in Δ must be contained entirely in Δ_1 or Δ_2, leading to a contradiction.

Finally, it is clear from the definition that $\Delta^* = (\Delta_2 \cup \mathcal{C})^*$, and so is minimal.

□

11.2 Definition of the minimal counterexample and cloves

The objective of the rest of this section is to exhibit an upper bound for the G-weight of a weakly minimal diagram in terms of its perimeter. In particular, we will prove that for any weakly minimal diagram Γ, the inequality

$$\mathrm{wt}_G(\Gamma) \leq N_4(|\partial\Gamma| + \sigma_\lambda(\Gamma^*))^2 + N_3\mu(\Gamma)$$

holds for large enough choices of the parameters N_4 and N_3. The proof of this follows a similar path as that presented in Section 7 of [25] and Section 9 of [18] (taking $F(x) = x^2$ and $g(x) = x$ in that setting).

Let Δ be a 'minimal counterexample' diagram with respect to $|\partial\Delta| + \sigma_\lambda(\Delta^*)$, i.e a weakly minimal diagram satisfying

$$\mathrm{wt}_G(\Delta) > N_4(|\partial\Delta| + \sigma_\lambda(\Delta^*))^2 + N_3\mu(\Delta)$$

while for any weakly minimal diagram Γ such that $|\partial\Gamma| + \sigma_\lambda(\Gamma^*) < |\partial\Delta| + \sigma_\lambda(\Delta^*)$, we have

$$\mathrm{wt}_G(\Gamma) \leq N_4(|\partial\Gamma| + \sigma_\lambda(\Gamma^*))^2 + N_3\mu(\Gamma)$$

As with Lemma 9.13, the following statement is an immediate consequence of the inductive hypothesis.

Lemma 11.2. *Let π be an a-cell contained in Δ. Suppose $\partial\pi$ has a subpath* **s** *shared with $\partial\Delta$. Then $\|\mathbf{s}\| \leq \frac{2}{3}\|\partial\pi\|$.*

Since Δ^* contains every disk of Δ and is minimal, Δ is a D-minimal diagram. So, Lemma 10.6 guarantees that it contains a disk Π with $L-4$ consecutive t-spokes $\mathcal{Q}_1, \dots, \mathcal{Q}_{L-4}$ ending on $\partial\Delta$ and bounding $L-5$ diskless subdiagrams (see Figure 10.3).

For $1 \leq i < j \leq L-4$, the subdiagram of Δ bounded by $\partial\Pi$, $\mathcal{Q}_i$, and $\mathcal{Q}_j$ (and not containing Π) is called a *clove* and is denoted Ψ_{ij}. The maximal clove $\Psi_{1,L-4}$ is simply denoted Ψ.

Lemma 11.3. *(Compare to Lemma 9.5 of [18] and Lemma 7.19 of [25]) Let $\mathcal{T}$ be a quasi-rim θ-band in Δ. Then the base of $\mathcal{T}$ has length $s > K$.*

Proof. Assume toward contradiction that $\mathcal{T}$ is a quasi-rim θ-band with base of length $s \leq K$. Then, define Δ' and Δ'' as in the proof of Lemma 9.16. As in that setting, Δ'' satisfies (WM1) and (WM2) and $|\partial\Delta''| \leq |\partial\Delta| - 1$.

As Δ' is a subdiagram of Δ, Lemma 11.1(a) implies that it is weakly minimal with $\sigma_\lambda((\Delta')^*) \leq \sigma_\lambda(\Delta^*)$.

Since the diagram Δ'' is formed from Δ' through the addition of a-cells, the 2-signatures of $(\Delta')^*$ and $(\Delta'')^*$ are equal. Hence, Δ'' is a weakly minimal diagram.

Further, every λ-shaft of $(\Delta'')^*$ is at most as long as the corresponding λ-shaft of Δ^*, so that $\sigma_\lambda((\Delta'')^*) \leq \sigma_\lambda(\Delta^*)$. Consequently, $|\partial\Delta''|+\sigma_\lambda((\Delta'')^*) \leq |\partial\Delta|+\sigma_\lambda(\Delta^*)-1$, and so the inductive hypothesis may be applied to Δ''.

Thus, the proof of Lemma 9.16 adapts naturally to this setting, providing a contradiction.

□

11.3 Properties of the cloves of Δ

The following statement is an adaptation of Lemma 9.14 to this setting and is proved in exactly the same way.

Lemma 11.4. *(1) Δ has no two disjoint subcombs Γ_1 and Γ_2 contained in Ψ with basic widths at most K and handles $\mathcal{B}_1$ and $\mathcal{B}_2$ such that some ends of these handles are connected by a subpath $\boldsymbol{x}$ of $\partial\Delta$ with $|\boldsymbol{x}|_q \leq c_0$.*

(2) If Γ is a subcomb of Δ contained in Ψ with basic width $s \leq K$, $|\partial\Gamma|_q = 2s$.

Lemma 11.5. *Any subcomb of Δ contained in Ψ has basic width at most K_0.*

Proof. Assume toward contradiction that there exists a subcomb of Δ contained in Ψ with basic width $s > K_0$. Then, using Lemma 11.3, an identical proof to the one presented in Lemma 9.9 implies that there exists a tight subcomb Γ of Δ contained in Ψ.

Further, an analogous proof to that presented in Lemma 9.15 implies that any subcomb of Γ has height greater than $\ell'/2$. Indeed, other than switching the parameters and using Lemma 11.4 in place of Lemma 9.16, the only necessary alteration to the proof of Lemma 9.15 is in the application of the inductive hypothesis, where we must use the inequality $\sigma_\lambda((\Delta')^*) \leq \sigma_\lambda(\Delta^*)$ arising from Lemma 11.1(a).

But then similar analogues of Lemmas 9.17-9.19 yield a contradiction in the same way. Only one major alteration is needed: In the adaptation of Lemma 9.19, the diagram Δ_0 is weakly minimal by Lemma 11.1(d) and satisfies $\sigma_\lambda(\Delta_0^*) = \sigma_\lambda(\Delta^*)$ since the handle of the tight subcomb Γ is a non-stem cutting q-band.

□

Remark 11.1. *The reason for our consideration of weakly minimal diagrams in this section is revealed in the proof of Lemma 11.5: The adaptation of Lemma 9.19 relies on Lemma 11.1(d), whose statement fails if one replaces 'weakly minimal' with 'minimal'.*

Lemma 11.6. *(Compare with Lemma 9.8 of [18] and Lemma 7.22 of [25])*

(1) Every maximal θ-band of Ψ crosses either $\mathcal{Q}_1$ or $\mathcal{Q}_{L-4}$

(2) There exists an r satisfying $(L-1)/2 - 3 \leq r \leq (L-1)/2$ such that the θ-bands of Ψ crossing $\mathcal{Q}_{L-4}$ do not cross $\mathcal{Q}_r$ and the θ-bands of Ψ crossing $\mathcal{Q}_1$ do not cross $\mathcal{Q}_{r+1}$

Proof. (1) Suppose there exists a maximal θ-band $\mathcal{T}$ of Ψ crossing neither $\mathcal{Q}_1$ nor $\mathcal{Q}_{L-4}$. As θ-bands cannot cross, we may assume that $\mathcal{T}$ is a quasi-rim θ-band. By Lemma 11.3, $\mathcal{T}$ must cross more than K maximal q-bands of Ψ.

Taking $K > 11L + 2K_0$, there exists a non-stem cutting q-band $\mathcal{C}'$ crossing $\mathcal{T}$ such that for Γ' the subdiagram of Ψ consisting of $\mathcal{C}'$ and the corresponding crown, Γ' contains no cells of the spokes of Π and at least K_0 maximal q-bands crossing $\mathcal{T}$ (see Figure 11.3).

By Lemma 11.5, Γ' cannot be a comb with handle $\mathcal{C}'$, and so must contain a maximal θ-band $\mathcal{T}'$ not crossing $\mathcal{C}'$. As above, we may assume $\mathcal{T}'$ is a quasi-rim θ-band containing greater than K (θ, q)-cells, yielding a non-stem cutting q-band $\mathcal{C}''$ crossing $\mathcal{T}'$ such that the corresponding subdiagram Γ'' does not contain any cell of $\mathcal{C}'$ but contains at least K_0 q-bands crossing $\mathcal{T}'$.

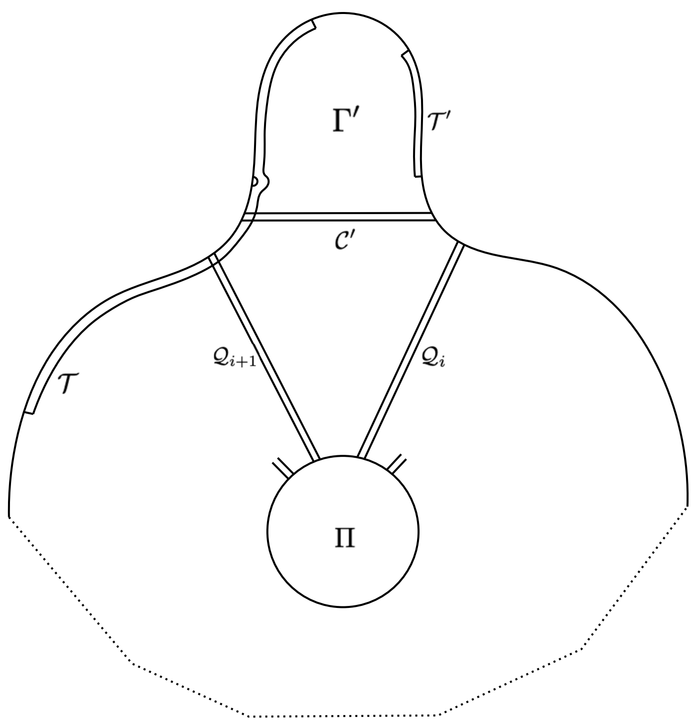


Figure 11.3: Lemma 11.6(1)

Iterating this process, we obtain a series of subdiagrams $\Gamma', \Gamma'', \dots$ satisfying

$$\text{wt}(\Gamma') > \text{wt}(\Gamma'') > \dots$$

Since these diagrams are finite, this process must terminate. But then the resulting subdiagram is a subcomb of Δ contained in Ψ with basic width at least K_0, contradicting Lemma 11.5.

(2) Let $\mathcal{T}$ be the maximal θ-band of Ψ crossing the t-spoke $\mathcal{Q}_1$ closest to Π, i.e the intersection of $\mathcal{T}$ and $\mathcal{Q}_1$ is the first cell of $\mathcal{Q}_1$.

Note that all spokes of a disk in Δ must lie in the minimal diagram Δ^*. So, if $\mathcal{Q}_1, \dots, \mathcal{Q}_\ell$ are the t-spokes crossed by $\mathcal{T}$, then $\ell \leq (L-1)/2$ by Lemma 10.7(1). Since $\mathcal{T}$ does not cross $\mathcal{Q}_{\ell+1}$, no other maximal θ-band of Ψ crossing $\mathcal{Q}_1$ can either. Similarly, no maximal θ-band crossing $\mathcal{Q}_{L-4}$ can cross $\mathcal{Q}_\ell$.

By the symmetric argument, if $\mathcal{Q}_{s+1}, \dots, \mathcal{Q}_{L-4}$ are the spokes crossed by the maximal θ-band crossing $\mathcal{Q}_{L-4}$ closest to Π, then no θ-band crossing $\mathcal{Q}_1$ can cross $\mathcal{Q}_{s+1}$ and $(L-4)-s \leq (L-1)/2$, i.e $s \geq (L-1)/2 - 3$.

Thus, the statement follows for $r = \max(\ell, (L-1)/2 - 3)$.

□

11.4 Paths in the cloves

For $1 \leq i < j \leq L-4$, denote $\mathbf{p}_{ij}$ as the shared subpath of $\partial\Psi_{ij}$ and $\partial\Delta$. For simplicity, denote the path $\mathbf{p}_{1,L-4}$ associated to the maximal clove simply as $\mathbf{p}$.

Let $\bar{\Delta}$ be the subdiagram of Δ consisting of Π and Ψ. Then, let $\bar{\mathbf{p}} = \mathbf{bot}(\mathcal{Q}_1)^{-1}\mathbf{u}^{-1}\mathbf{top}(\mathcal{Q}_{L-4})$ where $\mathbf{u}$ is a subpath of $\partial\Pi$ and such that cutting along $\bar{\mathbf{p}}$ separates Δ into two components, one of which is $\bar{\Delta}$. Denote the other component Ψ'.

Similarly, for $1 \leq i < j \leq L-4$, define the the path $\bar{\mathbf{p}}_{ij} = \mathbf{bot}(\mathcal{Q}_i)^{-1}\mathbf{u}_{ij}^{-1}\mathbf{top}(\mathcal{Q}_j)$ and the subdiagrams $\bar{\Delta}_{ij}$ and Ψ'_{ij} (see Figure 11.4).

Let $H_1, \dots, H_{L-4}$ be the histories of the spokes $\mathcal{Q}_1, \dots, \mathcal{Q}_{L-4}$, respectively, read starting from the disk Π. Further, let $h_i = \|H_i\|$ for all i. Lemma 11.6 then implies the inequalities

$$h_1 \geq h_2 \geq \dots \geq h_r; \quad h_{r+1} \leq \dots \leq h_{L-4}$$

where $(L-1)/2-3 \leq r \leq (L-1)/2$. It then follows that H_{i+1} is a prefix of H_i for $i = 1, \ldots, r-1$ while H_j is a prefix of H_{j+1} for $j = r+1, \ldots, L-5$.

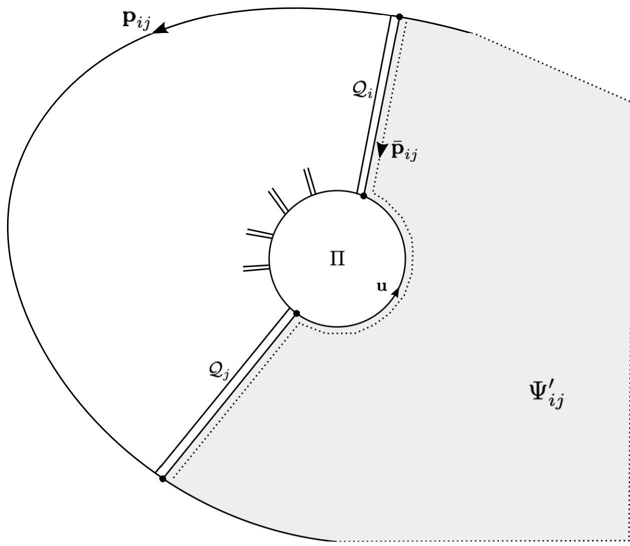


Figure 11.4: Subdiagrams and paths in Δ

Let W be the accepted configuration corresponding to Lab$(\partial\Pi)$. Then, using the notation of Section 6.3, $W \equiv W(1)W(2)\ldots W(L)$, where $W(2), \ldots, W(L)$ are all copies of the same configuration V of $\mathbf{M}_4$. Further, by Lemma 6.16, $|W(1)|_a \leq 2|V|_a$.

Lemma 11.7. *(Compare with Lemma 9.9 of [18] and Lemma 7.23 of [25]) For* $1 \leq i \leq L-5$, $|\boldsymbol{p}_{i,i+1}|_q < 3K_0$.

Proof. Suppose there exists a maximal q-band $\mathcal{B}$ which is not a spoke of Π and has one end on $\mathbf{p}_{i,i+1}$. Then, since q-bands cannot cross, $\mathcal{B}$ must have two ends on $\mathbf{p}_{i,i+1}$. So, $\mathcal{B}$ is a non-stem cutting q-band. Let Γ be the subdiagram of Ψ consisting of $\mathcal{B}$ and the corresponding crown. By Lemma 11.6(1), any maximal θ-band in Γ must cross $\mathcal{B}$, so that Γ is a comb with handle $\mathcal{B}$.

Note that any subcomb whose handle ends twice on $\mathbf{p}_{i,i+1}$ lies in a maximal subcomb with this condition.

Suppose $\Psi_{i,i+1}$ contains two such maximal subcombs. Then, let Γ_1 and Γ_2 be two adjacent such subcombs with handles $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively. By Lemma 11.5(1), Γ_1 and Γ_2 are disjoint subcombs of Δ contained in Ψ with basic widths at most K_0. Moreover, as we assume these subcombs are adjacent, there exists a subpath **x** of $\mathbf{p}_{i,i+1}$ connecting $\mathcal{B}_1$ and $\mathcal{B}_2$ such that any q-edge of **x** is the end a spoke of Π. As at most 23 spokes of Π can end on $\mathbf{p}_{i,i+1}$, $|\mathbf{x}|_q \leq c_0$. But this contradicts Lemma 11.4(1).

As a result, $\Psi_{i,i+1}$ contains at most one maximal subcomb whose handle ends twice on $\mathbf{p}_{i,i+1}$. By Lemmas 11.4(2) and 11.5(1), such a subcomb contributes at most $2K_0$ q-edges to $\mathbf{p}_{i,i+1}$.

Thus, $|\mathbf{p}_{i,i+1}|_q \leq 2K_0 + 23 < 3K_0$ by a parameter choice for K_0.

□

Lemma 11.8. *(Compare with Lemma 9.10 of [18] and Lemma 7.24 of [25])*

(1) If $i \leq r$ and $j \geq r+1$, then $|\boldsymbol{p}_{ij}| \geq |\boldsymbol{p}_{ij}|_\theta + |\boldsymbol{p}_{ij}|_q \geq h_i + h_j + 11(j-i) + 1$

(2) $|\bar{\boldsymbol{p}}_{ij}| \leq h_i + h_j + 11(L-j+i) + (L-j+i+1)\delta|V|_a - 1$

Proof. (1) Lemma 11.6(2) implies that $\mathbf{p}_{ij}$ contains $h_i + h_j$ θ-edges. Further, as q-bands cannot cross, every spoke starting on the complement $\bar{\mathbf{u}}_{ij}$ of $\mathbf{u}_{ij}$ in $\partial\Pi$ must end on $\mathbf{p}_{ij}$, so that $\mathbf{p}_{ij}$ contains at least $11(j-i)+1$ q-edges. The inequality thus follows.

(2) By parts (b) and (c) of Lemma 8.1, it suffices to show that

$$|\mathbf{u}_{ij}| \leq 11(L-j+i) + (L-j+i+1)\delta|V|_a - 1$$

As $\partial\Pi$ consists only of q-edges and a-edges, $|\mathbf{u}_{ij}| = |\partial\Pi| - |\bar{\mathbf{u}}_{ij}|$.

By Lemma 6.16, $|\partial\Pi| = 11L + \delta\sum_{i=1}^L |W(i)|_a \leq 11L + \delta(L+1)|V|_a$. Further, $\mathrm{Lab}(\bar{\mathbf{u}}_{ij})$ consists of at least $j-i$ copies of $V^{\pm 1}$ and one more t-letter, so that $|\bar{\mathbf{u}}_{ij}| \geq 11(j-i) + (j-i)\delta|V|_a + 1$. Thus, the inequality follows.

□

Lemma 11.9. *(Compare with Lemma 9.11 of [18] and Lemma 7.25 of [25])*

If $1 \leq i < j \leq L-4$ such that $j-i \geq L/2$, then

$$\mu(\Delta) - \mu(\Psi'_{ij}) > -2J|\partial\Delta|(h_i+h_j) \geq -2J|\partial\Delta||\boldsymbol{p}_{ij}|$$

Proof. As $j - i \geq L/2$, the path $\bar{\mathbf{u}}_{ij}$ contains at least $11L/2 + 1$ q-edges. So, since every spoke of Π starting on $\bar{\mathbf{u}}_{ij}$ must end on $\mathbf{p}_{ij}$, we have $|\mathbf{p}_{ij}|_q \geq |\bar{\mathbf{p}}_{ij}|_q$.

Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be a pair of θ-edges of $\partial\Delta$ such that neither is part of $\mathbf{p}_{ij}$. Further, let $\mathbf{x}$ be the subarc of $\partial\Delta$ connecting $\mathbf{e}_1$ and $\mathbf{e}_2$ and containing $\mathbf{p}_{ij}$. Then, replacing the subpath $\mathbf{p}_{ij}$ with $\bar{\mathbf{p}}_{ij}$ produces a subarc $\bar{\mathbf{x}}$ of $\partial\Psi'_{ij}$ connecting $\mathbf{e}_1$ and $\mathbf{e}_2$. Note that there are at least as many q-edges in $\mathbf{x}$ as in $\bar{\mathbf{x}}$. Hence, since the complement of $\mathbf{x}$ in $\partial\Delta$ is a subpath of $\partial\Psi'_{ij}$, both ordered pairs of white edges corresponding to $\mathbf{e}_1$ and $\mathbf{e}_2$ contributes at least as much to $\mu(\Delta)$ as it does to $\mu(\Psi'_{ij})$.

So, we need only consider the contribution to $\mu(\Psi'_{ij})$ from pairs of θ-edges where at least one is part of $\bar{\mathbf{p}}_{ij}$. As $\bar{\mathbf{p}}_{ij}$ consists of $h_i + h_j$ θ-edges from the sides of $\mathcal{Q}_i$ and $\mathcal{Q}_j$, there are at most $|\partial\Delta|(h_i + h_j)$ such unordered pairs. By definition, each ordered such pair contributes at most J to $\mu(\Psi'_{ij})$. Hence, $\mu(\Delta) - \mu(\Psi'_{ij}) \geq -2J|\partial\Delta|(h_i + h_j)$.

Since $j - i \geq L/2$, Lemma 11.6(2) implies that $i \leq r < r+1 \leq j$ and so every θ band crossing $\mathcal{Q}_i$ or $\mathcal{Q}_j$ ends on $\mathbf{p}_{ij}$. So, Lemma 8.1(a) yields $|\mathbf{p}_{ij}| \geq h_i + h_j$, implying the statement.

□

Lemma 11.10. *(Compare with Lemma 9.12 of [18] and Lemma 7.26 of [25])*

If $1 \leq i < j \leq L - 4$ *such that* $j - i \geq L/2$, *then*

$$|\boldsymbol{p}_{ij}| + \sigma_\lambda(\bar{\Delta}^*_{ij}) \leq |\boldsymbol{p}_{ij}| + \sigma_\lambda(\Delta^*) - \sigma_\lambda((\Psi'_{ij})^*) < (1+\varepsilon)|\bar{\boldsymbol{p}}_{ij}|$$

for $\varepsilon = 1/\sqrt{N_4}$.

Proof. Set $y = |\mathbf{p}_{ij}| + \sigma_\lambda(\Delta^*) - \sigma_\lambda((\Psi'_{ij})^*)$ and $d = y - |\bar{\mathbf{p}}_{ij}|$. Suppose $d \geq \varepsilon|\bar{\mathbf{p}}_{ij}| > 0$.

Then $d \geq y - \varepsilon^{-1}d$, so that $d \geq (1+\varepsilon^{-1})^{-1}y \geq \frac{\varepsilon y}{2}$ as $N_4 \geq 1$.

As Ψ'_{ij} and $\bar{\Delta}_{ij}$ are disjoint, the definition of the design on a minimal diagram implies

$$\sigma_\lambda(\bar{\Delta}^*_{ij}) + \sigma_\lambda((\Psi'_{ij})^*) \leq \sigma_\lambda(\Delta^*)$$

Let $\mathbf{s}$ be the complement of $\mathbf{p}_{ij}$ in $\partial\Delta$. As $\mathbf{p}_{ij}$ starts and ends with q-edges, $|\partial\Delta| = |\mathbf{p}_{ij}| + |\mathbf{s}|$.

Further, by Lemma 8.1(c), $|\partial \Psi'_{ij}| \leq |\bar{\mathbf{p}}_{ij}| + |\mathbf{s}|$.

So, these relations imply

$$\begin{aligned}(|\partial\Delta| + \sigma_\lambda(\Delta^*)) - (|\partial\Psi'_{ij}| + \sigma_\lambda((\Psi'_{ij})^*)) &\geq |\partial\Delta| - |\partial\Psi'_{ij}| + \sigma_\lambda(\Delta^*) - \sigma_\lambda((\Psi'_{ij})^*) \\ &\geq |\mathbf{p}_{ij}| - |\bar{\mathbf{p}}_{ij}| + \sigma_\lambda(\Delta^*) - \sigma_\lambda((\Psi'_{ij})^*) \\ &= d > 0\end{aligned}$$

Hence, if Ψ'_{ij} contains a disk, then we may apply the inductive hypothesis to it. Conversely, if Ψ'_{ij} contains no disks, then we may apply Lemma 9.19 to it. So, setting $x = |\partial\Delta| + \sigma_\lambda(\Delta^*)$, we have

$$\mathrm{wt}_G(\Psi'_{ij}) \leq N_4(x-d)^2 + N_3\mu(\Psi'_{ij})$$

Noting that $d \leq x$, Lemma 11.9 then implies

$$\mathrm{wt}_G(\Psi'_{ij}) \leq N_4x^2 - N_4xd + N_3\mu(\Delta) + 2N_3J|\partial\Delta||\mathbf{p}_{ij}| \tag{11.1}$$

Note that $|\partial\Pi| \leq (L+1)|\bar{\mathbf{p}}_{ij}| \leq (L+1)y$, so that

$$\mathrm{wt}(\Pi) \leq C_1(L+1)^2y^2 \tag{11.2}$$

Further, as $j - i \geq L/2$, we must have $i \leq r < r+1 \leq j$ by Lemma 11.6(2). So, Lemma 11.8 implies

$$|\bar{\mathbf{p}}_{ij}| < |\mathbf{p}_{ij}| + 11L + (L-1)\delta|V|_a \leq |\mathbf{p}_{ij}| + |\partial\Pi|$$

and hence $|\partial\Psi_{ij}| < 2|\mathbf{p}_{ij}| + |\partial\Pi| \leq (L+3)y$ by Lemma 8.1(c). Thus, by Lemma 9.19,

$$\mathrm{wt}_G(\Psi_{ij}) \leq N_2(L+3)^2y^2 + N_1\mu(\Psi_{ij}) \tag{11.3}$$

By (11.1), (11.2), and (11.3), Lemma 9.10 implies

$$\mathrm{wt}_G(\Delta) \leq N_4x^2 - N_4xd + N_3\mu(\Delta) + 2N_3Jy|\partial\Delta| + N_2(L+3)^2y^2 + N_1\mu(\Psi_{ij}) + C_1(L+1)^2y^2$$

Hence, in order to reach a contradiction, it suffices to show that

$$N_4xd \geq 2N_3Jy|\partial\Delta| + (N_2+C_1)(L+3)^2y^2 + N_1\mu(\Psi_{ij}) \tag{11.4}$$

Note that $x = |\partial\Delta| + \sigma_\lambda(\Delta^*) \geq |\mathbf{p}_{ij}| + \sigma_\lambda(\Delta^*)$, so that $x \geq \max(|\partial\Delta|, y)$. Hence,

$$N_4xd \geq \frac{\varepsilon}{2}N_4y\max(|\partial\Delta|, y) = \frac{1}{2}\sqrt{N_4}\max(y|\partial\Delta|, y^2)$$

The parameter choices $N_4 >> N_3 >> N_2 >> C_1 >> J >> L$ then allow us to assume

$$\frac{1}{2}N_4xd \geq 2N_3Jy|\partial\Delta| + (N_2+C_1)(L+3)^2y^2$$

Hence, by (11.4), it suffices to show that

$$N_4xd \geq 2N_1\mu(\Psi_{ij}) \tag{11.5}$$

As each θ-edge of $\partial\Psi_{ij}$ must be in its own factor of any decomposition of $\mathrm{Lab}(\Psi_{ij})$, the number of white beads on the necklace corresponding to $\partial\Psi_{ij}$ is at most $|\partial\Psi_{ij}| \leq (L+3)y$. So, Lemma 8.3(a) implies

$$\mu(\Psi_{ij}) \leq J(L+3)^2y^2$$

But as above, $N_4xd \geq \frac{1}{2}\sqrt{N_4}y^2$, so that (11.5) follows from the parameter choices $N_4 >> J >> L$.

□

For $i = 1, \dots, L-5$, if the pair of adjacent t-letters associated to $\mathcal{Q}_i$ and to $\mathcal{Q}_{i+1}$ are $t(L)$

and $t(2)$ (or vice versa), then $\Psi_{i,i+1}$ is called the *distinguished clove*. As q-bands cannot cross, the distinguished clove contains a cutting q-band $\mathcal{Q}_i'$ formed by the q-spoke of Π corresponding to the base letter $t(1)$. Let $\Lambda_{i,i+1}'$ (respectively $\Lambda_{i,i+1}''$) be the subdiagram bounded by $\mathcal{Q}_i'$ and the t-spoke corresponding to $t(L)$ (respectively $t(2)$). Define $\mathbf{p}_{i,i+1}'$ (respectively $\mathbf{p}_{i,i+1}''$) as the subpath of $\partial\Lambda_{i,i+1}'$ (respectively $\partial\Lambda_{i,i+1}''$) shared with $\partial\Delta$, so that $\mathbf{p}_{i,i+1}$ is the concatenation of these two paths along a shared q-edge.

Suppose $\Psi_{i,i+1}$ is not the distinguished clove. Then, let $\mathbf{q}_{i,i+1}$ be the shortest path in $\Psi_{i,i+1}$ homotopic to $\mathbf{p}_{i,i+1}$ and having the same first and last edges.

If $\Psi_{i,i+1}$ is the distinguished clove, then define $\mathbf{q}_{i,i+1}'$ and $\mathbf{q}_{i,i+1}''$ as the analgous shortest paths in $\Lambda_{i,i+1}'$ and $\Lambda_{i,i+1}''$. Then, let $\mathbf{q}_{i,i+1}$ be the concatenation of $\mathbf{q}_{i,i+1}'$ and $\mathbf{q}_{i,i+1}''$ along their shared q-edge.

For $1 \leq i < j \leq L-4$, let $\mathbf{q}_{ij}$ be the concatenation of the paths $\mathbf{q}_{i,i+1}, \dots, \mathbf{q}_{j-1,j}$ along their shared q-edges.

Then, let Ψ_{ij}^0 be the diagram obtained from Ψ_{ij} by replacing $\mathbf{p}_{ij}$ in the contour with $\mathbf{q}_{ij}$, i.e by removing any cells between $\mathbf{q}_{ij}$ and $\mathbf{p}_{ij}$. Similarly define Ψ^0, $(\Lambda_{i,i+1}')^0$, and $(\Lambda_{i,i+1}'')^0$.

The following is the direct analogue of Lemma 11.8(1) and is proved in exactly the same way.

Lemma 11.11. *(Compare with Lemma 9.13 of [18] and Lemma 7.27 of [25])*
If $i \leq r$ and $j \geq r+1$, then $|\mathbf{q}_{ij}| \geq h_i + h_j + 11(j-i) + 1$.

Lemma 11.12. *(Compare with Lemma 9.14 of [18] and Lemma 7.28 of [25])*

(1) Every maximal q-band of Ψ^0 corresponds to a spoke of Π.

(2) No two θ-edges of $\mathbf{q}_{i,i+1}$ are part of the same θ-band of $\Psi_{i,i+1}$.

Proof. (1) Assuming there exists a maximal q-band $\mathcal{Q}$ in Ψ^0 not corresponding to a spoke of Π, $\mathcal{Q}$ must end twice on $\mathbf{q}$. In fact, as q-bands cannot cross, there exists $1 \leq i \leq L-4$ such that $\mathcal{Q}$ ends twice on $\mathbf{q}_{i,i+1}$.

Let $\mathbf{x}$ be the subpath of $\mathbf{q}_{i,i+1}$ starting and ending with the ends of $\mathcal{Q}$. By Lemmas 9.1 and 10.7, any maximal θ-band crossing $\mathcal{Q}$ must have one end on $\mathbf{x}$. So, for ℓ the length of $\mathcal{Q}$, $|\mathbf{x}| \geq \ell + 2$.

By Lemma 8.1(b), $|\mathbf{bot}(\mathcal{Q})| = |\mathbf{top}(\mathcal{Q})| = \ell$. But then replacing **x** in $\mathbf{q}_{i,i+1}$ with a side of $\mathcal{Q}$ produces a homotopic path with shorter length, contradicting the definition of $\mathbf{q}_{i,i+1}$.

(2) Assuming the statement is false, there exists a θ-band $\mathcal{T}$ in $\Psi_{i,i+1}$ connecting θ-edges **e** and **f** of $\mathbf{q}_{i,i+1}$. Perhaps passing to a subband, we may assume that no other θ-edge comprising $\mathcal{T}$ is part of $\mathbf{q}_{i,i+1}$.

Let **y** be the subpath of $\mathbf{q}_{i,i+1}$ bounded by **e** and **f**. As θ-bands cannot cross, we may assume that **e** and **f** are the only θ-edges of **y**. So, every cell between a side of $\mathcal{T}$, say $\mathbf{top}(\mathcal{T})$, and **y** is an a-cell.

Suppose one of the following holds:

(i) $\Psi_{i,i+1}$ is not the distinguished clove (see Figure 11.5(i)),

(ii) $\Psi_{i,i+1}$ is the distinguished clove and **y** is a subpath of $\mathbf{q}'_{i,i+1}$ (see Figure 11.5(ii)), or

(iii) $\Psi_{i,i+1}$ is the distinguished clove and **y** is a subpath of $\mathbf{q}''_{i,i+1}$ (see Figure 11.5(iii)).

Note that every q-edge of $\mathbf{top}(\mathcal{T})$ must be shared with **y**. So, by (1), every q-band crossing $\mathcal{T}$ must be a spoke of Π in $\Psi_{i,i+1}$ (respectively $\Lambda'_{i,i+1}$, $\Lambda''_{i,i+1}$) in case (i) (respectively (ii), (iii)). As a result, the base of $\mathcal{T}$ is a subword of $B_3(j)^{\pm 1}$ for some j.

If there exists an a-cell π between $\mathbf{top}(\mathcal{T})$ and **y**, then let b_π be the number of edges of $\partial\pi$ which are on the boundary of a (θ, q)-cell of $\mathcal{T}$. By (WM1), at most $\frac{1}{2}\|\partial\pi\| + b_\pi$ of the edges of $\partial\pi$ are shared with $\mathbf{top}(\mathcal{T})$ while all other edges are shared with **y**.

By the definition of the rules of **M**, at most one edge on the boundary of a (θ, q)-cell of $\mathcal{T}$ is labelled with a letter from the alphabet of the 'special' input sector (on a cell corresponding to the base letter $Q_0(1)^{\pm 1}$). As a result, $\sum b_\pi \leq 1$, so that Lemma 8.1(c) implies $|\mathbf{y}| \geq 2 + |\mathbf{top}(\mathcal{T})| - 4\delta$.

As the base of $\mathcal{T}$ has length at most 11, Lemma 4.1 implies $|\mathbf{bot}(\mathcal{T})| - |\mathbf{top}(\mathcal{T})| \leq 22\delta$. So, $|\mathbf{y}| - |\mathbf{bot}(\mathcal{T})| \geq 2 - 26\delta \geq 1$ by a parameter choice for δ^{-1}.

But then replacing **y** in $\mathbf{q}_{i,i+1}$ with $\mathbf{bot}(\mathcal{T})$ contradicts the definition of $\mathbf{q}_{i,i+1}$.

Hence, it suffices to assume that (i), (ii), and (iii) all do not hold.

So, $\Psi_{i,i+1}$ is the distinguished clove, **e** is an edge of $\mathbf{q}'_{i,i+1}$, and **f** is an edge of $\mathbf{q}''_{i,i+1}$. Further, since $\mathbf{q}_{i,i+1}$ contains the q-edge of $\mathcal{Q}'$ shared with $\partial\Delta$, $\mathcal{T}$ must be contained in $\Psi^0_{i,i+1}$.

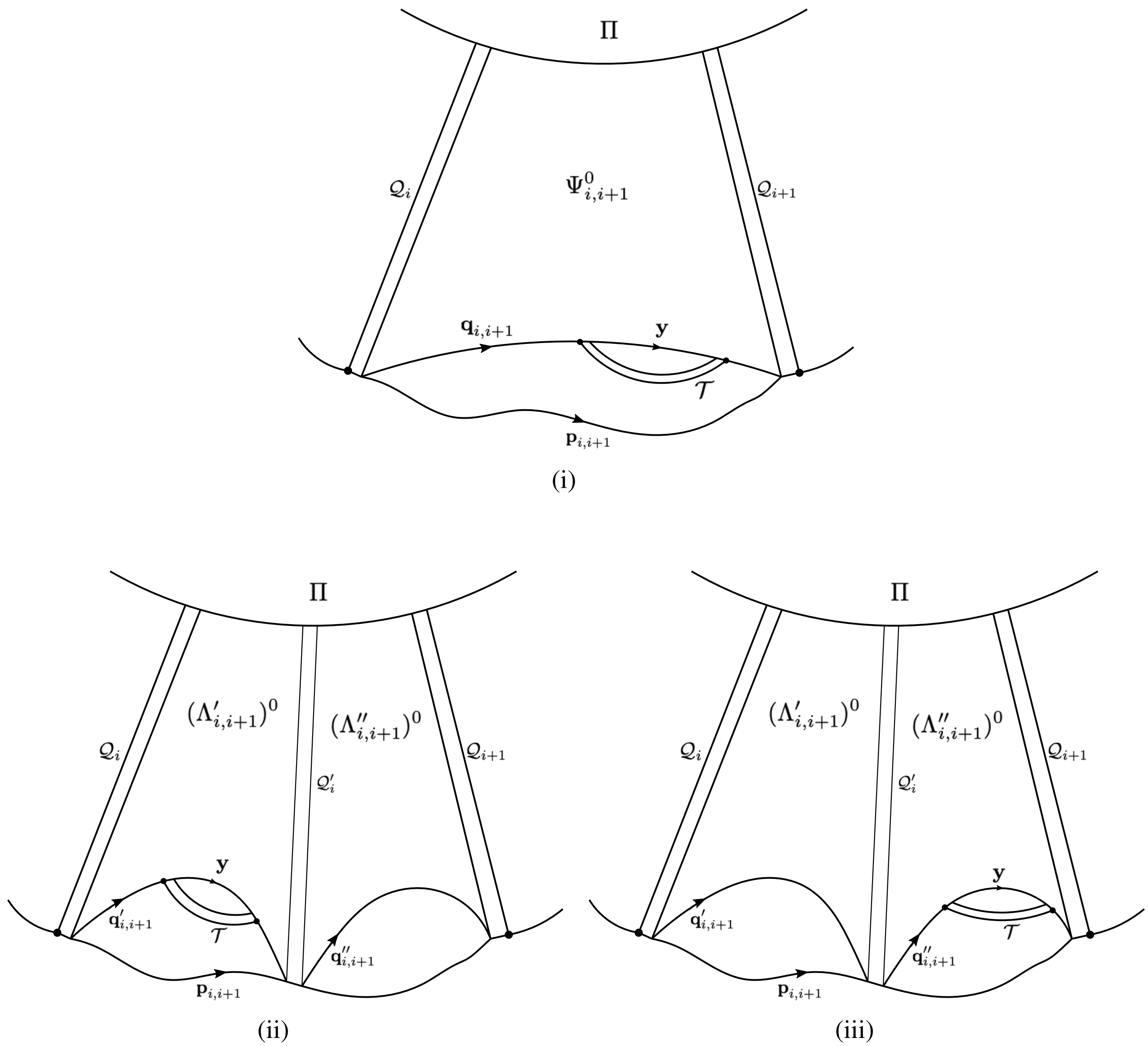


Figure 11.5: Lemma 11.12(2)

By Lemma 11.6(1), the maximal θ-band containing $\mathcal{T}$ must cross $\mathcal{Q}_i$ or $\mathcal{Q}_{i+1}$, so that it must contain another θ-edge of $\mathbf{q}'_{i,i+1}$ or $\mathbf{q}''_{i,i+1}$.

But then there exists a θ-band satisfying (ii) or (iii), so that a similar contradiction can be reached.

□

Lemma 11.13. *(1) If $\Psi_{i,i+1}$ is not the distinguished clove, then $\Psi_{i,i+1}^0$ contains no a-cells.*

(2) If $\Psi_{i,i+1}$ is the distinguished clove, then $(\Lambda'_{i,i+1})^0$ contains no a-cells.

Proof. (1) Suppose π is an a-cell contained in $\Psi_{i,i+1}^0$.

By Lemma 11.12(1), no maximal q-band of $\Psi_{i,i+1}^0$ corresponds to a base letter with coordinate 1. So, the contour of any (θ, a)- or (θ, q)-cell has no a-edge labelled by a letter from the alphabet of the 'special' input sector.

As a result, any edge of $\partial\pi$ must be shared with $\partial\Psi_{i,i+1}^0$. In particular, $\partial\pi$ must be a subpath of $\mathbf{q}_{i,i+1}$.

But then removing this subpath produces a path homotopic to $\mathbf{q}_{i,i+1}$ that contradicts its definition.

(2) is proved analogously, as the only base letter with coordinate 1 present in $(\Lambda'_{i,i+1})^0$ is $\{t(1)\}$.

□

11.5 Trapezia and combs in the cloves

For $1 \leq i \leq r-1$, suppose $\Psi_{i,i+1}$ is not the distinguished clove. Then Lemma 11.6 implies that all maximal θ-bands of $\Psi_{i,i+1}$ crossing $\mathcal{Q}_{i+1}$ must also cross $\mathcal{Q}_i$. So, these θ-bands bound an a-trapezium Γ_i in $\Psi_{i,i+1}^0$ with height h_{i+1}. The base of Γ_i (or its inverse) is $\{t(\ell)\}B_3(\ell)\{t(\ell+1)\}$ for some $2 \leq \ell \leq L-1$. Lemma 9.4 then implies that Γ_i is a trapezium. Set $\mathbf{y}_i = \mathbf{bot}(\Gamma_i)$ and $\mathbf{z}_i = \mathbf{top}(\Gamma_i)$. Note that $\mathbf{y}_i^{-1}$ is shared with $\partial\Pi$.

For $2 \leq i \leq r-1$, suppose neither $\Psi_{i-1,i}$ nor $\Psi_{i,i+1}$ is the distinguished clove. Then Lab$(\mathbf{y}_{i-1})$ and Lab$(\mathbf{y}_i)$ are coordinate shifts of one another while H_{i+1} is a prefix of H_i. So, h_{i+1} θ-bands of Γ_{i-1} form a copy of Γ_i, Γ'_i, contained in Γ_{i-1}. Set $\mathbf{y}'_i = \mathbf{bot}(\Gamma'_i)$ and $\mathbf{z}'_i = \mathbf{top}(\Gamma'_i)$. Note that $\mathbf{y}'_i = \mathbf{y}_{i-1}$.

For $1 \leq i \leq r-1$, if $\Psi_{i,i+1}$ is not the distinguished clove, then denote by E_i (respectively E_i^0) the maximal comb in $\Psi_{i,i+1}$ (respectively $\Psi_{i,i+1}^0$) containing the maximal θ-bands that cross the t-spoke $\mathcal{Q}_i$ but not the t-spoke $\mathcal{Q}_{i+1}$. The handle $\mathcal{C}_i$ of these combs has height $h_i - h_{i+1}$ and is contained in $\mathcal{Q}_i$. Any cell of $\Psi_{i,i+1}$ (respectively $\Psi_{i,i+1}^0$) not contained in Γ_i or E_i (respectively E_i^0) must be an a-cell attached to either $\mathbf{z}_i$ or $\mathcal{Q}_{i+1}$. By the structure of the relations, such an a-cell must share every boundary edge with $\partial\Delta$. But this contradicts Lemma 11.2. Hence, E_i (respectively E_i^0) is the complement of Γ_i in $\Psi_{i,i+1}$ (respectively $\Psi_{i,i+1}^0$).

Now suppose $\Psi_{i,i+1}$ is the distinguished clove for $1 \leq i \leq r-1$.

First, suppose $\mathcal{Q}_i$ corresponds to the base letter $\{t(L)\}$, so that the subdiagram $\Lambda'_{i,i+1}$ is bounded by $\mathcal{Q}_i$ and $\mathcal{Q}'_i$ (see Figure 11.7(a)). By Lemma 11.6, every maximal θ-band of $\Lambda'_{i,i+1}$ crossing $\mathcal{Q}'_i$ must also cross $\mathcal{Q}_i$. So, these θ-bands bound an a-trapezia Γ_i contained in $(\Lambda'_{i,i+1})^0$. As above, Lemma 9.4 implies that Γ_i must be a trapezium. The base of Γ_i (or its inverse) is $\{t(L)\}B_3(L)\{t(1)\}$, while the height is the length h'_i of the band $\mathcal{Q}'_i$.

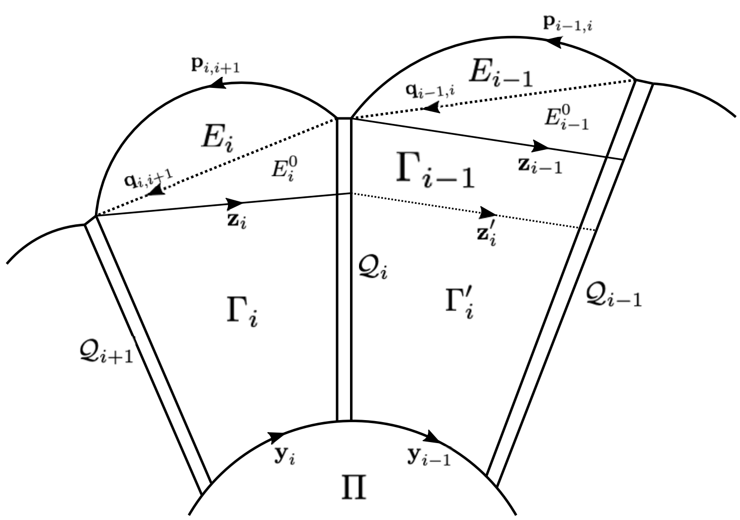


Figure 11.6: Trapezia and combs if neither $\Psi_{i-1,i}$ and $\Psi_{i,i+1}$ are distinguished

Otherwise, $\mathcal{Q}_{i+1}$ corresponds to the base letter $\{t(L)\}$, so that the subdiagram $\Lambda'_{i,i+1}$ is bounded by $\mathcal{Q}'_i$ and $\mathcal{Q}_{i+1}$ (see Figure 11.7(2)). Lemma 11.6 then implies that every maximal θ-band of $\Lambda'_{i,i+1}$ crossing $\mathcal{Q}_{i+1}$ must also cross $\mathcal{Q}'_i$, so that these θ-bands bound an a-trapezium Γ_i contained in $(\Lambda'_{i,i+1})^0$. Again, Γ_i must be a trapezium whose base (or its inverse) is $\{t(L)\}B_3(L)\{t(1)\}$. In this case, the height of Γ_i is h_{i+1}.

In either case, we define $\mathbf{y}_i = \mathbf{bot}(\Gamma_i)$ and $\mathbf{z}_i = \mathbf{top}(\Gamma_i)$. If $i \geq 2$, then again $\text{Lab}(\mathbf{y}_i)$ is a coordinate shift of $\text{Lab}(\mathbf{y}_{i-1})$ and there exists a copy Γ'_i of Γ_i in Γ_{i-1} with $\mathbf{bot}(\Gamma'_i) = \mathbf{y}'_i = \mathbf{y}_{i-1}$.

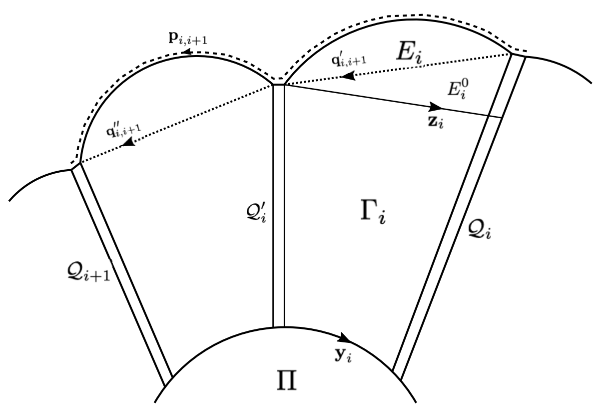
Similarly, if $i \leq r-2$, then $\text{Lab}(\mathbf{y}_i)$ is a coordinate shift of $\text{Lab}(\mathbf{y}_{i+1})$ and there exists a copy Γ'_{i+1} of Γ_{i+1} in Γ_i.

Suppose $\Lambda'_{i,i+1}$ is bounded by $\mathcal{Q}_i$ and $\mathcal{Q}'_i$. Then denote by E_i (respectively E_i^0) the maximal comb in $\Lambda'_{i,i+1}$ (respectively $(\Lambda'_{i,i+1})^0$) containing the maximal θ-bands that cross the t-spoke $\mathcal{Q}_i$ but not the q-spoke $\mathcal{Q}'_i$. The handle $\mathcal{C}_i$ of these combs has height $h_i - h'_i$ and is contained in $\mathcal{Q}_i$. As above, E_i (respectively E_i^0) is the complement of Γ_i in $\Lambda'_{i,i+1}$ (respectively $(\Lambda'_{i,i+1})^0$).

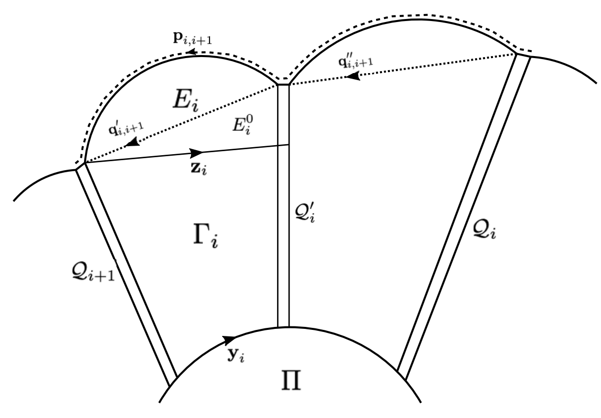
Otherwise, $\Lambda'_{i,i+1}$ is bounded by $\mathcal{Q}'_i$ and $\mathcal{Q}_{i+1}$. In this case denote by E_i (respectively E_i^0) the maximal comb in $\Lambda'_{i,i+1}$ (respectively $(\Lambda'_{i,i+1})^0$) containing the maximal θ-bands that cross the q-spoke $\mathcal{Q}'_i$ but not the t-spoke $\mathcal{Q}_{i+1}$. The handle $\mathcal{C}_i$ of these combs has height $h'_i - h_{i+1}$ and is contained in $\mathcal{Q}'_i$. Again, E_i (respectively E_i^0) is the complement of Γ_i in $\Lambda'_{i,i+1}$ (respectively $(\Lambda'_{i,i+1})^0$).

Note that no a-trapezium or comb has been defined in the subdiagram $\Lambda''_{i,i+1}$. Though such subdiagrams exist, their consideration is not necessary for the rest of the proof. As a result, one may view the indexing as 'skipping over' the portion of the clove between the base letters $\{t(1)\}$ and $\{t(2)\}$.

For $r+1 \leq i \leq L-5$, the trapezium Γ_i, the combs E_i and E_i^0, and the paths $\mathbf{y}_i$ and $\mathbf{z}_i$ are defined symmetrically.



(a) $\mathcal{Q}_i$ and $\mathcal{Q}'_i$ bound $\Lambda'_{i,i+1}$



(b) $\mathcal{Q}_i$ and $\mathcal{Q}'_i$ bound $\Lambda''_{i,i+1}$

Figure 11.7: Trapezia in the distinguished clove

Lemma 11.14. *(Compare with Lemma 9.15 of [18] and Lemma 7.29 of [25])*

For $i \in \{2, \dots, r-1\}$, suppose a maximal a-band $\mathcal{B}$ of E_i^0 starts on $\mathbf{z}_i$ and ends on a side of a maximal q-band $\mathcal{C}$. Let ∇ be the comb bounded by $\mathcal{B}$, a part of $\mathcal{C}$, and a subpath $\mathbf{x}$ of $\mathbf{z}_i$. Then there is a copy of the comb ∇ in the trapezium $\Gamma = \Gamma_{i-1} \setminus \Gamma_i'$.

Proof. By Lemma 11.13, ∇ contains no a-cells.

Let the a-edge $\mathbf{e}$ and the q-edge $\mathbf{f}$ be the first and last edge of $\mathbf{x}$, respectively. Since $\mathbf{z}_i'$ is a copy of $\mathbf{z}_i$ in the trapezium Γ_{i-1}, it contains a subpath $\mathbf{x}'$ that is a copy of $\mathbf{x}$ and starts with an a-edge $\mathbf{e}'$ and ends with a q-edge $\mathbf{f}'$. If π is the (θ, q)-cell attached to $\mathbf{f}$ in ∇, then the (θ, q)-cell π' attached to $\mathbf{f}'$ is a copy since it corresponds to the same letter of the history. Moving from $\mathbf{f}$ to $\mathbf{e}$, the whole maximal θ-band of ∇ containing π has a copy in Γ_{i-1}. Moving up, we find a copy of every maximal θ-band of ∇ in Γ_{i-1}, forming a copy of ∇ in Γ_{i-1}.

□

Lemma 11.15. *(Compare with Lemma 9.16 of [18] and Lemma 7.30 of [25])*
At most 6 a-bands starting on the path $\mathbf{y}_i$ (or $\mathbf{z}_i$) can end on (θ, q)-cells of the same θ-band.

Proof. Assume each of the a-bands $\mathbf{A}_1, \dots, \mathbf{A}_m$ starts from an edge of $\mathbf{y}_i$ and ends on some (θ, q)-cell of a θ-band $\mathcal{T}$. Let $\mathcal{T}_0$ be the minimal subband of $\mathcal{T}$ such that the a-bands $\mathbf{A}_2, \dots, \mathbf{A}_{m-1}$ end on $\mathcal{T}_0$. Then, let $\bar{\mathbf{y}}_i$ the minimal subpath of $\mathbf{y}_i$ where the a-bands $\mathbf{A}_1, \dots, \mathbf{A}_m$ start (see Figure 11.8).

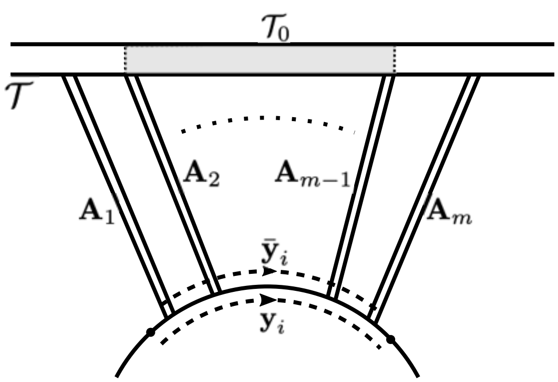


Figure 11.8:

By Lemma 9.1, each q-band starting on $\bar{\mathbf{y}}_i$ has to cross $\mathcal{T}_0$ and vice versa. So, the base of $\mathcal{T}_0$ is a subword of a reduced pararevolving base not containing the 'special' input sector.

As a result, we can identify this base with a subword of the standard base of $\mathbf{M}_4$ (or its inverse). By the structure of the rules of $\mathbf{M}_4$, an application of any rule inserts/deletes at most 4 a-letters in a configuration. Thus, $m-2\leq 4$, so that the statement follows.

An analogous argument applies for a-bands starting from $\mathbf{z}_i$.

□

By the parameter choice $L >> L_0$ and Lemma 11.6, we may assume that $L_0+1\leq r$ and $L-L_0-4\geq r+1$. Then, suppose without loss of generality that $h:=h_{L_0+1}\geq h_{L-L_0-4}$.

Lemma 11.16. *(Compare with Lemma 9.17 of [18] and Lemma 7.31 of [25])*
Let I be the subset of the set of indices $i\in[L_0+1,r-1]\cup[r+1,L-L_0-5]$ such that $|\mathbf{z}_i|_a\geq|V|_a/8c_3$. If $h\leq L_0^2|V|_a$, then $\#I\leq L/5$.

Proof. For any $i\in[L_0+1,r-1]\cup[r+1,L-L_0-5]$, denote the set of maximal a-bands of E_i^0 starting at $\mathbf{z}_i$ by $\mathbf{A}_i$. Then set $\mathbf{A}=\cup\mathbf{A}_i$.

As no base letter of Γ_i is of the form $Q_0(1)^{\pm1}$ or $P_0(1)^{\pm1}$, $\mathbf{z}_i$ has no a-letters from the 'special' input sector. So, any a-band of $\mathbf{A}$ either ends on a (θ,q)-cell or on $\mathbf{q}_{i,i+1}$.

Letting g_i be the length of the handle of E_i^0, then $\sum g_i\leq 2h$, where the sum is taken over the integers in $[L_0+1,r-1]\cup[r+1,L-L_0-5]$. So, by Lemma 11.15, at most $12h$ maximal a-bands of $\mathbf{A}$ end on (θ,q)-cells.

Assuming the statement is false, $\mathbf{A}$ contains at least $L|V|_a/40c_3$ a-bands. As a result, at least $\max(0,L|V|_a/40c_3-12h)$ bands from $\mathbf{A}$ must end on the subpaths $\mathbf{q}_{i,i+1}$. Since $\mathbf{q}_{i,i+1}$ has at most $2h$ θ-edges by Lemma 11.12(2), at least $\max(0,L|V|_a/40c_3-14h)$ a-edges contribute δ to $|\mathbf{q}_{i,i+1}|$.

By assumption, $14h\leq 14L_0^2|V|_a$, so that the parameter choices $L>>L_0>>c_3$ imply that $14h\leq L|V|_a/80c_3$.

It follows from Lemma 11.11 that

$$\begin{aligned}|\mathbf{p}_{L_0+1,L-L_0-4}|\geq|\mathbf{q}_{L_0+1,L-L_0-4}|&\geq h_{L_0+1}+h_{L-L_0-4}+11L/2+\delta(L|V|_a/40c_3-14h)\\&\geq h_{L_0+1}+h_{L-L_0-4}+11L/2+\delta L|V|_a/80c_3\end{aligned}$$

Also, by Lemma 11.8, we have

$$\begin{aligned}|\bar{\mathbf{p}}_{L_0+1,L-L_0-4}| &\le h_{L_0+1} + h_{L-L_0-4} + 11(3L_0) + 3L_0\delta|V|_a \\ &\le h_{L_0+1} + h_{L-L_0-4} + 11(3L_0) + \delta L|V|_a/160c_3\end{aligned}$$

as $L >> L_0 >> c_3$. These inequalities imply

$$|\mathbf{p}_{L_0+1,L-L_0-4}| - |\bar{\mathbf{p}}_{L_0+1,L-L_0-4}| \ge 11L/3 + \delta L|V|_a/160c_3 \tag{11.6}$$

Since $h_{L_0+1} + h_{L-L_0-4} \le 2h \le 2L_0^2|V|_a < \frac{1}{2}L|V|_a$, it follows that

$$|\bar{\mathbf{p}}_{L_0+1,L-L_0-4}| < \frac{1}{2}L|V|_a + 11(3L_0) + \delta L|V|_a/160c_3 \le 11(3L_0) + L|V|_a$$

which implies that

$$\frac{|\mathbf{p}_{L_0+1,L-L_0-4}| - |\bar{\mathbf{p}}_{L_0+1,L-L_0-4}|}{|\bar{\mathbf{p}}_{L_0+1,L-L_0-4}|} \ge \min\left(\frac{11L/3}{11(3L_0)}, \frac{\delta L|V|_a/160c_3}{L|V|_a}\right) = \delta/160c_3$$

since we have $L >> L_0$. Finally, $\delta/160c_3 > \varepsilon = 1/\sqrt{N_4}$ for sufficiently large N_4, and so

$$\frac{|\mathbf{p}_{L_0+1,L-L_0-4}|}{|\bar{\mathbf{p}}_{L_0+1,L-L_0-4}|} > 1 + \varepsilon$$

But $L - L_0 - 4 - (L_0 + 1) = L - 2L_0 - 5 \ge L - 3L_0 > L/2$ since $L >> L_0$, so that the above inequality contradicts Lemma 11.10.

□

Lemma 11.17. *(Compare with Lemma 9.18 of [18] and Lemma 7.32 of [25])*
If $h \le L_0^2|V|_a$, then the histories H_1 and H_{L-4} have different first letters.

Proof. Let $\mathcal{T}$ and $\mathcal{T}'$ be the maximal θ-bands of Ψ crossing $\mathcal{Q}_1$ and $\mathcal{Q}_{L-4}$, respectively, closest to the disk Π.

Set ℓ, ℓ' as the maximal integers such that $\mathcal{T}$ crosses the t-spokes $\mathcal{Q}_1, \dots, \mathcal{Q}_\ell$ and $\mathcal{T}'$ crosses the t-spokes $\mathcal{Q}_{L-\ell'-3}, \dots, \mathcal{Q}_{L-4}$. Note that $\ell \leq r \leq L - \ell' - 4$.

For any $\ell + 1 \leq i \leq L - \ell' - 5$, $\mathbf{z}_i$ is a subpath of $\partial\Pi$. As a result, $|\mathbf{z}_i| = |V|_a \geq |V|_a/8c_3$. Hence, if also $i \in [L_0+1, r-1] \cup [r+1, L-L_0-5]$, then $i \in I$ (for I as defined in the statement of Lemma 11.16).

If $\ell \leq L_0$, then $\#I \geq (r-1) - (L_0+1) \geq (L-1)/2 - L_0 - 5 \geq L/3$. Similarly, if $\ell' \leq L_0$, then $\#I \geq L/3$. But these inequalities contradict Lemma 11.16, so that $\ell, \ell' > L_0$.

This implies $\#I \geq (L - \ell' - 5) - (\ell + 1) - 2 \geq L - (\ell + \ell') - 8$, so that Lemma 11.16 yields $\ell + \ell' \geq L - L/5 - 8 \geq 3L/4$.

Thus, if the rules corresponding to $\mathcal{T}$ and $\mathcal{T}'$ are same, then the minimality of Δ^* contradicts Lemma 10.7(2).

□

Lemma 11.18. *(Compare with Lemma 7.33 of [25]) If $h \leq L_0^2|V|_a$, then $|V|_a > \dfrac{11L}{4\delta L_0}$.*

Proof. Assume that $|V|_a \leq 11L/4\delta L_0$. Then Lemma 11.8 implies the inequalities

$$|\mathbf{p}_{L_0+1,L-L_0-4}| \geq h_{L_0+1} + h_{L-L_0-4} + 11(L - 3L_0)$$

$$|\bar{\mathbf{p}}_{L_0+1,L-L_0-4}| \leq h_{L_0+1} + h_{L-L_0-4} + 3L_0(11 + \delta|V|_a)$$

Hence, as $L >> L_0$,

$$|\mathbf{p}_{L_0+1,L-L_0-4}| - |\bar{\mathbf{p}}_{L_0+1,L-L_0-4}| \geq 11(L - 6L_0) - 3L_0\delta|V|_a > 11(L - 6L_0) - 33L/4 > 11L/5$$

The inequality $h_{L_0+1} + h_{L-L_0-4} \leq 2h$ then implies

$$|\bar{\mathbf{p}}_{L_0+1,L-L_0-4}| \leq 2h + 3L_0(11 + 11L/4L_0) \leq 2L_0^2\frac{11L}{4\delta L_0} + 11L < 11L_0L/\delta$$

So, since $N_4 >> \delta^{-1} >> L_0$, we have

$$\frac{|\mathbf{p}_{L_0+1,L-L_0-4}| - |\bar{\mathbf{p}}_{L_0+1,L-L_0-4}|}{|\bar{\mathbf{p}}_{L_0+1,L-L_0-4}|} > \frac{\delta}{5L_0} > \varepsilon$$

But $L - L_0 - 4 - (L_0 + 1) \geq L/2$, so that the above inequality contradicts Lemma 11.10.

□

Lemma 11.19. *(Compare with Lemma 9.19 of [18] and Lemma 7.34 of [25])*
The inequality $h > L_0^2|V|_a$ must be true.

Proof. Assuming the statement is false, Lemma 11.16 implies that for at least $L-5-L/5-2L_0 > 3L/4$ indices $j \in \{1, \dots, L-5\}$, $|\mathbf{z}_j|_a < |V|_a/8c_3$. So, we can choose two such indices, i and j, such that $L_0 + 1 \leq i \leq r < r + 1 \leq j \leq L - L_0 - 5$, $j - i \geq 3L/5$, and neither $\Psi_{i,i+1}$ nor $\Psi_{j,j+1}$ is the distinguished clove.

Since H_{i+1} (respectively H_j) is a prefix of H_1 (respectively H_{L-4}), it follows from Lemma 11.17 that the first letters of H_{i+1} and H_j are different.

Since Lab($\mathbf{y}_i$) and Lab($\mathbf{y}_j$) are coordinate shifts of one another (and are copies of V), we can construct an auxiliary trapezium E by pasting the mirror of a coordinate shift of Γ_j to Γ_i along $\mathbf{y}_i$. The history of E is $H_j^{-1}H_{i+1}$, which is a reduced word since the first letter of H_{i+1} is different from the first letter of H_j.

The top and the bottom of E are copies of $\mathbf{z}_i$ and $\mathbf{z}_j$, respectively, and so have a-lengths less than $|V|_a/8c_3$. Without loss of generality, assume $h_{i+1} \geq h_j$, and so $h_{i+1} \geq t/2$ for t the height of E.

Note that $|V|_a - |V|_a/8c_3 > |V|_a/2$, and so $h_{i+1}, h_j > |V|_a/8$ since any rule of $\mathbf{M}_4$ alters the a-length of a configuration by at most four.

By Lemma 11.18, $|V|_a/8 > \frac{11L}{32\delta L_0} \geq 12c_3$ since $\delta^{-1} >> L >> L_0 >> c_3$. Further, letting W_0 and W_t be the bottom and top labels of E, $|V|_a/8 > c_3 \max(|W_0|_a, |W_t|_a)$.

As a result,

$$t = h_{i+1} + h_j > |V|_a/4 > c_3 \max(|W_0|_a, |W_t|_a) + 12c_3 \geq c_3 \max(\|W_0\|, \|W_t\|)$$

Let $\mathcal{C}$ be the computation associated to E through Lemma 7.4. Then the restriction of $\mathcal{C}$ (or its inverse) to $\{t(\ell)\}B_3(\ell)$ for the appropriate $\ell \geq 2$ satisfies the hypotheses of Lemma 6.20.

Setting $\lambda < 1/10$, every factorization $H'H''H'''$ of H_{i+1} with $\|H'\| + \|H'''\| \leq \lambda h_{i+1}$ satisfies $\|H''\| > 0.4t$. So, applying Lemma 6.20, H'' contains a controlled subword. Further, since all θ-bands crossing $\mathcal{Q}_{i+1}$ must cross $\mathcal{Q}_i$, $W(\ell)$ is H_{i+1}-admissible. Hence, $\mathcal{Q}_{i+1}$ is a λ-shaft.

Lemma 11.8(1) then implies that $|\mathbf{p}_{i+1,j}| + \sigma_\lambda(\bar{\Delta}^*_{i+1,j}) \geq 2h_{i+1} + h_j$.

As $h_{i+1} > |V|_a/8$, it follows that $\delta(L+1)|V|_a \leq 8\delta(L+1)h_{i+1} < \frac{1}{4}h_{i+1}$ by the parameter choice $\delta^{-1} >> L$. Similarly, by Lemma 11.18 and $\delta^{-1} >> L_0$, $11L < 4\delta L_0|V|_a < 32\delta L_0 h_{i+1} \leq \frac{1}{4}h_{i+1}$.

So, Lemma 11.8(2) yields $|\bar{\mathbf{p}}_{i+1,j}| \leq \frac{3}{2}h_{i+1} + h_j$.

Hence,

$$\frac{|\mathbf{p}_{i+1,j}| + \sigma_\lambda(\bar{\Delta}^*_{i+1,j})}{|\bar{\mathbf{p}}_{i+1,j}|} \geq \frac{2h_{i+1} + h_j}{\frac{3}{2}h_{i+1} + h_j} \geq \frac{6}{5}$$

since $h_j \leq h_{i+1}$.

Taking N_4 sufficiently large, $\varepsilon = 1/\sqrt{N_4} < 0.2$. However, as $j - (i+1) \geq 3L/5 - 1 \geq L/2$, the above inequality contradicts Lemma 11.10.

□

Lemma 11.20. *(Compare with Lemma 9.20 of [18] and Lemma 7.35 of [25]) For* $i = 1, \dots, L_0$, *we have* $h_i > \delta^{-1}$.

Proof. For such i, note that $h_i \geq h \geq h_{L-L_0-4}$. Assuming toward contradiction that $h_i \leq \delta^{-1}$, Lemma 11.19 implies that $\delta^{-1} > L_0^2|V|_a$, and so $\delta|V|_a < 1/L_0^2$.

Note that $L - L_0 - 4 - i \geq L - L_0 - 4 - L_0 \geq L - (2L_0 + 4)$. Taking $L >> L_0$, Lemma 11.8

then yields the inequalities

$$|\bar{\mathbf{p}}_{i,L-L_0-4}| \leq h_i + h_{L-L_0-4} + 3L_0(11 + \delta|V_a|) \leq h_i + h_{L-L_0-4} + 11(4L_0)$$

$$|\mathbf{p}_{i,L-L_0-4}| \geq h_i + h_{L-L_0-4} + 11(L - 2L_0 - 4) \geq h_i + h_{L-L_0-4} + 11L/2$$

But then $h_i + h_{L-L_0-4} \leq 2h_i \leq 2\delta^{-1}$ and $4L_0 < L/4$, so that

$$\frac{|\mathbf{p}_{i,L-L_0-4}|}{|\bar{\mathbf{p}}_{i,L-L_0-4}|} \geq \frac{h_i + h_{L-L_0-4} + 11L/2}{h_i + h_{L-L_0-4} + 11L/4} \geq \frac{8\delta^{-1} + 22L}{8\delta^{-1} + 11L} = 1 + \delta\frac{L}{\delta L + 8/11} > 1 + \delta$$

As $N_4 >> \delta^{-1}$, we may take $1 + \delta > 1 + \varepsilon$. But then noting that $L - L_0 - 4 - i \geq L/2$, the above inequality contradicts Lemma 11.10.

□

Lemma 11.21. *(Compare with Lemma 9.21 of [18] and Lemma 7.36 of [25])*
For $i = 1, \dots, L_0$, the spoke $\mathcal{Q}_i$ does not contain a λ-shaft of Π of length at least δh.

Proof. Let $j = L_0 + 1$ and $\ell = L - L_0 - 4$.

Since Π is removed when passing from Δ to $\Psi'_{j,\ell}$, $\mathcal{Q}_i$ is a cutting q-band of $\Psi'_{j,\ell}$. So, $\mathcal{Q}_i$ contains no λ-shaft in $\Psi'_{j,\ell}$.

As Lemma 11.1(1) implies $(\Psi'_{j,\ell})^* \subset \Delta^*$, we then have

$$\sigma_\lambda(\Delta^*) - \sigma_\lambda((\Psi'_{j,\ell})^*) \geq \delta h$$

Lemma 11.8 then yields the inequalities

$$|\mathbf{p}_{j,\ell}| \geq h_j + h_\ell + 11(L - 3L_0)$$

$$|\bar{\mathbf{p}}_{j,\ell}| \leq h_j + h_\ell + 3L_0(11 + \delta|V|_a)$$

By Lemma 11.19, $\delta|V|_a < \delta h/L_0^2$, so that

$$|\bar{\mathbf{p}}_{j,\ell}| < h_j + h_\ell + 11(3L_0) + 3\delta h/L_0$$

So, taking $L >> L_0 \geq 6$,

$$|\mathbf{p}_{j,\ell}| + \sigma_\lambda(\Delta^*) - \sigma_\lambda((\Psi'_{j,\ell})^*) - |\bar{\mathbf{p}}_{j,\ell}| \geq 11(L - 6L_0) + \delta h(1 - 3/L_0) \geq \frac{11L + \delta h}{2}$$

Hence, noting that $h_j \leq h = h_\ell$, we have:

$$\begin{aligned}\frac{|\mathbf{p}_{j,\ell}| + \sigma_\lambda(\Delta^*) - \sigma_\lambda((\Psi'_{j,\ell})^*) - |\bar{\mathbf{p}}_{j,\ell}|}{|\bar{\mathbf{p}}_{j,\ell}|} &\geq \frac{11L + \delta h}{2(2h + 11(3L_0) + 3\delta h/L_0)} \\ &\geq \frac{11L + \delta h}{11(6L_0) + 5h} \\ &\geq \min(L/6L_0, \delta/5) = \delta/5\end{aligned}$$

since $L >> L_0$. Taking $N_4 >> \delta^{-1}$ implies $\varepsilon < \delta/5$. But $\ell - j \geq L - 3L_0 \geq L/2$, so that the above inequality contradicts Lemma 11.10.

□

Lemma 11.22. *(Compare with Lemma 9.22 of [18] and Lemma 7.37 of [25]) For* $i = 1, \dots, L_0 - 1$, $|\mathbf{z}_i|_a > h_{i+1}/2c_3$.

Proof. Suppose to the contrary that $|\mathbf{z}_i|_a \leq h_{i+1}/2c_3$.

Then $\|\mathbf{z}_i\| = |\mathbf{z}_i|_a + 12 \leq (h_{i+1}/2c_3) + 12$.

Taking $\delta^{-1} >> c_3$, Lemma 11.20 yields $h_{i+1}/2c_3 > \delta^{-1}/2c_3 > 12$. So, $\|\mathbf{z}_i\| < h_{i+1}/c_3$.

Further, taking $\delta^{-1} >> L_0 >> c_3$, Lemma 11.19 yields

$$\|\mathbf{y}_i\| = 12 + |V|_a < 12 + h/L_0^2 \leq h_{i+1}/2c_3 + h_{i+1}/L_0^2 \leq h_{i+1}/c_3$$

Recall that by Lemma 11.13, Γ_i contains no a-cells, and so is a trapezium. By Lemma 7.4, there exists a reduced computation $\mathcal{C}'$ corresponding to Γ_i with base $(\{t(\ell)\}B_3(\ell)\{t(\ell+1)\})^{\pm 1}$ for some $2 \leq \ell \leq L$, where $L+1$ is taken to be 1. Let $\mathcal{C}$ be the restriction of $\mathcal{C}^{\pm 1}$ to the base $\{t(\ell)\}B_3(\ell)$.

Then, the history of $\mathcal{C}$ has length at least $h_{i+1} > c_3 \max(\|\mathbf{y}_i\|, \|\mathbf{z}_i\|)$. As a result, $\mathcal{C}$ satisfies the hypotheses of Lemma 6.20. Further, since every θ-band crossing $\mathcal{Q}_{i+1}$ also crosses $\mathcal{Q}_i$, $W(\ell)$ is H_{i+1}-admissible for some $\ell \geq 2$. So, $\mathcal{Q}_i$ contains a λ-shaft of length at least h_{i+1}.

But then $h_{i+1} \geq h > \delta h$, so that this contradicts Lemma 11.21.

□

Lemma 11.23. *(Compare with Lemma 9.23 of [18] and Lemma 7.38 of [25]) For* $i = 1, \dots, L_0 - 1$, $h_{i+1} < (1 - \frac{1}{30c_3})h_i$.

Proof. Assuming $h_{i+1} \geq (1 - \frac{1}{30c_3})h_i$, the handle of E_i has height at most $h_i - h_{i+1} \leq h_i/30c_3$. So, by Lemma 11.15, at most $h_i/5c_3$ maximal a-bands of E_i starting on $\mathbf{z}_i$ can end on (θ, q)-cells of E_i. Hence, at least $\max(0, |\mathbf{z}_i|_a - h_i/5c_3)$ of these bands end on $\mathbf{p}_{i,i+1}$.

Lemma 11.22 implies that $|\mathbf{z}_i|_a > h_{i+1}/2c_3$, so that

$$|\mathbf{z}_i|_a - h_i/5c_3 \geq h_{i+1}/2c_3 - h_i/5c_3 \geq \Big(1 - \frac{1}{30c_3}\Big)h_i/2c_3 - h_i/5c_3 > h_i/15c_3$$

By Lemma 11.6(2), $\mathbf{p}_{i,i+1}$ also has $h_i - h_{i+1} \leq h_i/30c_3$ θ-edges. So, Lemma 11.8(1) implies the inequalities

$$|\mathbf{p}_{i,i+1}| \geq h_i - h_{i+1} + \delta h_i/30c_3$$

$$|\mathbf{p}_{i+1,L-L_0-4}| \geq h_{i+1} + h_{L-L_0-4} + 22L/3$$

As these paths have an overlap of one q-edge, this implies

$$|\mathbf{p}_{i,L-L_0-4}| = |\mathbf{p}_{i,i+1}| + |\mathbf{p}_{i+1,L-L_0-4}| - 1 > h_i + h_{L-L_0-4} + 11L/2 + \delta h_i/30c_3$$

Meanwhile, Lemma 11.8(2) gives us

$$|\bar{\mathbf{p}}_{i,L-L_0-4}| \leq h_i + h_{L-L_0-4} + 11(3L_0) + 3L_0\delta|V|_a$$

As Lemma 11.19 implies $|V|_a < h/L_0^2 \leq h_i/L_0^2$, we then have

$$|\bar{\mathbf{p}}_{i,L-L_0-4}| \leq h_i + h_{L-L_0-4} + 11(3L_0) + 3\delta h_i/L_0$$

Hence, since $L >> L_0 >> c_3$ and $h_{L-L_0-4} \leq h \leq h_i$,

$$\begin{aligned}\frac{|\mathbf{p}_{i,L-L_0-4}| - |\bar{\mathbf{p}}_{i,L-L_0-4}|}{|\bar{\mathbf{p}}_{i,L-L_0-4}|} &\geq \frac{11(L/2 - 3L_0) + \delta h_i(1/30c_3 - 3/L_0)}{h_i + h_{L-L_0-4} + 11(3L_0) + 3\delta h_i/L_0} \\ &\geq \frac{11L/3 + \delta h_i/60c_3}{11(3L_0) + 3h_i} \\ &\geq \min(L/9L_0, \delta/180c_3) = \delta/180c_3\end{aligned}$$

However, taking $N_4 >> \delta^{-1} >> c_3$ yields $\varepsilon < \delta/180c_3$, so that the above inequality contradicts Lemma 11.10.

□

Lemma 11.24. *(Compare with Lemma 9.24 of [18] and Lemma 7.39 of [25]) For* $i = 1, \ldots, L_0-1$, $|\mathbf{z}_i|_a \leq 8h_i$.

Proof. Assume $|\mathbf{z}_i|_a > 8h_i$. By Lemma 11.15, at most $6h_i$ maximal a-bands of E_i^0 starting on $\mathbf{z}_i$ can end on the (θ, q)-cells of E_i. So, since a-bands cannot cross q-bands, Lemma 11.13 implies that at least $|\mathbf{z}_i|_a - 6h_i > 2h_i$ maximal a-bands of E_i^0 starting on $\mathbf{z}_i$ must end on the path $\mathbf{q}_{i,i+1}$. Hence, $|\mathbf{q}_{i,i+1}|_a > 2h_i$.

By Lemma 11.6(2), $\mathbf{q}_{i,i+1}$ has at most h_i θ-edges. As a result, Lemma 8.1 implies that at least h_i a-edges of $\mathbf{q}_{i,i+1}$ contribute δ to $|\mathbf{q}_{i,i+1}|$, and so also to $|\mathbf{q}_{i,L-L_0-4}|$. So, Lemmas 11.8 and 11.11 give the inequalities

$$|\mathbf{p}_{i,L-L_0-4}| \geq |\mathbf{q}_{i,L-L_0-4}| \geq h_i + h_{L-L_0-4} + 11L/2 + \delta h_i$$

$$|\bar{\mathbf{p}}_{i,L-L_0-4}| \leq h_i + h_{L-L_0-4} + 11(3L_0) + 3L_0\delta|V|_a$$

Taking $|V|_a < h/L_0^2 \leq h_i/L_0^2$ by Lemma 11.19 then gives

$$|\bar{\mathbf{p}}_{i,L-L_0-4}| \leq h_i + h_{L-L_0-4} + 11(3L_0) + 3\delta h_i/L_0$$

so that

$$|\mathbf{p}_{i,L-L_0-4}| - |\bar{\mathbf{p}}_{i,L-L_0-4}| \geq 11(L/2 - 3L_0) + \delta h_i(1 - 3/L_0) \geq 11L/3 + \delta h_i/2$$

Then, since $h_{L-L_0-4} \leq h \leq h_i$,

$$\begin{aligned}\frac{|\mathbf{p}_{i,L-L_0-4}| - |\bar{\mathbf{p}}_{i,L-L_0-4}|}{|\bar{\mathbf{p}}_{i,L-L_0-4}|} &\geq \frac{11L/3 + \delta h_i/2}{11(3L_0) + 3h_i} \\ &\geq \min(L/9L_0, \delta/6) = \delta/6\end{aligned}$$

However, again taking $N_4 >> \delta^{-1}$, $\varepsilon < \delta/6$ so that the above inequality contradicts Lemma 11.10.

□

Note that if $\Psi_{i,i+1}$ is the distinguished clove for $i \leq r-1$, then H_{i+1} need not be the history of Γ_i. To account for this, let H'_{i+1} be the history of Γ_i. Note that H_{i+1} is always a prefix of H'_{i+1}.

The following is the analogue of Lemma 9.25 of [18] and Lemma 7.40 of [25].

Lemma 11.25. *For $2 \leq i \leq L_0 - 2$, let $H'_i = H'_{i+1}H' = H'_{i+2}H''H'$ and $\mathcal{C}$ be the computation corresponding to the trapezium Γ_{i-1}. Suppose the subcomputation $\mathcal{D}$ of $\mathcal{C}$ with history $H''H'$ has step history of length 1. Then there is no two-letter subword $Q'Q$ of the base of Γ_{i-1} such that every rule of $\mathcal{D}$ inserts one letter to the left of Q.*

Proof. Let $\mathcal{Q}$ be the maximal q-band of E_i^0 that is a subband of the q-spoke of Π corresponding to a coordinate shift the state letter Q. Similarly, let $\mathcal{Q}'$ be the maximal q-band corresponding to a coordinate shift of Q', so that $\mathcal{Q}'$ and $\mathcal{Q}$ are neighbor q-bands. Let $\mathbf{x}$ be the subpath of $\mathbf{z}_i$ between $\mathcal{Q}'$ and $\mathcal{Q}$.

Since Γ_i contains a copy Γ'_{i+1} of the trapezium Γ_{i+1}, the bottom of the trapezium $\Gamma_i \setminus \Gamma'_{i+1}$ is a copy $\mathbf{z}'_{i+1}$ of $\mathbf{z}_{i+1}$, while the top is $\mathbf{z}_i$. This trapezium has history H'', so that the corresponding computation inserts one a-letter to the left of the state letter corresponding to $\mathcal{Q}$ at each transition. As a result, $|\mathbf{x}|_a \geq \|H''\| \geq h_{i+1} - h_{i+2}$.

By Lemma 11.23, $h_{i+1} - h_{i+2} > \frac{1}{30c_3}h_{i+1}$. As $h_{i+1} \geq h$, Lemma 11.19 and the parameter choice $L_0 >> c_3$ imply

$$|\mathbf{x}|_a \geq \frac{h}{30c_3} > \frac{L_0^2|V|_a}{30c_3} > 10L_0|V|_a$$

If an a-band starting on **x** ended on a (θ, q)-cell of $\mathcal{Q}$, then Lemma 11.10 implies that there is a copy of this in the trapezium $\Gamma_{i-1} \setminus \Gamma'_i$. By Lemma 7.4, though, this would contradict the assumption that rules of $\mathcal{D}$ only write letters in the sector.

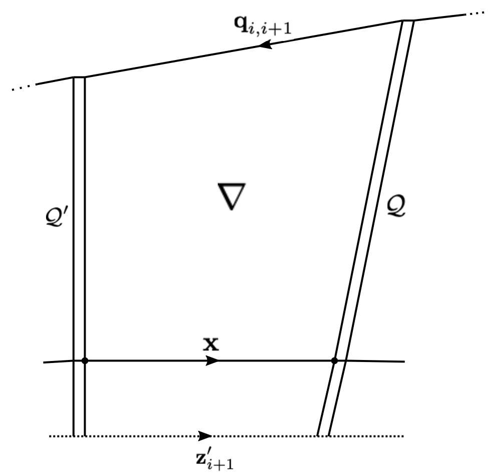


Figure 11.9:

Now, consider the comb ∇ contained in E_i^0 bounded by $\mathcal{Q}'$, $\mathcal{Q}$, **x**, and $\mathbf{q}_{i,i+1}$ (see Figure 11.9). Set s and s' as the lengths of $\mathcal{Q}$ and $\mathcal{Q}'$, respectively. Lemmas 11.6(2) and 11.12(2) imply $s' \leq s$. So, by Lemma 11.12(1), there are $|\mathbf{x}|_a + s$ maximal a-bands starting on **x** or $\mathcal{Q}$ and ending on $\mathcal{Q}'$ or on $\mathbf{q}_{i,i+1}$. Since only s' a-bands can end on $\mathcal{Q}'$, at least $|\mathbf{x}|_a + s - s'$ of them end on the segment of $\mathbf{q}_{i,i+1}$ between $\mathcal{Q}$ and $\mathcal{Q}'$. By Lemmas 11.6(2) and 11.12(2), the same segment contains $s - s'$ θ-edges, meaning at least $|\mathbf{x}|_a$ of them contribute δ to its length. So, by Lemma 11.8(1),

$$|\mathbf{p}_{i,L-L_0-4}| \geq h_i + h_{L-L_0-4} + 11L/2 + \delta\frac{h_{i+1}}{30c_3}$$

$$\geq h_i + h_{L-L_0-4} + 11L/2 + 10\delta L_0|V|_a$$

Also by Lemma 11.8(2) and 11.19,

$$|\bar{\mathbf{p}}_{i,L-L_0-4}| \leq h_i + h_{L-L_0-4} + 11(3L_0) + 3\delta L_0|V|_a \leq h_i + h_{L-L_0-4} + 11(3L_0) + 3\delta h/L_0$$

So,

$$\begin{aligned}|\mathbf{p}_{i,L-L_0-4}| - |\bar{\mathbf{p}}_{i,L-L_0-4}| &\geq 11(L/2 - 3L_0) + \delta(h_{i+1}/30c_3 - 3h/L_0)\\ &> \delta h_{i+1}(1/30c_3 - 3/L_0)\\ &> \delta h_{i+1}/50c_3\end{aligned}$$

by again taking $L_0 >> c_3$.

Let $\mathbf{s}$ be the complement of the $\mathbf{p}_{i,L-L_0-4}$ in $\partial\Delta$. Then, since $\mathbf{p}_{i,L-L_0-4}$ starts and ends with q-edges, Lemma 8.1(c) implies

$$\begin{aligned}|\partial\Psi'_{i,L-L_0-4}| &\leq |\mathbf{s}| + |\bar{\mathbf{p}}_{i,L-L_0-4}|\\ &< |\mathbf{s}| + |\mathbf{p}_{i,L-L_0-4}| - \delta h_{i+1}/50c_3\\ &= |\partial\Delta| - \delta h_{i+1}/50c_3 \qquad (11.7)\end{aligned}$$

Lemma 11.1(1) implies that $\Psi'_{i,L-L_0-4}$ is weakly minimal with $\sigma_\lambda((\Psi'_{i,L-L_0-4})^*) \leq \sigma_\lambda(\Delta^*)$.

Hence, if $\Psi'_{i,L-L_0-4}$ contains a disk, then we may apply the inductive hypothesis to it. Otherwise, we may apply Lemma 9.19 to $\Psi'_{i,L-L_0-4}$. In either case, this implies

$$\text{wt}_G(\Psi'_{i,L-L_0-4}) \leq N_4(|\Psi'_{i,L-L_0-4}| + \sigma_\lambda((\Psi'_{i,L-L_0-4})^*))^2 + N_3\mu(\Psi'_{i,L-L_0-4})$$

Taking $x = |\partial\Delta| + \sigma_\lambda(\Delta^*)$, we have

$$0 \le |\Psi'_{i,L-L_0-4}| + \sigma_\lambda((\Psi'_{i,L-L_0-4})^*) \le x - \delta h_{i+1}/50c_3$$

So, since $\delta h_{i+1}/50c_3 \le x$,

$$\mathrm{wt}_G(\Psi'_{i,L-L_0-4}) < N_4x^2 - N_4\delta x h_{i+1}/50c_3 + N_3\mu(\Psi'_{i,L-L_0-4}) \tag{11.8}$$

Next, note that $|V|_a \le h_i/L_0^2$ by Lemma 11.19, $h_i > \delta^{-1} > 11(100L_0)$ by Lemma 11.20, and $h_{L-L_0-4} \le h \le h_i$. So, for sufficiently large L_0, we have

$$|\bar{\mathbf{p}}_{i,L-L_0-4}| \le 2h_i + 3h_i/100 + 3\delta h_i/L_0 \le 2.1h_i$$

As Lemma 11.10 implies $|\mathbf{p}_{i,L-L_0-4}| \le (1+\varepsilon)|\bar{\mathbf{p}}_{i,L-L_0-4}|$, taking N_4 sufficiently large yields

$$|\partial\Psi_{i,L-L_0-4}| \le |\mathbf{p}_{i,L-L_0-4}| + |\bar{\mathbf{p}}_{i,L-L_0-4}| + |\partial\Pi| \le 4.5h_i + |\partial\Pi|$$

Taking $\delta^{-1} >> L$, note that

$$|\partial\Pi| \le 11L + (L+1)\delta|V|_a \le \delta^{-1}/4 + h_i/L_0^2 \le h_i/2$$

Since $\Psi_{i,L-L_0-4}$ contains no disks, Lemma 9.19 implies

$$\mathrm{wt}_G(\Psi_{i,L-L_0-4}) \le 25N_2h_i^2 + N_1\mu(\Psi_{i,L-L_0-4})$$

while the assignment of weight implies

$$\mathrm{wt}(\Pi) = C_1|\partial\Pi|^2 \le C_1h_i^2$$

Combining these two inequalities, Lemma 9.10 implies

$$\text{wt}_G(\bar{\Delta}_{i,L-L_0-4}) \leq 25N_2h_i^2 + N_1\mu(\Psi_{i,L-L_0-4}) + C_1h_i^2 < 26N_2h_i^2 + N_1\mu(\Psi_{i,L-L_0-4})$$

By Lemma 11.6, $|\partial\Psi_{i,L-L_0-4}|_\theta = 2(h_i + h_{L-L_0-4}) \leq 4h_i$. So, by Lemma 8.3(1),

$$\mu(\Psi_{i,L-L_0-4}) \leq 16Jh_i^2$$

Hence, taking $N_2 >> N_1 >> J$,

$$\text{wt}_G(\bar{\Delta}_{i,L-L_0-4}) \leq 26N_2h_i^2 + 16N_1Jh_i^2 \leq 30N_2h_i^2 \tag{11.9}$$

Thus, combining (11.8) and (11.9), Lemma 9.10 implies

$$\text{wt}_G(\Delta) \leq N_4x^2 - N_4\delta x h_{i+1}/50c_3 + N_3\mu(\Psi'_{i,L-L_0-4}) + 30N_2h_i^2$$

Hence, to reach a contradiction, it suffices to show that

$$N_4\delta x h_{i+1}/50c_3 \geq N_3\mu(\Psi'_{i,L-L_0-4}) - N_3\mu(\Delta) + 30N_2h_i^2 \tag{11.10}$$

Now consider the diagram $\Psi'_{i+1,L-L_0-4}$. When passing from $\partial\Psi'_{i+1,L-L_0-4}$ to $\partial\Psi'_{i,L-L_0-4}$, a subpath **t** is replaced with $\textbf{bot}(\mathcal{Q}_i)^{-1}$ (see Figure 11.10). The subpath **t** consists of:

- the subpath $\mathbf{p}'_{i,i+1}$ of $\partial\Delta$ obtained from $\mathbf{p}_{i,i+1}$ by removing the end of $\mathcal{Q}_{i+1}$,
- $\textbf{bot}(\mathcal{Q}_{i+1})^{-1}$, and
- a subpath of the inverse of $\partial\Pi$

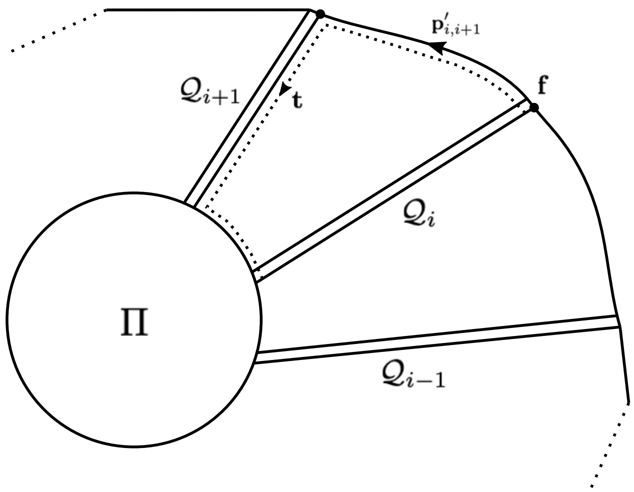


Figure 11.10:

By Lemma 11.6, there is a correspondence between the θ-edges of **t** and those of $\mathbf{bot}(\mathcal{Q}_i)$. So, since $\mathbf{bot}(\mathcal{Q}_i)$ contains no q-edges, the necklace corresponding to $\Psi'_{i,L-L_0-4}$ may be obtained from that of $\Psi'_{i+1,L-L_0-4}$ by the removal of the black beads corresponding to the q-edges of **t**.

Consider the q-edge **f** on the end of $\mathcal{Q}_i$. In $\partial\Psi'_{i+1,L-L_0-4}$, **f** separates the $h_{i-1} - h_i$ θ-edges of $\mathbf{p}_{i-1,i}$ from the $h_i + h_{L-L_0-4}$ θ-edges of the path $\mathbf{p}'_{i,i+1}\bar{\mathbf{p}}_{i+1,L-L_0-4}$.

By Lemma 11.7, $|\mathbf{p}'_{i,i+1}|_q < 3K_0$, while $|\bar{\mathbf{p}}_{i+1,L-L_0-4}|_q \leq 11L$. So, as $J >> K >> L$, there are at most J q-edges of $\partial\Psi'_{i+1,L-L_0-4}$ between any pair of θ-edges separated by **f** mentioned above. As such, Lemma 8.3(d) implies

$$\mu(\Psi'_{i+1,L-L_0-4}) - \mu(\Psi'_{i,L-L_0-4}) \geq (h_{i-1} - h_i)(h_i + h_{L-L_0-4})$$

Meanwhile, Lemma 11.9 implies

$$\mu(\Delta) - \mu(\Psi'_{i+1,L-L_0-4}) \geq -2J|\partial\Delta|(h_{i+1} + h_{L-L_0-4})$$

Combining these and noting that $h_{L-L_0-4} \leq h \leq h_{i+1} \leq h_i$, we have

$$\mu(\Delta) - \mu(\Psi'_{i,L-L_0-4}) \geq 2h_i(h_{i-1} - h_i) - 4J|\partial\Delta|h_{i+1}$$

Hence by (11.10), it suffices to show that

$$N_4\delta x h_{i+1}/50c_3 + 2N_3 h_i(h_{i-1} - h_i) \geq 4N_3 J|\partial\Delta|h_{i+1} + 30N_2 h_i^2 \tag{11.11}$$

As $x = |\partial\Delta| + \sigma_\lambda(\Delta^*) \geq |\partial\Delta|$, the parameter choices $N_4 >> N_3 >> \delta^{-1} >> J >> c_3$ imply

$$N_4\delta x h_{i+1}/50c_3 \geq \frac{N_4\delta}{50c_3}|\partial\Delta|h_{i+1} \geq 4N_3 J|\partial\Delta|h_{i+1}$$

Moreover, by Lemma 11.23, $h_{i-1} - h_i > h_{i-1}/30c_3 \geq h_i/30c_3$. Hence, the parameter choices $N_3 >> N_2 >> c_3$ imply

$$2N_3 h_i(h_{i-1} - h_i) > \frac{N_3}{15c_3}h_i^2 \geq 30N_2 h_i^2$$

Thus, the statement is proved.

□

Remark 11.2. *Recall that we have assumed without loss of generality that $h_{L-L_0-4} \leq h_{L_0+1}$. If $h_{L-L_0-4} > h_{L_0+1}$, then the symmetric statement to Lemma 11.25 will be needed for $L - L_0 - 2 \leq i \leq L - 1$. This statement can be proved analogously.*

Finally, we reach the final contradiction of this section, the analogue of Lemma 9.26 of [18] and Lemma 7.41 of [25].

Lemma 11.26. *The counterexample diagram Δ cannot exist.*

Proof. First, fix an integer $\eta \geq 2$ dependant on c_3 such that $(1 - \frac{1}{30c_3})^\eta < \frac{1}{64c_3}$. Note that, although η is not listed as one of the parameters of Section 4.3, we may take $L_0 >> \eta$ since L_0 is chosen after c_3.

For $i = 1, \ldots, L_0 - 1$, Lemma 11.23 implies $h_{i+1} < (1 - \frac{1}{30c_3})h_i$. So, if $1 \leq i < j \leq L_0 - 1$ with $j - i - 1 \geq \eta$, then $h_j < (1 - \frac{1}{30c_3})^\eta h_{i+1} < \frac{1}{64c_3}h_{i+1}$.

For such i, j, Lemma 11.22 then implies that $|\mathbf{z}_i|_a \geq h_{i+1}/2c_3 > 32h_j$. As Lemma 11.24 implies $8h_j \geq |\mathbf{z}_j|_a$, we then have $|\mathbf{z}_i|_a > 4|\mathbf{z}_j|_a$.

Now, as $L_0 >> \eta$ and $L_0 >> c_0$, there exist indices $2 \leq j_1 < j_2 < \cdots < j_m \leq L_0 - 1$ such that $m \geq c_0$ and $j_{i+1} - j_i - 1 \geq \eta$. So, $|\mathbf{z}_{j_i}|_a > 4|\mathbf{z}_{j_{i+1}}|_a$ and $h_{j_i+1} \geq 64c_3h_{j_{i+1}}$.

Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be the computation corresponding to the trapezium Γ_{j_2} by Lemma 7.4. As Γ_{j_2} contains a copy of Γ_{j_2+1}, which in turn contains a copy of Γ_{j_2+2} and so on, there exist words V_i in $\mathcal{C}$ for $i = 1, \ldots, m$ that are coordinate shifts of the labels of $\mathbf{z}_{j_i}$. By the inequalities above, $|V_{i+1}|_a > 4|V_i|_a$.

If for some i the subcomputation $V_{i+2} \to \cdots \to V_i$ is a one-step computation, then by Lemma 6.21 there exists a right-active (or left-active if $h = h_{L-L_0-4}$) sector $Q'Q$ such that the sector's length increases with each transition of the subcomputation. But since $\eta \geq 2$, there must exist a subcomputation contradicting Lemma 11.25.

Hence, the subcomputation $\mathcal{C}' : V_m \to \cdots \to W_t$ of $\mathcal{C}$ must contain at least $c_0/2 \geq 8n$ distinct one-step computations. Lemma 6.3 then implies that the step history of $\mathcal{C}'$ contains a subword of the form $(4n-2, 4n-1)_j(4n-1)_j(4n-1, 4n)_j$ or $(4n, 4n-1)_j(4n-1)_j(4n-1, 4n-2)_j$. Let $\mathcal{C}''$ be the subcomputation of $\mathcal{C}'$ with this step history.

Then, we may factor $H'_{j_2+1} \equiv H'H''H'''$ where H'' is a controlled history. Further, since the subcomputation $\mathcal{C}''$ repeats k copies of a controlled history, taking $k \geq 3$ allows us to assume $\|H''\| \leq \|H'\|$.

Since $h_{j_1+1} > 64c_3h_{j_2}$, H'_{j_1+1} has prefix $K \equiv H'H''H_1'''$ where $\|H_1'''\| = \|H'\| \geq \|H''\|$. Set $\mathcal{B}$ as the subband of the spoke $\mathcal{Q}_{j_1}$ with history K. Then, for any factorization $\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3$ such that the sum of the lengths of $\mathcal{B}_1$ and $\mathcal{B}_3$ is at most $\frac{1}{3}\|K\|$, the history of $\mathcal{B}_2$ must contain H''. So, since all θ-bands crossing $\mathcal{Q}_{j_1}$ must cross $\mathcal{Q}_{j_1-1}$, taking $\lambda < 1/3$ implies $\mathcal{B}$ is a λ-shaft with length $\|K\|$.

However, note that the subcomputation $W_0 \to \cdots \to V_m$ has length at least $h_{L_0-1} \geq h$, so that $\|K\| \geq \|H'\| \geq h > \delta h$. Thus, the existence of $\mathcal{B}$ in $\mathcal{Q}_{j_1}$ contradicts Lemma 11.21.

□

Chapter 12

Proof of Theorem A

We now complete the proof of the main theorem.

The first step toward this is to justify the assignments made throughout the construction.

12.1 Assignment of a-relations and weights

As mentioned in the introduction to the groups of interest in Chapter 6, the set of a-relators of interest in this chapter, $\mathcal{S}$, is the set of words in the letters $\mathcal{A} \cup \mathcal{A}^{-1}$ whose value in the free Burnside group $B(\mathcal{A}, n)$ is 1.

The following Lemma sheds some light on why these particular relations are adjoined to the group presentation.

Lemma 12.1. *For any word $u \in F(\mathcal{A})$, the relation $u^n = 1$ holds in the group $G(\boldsymbol{M})$.*

Proof. Lemmas 6.6 and 8.2 imply that the words corresponding to the configurations $I(u^n)$ and $J(u^n)$ are trivial over the group $G(\mathbf{M})$. These two words differ only by the insertion of the word u^n in the 'special' input sector, so that $u^n = 1$ in $G(\mathbf{M})$.

□

Lemma 12.2. *The groups $G(\boldsymbol{M})$ and $G_{\mathcal{S}}(\boldsymbol{M})$ are isomorphic.*

Proof. Identify $B(\mathcal{A}, n)$ with the presentation $\langle \mathcal{A} \mid w = 1, w \in \mathcal{L} \rangle$.

Then let $\varphi : \mathcal{A} \to G(\mathbf{M})$ be the map sending each letter to its natural copy in the tape alphabet of the 'special' input sector. By the theorem of von Dyck (Theorem 4.5 in [16]), Lemma 12.1 implies that φ extends to a homomorphism $B(\mathcal{A}, n) \to G(\mathbf{M})$. So, for any word w corresponding to an a-relation $w = 1$, the relation $w = 1$ holds in $G(\mathbf{M})$.

The theorem of von Dyck then implies that the map sending each generator of the canonical presentation of $G(\mathbf{M})$ to the corresponding generator of the disk presentation of $G_{\mathcal{S}}(\mathbf{M})$ extends to an isomorphism between the two groups.

□

Lemma 12.3. *The group $B(\mathcal{A},n)$ embeds in the group $G(\boldsymbol{M})$.*

Proof. Consider the natural map $\varphi:\mathcal{A}\to G_{\mathcal{S}}(\mathbf{M})$ sending the elements of $\mathcal{A}$ to their copies in the tape alphabet of the 'special' input sector. The theorem of von Dyck implies that this extends to a homomorphism $\varphi: B(\mathcal{A},n)\to G_{\mathcal{S}}(\mathbf{M})$.

Now suppose the reduced word w over $\mathcal{A}$ satisfies $\varphi(w)=1$. Then by Lemma 10.9, there exists a minimal diagram Δ over $G_{\mathcal{S}}(\mathbf{M})$ satisfying $\mathrm{Lab}(\partial\Delta)\equiv w$. By Lemmas 10.6 and 10.7, every cell of Δ must be an a-cell. But then this is a diagram over $B(\mathcal{A},n)$, so that $w=1$ in $B(\mathcal{A},n)$.

So, $\varphi: B(\mathcal{A},n)\to G_{\mathcal{S}}(\mathbf{M})$ is an embedding. Lemma 12.2 then implies the statement.

□

Now we wish to justify our assignment of weights to a-cells and disks over the disk presentation of $G_{\mathcal{S}}(\mathbf{M})$. To do so, we first study areas of a diagram over the canonical presentation of $G(\mathbf{M})$ with contour label corresponding to a disk relation.

Lemma 12.4. *(1) For any configuration W accepted by $\boldsymbol{M}$, there exists a reduced diagram Δ over the canonical presentation of $G(\boldsymbol{M})$ such that Lab$(\partial\Delta)\equiv W$ and Area$(\Delta)\leq C_1|W|^2$.*

(2) For any $u^n\in\mathcal{L}$, there exists a reduced diagram Δ over the canonical presentation of $G(\boldsymbol{M})$ with Lab$(\partial\Delta)\equiv u^n$ and Area$(\Delta)\leq C_1\|u\|^2$.

Proof. (1) By Lemma 6.17, there exists a computation $\mathcal{C}: W\equiv W_0\to\cdots\to W_t$ accepting W such that $t\leq c_2\|W(i)\|$ for all $i\geq 2$. Further, by Lemma 6.18, $\|W_j\|\leq c_2\|W\|$ for all j.

By Lemma 7.5, we can then build a trapezium Γ over $M(\mathbf{M})$ corresponding to $\mathcal{C}$, so that $\mathrm{Lab}(\mathbf{tbot}(\Gamma))\equiv W$ and $\mathrm{Lab}(\mathbf{ttop}(\Gamma))\equiv W_{ac}$.

Given a maximal θ-band $\mathcal{T}$ of Γ, $\|\mathbf{tbot}(\mathcal{T})\| \leq c_2\|W\|$. So, Area$(\Gamma) \leq 3c_2^2\|W\|^2$.

As the $Q_s(L)\{t(1)\}$-sector is locked by every rule, the sides of Γ are labelled identically and no trimming was necessary. So, we may glue these sides together and paste a hub into the middle of the diagram. This produces a reduced diagram Δ over the canonical presentation of $G(\mathbf{M})$ with Lab$(\partial\Delta) \equiv W$ and satisfying Area$(\Delta) \leq 3c_2^2\|W\|^2 + 1$.

The statement then follows as we choose the parameter C_1 after c_2 and δ.

(2) Clearly, we may assume that u^n is nontrivial in $F(\mathcal{A})$.

As in the previous case, we can build diagrams Δ_1 and Δ_2 over the canonical presentations of $G(\mathbf{M})$ where Δ_j is made of a hub and a trapezium satisfying:

- Lab$(\Delta_1) \equiv I(u^n)$ and Area$(\Delta_1) \leq 3c_2^2\|I(u^n)\|^2 + 1$
- Lab$(\Delta_2) \equiv J(u^n)$ and Area$(\Delta_2) \leq 3c_2^2\|J(u^n)\|^2 + 1$

Note that $\|I(u^n)\|, \|J(u^n)\| \leq L(11 + n\|u\|)$. So, since C_1 is chosen after c_2, L, and n, we can assume that Area$(\Delta_j) \leq \frac{1}{2}C_1\|u\|^2$ for $j = 1, 2$.

Gluing Δ_1 and Δ_2 along their common contours (and making any possible cancellations) then yields a diagram Δ satisfying the statement.

$\square$

Lemma 12.5. *If w is a reduced word over the alphabet $\mathcal{A}$ such that $w = 1$ in $B(\mathcal{A}, n)$, then there exists a reduced diagram Δ over the canonical presentation of $G(\boldsymbol{M})$ with Lab$(\partial\Delta) \equiv w$ and satisfying Area$(\Delta) \leq C_1\|w\|^2$.*

Proof. Let Δ_0 be a van Kampen diagram over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ of $B(\mathcal{A}, n)$ (see Section 3.7) with Lab$(\partial\Delta_0) \equiv w$. For each cell Π_0 in Δ_0, Lab$(\partial\Pi_0) \in \mathcal{R} \subset \mathcal{L}$. Setting Lab$(\partial\Pi_0) \equiv (u(\Pi_0))^n$, Lemma 12.4(2) then implies that there exists a diagram Π over the canonical presentation of $G(\mathbf{M})$ satisfying Lab$(\partial\Pi) \equiv (u(\Pi_0))^n$ and Area$(\Pi) \leq C_1\|u(\Pi_0)\|^2$.

Pasting Π in place of Π_0 for each cell of Δ_0 then produces a diagram Δ over the canonical presentation of $G(\mathbf{M})$ satsifying $\text{Lab}(\partial\Delta) \equiv w$ and

$$\text{Area}(\Delta) = \sum \text{Area}(\Pi) \leq \sum_{\Pi_0 \in \Delta_0} C_1 \|u(\Pi_0)\|^2$$

But defining $\rho(\Pi_0) = \|u(\Pi_0)\|^2$ as in the definition of mass in Section 3.8, Lemma 3.8 implies

$$\sum_{\Pi_0 \in \Delta_0} \|u(\Pi_0)\|^2 = \sum_{\Pi_0 \in \Delta_0} \rho(\Pi_0) := \rho(\Delta_0) \leq \|\partial\Delta_0\|^2$$

Hence, $\text{Area}(\Delta) \leq C_1 \|\partial\Delta_0\|^2 = C_1 \|w\|^2$.

$\square$

Note that the proof of Lemma 12.5 relies on the assumption that $n \in \mathbb{N}^*$.

12.2 Assignment of G-weight

Lemma 12.6. *Let Δ be an impeding a-trapezium. Then there exists a reduced diagram $\tilde{\Delta}$ over $G(\boldsymbol{M})$ such that Lab$(\partial\tilde{\Delta}) \equiv$ Lab$(\partial\Delta)$ and Area$(\tilde{\Delta}) \leq 2$wt$_G(\Delta)$.*

Proof. Suppose $\text{wt}_G(\Delta) = \frac{1}{2}\text{wt}(\Delta)$. Then, let $\tilde{\Delta}$ be the diagram constructed from Δ by replacing each a-cell with the corresponding reduced diagram over $G(\mathbf{M})$ constructed in Lemma 12.4. Then $\text{Area}(\tilde{\Delta}) \leq \text{wt}(\Delta) = 2\text{wt}_G(\Delta)$.

So, it suffices to assume that:

$$\text{wt}_G(\Delta) = 3h \max(\|\mathbf{tbot}(\Delta)\|, \|\mathbf{ttop}(\Delta)\|) + 3C_1 h\eta + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a + 2\eta)^2$$

for $\eta = \|H_1\| + n\|H_2\| + \|H_3\|$.

Let $\mathcal{T}_0$ be the maximal θ-band of Δ such that $\mathbf{bot}(\mathcal{T}_0) = \mathbf{bot}(\Delta)$. Letting θ_0 be the rule corresponding to $\mathcal{T}_0$, Lemma 7.2 implies that the admissible word $V_0 \equiv \text{Lab}(\mathbf{tbot}(\Delta))$ is θ_0-admissible.

If the base of Δ is $(P_0(1)Q_0(1))^{\pm 1}$, then V_0 is H-admissible by the definition of the rules.

Otherwise, the base of Δ is $Q_0(1)^{-1}Q_0(1)$. As V_0 is θ_0-admissible, its tape word must be nonempty. The application of each rule conjugates the tape word of this sector, so that V_0 must again be H-admissible.

Suppose $\ell \leq n$. Let $\mathcal{C}$ be the reduced computation starting with V_0 and with history H. Then, let Γ be the trapezium corresponding to $\mathcal{C}$ by Lemma 7.5. By Lemmas 4.3 and 4.4, for any maximal θ-band $\mathcal{T}$ of Γ, $\|\mathbf{tbot}(\mathcal{T})\| \leq \max(\|\mathbf{tbot}(\Gamma)\|, \|\mathbf{ttop}(\Gamma)\|)$, so that $\mathrm{wt}(\Gamma) \leq 3h\max(\|\mathbf{tbot}(\Gamma)\|, \|\mathbf{ttop}(\Gamma)\|)$.

Note that $|\mathbf{ttop}(\Gamma)|_a \leq |V_0|_a + 2h$ and $h \leq \|H_1\| + n\|H_2\| + \|H_3\| = \eta$, so that

$$\mathrm{wt}(\Gamma) \leq 3h(\|\mathbf{tbot}(\Delta)\| + 2\eta)$$

Further, the bottom and side labels of Γ are the same as those of Δ, while the top labels differ only by a word w from the 'special' input sector with $\|w\| \leq |\mathbf{ttop}(\Gamma)|_a + |\mathbf{ttop}(\Delta)|_a$. Pasting Γ and Δ along their shared contour then yields a diagram over $M_a(\mathbf{M})$ with contour label w. By Lemma 12.3, w must be an a-relation. So, we may paste an a-cell corresponding to w to the top of Γ to produce a diagram $\tilde{\Gamma}$ with the same contour label as Δ and

$$\mathrm{wt}(\tilde{\Gamma}) \leq 3h(\|\mathbf{tbot}(\Delta)\| + 2\eta) + C_1(|\mathbf{tbot}(\Delta)|_a + 2\eta + |\mathbf{ttop}(\Delta)|_a)^2 \leq \mathrm{wt}_G(\Delta)$$

Now suppose $\ell > n$.

Let $\mathcal{C}_1$ be the reduced computation starting with V_0 and having history H_1 and Γ_1 be the trapezium corresponding to $\mathcal{C}_1$ by Lemma 7.5. Set $V_1 \equiv V_0 \cdot H_1 \equiv \mathrm{Lab}(\mathbf{ttop}(\Gamma_1))$.

As V_0 is H-admissible, there exists a reduced computation $\mathcal{C}_2$ starting with V_1 and having history H_2^n. For $q \in \mathbb{N}$ such that $\ell = qn + r$ with $0 \leq r < n$, let $\Gamma_2(1), \dots, \Gamma_2(q)$ be q copies of the trapezium corresponding to $\mathcal{C}_2$ by Lemma 7.5.

Let v_1 be the tape word of V_1. If the base of Δ is $P_0(1)Q_0(1)$ (or $Q_0(1)^{-1}P_0(1)^{-1}$), then the tape word written on $\mathbf{ttop}(\Gamma_2(i))$ is equal to v_1u^{-n} (or u^nv_1) in $F(\mathcal{A})$, where u is the natural copy of H_2 over the alphabet $\mathcal{A}$. Otherwise, the base of Δ is $Q_0(1)^{-1}Q_0(1)$, so that the tape word written on $\mathbf{ttop}(\Gamma_2(i))$ is equal in $F(\mathcal{A})$ to $u^nv_1u^{-n}$.

In each case, the projection of Lab(**ttop**$(\Gamma_2(i))$) onto $F(\mathcal{A})$ is equivalent to v_1 in $B(\mathcal{A}, n)$. In particular, for $1 \leq i \leq q$ we may attach a-cells corresponding to $u^{\pm n}$ to **top**$(\Gamma_2(i))$ so that the top of the resulting diagram $\tilde{\Gamma}_2(i)$ is equivalent to V_1. By the assignment of weights, each of these a-cells has weight at most $C_1 n^2 \|H_2\|^2$.

Finally, let $\tilde{\Gamma}_2(q+1)$ be the trapezium corresponding to the reduced computation starting with V_1 and having history H_2^r.

Then, we paste the top of $\tilde{\Gamma}_2(i)$ to the bottom of $\tilde{\Gamma}_2(i+1)$ for each $1 \leq i \leq q$ to form the diagram Γ_2.

Let $V_2 \equiv$ Lab(**ttop**$(\tilde{\Gamma}_2(q+1))$). Note that V_2 is θ_2-admissible, where θ_2 is the first rule of H_2. So as above, V_2 is H_3-admissible, i.e there exists a reduced computation $\mathcal{C}_3$ starting with V_2 and having history H_3. Let Γ_3 be the corresponding trapezium.

Finally, we form the the diagram Γ by pasting together Γ_1, Γ_2, and Γ_3.

Note that $\|V_1\| \leq \|V_0\| + 2\|H_1\|$, while $\|V_2\| \leq \|V_1\| + 2r\|H_2\| \leq \|V_0\| + 2\|H_1\| + 2r\|H_2\|$. So, for any maximal θ-band $\mathcal{T}$ of Γ, $|$**tbot**$(\mathcal{T})|_a \leq |V_0|_a + 2\eta$.

In particular, $|$**ttop**$(\Gamma)|_a \leq |V_0|_a + 2\eta$ and

$$\text{wt}(\Gamma) \leq 3h(\|\textbf{tbot}(\Delta)\| + 2\eta) + 2C_1 q n^2 \|H_2\|^2$$

As above, there exists an a-relation w with $\|w\| \leq |$**ttop**$(\Gamma)|_a + |$**ttop**$(\Delta)|_a$ such that if we paste the a-cell corresponding to w to the top of Γ, we obtain a diagram $\tilde{\Gamma}$ with Lab$(\partial\tilde{\Gamma}) \equiv$ Lab$(\partial\Delta)$ and

$$\text{wt}(\tilde{\Gamma}) \leq 3h(\|\textbf{tbot}(\Delta)\| + 2\eta) + 2C_1 q n^2 \|H_2\|^2 + C_1(|\textbf{tbot}(\Delta)|_a + 2\eta + |\textbf{ttop}(\Delta)|_a)^2$$

As $qn\|H_2\| \leq \ell\|H_2\| \leq h$, we also have $2C_1 q n^2 \|H_2\|^2 \leq 2C_1 h n \|H_2\| \leq 2C_1 h\eta$. Hence, taking $C_1 \geq 6$,

$$\text{wt}(\tilde{\Gamma}) \leq 3h\|\textbf{tbot}(\Delta)\| + (2C_1 + 6)h\eta + C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a + 2\eta)^2 \leq \text{wt}_G(\Delta)$$

Thus, the reduced diagram $\tilde{\Delta}$ obtained from $\tilde{\Gamma}$ by replacing any a-cell with the diagram constructed in Lemma 12.4 (and making any necessary cancellations) satisfies the statement.

$\square$

Lemma 12.7. *For every big a-trapezium Δ, there is a reduced diagram $\tilde{\Delta}$ over the finite presentation of $G(\boldsymbol{M})$ such that Lab$(\partial\tilde{\Delta}) \equiv$ Lab$(\partial\Delta)$ and Area$(\tilde{\Delta}) \leq 2wt_G(\Delta)$.*

Proof. As in the proof of Lemma 12.6, if $\text{wt}_G(\Delta) = \frac{1}{2}\text{wt}(\Delta)$, then we may construct $\tilde{\Delta}$ simply by replacing all a-cells with the corresponding diagram constructed in Lemma 12.4. Hence, it suffices to assume that

$$\text{wt}_G(\Delta) = c_5 h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|) + 4C_1(|\textbf{tbot}(\Delta)|_a + |\textbf{ttop}(\Delta)|_a)^2$$

Without loss of generality, suppose the base of Δ begins and ends with $\{t(1)\}$. Then, let $\Delta_1, \dots, \Delta_L$ be the maximal subdiagrams of Δ bounded by the maximal t-bands, so that each is an a-trapezium with pararevolving base.

By Lemma 9.4, only Δ_1 may contain a-cells, so that $\Delta_2, \dots, \Delta_L$ are trapezia.

For $\mathcal{T}$ a maximal θ-band of Δ, let $\mathcal{T}_i$ be the subband which is a maximal θ-band of Δ_i.

Suppose $h \leq c_3 \max(\|\textbf{tbot}(\Delta_2)\|, \|\textbf{ttop}(\Delta_2)\|)$. Then, the parameter choice $c_4 >> c_3$ implies $\|\textbf{tbot}(\mathcal{T}_2)\| \leq c_4 \max(\|\textbf{tbot}(\Delta_2)\|, \|\textbf{ttop}(\Delta_2)\|)$.

Since Δ is big, its history contains a controlled subword. So, by Lemma 6.7, each Δ_i must be a coordinate shift of Δ_2 for $i \geq 2$, so that $h \leq c_3 \max(\|\textbf{tbot}(\Delta_i)\|, \|\textbf{ttop}(\Delta_i)\|)$ and $\|\textbf{tbot}(\mathcal{T}_i)\| \leq c_4 \max(\|\textbf{tbot}(\Delta_i)\|, \|\textbf{ttop}(\Delta_i)\|)$. As $L >> c_3$, this implies $h \leq \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|)$.

Moreover, the only sector of Δ_1 that may not be a coordinate shift of the corresponding sector of Δ_2 is the ‘special’ input sector.

For any a-edge **e** of $\textbf{tbot}(\mathcal{T}_1)$ in the ‘special’ input sector, let $\mathcal{B}$ be the maximal a-band containing **e**. Then Lemma 9.2 and (MM3) imply that $\mathcal{B}$ must have one end on $\textbf{ttop}(\Delta_1)$, on $\textbf{tbot}(\Delta_1)$, or on the q-band corresponding to $Q_0(1)$.

So, $\|\textbf{tbot}(\mathcal{T}_1)\| \leq \|\textbf{tbot}(\mathcal{T}_2)\| + \|\textbf{tbot}(\Delta_1)\| + \|\textbf{ttop}(\Delta_1)\| + h$. It then follows that:

$$\begin{aligned}
\|\textbf{tbot}(\mathcal{T})\| &\leq \sum_{i=1}^{L} \|\textbf{tbot}(\mathcal{T}_i)\| \leq \left(\sum_{i=2}^{L} 2\|\textbf{tbot}(\mathcal{T}_i)\| \right) + \|\textbf{tbot}(\Delta_1)\| + \|\textbf{ttop}(\Delta_1)\| + h \\
&\leq \left(\sum_{i=2}^{L} 2c_4 \max(\|\textbf{tbot}(\Delta_i)\|, \|\textbf{ttop}(\Delta_i)\|) \right) + \|\textbf{tbot}(\Delta_1)\| + \|\textbf{ttop}(\Delta_1)\| + h \\
&\leq 2c_4 \left(\sum_{i=1}^{L} \|\textbf{tbot}(\Delta_i)\| + \|\textbf{ttop}(\Delta_i)\| \right) + h \\
&\leq 4c_4(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|) + h \leq (8c_4+1)\max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|)
\end{aligned}$$

So, by the parameter choice $c_5 >> c_3$, the sum of the lengths, and so the weights, of the maximal θ-bands of Δ is at most $c_5 h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|)$.

Any other cell of Δ is an a-cell. As above, Lemma 9.2 and (MM2) then imply that each maximal a-band starting from an a-cell must end on the q-band corresponding to $Q_0(1)$, on $\textbf{ttop}(\Delta)$, or on $\textbf{tbot}(\Delta)$. So, the sum of the combinatorial perimeters of the a-cells of Δ is at most $h + |\textbf{ttop}(\Delta)|_a + |\textbf{tbot}(\Delta)|_a$, i.e at most $2(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|)$. As a result,

$$\mathrm{wt}(\Delta) \leq c_5 h \max(\|\textbf{tbot}(\Delta)\|, \|\textbf{ttop}(\Delta)\|) + 4C_1(\|\textbf{tbot}(\Delta)\| + \|\textbf{ttop}(\Delta)\|)^2 = \mathrm{wt}_G(\Delta)$$

The reduced diagram $\tilde{\Delta}$ constructed from Δ by replacing any a-cell with the corresponding reduced diagram from Lemma 12.4 then satisfies the statement.

Hence, we may assume that $h > c_3 \max(\|\textbf{tbot}(\Delta_2)\|, \|\textbf{ttop}(\Delta_2)\|)$.

As Δ is big, its history must contain a controlled subword H'. Let Δ' be the subtrapezium whose history is H'.

Let Δ_- be the subdiagram of Δ obtained by removing the maximal q-band $\mathcal{Q}$ corresponding to the final letter of the base of Δ. So, Δ_- is an a-trapezium with the standard base. Similarly define Δ'_- as the corresponding subdiagram of Δ'. Lemma 6.7 then implies that $\mathrm{Lab}(\textbf{tbot}(\Delta'_-)) = \mathrm{Lab}(\textbf{bot}(\Delta'_-))$ is an accepted configuration.

As Δ_2 is a trapezium, Lemma 7.4 yields a corresponding computation $\mathcal{C}_2 : V_0 \to \cdots \to V_h$

with base $\{t(2)\}B_3(2)$ satisfying $h > c_3 \max(\|V_0\|, \|V_h\|)$. Hence, $\mathcal{C}_2$ satisfies the hypotheses of Lemma 6.19, so that there exist accepted configurations W_0' and W_h' with $W_0'(2) \equiv V_0$ and $W_h'(2) \equiv V_h$.

Let Δ_0 be the subdiagram of Δ_- which is an a-trapezium with $\mathbf{bot}(\Delta_0) = \mathbf{bot}(\Delta_-)$ and $\mathbf{top}(\Delta_0) = \mathbf{bot}(\Delta_-')$. Then Δ_0 is an a-trapezium with sides labelled identically and top labelled by an accepted configuration. So, $W_0 \equiv \text{Lab}(\mathbf{bot}(\Delta_-))$ must be a configuration which is trivial in $G_\mathcal{S}(\mathbf{M})$.

Similarly, $W_h \equiv \text{Lab}(\mathbf{top}(\Delta_-))$ is a configuration which is trivial in $G_\mathcal{S}(\mathbf{M})$.

By Lemma 9.4, any sector of Δ_0 other than the 'special' input sector is a trapezium. So, since $\text{Lab}(\mathbf{top}(\Delta_0))$ is an accepted configuration, the parallel nature of the rules implies that $W_0(i)$ and $W_0(j)$ are coordinate shifts of one another for $i, j \geq 2$ while the corresponding coordinate shift of $W_0(i)$ differs from $W_0(1)$ only in the 'special' input sector. Hence, since W_0' is an accepted configuration with $W_0(2) \equiv W_0'(2)$, W_0 and W_0' can differ only in the 'special' input sector.

Let w_0 and w_0' be the tape words of W_0 and W_0', respectively, in this sector. Then, since W_0 and W_0' are each trivial over $G_\mathcal{S}(\mathbf{M})$, it follows that $w_0^{-1}w_0'$ is trivial over $G_\mathcal{S}(\mathbf{M})$. Lemma 12.3 then implies that $w_0^{-1}w_0'$ is trivial over $B(\mathcal{A}, n)$, so that it corresponds to an a-relation.

Let H_0 be the history of an accepting computation $\mathcal{C}_0$ of W_0' with $\ell(\mathcal{C}_0) = \ell(W_0')$. Then using Lemmas 6.17 and 7.5, we may construct a reduced diagram Ψ_0' over $G(\mathbf{M})$ with $\text{Lab}(\partial\Psi_0') \equiv W_0'$ consisting of one hub and a trapezium with $\text{Area}(\Psi_0') \leq c_3\|H_0\|\|W_0'\|$ ($c_3 >> c_2$).

By Lemma 12.5, we also construct a reduced diagram Ψ_0'' over $G(\mathbf{M})$ with $\text{Lab}(\partial\Psi_0'') \equiv w_0^{-1}w_0'$ with $\text{Area}(\Psi_0'') \leq C_1(\|w_0\| + \|w_0'\|)^2$.

Lemma 6.16 implies that $\|w_0'\| \leq |W_0'(1)|_a \leq 2|W_0'(2)|_a = 2|W_0(2)|_a$. So,

$$\|W_0'\| \leq \|W_0\| + \|w_0'\| \leq \|W_0\| + |W_0(2)|_a + |W_0(3)|_a \leq 2\|W_0\|$$

Further, $\|w_0\| \leq |\mathbf{tbot}(\Delta_1)|_a$ and, since $L \geq 3$, $\|w_0'\| \leq |\mathbf{tbot}(\Delta_2)|_a + |\mathbf{tbot}(\Delta_3)|_a$.

Let Ψ_0 be the diagram obtained from pasting Ψ_0' and Ψ_0'' along their common boundary labels.

Then, $\text{Lab}(\partial\Psi_0) \equiv W_0$ and $\text{Area}(\Psi_0) \leq 2c_3\|H_0\|\|W_0\| + C_1(|W_0|_a)^2$.

Similarly, we may construct a reduced diagram Ψ_h over the finite presentation of $G(\mathbf{M})$ satisfying $\text{Lab}(\partial\Psi_h) \equiv W_h$ and $\text{Area}(\Psi_h) \leq 2c_3\|H_h\|\|W_h\| + C_1(|W_h|_a)^2$, where H_h is the history of an accepting computation $\mathcal{C}_h$ of W_h' with $\ell(\mathcal{C}_h) = \ell(W_h')$.

Attaching the corresponding ends of $\mathcal{Q}$ to Ψ_0 and Ψ_h, we then obtain a reduced diagram $\tilde{\Delta}$ with $\text{Lab}(\partial\tilde{\Delta}) \equiv \text{Lab}(\partial\Delta)$ and

$$\begin{aligned}\text{Area}(\tilde{\Delta}) &\leq 2c_3(\|H_0\|\|W_0\| + \|H_h\|\|W_h\|) + C_1(|W_0|_a + |W_h|_a)^2 + h \\ &\leq 2c_3(\|H_0\| + \|H_h\|)(\|\mathbf{tbot}(\Delta)\| + \|\mathbf{ttop}(\Delta)\|) + C_1(|\mathbf{tbot}(\Delta)|_a + |\mathbf{ttop}(\Delta)|_a)^2 + h\end{aligned}$$

By Lemma 6.19, $\|H_0\| + \|H_h\| \leq h$.

Thus, the parameter choice $c_5 >> c_3$ implies $\text{Area}(\tilde{\Delta}) \leq \text{wt}_G(\Delta)$.

□

12.3 Quadratic upper bound

Finally, we complete the proof of Theorem A.

As Lemma 12.3 implies that $G(\mathbf{M})$ contains an infinite torsion subgroup, the Dehn function of $G(\mathbf{M})$ is at least quadratic. Thus, it suffices to prove a quadratic upper bound bound.

Let $w \in F(\mathcal{X})$ such that $w = 1$ in $G(\mathbf{M})$. By Lemma 12.2, w is also trivial over the group $G_{\mathcal{S}}(\mathbf{M})$, so that Lemma 10.9 yields a minimal diagram Δ_a over $G_{\mathcal{S}}(\mathbf{M})$ with $\text{Lab}(\partial\Delta_a) \equiv w$. By Lemma 11.26, we have

$$\text{wt}_G(\Delta_a) \leq N_4(|w| + \sigma_\lambda(\Delta_a^*))^2 + N_3\mu(\Delta_a)$$

Lemma 10.15 implies that $\sigma_\lambda(\Delta_a^*) \leq C_1|w|$, while Lemma 8.3(a) implies $\mu(\Delta_a) \leq J|w|^2$. So, as

$|w| \leq \|w\|$, we can choose N_5 large enough so that

$$\text{wt}_G(\Delta_a) \leq \frac{1}{2} N_5 \|w\|^2$$

Now, let **P** be a minimal covering of Δ_a and construct the reduced diagram Δ over the canonical presentation of $G(\mathbf{M})$ by:

- excising any impeding a-trapezium $P \in \mathbf{P}$ and pasting in its place the reduced diagram given in Lemma 12.6 with the same contour label and area at most $2\text{wt}_G(P)$
- excising any big a-trapezium $P \in \mathbf{P}$ and pasting in its place the reduced diagram given in Lemma 12.7 with the same contour label and area at most $2\text{wt}_G(P)$
- excising any disk $\Pi \in \mathbf{P}$ and pasting in its place the reduced diagram given in Lemma 12.4 with the same contour label and area at most $C_1|\partial\Pi|^2$
- excising any a-cell $\pi \in \mathbf{P}$ and pasting in its place the reduced diagram given in Lemma 12.5 with the same contour label and area at most $C_1\|\partial\pi\|^2$

By the definition of G-weight, it follows that $\text{Area}(\Delta) \leq 2\text{wt}_G(\Delta_a) \leq N_5\|w\|^2$.

Therefore, the Dehn function of $G(\mathbf{M})$ is at most quadratic, and so the proof of Theorem A is complete.

Chapter 13

Proof of Theorem B

13.1 g-diagrams and g-minimal diagrams

By Lemma 12.3, every $g \in B(\mathcal{A}, n)$ can be identified with an element of $G_{\mathcal{S}}(\mathbf{M})$, namely $\varphi(g)$. For $g \in B(\mathcal{A}, n)$, define $|g|_{\mathcal{A}}$ as the smallest number of letters comprising a word over $\mathcal{A}$ whose value in $B(\mathcal{A}, n)$ is g.

For $g \in B(\mathcal{A}, n)$, a minimal diagram Δ is called a *g-diagram* if $\partial\Delta = \mathbf{st}$, Lab($\mathbf{t}$) is a word over $\mathcal{A}$ whose value in $B(\mathcal{A}, n)$ is g^{-1}, and $\|\mathbf{t}\| = |g|_{\mathcal{A}}$.

A g-diagram Δ is called *g-minimal* if $|\partial\Delta| + \sigma_\lambda(\Delta)$ is minimal amongst all g-diagrams.

Lemma 13.1. *For $g \in B(\mathcal{A}, n)$, if Δ is a g-minimal diagram, then $|\partial\Delta| + \sigma_\lambda(\Delta) \leq 2\delta|g|_{\mathcal{A}}$.*

Proof. Let v be a word over $\mathcal{A}$ whose value in $B(\mathcal{A}, n)$ is g and such that $\|v\| = |g|_{\mathcal{A}}$.

By van Kampen's Lemma, there exists a diagram Φ over $F(\mathcal{A})$ (in which every cell is a 0-cell) with Lab($\partial\Phi$) $\equiv vv^{-1}$. Viewing Φ as a diagram over $G_{\mathcal{S}}(\mathbf{M})$, it is clear that Φ is a g-diagram containing no disks, so that $\sigma_\lambda(\Phi) = 0$.

Hence, for any g-minimal diagram Δ,

$$|\partial\Delta| + \sigma_\lambda(\Delta) \leq |\partial\Phi| = 2\delta\|v\| = 2\delta|g|_{\mathcal{A}}$$

□

Lemma 13.2. *If Δ is a g-minimal diagram for some $g \in B(\mathcal{A}, n)$, then no q-band of Δ has two ends on $\partial\Delta$.*

Proof. Decompose $\partial\Delta = \mathbf{st}$ as in the definition of g-diagram.

Suppose $\mathcal{Q}$ is a q-band with two ends on $\partial\Delta$. Then, since $\mathbf{t}$ consists entirely of a-edges, both ends of $\mathcal{Q}$ must be edges of $\mathbf{s}$.

Let $\mathbf{s}_0$ be the subpath of $\mathbf{s}$ bounded by the two ends of $\mathcal{Q}$. So, $\mathbf{s}_0$ and a side of $\mathcal{Q}$, say $\mathbf{top}(\mathcal{Q})$, bound a subdiagram Δ_0 of Δ containing $\mathcal{Q}$.

By Lemma 8.1(b), $|\mathbf{top}(\mathcal{Q})| = \ell$, where ℓ is the length of $\mathcal{Q}$.

Further, by Lemma 10.8, every maximal θ-band of Δ_0 must have two ends on $\partial\Delta_0$. Lemma 9.1 implies that no θ-band can end twice on $\mathbf{top}(\mathcal{Q})$. As a result, each of the ℓ θ-edges of $\mathbf{top}(\mathcal{Q})$ correspond to a θ-edge of $\mathbf{s}_0$, so that Lemma 8.1(a) implies $|\mathbf{s}_0| \geq \ell + 2$.

Consider the diagram Δ' obtained from Δ by cutting off Δ_0. As Δ' is a subdiagram of Δ, it is minimal. Moreover, $\partial\Delta'$ can be decomposed as $\mathbf{s}_1(\mathbf{top}(\mathcal{Q}))\mathbf{s}_2\mathbf{t}$, where $\mathbf{s} = \mathbf{s}_1\mathbf{s}_0\mathbf{s}_2$. As a result, Δ' is a g-diagram.

By Lemma 8.1(c),

$$|\partial\Delta'| \leq |\mathbf{s}_1| + |\mathbf{top}(\mathcal{Q})| + |\mathbf{s}_2| + |\mathbf{t}| = |\mathbf{s}_1| + \ell + |\mathbf{s}_2| + |\mathbf{t}|$$
$$\leq |\mathbf{s}_1| + |\mathbf{s}_0| - 2 + |\mathbf{s}_2| + |\mathbf{t}| \leq |\partial\Delta| - (2 - 4\delta)$$

A parameter choice for δ then implies $|\partial\Delta'| \leq |\partial\Delta| - 1$.

Finally, as Δ_0 and Δ' are disjoint, $\sigma_\lambda(\Delta) \geq \sigma_\lambda(\Delta_0) + \sigma_\lambda(\Delta')$. In particular, this implies

$$|\partial\Delta'| + \sigma_\lambda(\Delta') \leq |\partial\Delta| + \sigma_\lambda(\Delta) - 1$$

But this contradicts the assumption that Δ is a g-minimal diagram.

□

Let Δ be a g-minimal diagram for some $g \in B(\mathcal{A}, n)$ and decompose $\partial\Delta = \mathbf{st}$ as in the definition of g-diagram. Suppose Δ contains a quasi-rim θ-band $\mathcal{T}$. Since $\mathbf{t}$ is comprised entirely of a-edges, $\mathcal{T}$ must end twice on $\mathbf{s}$. Let $\mathbf{s}_0$ be the subpath of $\partial\Delta$ bounded by the two ends of $\mathcal{T}$ such that, per the definition of quasi-rim θ-band, any cell between $\mathbf{bot}(\mathcal{T})$ (or $\mathbf{top}(\mathcal{T})$) and $\mathbf{s}_0$ is an a-cell. If $\mathbf{s}_0$ is a subpath of $\mathbf{s}$, then $\mathcal{T}$ is called a *g-rim θ-band*.

Lemma 13.3. *Let Δ be a g-minimal diagram for some $g \in B(\mathcal{A}, n)$. If $\mathcal{T}$ is an g-rim θ-band in*

Δ, then the base of $\mathcal{T}$ has length $s > K$.

Proof. Suppose to the contrary that the base of $\mathcal{T}$ has length $s \leq K$.

Decompose $\partial\Delta = \mathbf{st}$ as in the definition of g-diagram and let $\mathbf{s}_0$ be the subpath of $\mathbf{s}$ as in the definition of g-rim θ-band.

Suppose every cell between $\mathbf{bot}(\mathcal{T})$ and $\mathbf{s}_0$ is an a-cell. As in the proof of Lemma 13.2, let Δ_0 be the subdiagram bounded by $\mathbf{top}(\mathcal{T})$ and $\mathbf{s}_0$ and let Δ' be the diagram obtained from Δ by cutting off Δ_0.

Then, any cell of Δ_0 not comprising $\mathcal{T}$ is an a-cell. Letting π be such an a-cell, for any edge $\mathbf{e}$ of $\partial\pi$, either $\mathbf{e}$ is shared with $\mathbf{s}_0$ or $\mathbf{e}^{-1}$ is shared with $\mathbf{bot}(\mathcal{T})$. Similarly, any edge of $\mathbf{bot}(\mathcal{T})$ is either shared with $\mathbf{s}_0$ or its inverse is on the boundary of an a-cell in Δ_0.

For an a-cell π in Δ_0, let $\partial\pi = \mathbf{p}_\pi\mathbf{q}_\pi$, where $\mathbf{p}_\pi$ is a maximal subpath shared with $\partial\Delta$. By (M1), at most $\frac{1}{2}\|\partial\pi\|$ edges of $\partial\pi$ are shared with the boundary of a (θ, a)-cell of $\mathcal{T}$. So, $\|\mathbf{q}_\pi\| \leq \frac{1}{2}\|\partial\pi\| + b_\pi$, where b_π is the number of edges of $\partial\pi$ shared with the boundary of a (θ, q)-cell of $\mathcal{T}$. As a result, $\|\mathbf{p}_\pi\| \geq \frac{1}{2}\|\partial\pi\| - b_\pi \geq \|\mathbf{q}_\pi\| - 2b_\pi$.

Since $\mathcal{T}$ contains s (θ, q)-cells, the boundary of any of which contains at most one a-edge labelled by a letter from the 'special' input sector, $\sum b_\pi \leq s$. Hence, $|\mathbf{s}_0|_a \geq |\mathbf{bot}(\mathcal{T})|_a - 2s \geq |\mathbf{bot}(\mathcal{T})|_a - 2K$.

As every q-edge of $\mathbf{bot}(\mathcal{T})$ is shared with $\mathbf{s}_0$, it then follows from Lemma 8.1 that

$$\begin{aligned} |\mathbf{s}_0| &\geq 2 + |\mathbf{s}_0|_q + \delta(|\mathbf{s}_0|_a - 2) \geq 2 - 2\delta + |\mathbf{bot}(\mathcal{T})|_q + \delta(|\mathbf{bot}(\mathcal{T})|_a - 2K) \\ &\geq 2 + |\mathbf{bot}(\mathcal{T})| - (2K+2)\delta \end{aligned}$$

Further, by Lemma 4.1, $|\mathbf{top}(\mathcal{T})|_a \leq |\mathbf{bot}(\mathcal{T})|_a + 2s \leq |\mathbf{bot}(\mathcal{T})|_a + 2K$.

Thus, $|\mathbf{s}_0| \geq 2 - (2K+2)\delta + |\mathbf{top}(\mathcal{T})| - 2K\delta \geq |\mathbf{top}(\mathcal{T})| + 2 - (4K+2)\delta \geq |\mathbf{top}(\mathcal{T})| + 1$ by the parameter choice $\delta^{-1} >> K$.

But then Lemma 8.1 and a parameter choice for δ implies that Δ' is a g-diagram satisfying $|\partial\Delta'| + \sigma_\lambda(\Delta') < |\partial\Delta| + \sigma_\lambda(\Delta)$, contradicting the assumption that Δ is g-minimal.

□

Lemma 13.4. *Let Δ be a g-minimal diagram for some $g \in B(\mathcal{A}, n)$ and decompose $\partial\Delta = $ **st** as in the definition of g-diagram. If Δ contains no disks, then $|$**s**$| = |$**t**$| = \delta|g|_{\mathcal{A}}$.*

Proof. By the definition of the design, $\sigma_\lambda(\Delta) = 0$. Further, Lemma 13.2 implies that Δ contains no q-bands, so that Lemma 9.1 implies that $\partial\Delta$ contains no q-edges.

If Δ contains a θ-band, then Lemma 9.3 implies that Δ contains a g-rim θ-band $\mathcal{T}$. But then $\mathcal{T}$ has base of length zero, contradicting Lemma 13.3.

Hence, any cell of Δ is an a-cell and every edge of $\partial\Delta$ is an a-edge.

Suppose **e** is an edge of **s** that is not on the boundary of an a-cell. Then the maximal a-band of Δ starting at **e** must be of length zero, i.e $\mathbf{e}^{-1}$ is an edge of $\partial\Delta$. If $\mathbf{e}^{-1}$ is part of **s**, then deleting the subpath of **s** bounded by **e** and $\mathbf{e}^{-1}$ results in a g-diagram with smaller perimeter, contradicting the assumption that Δ is g-minimal. So, $\mathbf{e}^{-1}$ must be an edge of **t**.

As a result, Lab$(\partial\Delta)$ must be a word over $\mathcal{A}$.

Now, as in the proof of Lemma 12.5, excise any a-cell of Δ and paste in its place an appropriate reduced diagram over the presentation $\langle \mathcal{A} \mid \mathcal{R} \rangle$ of $B(\mathcal{A}, n)$. This produces a reduced diagram Ψ over $B(\mathcal{A}, n)$ with Lab$(\partial\Psi) \equiv$ Lab$(\partial\Delta)$. Hence, Lab$(\mathbf{s}) = g$ in $B(\mathcal{A}, n)$.

By the definition of the word norm, this implies that $|\mathbf{s}| = \delta\|\mathbf{s}\| \geq \delta|g|_{\mathcal{A}}$.

Hence, $|\partial\Delta| = |\mathbf{s}| + |\mathbf{t}| \geq 2\delta|g|_{\mathcal{A}}$.

But then Lemma 13.1 implies $|\partial\Delta| = 2\delta|g|_{\mathcal{A}}$, so that $|\mathbf{s}| = \delta|g|_{\mathcal{A}}$.

□

13.2 g-minimal diagrams containing disks

Lemma 13.5. *Let Δ be a g-minimal diagram for some $g \in B(\mathcal{A}, n)$ containing at least one disk. Decompose $\partial\Delta = $ **st** as in the definition of g-diagram. Then Δ contains a disk Π such that:*

(a) $L - 6$ consecutive t-spokes $\mathcal{Q}_1, \dots, \mathcal{Q}_{L-6}$ of Π end on $\partial\Delta$

(b) for $i = 1, \ldots, L-7$*, the subdiagram* $\Psi_{i,i+1}$ *of* Δ *bounded by* $\mathcal{Q}_i$*,* $\mathcal{Q}_{i+1}$*,* $\partial\Pi$*, and* $\partial\Delta$ *contains no disks, and*

(c) $\mathbf{t}$ *is not a subpath of* $\partial\Psi_{i,i+1}$ *for any* $i = 1, \ldots, L-7$*.*

Proof. Let Π_1 be a disk in Δ guaranteed by Lemma 10.6. As **t** consists entirely of a-edges, every t-spoke of Π ending on $\partial\Delta$ must end on **s**.

Let Ψ_1 be the subdiagram of Δ bounded by $\partial\Pi_1$, $\partial\Delta$, and the $L-4$ consecutive t-spokes of Π_1 ending on $\partial\Delta$. Then, one may assume that $\partial\Psi_1$ contains **t** as a subpath, as otherwise choosing $\Pi = \Pi_1$ satisfies the statement. Similarly, one may assume that the complement of $\Psi_1 \cup \Pi_1$ in Δ, Ψ_1', contains a disk, as otherwise Π_1 is the only disk in Δ.

Then, decompose $\partial\Psi_1' = \mathbf{s}_1\mathbf{t}_1$ where $\mathbf{s}_1$ is the maximal subpath of **s**.

Next, apply Lemma 10.6 to Ψ_1', yielding a disk Π_2. Let Ψ_2 be the subdiagram of Δ bounded by $\partial\Pi_2$, $\partial\Psi_2$, and the $L-4$ consecutive t-spokes of Π_2 ending on $\partial\Psi_1'$.

Suppose $\partial\Psi_2$ does not contain $\mathbf{t}_1$. Then, by Lemma 10.5, at most two of the $L-4$ t-spokes of Π_2 ending on $\partial\Psi_1'$ end on $\mathbf{t}_1$, in which case such a spoke is the first or last in the sequence. As such, choosing $\Pi = \Pi_2$ satisfies the statement. So, one may assume that $\partial\Psi_2$ contains $\mathbf{t}_1$.

Similarly, one may assume that the complement of $\Psi_2 \cup \Pi_2$ in Ψ_1', Ψ_2', contains a disk, as otherwise Π_2 is the only disk in Ψ_1'.

Now decompose $\partial\Psi_2' = \mathbf{s}_2\mathbf{t}_2$ where $\mathbf{s}_2$ is a maximal subpath of **s** and apply Lemma 10.6 to Ψ_2', yielding a disk Π_3.

Continuing in this way, the finiteness of Δ implies that the process must terminate. Hence, there exists ℓ such that Π_ℓ satisfies the statement.

□

The goal throughout the rest of this section is to prove that for any $g \in B(\mathcal{A}, n)$, a g-minimal diagram must be diskless. This is done by arguing toward contradiction in much the same way as proof presented in Chapter 11.

As such, we adopt much of the same notation of Chapter 11 for a g-minimal diagram Δ containing a disk Π satisfying Lemma 13.5. So, for $1 \leq i < j \leq L-6$, we define the the subdiagrams Ψ_{ij}, Ψ'_{ij}, and $\bar{\Delta}_{ij}$ as well as the paths $\mathbf{p}_{ij}$ and $\bar{\mathbf{p}}_{ij}$. As in Chapter 11, the subscripts are suppressed in the case that $j-i$ is maximal; in other words, $\Psi = \Psi_{1,L-6}$, $\mathbf{p} = \mathbf{p}_{1,L-6}$, etc.

Further, for each $i \in \{1, \dots, L-6\}$, define H_i as the history of the t-spoke $\mathcal{Q}_i$ and define $h_i = \|H_i\|$.

Finally, let W be the accepted configuration of $\mathbf{M}$ corresponding to Lab$(\partial\Pi)$ and let V be the accepted configuration of $\mathbf{M}_4$ such that $W(i)$ is a copy of V for each $i \geq 2$.

Note that Lemma 13.2 can function as an analogue of Lemma 11.5 in this setting. The following statement can similarly be viewed as an analogue of Lemma 11.6.

Lemma 13.6.

(1) Every maximal θ-band of Ψ crosses either $\mathcal{Q}_1$ or $\mathcal{Q}_{L-6}$.

(2) There exists an r satisfying $(L-1)/2 - 5 \leq r \leq (L-1)/2$ such that the θ-bands of Ψ crossing $\mathcal{Q}_{L-6}$ do not cross $\mathcal{Q}_r$ and the θ-bands of Ψ crossing $\mathcal{Q}_1$ do not cross $\mathcal{Q}_{r+1}$.

Proof. (1) Suppose there exists a θ-band $\mathcal{T}$ of Ψ not crossing either $\mathcal{Q}_1$ or $\mathcal{Q}_{L-6}$. Then $\mathcal{T}$ must have both ends on $\mathbf{p}$ which is a subpath of $\mathbf{s}$. So, $\mathcal{T}$ is a maximal θ-band of Δ.

As Ψ is diskless, perhaps passing to θ-bands contained in the subdiagram bounded by $\mathcal{T}$ and $\mathbf{p}$, it is no loss of generality to assume that $\mathcal{T}$ is a quasi-rim θ-band. By definition, this means that $\mathcal{T}$ is a g-rim θ-band.

By Lemma 13.2, every q-band crossing $\mathcal{T}$ must be a spoke of Π. But then the base of $\mathcal{T}$ has length at most $11(L-7)+1 < K$ by the parameter choice $K >> L$, contradicting Lemma 13.3.

(2) This is proved in much the same way as Lemma 11.6(2). Note that the difference in the bounds is due to the difference in the number of relevant t-spokes.

□

By Lemma 13.6(2), it follows that, as in Chapter 11,

$$h_1 \geq h_2 \geq \cdots \geq h_{r-1} \geq h_r \ ;$$

$$h_{r+1} \leq h_{r+2} \leq \cdots \leq h_{L-7} \leq h_{L-6}$$

The next statement is the analogue of Lemma 11.8 in this setting. Its proof is an exact copy of the one provided in Chapter 11.

Lemma 13.7.

(1) If $i \leq r$ *and* $j \geq r+1$*, then*

$$\begin{aligned} |\boldsymbol{p}_{ij}| &\geq |\boldsymbol{p}_{ij}|_\theta + |\boldsymbol{p}_{ij}|_q + \delta(|\boldsymbol{p}_{ij}|_a - |\boldsymbol{p}_{ij}|_\theta) \\ &\geq h_i + h_j + 11(j-i) + \delta(|\boldsymbol{p}_{ij}|_a - h_i - h_j) + 1 \end{aligned}$$

(2) $|\bar{\boldsymbol{p}}_{ij}| \leq h_i + h_j + 11(L-j+i) + (L-j+i+1)\delta|V|_a - 1.$

The following statement is the analogue of Lemma 11.10 in this setting.

Lemma 13.8. *For any* $1 \leq i < j \leq L-6$, $|\boldsymbol{p}_{ij}| + \sigma_\lambda(\bar{\Delta}_{ij}) \leq |\boldsymbol{p}_{ij}| + \sigma_\lambda(\Delta) - \sigma_\lambda(\Psi'_{ij}) \leq |\bar{\boldsymbol{p}}_{ij}|$.

Proof. As $\bar{\Delta}_{ij}$ is the complement of Ψ'_{ij} in Δ, $\sigma_\lambda(\Psi'_{ij}) + \sigma_\lambda(\bar{\Delta}_{ij}) \leq \sigma_\lambda(\Delta)$. Hence, it suffices to show that $|\mathbf{p}_{ij}| + \sigma_\lambda(\Delta) \leq |\bar{\mathbf{p}}_{ij}| + \sigma_\lambda(\Psi'_{ij})$.

Let $\mathbf{p}'_{ij}$ be the complement of $\mathbf{p}_{ij}$ in $\partial\Delta$. As $\mathbf{p}_{ij}$ starts and ends with t-edges, $|\partial\Delta| = |\mathbf{p}_{ij}| + |\mathbf{p}'_{ij}|$. Further, since $\partial\Psi'_{ij} = \bar{\mathbf{p}}_{ij}\mathbf{p}'_{ij}$, Lemma 8.1(c) implies $|\partial\Psi'_{ij}| \leq |\bar{\mathbf{p}}_{ij}| + |\mathbf{p}'_{ij}|$.

Since $\partial\Psi'_{ij}$ contains **t** as a subpath, Ψ'_{ij} is a g-diagram. But then Δ is g-minimal, so that $|\partial\Delta| + \sigma_\lambda(\Delta) \leq |\partial\Psi'_{ij}| + \sigma_\lambda(\Psi'_{ij})$. Combining these inequalities then yields

$$|\mathbf{p}_{ij}| + |\mathbf{p}'_{ij}| + \sigma_\lambda(\Delta) \leq |\bar{\mathbf{p}}_{ij}| + |\mathbf{p}'_{ij}| + \sigma_\lambda(\Psi'_{ij})$$

yielding the desired inequality.

□

Lemma 13.9. *The inequality $|V|_a > \frac{11L}{16\delta}$ must be true.*

Proof. As $\sigma_\lambda(\bar{\Delta}) \geq 0$, Lemma 13.8 implies that $|\mathbf{p}| \leq |\bar{\mathbf{p}}|$.

So, by Lemma 13.7,

$$h_1 + h_{L-6} + 11(L-7) \leq h_1 + h_{L-6} + 11(7) + 8\delta|V|_a$$

As a result, $|V|_a \geq \frac{11(L-14)}{8\delta} > \frac{11L}{16\delta}$ by a parameter choice for L.

□

Note that the previous statement serves as an analogue (though a strengthened version) of Lemma 11.18.

Lemma 13.10. *Let $\mathcal{T}$ (resp $\mathcal{T}'$) be the maximal θ-band of Ψ crossing $\mathcal{Q}_1$ (resp $\mathcal{Q}_{L-6}$) closest to Π. Let ℓ and ℓ' be the maximal integers for which $\mathcal{T}$ crosses $\mathcal{Q}_1, \dots, \mathcal{Q}_\ell$ and $\mathcal{T}'$ crosses $\mathcal{Q}_{L-5-\ell'}, \dots, \mathcal{Q}_{L-6}$. Then:*

(1) $\ell + \ell' > L - L_0$

(2) H_1 and H_{L-6} have different first letters.

Proof. (1) By Lemma 13.3, for $\ell \leq i \leq L-6-\ell'$, $\Psi_{i,i+1}$ contains no cells. In particular, $\mathbf{p}_{i,i+1}$ is shared with $\partial\Pi$, and so contains $|V|_a$ a-edges and no θ-edges. Applying Lemma 8.1, this implies

$$|\mathbf{p}| \geq h_1 + h_{L-6} + 11(L-7) + (L-6-\ell-\ell')\delta|V|_a$$

By Lemma 13.7(2), $|\bar{\mathbf{p}}| \leq h_1 + h_{L-6} + 11(7) + 8\delta|V|_a$.

By Lemma 13.8, these inequalities yield

$$11(L-14) + (L-14)\delta|V|_a \leq (\ell+\ell')\delta|V|_a$$

As $L > 0$, this necessitates $|V|_a > 0$. But then $\ell + \ell' > L - 14$, so that a parameter choice for L_0 implies the statement.

(2) The parameter choice $L >> L_0$ allows one to assume that $L - L_0 > (L-1)/2$. But then if the statement were false, then (1) would contradict Lemma 10.7(2).

□

Note that Lemma 13.10(2) is the natural analogue of Lemma 11.17 in this setting.

Since Lemma 10.7(1) implies that $\ell, \ell' \leq (L-1)/2$, it follows from Lemma 13.10(1) and the parameter choice $L >> L_0$ that $\ell, \ell' > L_0$.

The distinguished clove $\Psi_{i,i+1}$ is defined here in the analogous way as it is in Chapter 11, as are its subdiagrams $\Lambda'_{i,i+1}$ and $\Lambda''_{i,i+1}$.

Further, for $i \in [1, r-1] \cup [r+1, L-7]$, define the trapezium Γ_i, the comb E_i, and the paths $\mathbf{y}_i$ and $\mathbf{z}_i$ in the same way they were defined in Section 11.5.

Lemma 13.11. *For each i, the comb E_i contains no a-cell.*

Proof. Suppose $\Psi_{i,i+1}$ is not the distinguished clove.

By Lemma 13.2, every q-band of E_i corresponds to a spoke of Π contained in $\Psi_{i,i+1}$. So, by the structure of the relations, every edge of any a-cell of E_i must be shared with $\mathbf{p}_{i,i+1}$. But then cutting this a-cell off of Δ produces a g-diagram with smaller inductive parameter, contradicting the assumption that Δ is g-minimal.

If $\Psi_{i,i+1}$ is the distinguished clove, then the same argument applies, as the spokes of Π contained in $\Lambda'_{i,i+1}$ cannot correspond to the 'special' input sector.

□

The following two statements are the analogues of Lemmas 11.14 and 11.15 and are proved in exactly the same way.

Lemma 13.12. *For $i \in \{2, \dots, r-1\}$, suppose a maximal a-band $\mathcal{B}$ of E_i starts on $\mathbf{z}_i$ and ends on a side of a maximal q-band $\mathcal{C}$. Let ∇ be the comb bounded by $\mathcal{B}$, a part of $\mathcal{C}$, and a subpath $\boldsymbol{x}$*

of $\mathbf{z}_i$. Then there is a copy of the comb ∇ in the trapezium $\Gamma = \Gamma_{i-1} \setminus \Gamma_i'$, where Γ_i' is the natural copy of Γ_i in Γ_{i-1}.

Lemma 13.13. *At most 6 a-bands starting on the path $\mathbf{y}_i$ (or $\mathbf{z}_i$) can end on (θ, q)-cells of the same θ-band.*

Lemma 13.14. *There exist no two indices $i \in \{1, \dots, r-1\}$ and $j \in \{r+1, \dots, L-7\}$ such that $|\mathbf{z}_i|_a, |\mathbf{z}_j|_a < |V|_a/8c_3$.*

Proof. Suppose neither $\Psi_{i,i+1}$ nor $\Psi_{j,j+1}$ is the distinguished clove.

As in the proof of Lemma 11.19, we may use Lemma 13.10(2) to construct a trapezium E with history $H_j^{-1}H_{i+1}$ by pasting the mirror of a coordinate shift of Γ_j to Γ_i. Without loss of generality, suppose $h_{i+1} \geq h_j$.

Note that $|V|_a - |V|_a/8c_3 > |V|_a/2$, so that $h_{i+1}, h_j > |V|_a/8$ since any rule of $\mathbf{M}_4$ alters the a-length of a configuration by at most four. By Lemma 13.9, $|V|_a/8 > \frac{11L}{128\delta} \geq 12c_3$ since $\delta^{-1} >> L >> c_3$. Further, for t the height of E, $W_0 \equiv \mathbf{tbot}(E)$, and $W_t \equiv \mathbf{ttop}(E)$, $|V|_a/8 > c_3 \max(|W_0|_a, |W_t|_a)$. So, $t > c_3 \max(\|W_0\|, \|W_t\|)$.

Taking $\lambda < 1/10$, Lemmas 7.4 and 6.20 imply that $\mathcal{Q}_i$ contains a λ-shaft of Π of length at least h_{i+1}. So, Lemma 13.7(1) yields the inequality $|\mathbf{p}_{i,j}| + \sigma_\lambda(\bar{\Delta}_{i,j}) \geq h_i + h_j + h_{i+1}$.

By Lemma 13.9 and a parameter choice for δ, $11L < 16\delta|V|_a < 128\delta h_{i+1} \leq \frac{1}{4}h_{i+1}$. Similarly, $(L+1)\delta|V|_a \leq 8\delta h_{i+1} < \frac{1}{4}h_{i+1}$. Hence, $|\bar{\mathbf{p}}_{i,j}| < h_i + h_j + \frac{1}{2}h_{i+1}$.

But then $|\bar{\mathbf{p}}_{i,j}| < |\mathbf{p}_{i,j}| + \sigma_\lambda(\bar{\Delta}_{i,j})$, contradicting Lemma 13.8.

Now suppose either $\Psi_{i,i+1}$ or $\Psi_{j,j+1}$ is the distinguished clove. Then letting h_{i+1}' and h_j' be the heights of Γ_i and Γ_j, respectively, we may construct the trapezium E as above. Assuming $h_{i+1}' \geq h_j'$, then the same arguments as above imply that $\mathcal{Q}_i$ contains a λ-shaft of Π of length at least $h_{i+1}' \geq h_{i+1}$. But then a contradiction is reached as above.

□

As a result of Lemma 13.14, we may assume without loss of generality that for all $i = 1, \dots, r-1$, $|\mathbf{z}_i|_a \geq |V|_a/8c_3$.

Lemma 13.15. *Let I be the subset of the indices $\{1, \dots, r-1\}$ defined by the property that for any $i \in I$, $h_i - h_{i+1} \leq |\mathbf{z}_i|_a/8$. Then $\#I \leq c_4$.*

Proof. Note that the length of the handle E_i is at most $h_i - h_{i+1}$. So, by Lemma 13.13, at most $6(h_i - h_{i+1})$ a-bands beginning on $\mathbf{z}_i$ can end on a (θ, q)-cell of E_i.

Since $\mathbf{p}_{i,i+1}$ consists of exactly $h_i - h_{i+1}$ θ-edges, for any $i \in I$, at least $|\mathbf{z}_i|_a - 7(h_i - h_{i+1}) \geq |\mathbf{z}_i|_a/8$ a-edges contribute δ to $|\mathbf{p}_{i,i+1}|$, and so to $|\mathbf{p}|$.

As $|\mathbf{z}_i|_a \geq |V|_a/8c_3$ for all $i \in I$, it follows that at least $|V|_a/64c_3$ a-edges contribute δ to $|\mathbf{p}|$.

Suppose $\#I > c_4$. Then, by Lemma 13.7 and the parameter choice $c_4 >> c_3$,

$$|\mathbf{p}| \geq h_1 + h_{L-6} + 11(L-7) + 8\delta|V|_a$$

$$|\bar{\mathbf{p}}| \leq h_1 + h_{L-6} + 11(7) + 8\delta|V|_a$$

But then $|\bar{\mathbf{p}}| < |\mathbf{p}|$, contradicting Lemma 13.8.

□

Lemma 13.16. *For $i = 1, \dots, L_0$, $h_i > L_0^2|V|_a$.*

Proof. As $h_i \geq h_{i+1} \geq \dots \geq h_{r-1} \geq h_r$, we have $h_i \geq h_i - h_r = \sum_{j=i}^{r-1}(h_j - h_{j+1})$.

By Lemma 13.15, the number of $j \in \{i, \dots, r-1\}$ such that $j \notin I$ is at least $r - 1 + i - c_4$. Since $r \geq (L-1)/2 - 5$ by Lemma 13.6, the parameter choices $L >> L_0 >> c_4$ imply that $r - 1 + i - c_4 > L_0^3$.

As a result, $h_i > L_0^3|\mathbf{z}_i|_a/8c_3 > L_0^2|V|_a$ by the parameter choice $L_0 >> c_3$.

□

Lemma 13.17. *For $i = 1, \dots, L_0$, $\mathcal{Q}_i$ contains no λ-shaft of length at least h_{L_0}.*

Proof. Assuming to the contrary that $\mathcal{Q}_i$ contains such a λ-shaft, Lemma 13.16 implies

$$\sigma_\lambda(\Delta) - \sigma_\lambda(\Psi'_{i,L-6}) \geq h_{L_0} \geq L_0|V|_a$$

Then, Lemma 13.7 and the parameter choices $L >> L_0$ yield the inequalities

$$
\begin{aligned}
|\mathbf{p}_{i,L-6}| + \sigma_\lambda(\Delta) - \sigma_\lambda(\Psi'_{i,L-6}) &> h_i + h_{L-6} + 11(L-i-6) + L_0|V|_a \\
&> h_i + h_{L-6} + 11L/2 + L_0|V|_a
\end{aligned}
$$

$$
\begin{aligned}
|\bar{\mathbf{p}}_{i,L-6}| &\leq h_i + h_{L-6} + 11(i+6) + (i+7)\delta|V|_a \\
&< h_i + h_{L-6} + 11(2L_0) + 2L_0\delta|V|_a
\end{aligned}
$$

But then a parameter choice for δ implies that $|\mathbf{p}_{i+1,L-6}| + \sigma_\lambda(\Delta) - \sigma_\lambda(\Psi'_{i+1,L-6}) > |\bar{\mathbf{p}}_{i+1,L-6}|$, contradicting Lemma 13.8.

□

Lemma 13.18. *For all* $i = 1, \dots, L_0 - 1$, $h_{i+1} \leq 2c_3|\mathbf{z}_i|_a$.

Proof. Assume toward contradiction that $h_{i+1} > 2c_3|\mathbf{z}_i|_a$.

Then, for the computation $\mathcal{C} : W_0 \to \cdots \to W_t$ corresponding to the trapezium Γ_i through Lemma 7.4, $t > 2c_3|W_t|_a$. Further, Lemma 13.16 implies $t > L_0|V|_a$. The parameter choice $L_0 >> c_3$ then implies $t > 2c_3 \max(|W_0|_a, |W_t|_a)$.

Lemma 13.9 then implies that $t \geq |V|_a \geq 11L/16\delta$, so that the parameter choices $\delta^{-1} >> L >> c_3$ implies $t > c_3 \max(\|W_0\|, \|W_t\|)$.

As in the proof of Lemma 13.14, this implies that $\mathcal{Q}_i$ contains a λ-shaft of length at least h_{i+1}. But this contradicts Lemma 13.17.

□

Lemma 13.19. *For* $i = 1, \dots, L_0 - 1$, $h_{i+1} < \left(1 - \frac{1}{30c_3}\right) h_i$.

Proof. Assuming the statement is false, $h_i - h_{i+1} \leq h_i/30c_3$. By Lemma 13.13, at most $h_i/5c_3$ maximal a-bands of E_i starting on $\mathbf{z}_i$ can end on (θ, q)-cells. So, at least $\max(0, |\mathbf{z}_i|_a - h_i/5c_3)$ of these bands end on $\mathbf{p}_{i,i+1}$.

Lemma 13.18 implies $|\mathbf{z}_i|_a \geq h_{i+1}/2c_3$, so that $|\mathbf{z}_i|_a - h_i/5c_3 \geq h_i/15c_3$.

By Lemma 13.6, $\mathbf{p}_{i,i+1}$ has $h_i - h_{i+1} \leq h_i/30c_3$ θ-edges. So, Lemma 13.7(1) implies

$$|\mathbf{p}_{i,i+1}| \geq h_i - h_{i+1} + 12 + \delta h_i/30c_3$$

$$|\mathbf{p}_{i+1,L-6}| \geq h_{i+1} + h_{L-6} + 11(L-i-7)$$

As in the proof of Lemma 11.23, this implies

$$\begin{aligned}|\mathbf{p}_{i,L-6}| &= |\mathbf{p}_{i,i+1}| + |\mathbf{p}_{i+1,L-6}| - 1 \\ &\geq h_i + h_{L-6} + 11(L-i-6) + \delta h_i/30c_3\end{aligned}$$

Meanwhile, Lemma 13.7 implies

$$|\bar{\mathbf{p}}_{i,L-6}| \leq h_i + h_{L-6} + 11(i+6) + (i+7)\delta|V|_a$$

Lemma 13.16 yields $|V|_a < h_i/L_0^2$, so that

$$|\bar{\mathbf{p}}_{i,L-6}| \leq h_i + h_{L-6} + 11(2L_0) + 2\delta h_i/L_0$$

But then the parameter choices $L >> L_0 >> c_3$ imply $|\bar{\mathbf{p}}_{i,L-L_0-6}| < |\mathbf{p}_{i,L-L_0-6}|$, contradicting Lemma 13.8.

□

Recall that if $\Psi_{i,i+1}$ is the distinguished clove for $i \leq r-1$, then the history of Γ_i need not be H_{i+1}, but rather could be a proper prefix. As in Chapter 11, set H'_{i+1} as the history of Γ_i.

Lemma 13.20. *For $2 \leq i \leq L_0 - 2$, let $H'_i = H'_{i+1}H' = H'_{i+2}H''H'$ and $\mathcal{C}$ be the computation corresponding to the trapezium Γ_{i-1}. Suppose the subcomputation $\mathcal{D}$ of $\mathcal{C}$ with history $H''H'$ has step history of length 1. Then there is no two-letter subword $Q'Q$ of the base of Γ_{i-1} such that* *every rule of $\mathcal{D}$ inserts one letter to the left of Q.*

Proof. Let $\mathcal{Q}$ be the maximal q-band of E_i that is a subband of the q-spoke of Π corresponding to a coordinate shift of the state letter Q. Similarly, let $\mathcal{Q}'$ be the q-band corresponding to a coordinate shift of Q', so that $\mathcal{Q}'$ and $\mathcal{Q}$ are neighbor q-bands. Let $\mathbf{x}$ be the subpath of $\mathbf{z}_i$ between $\mathcal{Q}'$ and $\mathcal{Q}$.

As in the proof of Lemma 11.25, $|\mathbf{x}|_a \geq \|H''\| \geq h_{i+1} - h_{i+2}$.

Further, Lemma 13.19 implies that $h_{i+1} - h_{i+2} > \frac{1}{30c_3}h_{i+1}$, so that Lemma 13.16 and the parameter choice $L_0 >> c_3$ yield $|\mathbf{x}|_a \geq 10L_0|V|_a$.

Consider the comb ∇ contained in E_i bounded by $\mathcal{Q}'$, $\mathcal{Q}$, $\mathbf{x}$, and $\mathbf{p}_{i,i+1}$. By Lemma 13.11, at least $|\mathbf{x}|_a$ a-edges contribute δ to $\mathbf{p}_{i,i+1}$.

Lemma 13.7 then gives the inequalities:

$$|\mathbf{p}_{i,L-6}| \geq h_i + h_{L-6} + 11L/2 + 10\delta L_0|V|_a$$

$$|\bar{\mathbf{p}}_{i,L-6}| \leq h_i + h_{L-6} + 11(2L_0) + 2\delta L_0|V|_a$$

But then the parameter choice $L >> L_0$ implies $|\bar{\mathbf{p}}_{i,L-6}| < |\mathbf{p}_{i,L-6}|$, contradicting Lemma 13.8.

□

Finally, the following statement yields the desired contradiction.

Lemma 13.21. *For any $g \in B(\mathcal{A}, n)$, a g-minimal diagram contains no disks.*

Proof. The proof follows the same outline as that of Lemma 11.26.

For $\eta \geq 2$ an integer such that $\left(1 - \frac{1}{30c_3}\right)^\eta < \frac{1}{64c_3}$, Lemma 13.19 implies that if $1 \leq i < j \leq L_0 - 1$ with $j - i - 1 \geq \eta$, then $h_j < \frac{1}{64c_3}h_{i+1}$.

For such i, j, Lemma 13.18 implies that $|\mathbf{z}_i|_a > 32h_j$, while Lemma 13.15 implies $8h_j \geq |\mathbf{z}_j|_a$. As a result, $|\mathbf{z}_i|_a > 4|\mathbf{z}_j|_a$.

Taking $L_0 >> \eta$ and $L_0 >> c_0$, there exist indices $2 \leq j_1 < j_2 < \cdots < j_m \leq L_0 - 1$ such that $m \geq c_0$ and $j_{i+1} - j_i - 1 \geq \eta$. So, $|\mathbf{z}_{j_i}|_a > 4|\mathbf{z}_{j_{i+1}}|_a$ and $h_{j_i+1} \geq 64c_3h_{j_{i+1}}$.

Let $\mathcal{C} : W_0 \to \cdots \to W_t$ be the computation corresponding to the trapezium Γ_{j_2}. As Γ_{j_2} contains a copy of Γ_{j_2+1}, which in turn contains a copy of Γ_{j_2+2} and so on, there exist words V_i in $\mathcal{C}$ for $i = 1, \dots, m$ that are coordinate shifts of $\text{Lab}(\mathbf{z}_{j_i})$. Note that $|V_{i+1}|_a > 4|V_i|_a$.

As in the proof of Lemma 11.26, the subcomputations $V_{i+2} \to \cdots \to V_i$ cannot be one-step, as an application of Lemma 6.21 would lead to a contradiction of Lemma 13.20.

But then this implies that $\mathcal{Q}_{j_1}$ contains a λ-shaft of length at least h_{L_0}, contradicting Lemma 13.17.

□

13.3 Upper bound

As referenced in the Introduction, to prove Theorem B, it suffices to find a constant $M > 0$ such that for any $g \in B(\mathcal{A}, n)$, $|g|_{\mathcal{A}} \leq M|g|_{\mathcal{X}}$.

For $g \in B(\mathcal{A}, n)$, let Δ be an g-minimal diagram.

Fix $w \in F(\mathcal{X})$ so that the value of w in $G_{\mathcal{S}}(\mathbf{M})$ is g and $|w|$ is minimal for all such words. Further, let Γ be a minimal diagram over $G_{\mathcal{S}}(\mathbf{M})$ such that $\text{Lab}(\partial\Gamma) \equiv wv^{-1}$ for $v \in F(\mathcal{A})$ such that the value of v in $B(\mathcal{A}, n)$ is g and $\|v\| = |g|_{\mathcal{A}}$.

Then, Γ is a g-diagram, so that $|\partial\Gamma| + \sigma_\lambda(\Gamma) \geq |\partial\Delta| + \sigma_\lambda(\Delta)$.

By Lemmas 13.21 and 13.4, $|\partial\Delta| + \sigma_\lambda(\Delta) = |\partial\Delta| = 2\delta|g|_{\mathcal{A}}$. Further, Lemma 8.1(c) implies $|\partial\Gamma| \leq |w| + \delta|g|_{\mathcal{A}}$. Hence, $|w| + \sigma_\lambda(\Gamma) \geq \delta|g|_{\mathcal{A}}$.

Further, by Lemma 10.15, $\sigma_\lambda(\Gamma) \leq C_1|\partial\Gamma|_\theta$. Noting that v consists entirely of a-letters, it follows that $|\partial\Gamma|_\theta \leq |w|_\theta \leq |w|$.

As a result, $|w| \geq \delta(C_1 + 1)^{-1}|g|_{\mathcal{A}}$.

But for any word u in the alphabet $\mathcal{X} \cup \mathcal{X}^{-1}$, $|u|_{\mathcal{X}} = \|u\| \geq \delta^{-1}|u|$. So, since $|w|$ is minimal for all words over $\mathcal{X}$ whose value in $G_{\mathcal{S}}(\mathbf{M})$ is g, $|g|_{\mathcal{X}} \geq (C_1 + 1)^{-1}|g|_{\mathcal{A}}$.

Thus, taking $M = C_1 + 1$ completes the proof of Theorem B.

Chapter 14

Proof of Theorem C

Let T be some subset of $B(\mathcal{A},n)$. Clearly, $\langle\langle T\rangle\rangle^{B(\mathcal{A},n)} \subseteq B(\mathcal{A},n)\cap\langle\langle T\rangle\rangle^{G_{\mathcal{S}}(\mathbf{M})}$, so that it suffices just to show the opposite inclusion.

Let Σ_1 be a set of words over $\mathcal{A}$ so that for each $g\in T$, there exists a word $w\in\Sigma_1$ such that the value of w in $B(\mathcal{A},n)$ is g. Then, set $\Sigma=\Sigma_1\cup\mathcal{S}$.

Then, $G_\Sigma(\mathbf{M})=G_{\mathcal{S}}(\mathbf{M})/\langle\langle\Sigma_1\rangle\rangle^{G_{\mathcal{S}}(\mathbf{M})}\cong G_{\mathcal{S}}(\mathbf{M})/\langle\langle T\rangle\rangle^{G_{\mathcal{S}}(\mathbf{M})}$.

Fix $g\in B(\mathcal{A},n)\cap\langle\langle T\rangle\rangle^{G_{\mathcal{S}}(\mathbf{M})}$ and let w be a word over $\mathcal{A}$ whose value in $G_{\mathcal{S}}(\mathbf{M})$ is g. Then w represents the trivial element of $G_\Sigma(\mathbf{M})$, so that Lemma 10.9 implies that there exists a minimal diagram Δ over $G_\Sigma(\mathbf{M})$ such that Lab$(\partial\Delta)\equiv w$.

If Δ were to contain a disk, then Lemma 10.6 implies that at least $L-4$ t-spokes end on $\partial\Delta$. But $\partial\Delta$ contains no t-edges, so that this is impossible. Hence, Δ must be an M-minimal diagram.

By Lemmas 9.1 and 9.3, each maximal q-band and each maximal θ-band of Δ end twice on $\partial\Delta$. But again, $\partial\Delta$ contains no q-edge or θ-edge, so that Δ can contain no q-band or θ-band.

As a result, each cell of Δ must be an a-cell, i.e Δ is a reduced diagram over the group with presentation $\langle\mathcal{A}\mid\Sigma\rangle\cong B(\mathcal{A},n)/\langle\langle T\rangle\rangle^{B(\mathcal{A},n)}$. So, the value of w in $B(\mathcal{A},n)$ is an element of $\langle\langle T\rangle\rangle^{B(\mathcal{A},n)}$. But by Lemma 12.3, the value of w in $B(\mathcal{A},n)$ is g, so that $g\in\langle\langle T\rangle\rangle^{B(\mathcal{A},n)}$.

Thus, $B(\mathcal{A},n)\cap\langle\langle T\rangle\rangle^{G_{\mathcal{S}}(\mathbf{M})}\subseteq\langle\langle T\rangle\rangle^{B(\mathcal{A},n)}$, and so $B(\mathcal{A},n)\leq_{CEP} G_{\mathcal{S}}(\mathbf{M})$.

BIBLIOGRAPHY

[1] S. I. Adian, *The Burnside Problem and Identities in Groups*, Springer-Verlag, (1979).

[2] S. I. Adian, *New estimates of odd exponents of infinite Burnside groups*, Selected issues of mathematics and mechanics, Collected papers. In commemoration of the 150th anniversary of Academician Vladimir Andreevich Steklov, Trudy MIAN, 289, MAIK Nauka/Interperiodica, Moscow, 2015, 41?82; Proc. Steklov Inst. Math., 289 (2015), 33–71

[3] J.-C. Birget, A. Yu. Ol'shanskii, E. Rips, M. Sapir, *Isoperimetric functions of groups and combinatorial complexity of the word problem*, Annals of Mathematics , 156 (2002), no. 2, 467–518.

[4] B. H. Bowditch, *Notes on Gromov's hyperbolicity criterion for path-metric spaces*, "Group theory from a geometrical viewpoint (Trieste, 1990)", (E Ghys, A Haefliger, A Verjovsky, editors), World Sci. Publ., River Edge, NJ (1991)

[5] M. Bridson, A. Haefliger, *Metric Spaces of Non-Positive Curvature*, Grundlehren der mathematischen Wissenschaften, Volume 319, Springer (1999).

[6] E. Ghys, P. de la Harpe, *Sur les Groupes Hyperboliques d'après Mikhael Gromov*, Springer, (1990).

[7] E. S. Golod, I. R. Shafarevich, *On the class field tower*, Izv. Akad. Nauk SSSR Ser. Mat., 28:2 (1964), 261-272

[8] M. Gromov, *Hyperbolic groups*, Essays in Group Theory (S.M.Gersten, ed.), M.S.R.I. Pub. 8, Springer, (1987), 75-263.

[9] S. V. Ivanov, *The free Burnside groups of sufficiently large exponents*, International Journal of Algebra and Computation, 4(1-2), (1994) ii+308.

[10] S. V. Ivanov, A. Yu. Ol'shanskii, *On finite and locally finite subgroups of free burnside groups of large even exponents*. Journal of Algebra, 195(1), (1997) 241-284.

[11] R. C. Lyndon and P. E. Schupp, *Combinatorial group theory*, Springer-Verlag, 1977.

[12] I. G. Lysenok, "Infinite Burnside groups of even exponent", Izv. RAN. Ser. Mat., 60:3 (1996), 3–224; Izv. Math., 60:3 (1996), 453–654

[13] P. S. Novikov, S. I. Adian, *Defining relations and the word problem for free periodic groups of odd order*, Izv. Akad. Nauk SSSR Ser. Mat., 32:4 (1968)

[14] A. Yu. Ol'shanskii, *Groups of bounded period with subgroups of prime order*, Algebra and Logic 21 (1983), 369–418; translation of Algebra i Logika 21 (1982)

[15] A. Yu. Ol'shanskii, *Hyperbolicity of groups with subquadratic isoperimetric inequality*, Internat. J. Algebra Comput. 1 (1991), no. 3, 281–289.

[16] A. Yu. Ol'shanskii, *Geometry of Defining Relations in Groups*, Springer Netherlands, (1991)

[17] A. Yu. Ol'shanskii, *On subgroup distortion in finitely presented groups* Mat. Sb., 188:11 (1997), 51–98; Sb. Math., 188:11 (1997), 1617–1664

[18] A. Yu. Ol'shanskii, *Polynomially-bounded Dehn functions of groups*, Journal of Combinatorial Algebra, 2. (2018) 311–433

[19] A. Yu. Ol'shanskii, M. V. Sapir, *Embeddings of relatively free groups into finitely presented groups*, (2000)

[20] A. Yu. Ol'shanskii, M. V. Sapir, *Length and area functions in groups and quasiisometric Higman embeddings*, Intern. J. Algebra and Comput., 11 (2001), no. 2, 137–170.

[21] A. Yu. Ol'shanskii, M. V. Sapir, *The Conjugacy Problem and Higman Embeddings*. Memoirs of the American Mathematical Society. 170. (2003).

[22] A. Yu. Ol'shanskii, M. V. Sapir, *Non-Amenable Finitely Presented Torsion-by-Cyclic Groups*, Publ. math., Inst. Hautes Étud. Sci. (2003)

[23] A. Yu. Ol'shanskii, M. V. Sapir, *Groups with Small Dehn functions and Bipartite Chord Diagrams* GAFA, Geom. funct. anal. 16 (2006), 1324

[24] A. Yu. Ol'shanskii, M. V. Sapir, *Groups with undecidable word problem and almost quadratic Dehn function*, Journal of Topology. 5. (2012) 785-886. 10.1112/jtopol/jts020

[25] A. Yu. Ol'shanskii, M. V. Sapir, *Conjugacy problem in groups with quadratic Dehn function*, (2018).

[26] M. V. Sapir, J. C. Birget, E. Rips, *Isoperimetric and Isodiametric Functions of Groups*, Annals of Mathematics, 156(2), second series, (1998), 345-466

[27] M. V. Sapir. *Algorithmic and asymptotic properties of groups*, International Congress of Mathematicians, ICM (2006).

[28] 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).

[29] V. L. Shirvanyan, *Embedding the group* $B(\infty, n)$ *in the group* $B(2, n)$, Izv. Akad. Nauk SSR Ser. Mat. 40 (1976), 190–208.

[30] D. Sonkin, *CEP-Subgroups of Free Burnside Groups of Large Odd Exponents*, Communications in Algebra Vol. 31. No. 10. (2003) 4687-4695. 10.1081/AGB-120023127.

[31] E. van Kampen, *On Some Lemmas in the Theory of Groups*, American Journal of Mathematics Vol. 55, No. 1 (1933), pp. 268-273.

[32] F. Wagner, *Torsion Subgroups of Groups with Quadratic Dehn Function*, https://arxiv.org/pdf/2010.05381.pdf