Word Measures on Unitary Groups

Michael Magee* and Doron Puder†

July 13, 2016

Abstract

We combine concepts from random matrix theory and free probability together with ideas from the theory of commutator length in groups and maps from surfaces, and establish new connections between the two.

More particularly, we study measures induced by free words on the unitary groups $\mathcal{U}(n)$. Every word w in the free group $\mathbf{F}_r$ on r generators determines a word map from $\mathcal{U}(n)^r$ to $\mathcal{U}(n)$, defined by substitutions. The w-measure on $\mathcal{U}(n)$ is defined as the pushforward via this word map of the Haar measure on $\mathcal{U}(n)^r$.

Let $\mathcal{T}r_w(n)$ denote the expected trace of a random unitary matrix sampled from $\mathcal{U}(n)$ according to the w-measure. It was shown by Voiculescu [Voi91] that for $w \neq 1$ this expected trace is $o(n)$ asymptotically in n. We relate the numbers $\mathcal{T}r_w(n)$ to the theory of commutator length of words and obtain a much stronger statement: $\mathcal{T}r_w(n) = O\left(n^{1-2g}\right)$, where g is the commutator length of w. Moreover, we analyze the number $\lim_{n\to\infty} n^{2g-1} \cdot \mathcal{T}r_w(n)$ and show it is an integer which, roughly, counts the number of (equivalence classes of) solutions to the equation $[u_1, v_1] \ldots [u_g, v_g] = w$ with $u_i, v_i \in \mathbf{F}_r$.

Similar results are obtained for finite sets of words and their commutator length, and we deduce that one can "hear" the stable commutator length of a word by "listening" to its unitary measures.

Contents

*Author Magee was partially supported by the National Science Foundation under agreement No. DMS-1128155.

†Author Puder was supported by the Rothschild fellowship and by the National Science Foundation under agreement No. CCF-1412958.

1 Introduction

1.1 The expected trace

Let $x_1, \ldots, x_r$ denote generators of the free group $\mathbf{F}_r$ on r generators. Consider a word $w \in \mathbf{F}_r$, given by $x_1, \ldots, x_r$

$$w = \prod_{1 \le j \le |w|} x_{i_j}^{\varepsilon_j}, \tag{1.1}$$

where each $\varepsilon_j \in \{\pm 1\}$ and[1] $i_j \in [r]$. Let $(\mathcal{U}(n), \mu_n)$ be the probability space of $n \times n$ unitary matrices, equipped with unit-normalized Haar measure. We consider a tuple $\{U_i^{(n)}\}_{i \in [r]}$ of r independent random matrices sampled from $(\mathcal{U}(n), \mu_n)$. For each n we can form the **word map**[2] $(\mathcal{U}(n), \mu_n)$

$$w : \mathcal{U}(n)^r \to \mathcal{U}(n), \quad w(u_1, \ldots, u_r) \equiv \prod_{1 \le j \le |w|} u_{i_j}^{\varepsilon_j} \tag{1.2}$$

where we abuse notation to identify w with the corresponding map and suppress the dependence on n. We call the pushforward by w of the Haar measure μ_n^r on $\mathcal{U}(n)^r$ the w**-measure** on $\mathcal{U}(n)$. In this paper we study word measures on $\mathcal{U}(n)$ and relate them to algebraic properties of the word w.

Word measures on unitary groups were studied mostly in the context of free probability. Let tr denote the standard trace on complex $n \times n$ matrices, and denote by $\mathcal{T}r_w(n)$ the expected value of the trace of a random unitary matrix in $\mathcal{U}(n)$ under the w-measure. It is a fundamental result of Voiculescu [Voi91, Theorem 3.8] that for $w \in \mathbf{F}_r$, tr $\mathcal{T}r_w(n)$

$$\mathcal{T}r_w(n) \stackrel{\text{def}}{=} \mathbb{E}\left[\operatorname{tr}\left(w\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right)\right] = \begin{cases} n & \text{if } w = 1 \\ o(n) & \text{else} \end{cases} \tag{1.3}$$

[1] We use the standard notation $[r]$ for $\{1, \ldots, r\}$.

[2] Unless we stick to reduced forms, every word $w \in \mathbf{F}_r$ has different expressions as products of the generators $x_1, \ldots, x_r$ and their inverses. However, the word map $w : \mathcal{U}(n)^r \to \mathcal{U}(n)$ is well-defined independently of the particular expression. Namely, omitting from the expression for w or adding to it subwords of the form $x_i x_i^{-1}$ or $x_i^{-1} x_i$ does not effect the resulting word map.

(the small o notation is in the regime $n \to \infty$). It follows that the random variables $U_1^{(n)}, (U_1^{(n)})^*, \ldots, U_r^{(n)}, (U_r^{(n)})^*$ are *asymptotically free*[3], referring to the fact that in the limit, as $n \to \infty$, the family $\{U_i^{(n)}, (U_i^{(n)})^*\}_{i\in[r]}$ can be modeled by the "Free Probability Theory" developed by Voiculescu (see, for example, [Voi85] and the monograph [VDN92]). Such asymptotic freeness results are known for broad families of ensembles[4], including general Gaussian random matrices (due to Voiculescu in the same paper [Voi91, Theorem 2.2]). In later works (1.3) is strengthened to $\mathcal{T}r_w(n) = O\left(\frac{1}{n}\right)$ whenever $w \neq 1$ [MŚS07, Răd06].

Although our results are more general, we first describe them in the special case of the expected trace $\mathcal{T}r_w(n)$, and defer the discussion of the general results to Section 1.2. The starting point for this paper is the intriguing observation that the w-measure on any compact group, and in particular, the w-measure on $\mathcal{U}(n)$ and the quantity $\mathcal{T}r_w(n)$, are invariant under $w \mapsto \theta(w)$ for any $\theta \in \mathrm{Aut}(\mathbf{F}_r)$ (see Section 2.2). It follows that this quantity is determined by some algebraic, $\mathrm{Aut}(\mathbf{F}_r)$-invariant, properties of the word w.

The first step in our analysis of $\mathcal{T}r_w(n)$ builds on results of Xu and of Collins and Śniady [Xu97, CŚ06]. In Section 3 we explain how it follows readily from these results that $\mathcal{T}r_w(n)$ is a rational function of n with coefficients in $\mathbb{Q}$ (which can be algorithmically computed)[5]. For example, this function is $\frac{-4}{n^3-n}$ for $w = [x_1, x_2]^2$ – see (3.5) below. This function can hence be written as a Laurent series in n^{-1} with rational coefficients. By (1.3), whenever $w \neq 1$ we may write $\mathcal{T}r_w(n)$ as a power series:

$$\mathcal{T}r_w(n) \in \mathbb{Q}\left[\frac{1}{n}\right].$$

Unlike previous works, we are not only interested in the limit $\lim_{n\to\infty} \mathcal{T}r_w(n)$. Rather, in this paper our aim is to explain the leading term of $\mathcal{T}r_w(n)$. That is, we give algebraic interpretation for the following two quantities:

Leading exponent The exponent of the leading order term of $\mathcal{T}r_w(n)$

Leading coefficient The coefficient of the leading order term of $\mathcal{T}r_w(n)$

The second of these two quantities is the more subtle[6].

In fact, an easy observation is that unless w is in the commutator subgroup $[\mathbf{F}_r, \mathbf{F}_r]$, the expected trace $\mathcal{T}r_w(n)$ vanishes for every n (Claim 3.1 below). The interesting case is, therefore, when $w \in [\mathbf{F}_r, \mathbf{F}_r]$. Every word in this subgroup is a product of commutators, and the **commutator length** $\mathrm{cl}(w)$ of the word w is the smallest g such that w is a product of g commutators. Namely, the smallest g for which $\mathrm{cl}(w)$

$$w = [u_1, v_1][u_2, v_2] \ldots [u_g, v_g] \tag{1.4}$$

for some $u_i, v_i \in \mathbf{F}_r$. The theory of commutator length suffices to explain the leading exponent of $\mathcal{T}r_w(n)$ (modulo the exceptional event mentioned in Footnote 6):

[3] This is sometimes called asymptotically $*$-freeness of $U_1^{(n)}, \ldots, U_r^{(n)}$. The statement of [Voi91, Theorem 3.8] is actually stronger: it involves additional deterministic matrices.

[4] In the case of unitary matrices, we analyze expressions with negative exponents because $(U_i^{(n)})^{-1} = (U_i^{(n)})^*$. In the general case, one does not allow negative exponents ε_j.

[5] Such a formula, in a slightly more restricted version, appears also in [Răd06].

[6] To be precise, there are degenerate cases where the coefficient we explain vanishes — see Example 4.14 and Section 8. In these cases we lose track of the leading coefficient and only obtain a lower bound for the leading exponent.

Theorem 1.1. *Let* $w \in [\mathbf{F}_r, \mathbf{F}_r]$ *and denote* $g = \mathrm{cl}(w)$. *Then,*

$$\mathcal{T}r_w(n) = O\left(\frac{1}{n^{2g-1}}\right).$$

(The big O *notation is in the regime* $n \to \infty$.*)*

The analysis of the leading coefficient necessitates a subtler study, not only of the commutator length of w, but also of the set of products of commutators of length $\mathrm{cl}(w)$ giving w. To formalize this, consider the following. Let $a_1, b_1, \ldots, a_g, b_g$ be generators of $\mathbf{F}_{2g}$, where $g = \mathrm{cl}(w)$ as above, and let $\delta_g = [a_1, b_1] \ldots [a_g, b_g]$. Solutions to (1.4) correspond to elements $\phi \in \mathrm{Hom}(\mathbf{F}_{2g}, \mathbf{F}_r)$ such that $a_1,b_1,\ldots,a_g,b_g$ δ_g

$$\phi(\delta_g) = w. \tag{1.5}$$

We write $\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$ for the set of homomorphisms $\mathbf{F}_{2g} \to \mathbf{F}_r$ satisfying (1.5). The group $\mathrm{Aut}(\mathbf{F}_{2g})$ acts on $\mathrm{Hom}(\mathbf{F}_{2g}, \mathbf{F}_r)$ by precomposition. We define $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ to be the stabilizer in $\mathrm{Aut}(\mathbf{F}_{2g})$ of δ_g. For example, the automorphism $a_1 \mapsto a_1 b_1$ (leaving all other generators unchanged) is in[7] $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ while $a_1 \longleftrightarrow b_1$ is not. $\mathrm{Hom}_w(\mathbf{F}_{2g},\mathbf{F}_r)$ $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$

Clearly, $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ acts on $\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$, the solution space to (1.4), for every w. We think of the orbits $\mathrm{Aut}_\delta(\mathbf{F}_{2g}) \backslash \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$ as equivalence classes of solutions. So the elements of $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ permute the solutions inside the same equivalence class. For instance, the automorphism $a_1 \mapsto a_1 b_1$ mentioned above shows that the solutions $[x_1, x_2]$ and $[x_1 x_2, x_2]$ belong to the same class. Occasionally, elements of $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ stabilize a solution. For example, consider the word $w = [x_1, x_2]^2$. It can be shown that its commutator length is $g = 2$, and that it has a single class of solutions. The solution $[x_1, x_2]\,[x_1, x_2]$ is stabilized by the automorphism[8]

$$a_1 \mapsto a_1 a_2 a_1 A_2 A_1 \quad b_1 \mapsto a_1 a_2 A_1 A_2 b_1 a_1^2 A_2 A_1 \quad a_2 \mapsto a_1 a_2 A_1 \quad b_2 \mapsto b_2 a_2 A_1,$$

which belongs to $\mathrm{Aut}_\delta(\mathbf{F}_4)$. For every class $[\phi] \in \mathrm{Aut}_\delta(\mathbf{F}_{2g}) \backslash \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$, the stabilizer of any representative ϕ belongs to a well-defined conjugacy class of subgroups of $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$.

As we show below, the leading coefficient of $\mathcal{T}r_w(n)$ is controlled by the set of equivalence classes of solutions to (1.4), and by the isomorphism type of the stabilizer in every class. The important invariant of the stabilizers is their *Euler characteristic*.

The Euler characteristic of a group is defined for a large class of groups of certain finiteness conditions (see [Bro82, Chapter IX]). The simplest case is when a group Γ admits a finite CW-complex as Eilenberg-MacLane space of type[9] $\mathrm{K}(\Gamma, 1)$. In this case, the Euler characteristic $\chi(\Gamma)$ coincides with the topological Euler characteristic of the $\mathrm{K}(\Gamma, 1)$-space, and, in particular, is an integer.

We can now state our main theorem regarding $\mathcal{T}r_w(n)$, which is a more detailed version of Theorem 1.1:

Theorem 1.2. *Let* $w \in [\mathbf{F}_r, \mathbf{F}_r]$ *and denote* $g = \mathrm{cl}(w)$. *Then,*

$$\mathcal{T}r_w(n) = \frac{1}{n^{2g-1}} \left[\sum_{[\phi] \in \mathrm{Aut}_\delta(\mathbf{F}_{2g}) \backslash \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)} \chi\left(\mathrm{Stab}_{\mathrm{Aut}_\delta(\mathbf{F}_{2g})}(\phi)\right) \right] + O\left(\frac{1}{n^{2g+1}}\right).$$

(Again, the big O *notation is in the regime* $n \to \infty$.*)*

[7] Our convention is that $[x, y] = xyx^{-1}y^{-1}$.

[8] We often use the handy convention that capital letters mark inverses. For example, A_1 is a_1^{-1}, the inverse of a_1.

[9] An Eilenberg-MacLane space of type $\mathrm{K}(\Gamma, 1)$, or simply a $\mathrm{K}(\Gamma, 1)$-space, is a path-connected topological space with fundamental group isomorphic to Γ and with a contractible universal cover (e.g. [Bro82, Section I.4]).

Remark 1.3. Note that when $\phi \in \mathrm{Hom}_w(F_{2g}, F_r)$ is injective, $\mathrm{Stab}_{\mathrm{Aut}_\delta(\mathbf{F}_{2g})}(\phi)$ is trivial (indeed, even $\mathrm{Stab}_{\mathrm{Aut}(\mathbf{F}_{2g})}(\phi)$ is trivial), and so its Euler characteristic is 1. This is the case precisely when $\{\phi(a_1), \phi(b_1), \ldots, \phi(a_g), \phi(b_g)\}$ is a free set in $\mathbf{F}_r$, which is in some sense the generic case. Therefore, one could say

"The leading coefficient of $\mathcal{T}r_w(n)$ counts the number of equivalence classes of solutions to (1.4), up to corrections for the existence of non-trivial stabilizers."

For instance, when $g = 1$, namely, when w is a commutator, $\phi(a_1)$ and $\phi(b_1)$ are necessarily free (otherwise they commute and $w = 1$). Hence, if $\mathrm{cl}(w) = 1$ and K marks the number of equivalence classes of solutions to $[u, v] = w$, then $\mathcal{T}r_w(n) = \frac{K}{n} + O\left(\frac{1}{n^3}\right)$. As an example[10], $\mathcal{T}r_{[x_1^k, x_2]}(n) = \frac{k}{n} + O\left(\frac{1}{n^3}\right)$, the different solution classes represented by $\left[x_1^k, x_2 x_1^j\right]$, $0 \le j \le k-1$.

The fact that $\mathrm{Stab}_{\mathrm{Aut}_\delta(\mathbf{F}_{2g})}(\phi)$ has a well-defined Euler characteristic, which is moreover an integer, follows from the following:

Theorem 1.4. *Let $w \in [\mathbf{F}_r, \mathbf{F}_r]$ and denote $g = \mathrm{cl}(w)$. For every $\phi \in \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$, the stabilizer*

$$G \stackrel{\mathrm{def}}{=} \mathrm{Stab}_{\mathrm{Aut}_\delta(\mathbf{F}_{2g})}(\phi) \le \mathrm{Aut}_\delta(\mathbf{F}_{2g})$$

admits a finite simplicial complex as a $\mathrm{K}(G, 1)$-space.

In particular, the stabilizer is finitely presented. The particular finite simplicial complex we construct as a $\mathrm{K}(G, 1)$-space for the stabilizer yields further properties such as solvability of the word problem. We elaborate more in Section 7.

1.2 Expected product of traces

For every finite set of words $w_1, \ldots, w_\ell \in \mathbf{F}_r$ consider the expected product of traces $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$

$$\mathcal{T}r_{w_1,\ldots,w_\ell}(n) \stackrel{\mathrm{def}}{=} \mathbb{E}\left[\mathrm{tr}\left(w_1\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right) \cdot \ldots \cdot \mathrm{tr}\left(w_\ell\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right)\right].$$

The results we described in Section 1.1 for single words generalize to finite sets of words.

The numbers $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ were studied before. Diaconis and Shahshahani [DS94] consider the joint distribution of $\mathrm{tr}\left(U^{(n)}\right), \mathrm{tr}((U^{(n)})^2), \mathrm{tr}((U^{(n)})^3), \ldots$ (here $U^{(n)} \in \mathcal{U}(n)$ is Haar random). They show that these random variables converge in distribution to independent variables, and as $n \to \infty$, $\mathrm{tr}((U^{(n)})^j)$ converges to $\sqrt{j}Z$, where Z is a standard complex normal variable. This work can be interpreted as the study of (limits of) word measures when the words are in $F_1 \cong \mathbb{Z}$. Later, Mingo, Śniady and Speicher [MŚS07], and independently Rădulescu [Răd06], generalized this result to words in $\mathbf{F}_r$, $r \ge 2$. (The main goal of [MŚS07] is to establish "second order freeness" of random unitary matrices.) Namely, given $w_1, \ldots, w_k \in \mathbf{F}_r$, they consider the random variables

$$\mathrm{tr}\left(w_1\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right), \ \ldots, \ \mathrm{tr}\left(w_k\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right),$$

and study their joint distribution in the limit as $n \to \infty$. All of [DS94], [MŚS07] and [Răd06] use the method of moments which translates the study of the joint limit distribution to the study of (limits as $n \to \infty$ of) expected products of traces, namely, of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ for all possible finite subsets $\{w_1, \ldots, w_\ell\}$ of $\mathbf{F}_r \setminus \{1\}$.

As in the case of $\mathcal{T}r_w(n)$ – the expected trace of a single word – $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ can also be written as a rational expression in n (see Theorem 3.7). As in (1.3), the main interest of

[10] In fact, in this particular case, $\mathcal{T}r_{[x_1^k, x_2]}(n) = \frac{k}{n}$ with no further terms.

[MŚS07] and [Răd06] is in $\lim_{n\to\infty} \mathcal{T}r_{w_1,\ldots,w_\ell}(n)$, namely, in the free coefficient of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ as a Laurent series in $\frac{1}{n}$. We explain their result in Example 1.13 below. Our goal is to explain the leading term (exponent and coefficient) of this rational expression, even when $\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = O\left(\frac{1}{n}\right)$.

Indeed, we establish parallels to Theorems 1.1, 1.2 and 1.4 for the more general object $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$. We introduce these general results in geometric terms rather than algebraic: the geometric language here is more natural both in terms of the statements of the results and in terms of the proofs.

The geometric interpretation of commutator length of words goes back to Culler [Cul81] and explains why $\mathrm{cl}(w)$ is often called "the genus of w". In the geometric approach, solutions to the commutator equation (1.4) are given in terms of maps[11] from surfaces with boundary to a wedge of circles. More concretely, we think of the free group $\mathbf{F}_r$ as the fundamental group of a bouquet of r cycles, denoted $\bigvee^r S^1$, pointed at the wedge point o. For the free group $\mathbf{F}_{2g} = \mathbf{F}(a_1, b_1, \ldots, a_g, b_g)$ we consider a different topological space: the oriented surface of genus g with one boundary component, which we denote by $\Sigma_{g,1}$. Let v_1 be a basepoint of $\Sigma_{g,1}$ at the boundary. Let $(S^1, 1)$ be a cycle pointed at 1, and let $\partial_1 : (S^1, 1) \to (\Sigma_{g,1}, v_1)$ be a fixed map which identifies the boundary of $\Sigma_{g,1}$ with S^1. Identify $a_1, b_1, \ldots, a_g, b_g$ with a suitable basis of $\pi_1(\Sigma_{g,1}, v_1)$ so that $\delta_g = [a_1, b_1] \ldots [a_g, b_g]$ is represented by $[\partial_1]$. It is shown in [Cul81] that every solution to (1.4) can be given by a map $f\colon (\Sigma_{g,1}, v_1) \to \left(\bigvee^r S^1, o\right)$ with $f_*([\partial_1]) = w$. In fact, there is a one-to-one correspondence between the solutions in $\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$ and homotopy classes of such maps $(\Sigma_{g,1}, v_1) \to \left(\bigvee^r S^1, o\right)$ (see Proposition 2.3).

$\bigvee^r S^1$

o

$\Sigma_{g,1}$

v_1

∂_1

We now describe the geometric analogue of $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$. For this sake, we first fix the map from the boundary of $\Sigma_{g,1}$ to the wedge. Formally, for every $w \in \mathbf{F}_r$ fix

f_w

$$f_w : (S^1, 1) \to \left(\bigvee^r S^1, o\right)$$

a map which describes a fixed loop in $\bigvee^r S^1$ representing w, namely, $[f_w] = w \in \pi_1\left(\bigvee^r S^1, o\right)$, and consider the set of maps

$$\left\{ f : (\Sigma_{g,1}, v_1) \to \left(\bigvee^r S^1, o\right) \,\middle|\, f \circ \partial_1 = f_w \right\}. \tag{1.6}$$

Let $\mathrm{Homeo}_\delta(\Sigma_{g,1})$ denote the group of homeomorphisms of $\Sigma_{g,1}$ that fix the boundary pointwise, and write $\mathrm{Homeo}_0(\Sigma_{g,1})$ for the normal subgroup of $\mathrm{Homeo}_\delta(\Sigma_{g,1})$ consisting of homeomorphisms isotopic to the identity. While $\mathrm{Homeo}_\delta(\Sigma_{g,1})$ acts on the set of maps in (1.6) by precomposition, the quotient by $\mathrm{Homeo}_0(\Sigma_{g,1})$ acts on homotopy classes of these maps. This quotient is precisely the **mapping class group** of $\Sigma_{g,1}$:

$\mathrm{MCG}(\Sigma_{g,1})$

$$\mathrm{MCG}(\Sigma_{g,1}) \stackrel{\mathrm{def}}{=} \mathrm{Homeo}_\delta(\Sigma_{g,1}) / \mathrm{Homeo}_0(\Sigma_{g,1}).$$

The Dehn-Nielsen-Baer theorem (Theorem 2.4 below) states there is a natural isomorphism $\mathrm{Aut}_\delta(\mathbf{F}_{2g}) \cong \mathrm{MCG}(\Sigma_{g,1})$. Through this isomorphism, the action of $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ on $\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$ is identical to the action of $\mathrm{MCG}(\Sigma_{g,1})$ on the homotopy classes of maps in (1.6). We summarize this algebra-geometry dictionary in Table 1. We give more details and further explanations in Section 2.1.

We can now describe our general results. To deal with multiple words, we need surfaces with multiple boundary components. More concretely,

Definition 1.5. Let Σ be a surface and $f\colon \Sigma \to \bigvee^r S^1$. We say that (Σ, f) is **admissible** for $w_1, \ldots, w_\ell \in \mathbf{F}_r$ if the following three conditions hold:

(Σ, f) admissible for $w_1, \ldots, w_\ell$

[11] All maps in this paper are assumed to be continuous.

$\mathbf{F}_r$	$\pi_1\left(\bigvee^r S^1, o\right)$
$\mathbf{F}_{2g} = \mathbf{F}(a_1, b_1, \ldots, a_g, b_g)$ $\delta_g = [a_1, b_1] \ldots [a_g, b_g]$	$\pi_1(\Sigma_{g,1}, v_1)$ with fixed loops representing $a_1, b_1, \ldots, a_g, b_g$ so that $[\partial_1] = \delta_g$
$\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$	homotopy classes of $\left\{f \colon (\Sigma_{g,1}, v_1) \to \left(\bigvee^r S^1, o\right) \,\middle|\, f \circ \partial_1 = f_w\right\}$
$\mathrm{cl}(w)$	$\min\left\{g \,\middle|\, \exists f \colon \Sigma_{g,1} \to \bigvee^r S^1 \text{ with } f \circ \partial_1 = f_w\right\}$
$\mathrm{Aut}_\delta(\mathbf{F}_{2g})$	$\mathrm{MCG}(\Sigma_{g,1})$
equivalence classes of solutions: $\mathrm{Aut}_\delta(\mathbf{F}_{2g}) \backslash \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$	$\mathrm{MCG}(\Sigma_{g,1}) \setminus \{$homotopy classes of maps as above$\}$

Table 1: Algebra-geometry dictionary.

1. Σ is compact, oriented, with ℓ boundary components, and contains no closed connected components (but is not necessarily connected).
2. Σ has ℓ marked points $v_1, \ldots, v_\ell$ on Σ, one point in every boundary component, and fixed identifications of the boundaries with S^1 with common orientation given by $\partial_1, \ldots, \partial_\ell$
$$\partial_1 : \left(S^1, 1\right) \to (\Sigma, v_1) \quad \ldots \quad \partial_\ell : \left(S^1, 1\right) \to (\Sigma, v_\ell)$$
3. f maps the boundary components to $w_1, \ldots, w_\ell$, namely,
$$f \circ \partial_1 = f_{w_1} \quad \ldots \quad f \circ \partial_\ell = f_{w_\ell}.$$

In particular, every admissible map sends the marked points $v_1, \ldots, v_\ell$ to o. The next definition captures the maximal possible Euler characteristic of an admissible surface:

Definition 1.6. For $w_1, \ldots, w_\ell \in \mathbf{F}_r$ define $\mathrm{chi}(w_1,\ldots,w_\ell)$

$$\mathrm{chi}(w_1, \ldots, w_\ell) \overset{\text{def}}{=} \max\left\{\chi(\Sigma) \,|\, (\Sigma, f) \text{ is admissible for } w_1, \ldots, w_\ell\right\},$$

where $\chi(\Sigma)$ is the Euler characteristic of Σ. If no such surface exists, define $\mathrm{chi}(w_1, \ldots, w_\ell) = -\infty$.

As we explain below, $\mathrm{chi}(w_1, \ldots, w_\ell) \neq -\infty$ (i.e. there exists an admissible map for $w_1, \ldots, w_\ell$), if and only if the product $w_1 w_2 \cdots w_\ell \in [\mathbf{F}_r, \mathbf{F}_r]$. Equivalently, this holds if and only if the sum of exponents of the letter x_i across $w_1, \ldots, w_\ell$ is zero for every $1 \leq i \leq r$. As for a single word, if $w_1 \cdots w_\ell \notin [\mathbf{F}_r, \mathbf{F}_r]$ then $\mathcal{T}r_{w_1,\ldots,w_\ell}(n) \equiv 0$ vanishes for every n (Claim 3.1).

Remark 1.7. For a single word, $\mathrm{chi}(w) = 1 - 2 \cdot \mathrm{cl}(w)$. More generally, the commutator length of a finite set of words $w_1, \ldots, w_\ell \in \mathbf{F}_r$, introduced by Calegari (e.g. [Cal09a, Definition 2.71]), is defined as the smallest number of commutators whose product is equal to an expression of the form

$$w_1 t_1 w_2 t_1^{-1} \ldots t_\ell w_\ell t_\ell^{-1}$$

with $t_2, \ldots, t_\ell \in \mathbf{F}_r$. This number, which can be denoted $\mathrm{cl}(w_1, \ldots, w_\ell)$, relates to $\mathrm{chi}(w_1, \ldots, w_\ell)$ by

$$\mathrm{chi}(w_1, \ldots, w_\ell) = 2 - \ell - 2 \cdot \mathrm{cl}(w_1, \ldots, w_\ell)$$

(when $w_1, \ldots, w_\ell \neq 1$). However, $\mathrm{chi}()$ is more natural then $\mathrm{cl}()$ in this general case: it simplifies the statement of our results below, and appears more directly in the proofs[12]. In fact, Calegari himself also mostly uses the geometric definition in his works.

[12] Another advantage of $\mathrm{chi}()$ compared with $\mathrm{cl}()$ is that with $\mathrm{chi}()$, the statements of our results remain valid when some of the words are the identity element $1 \in \mathbf{F}_r$. (Observe that $\mathrm{chi}(w_1, \ldots, w_\ell, 1) = \mathrm{chi}(w_1, \ldots, w_\ell) + 1$.)

With this definition, the leading exponent from Theorem 1.1 is simply $n^{\mathrm{chi}(w)}$. This generalizes to

Theorem 1.8. *For $w_1, \ldots, w_\ell \in \mathbf{F}_r$ we have*

$$\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = O\left(n^{\mathrm{chi}(w_1,\ldots,w_\ell)}\right).$$

Next, in order to state our result for the leading coefficient of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$, we define equivalence classes of "solutions", namely, of admissible maps of maximal Euler characteristic, for the words $w_1, \ldots, w_\ell$. We say that two admissible maps (Σ, f) and (Σ', f') are equivalent, and denote $(\Sigma, f) \sim (\Sigma', f')$, if there is an homeomorphism $\rho\colon \Sigma \to \Sigma'$ so that $f \simeq f' \circ \rho$ are homotopic relative $\partial\Sigma$ while the boundary components are identified pointwise, that is, $\partial'_i = \rho \circ \partial_i$ for $1 \le i \le \ell$. $(\Sigma,f)\sim(\Sigma',f')$

Definition 1.9. For $w_1, \ldots, w_\ell \in \mathbf{F}_r$ let $\mathrm{Solu}(w_1, \ldots, w_\ell)$ denote the set of equivalence classes of "solutions", or admissible maps of maximal Euler characteristic, for $w_1, \ldots, w_\ell$. Namely, $\mathrm{Solu}(w_1,\ldots,w_\ell)$

$$\mathrm{Solu}(w_1,\ldots,w_\ell) \stackrel{\mathrm{def}}{=} \left\{ (\Sigma, f) \,\middle|\, \begin{array}{c} (\Sigma, f) \text{ is admissible for } w_1, \ldots, w_\ell, \text{ and} \\ \chi(\Sigma) = \mathrm{chi}(w_1, \ldots, w_\ell) \end{array} \right\} / \, (\Sigma, f) \sim \left(\Sigma', f'\right).$$

We denote by $[(\Sigma, f)]$ the equivalence class of the admissible map (Σ, f). $[(\Sigma, f)]$

We can now state the more detailed version of Theorem 1.8 which generalizes Theorem 1.2:

Theorem 1.10. *Let $w_1, \ldots, w_\ell \in \mathbf{F}_r$. Then,*

$$\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = n^{\mathrm{chi}(w_1,\ldots,w_\ell)} \left[\sum_{[(\Sigma,f)] \in \mathrm{Solu}(w_1,\ldots,w_\ell)} \chi\left(\mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)\right) \right] + O\left(n^{\mathrm{chi}(w_1,\ldots,w_\ell)-2}\right),$$

where $\tilde{f}$ is the homotopy class of f (relative the boundary of Σ). $\tilde{f}$

As above, $\mathrm{MCG}(\Sigma)$ is the mapping class group of the surface Σ, consisting of mapping classes which fix the boundary pointwise. It acts on homotopy classes of maps from the surface by precomposition. $\mathrm{MCG}(\Sigma)$

Indeed, $\mathrm{Solu}(w_1, \ldots, w_\ell)$ is always a finite set (see Corollary 4.11). Finally, we need to justify our usage of the Euler characteristic[13] of the stabilizers of the maps in $\mathrm{Solu}(w_1, \ldots, w_\ell)$, namely, to give a generalized version of Theorem 1.4. It turns out that the crucial property of these maps is their being incompressible:

Definition 1.11. A map f from a surface Σ to a topological space is called **compressible** if there is a non-nullhomotopic simple closed curve γ in Σ such that $f(\gamma)$ is (freely) nullhomotopic. Otherwise, f is called **incompressible.**

This term is standard (see, e.g., [Cal09a]). It incorporates maps solving the commutator equation (1.4), and more generally, maps in $\mathrm{Solu}(w_1, \ldots, w_\ell)$: if (Σ, f) is admissible for $w_1, \ldots, w_\ell$ and f is compressible, one can cut Σ along the compressing simple closed curve γ, cap with two discs to obtain a new surface Σ' and extend f to a map f' from Σ'. But then (Σ', f') is also admissible for $w_1, \ldots, w_\ell$ with $\chi(\Sigma') = \chi(\Sigma) + 2$. So Σ cannot be of maximal Euler characteristic.

[13] Note that our results use two different instances of Euler characteristics. On the one hand, they use Euler characteristics of compact surfaces, and on the other hand the Euler characteristic of stabilizer subgroups, or of the corresponding $K(G,1)$-spaces. We try to ease the confusion by using the notation $\mathrm{chi}(w_1, \ldots, w_\ell)$ for the former (instead of, say, the more natural $\chi(w_1, \ldots, w_\ell)$).

Theorem 1.12. *Let $\tilde{f}\colon \Sigma \to \bigvee^r S^1$ be a homotopy class (relative $\partial\Sigma$) of incompressible maps from a compact oriented surface to the wedge. Then the stabilizer*

$$G = \mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$$

admits a finite simplicial complex as $K(G,1)$-space.

We remark the statement is void when Σ has a closed connected component of positive genus: there are no incompressible maps from a closed surface to the wedge[14].

The following special case of Theorem 1.10 is due to [MŚS07] and [Răd06]:

Example 1.13. Consider the limit

$$\lim_{n\to\infty} \mathcal{T}r_{w_1,\ldots,w_\ell}(n) \tag{1.7}$$

for $w_1,\ldots,w_\ell \neq 1$. Assume (Σ, f) is admissible for $w_1,\ldots,w_\ell$. By definition, every connected component of Σ has non-empty boundary, so its Euler characteristic is negative unless it is a disc or an annulus. But a disc is impossible as we assume $w_1,\ldots,w_\ell \neq 1$. Thus, the only case in which (1.7) is non-zero is when there is an admissible pair (Σ, f) with Σ is a disjoint union of (one or more) annuli. In every annulus A, if w and w' are the two words on the boundary components, then necessarily w^{-1} is conjugate to w'. Moreover, write $w = u^d$ with $u \in \mathbf{F}_r$ a non-power and $d \geq 1$, then the number of equivalence classes of maps h such that (A, h) is admissible for w, w' is exactly d. Since the mapping class group of the annulus is simple to analyze (isomorphic to $\mathbb{Z}$, generated by a Dehn twist), it is not hard to see the stabilizers $\mathrm{Stab}_{\mathrm{MCG}(A)}\left(\tilde{h}\right)$ are always trivial.

These considerations yield Theorem 4.1 in [Răd06][15]: (1.7) is non-zero if and only if $w_1,\ldots,w_\ell$ can be matched in pairs in which each word is conjugate to the inverse of its mate. In this case, the limit in (1.7) is equal to the number of such matchings, times the product of exponents of the words (one exponent for every pair).

Because of the degenerate case described in Footnote 6, it is not clear whether the commutator length $\mathrm{cl}(w)$ is determined by the w-measures on $\{\mathcal{U}(n)\}_{n\in\mathbb{N}}$, or, more generally, if $\mathrm{chi}(w_1,\ldots,w_\ell)$ is determined by the joint measures of $w_1,\ldots,w_\ell$ on unitary groups. However, the measures do determine a related number, the *stable commutator length* of w. This algebraic quantity is defined by

$$\mathrm{scl}(w) \equiv \lim_{m\to\infty} \frac{\mathrm{cl}(w^m)}{m}. \tag{1.8}$$

(There is an analogous definition for finite sets of words.) There is a deep theory behind this invariant, and for background we refer to the short survey [Cal08] and long one [Cal09a] by Calegari. Relying on the rationality result of Calegari [Cal09b] that shows, in particular, that scl takes on rational values in $\mathbf{F}_r$, we are able to show the following:

Corollary 1.14. *The stable commutator length of a word $w \in \mathbf{F}_r$ can be "read" from the measures it induces on unitary groups in the following way:*

$$\mathrm{scl}(w) = \inf_{\ell>0;\, j_1,\ldots,j_\ell>0} \frac{-\lim_{n\to\infty} \log_n \mathcal{T}r_{w^{j_1},\ldots,w^{j_\ell}}(n)}{2(j_1+\ldots+j_\ell)}. \tag{1.9}$$

A similar result is true for the stable commutator length of several words. We explain how Corollary 1.14 follows from Theorem 1.10 and Calegari's rationality theorem in Section 2.3.

[14] For example, this can be seen using the proof of Theorem 1.4 in [Cul81], by turning a map from a closed surface to a "tight" map.

[15] The same theorem is an immediate consequence of Theorem 2 in [MŚS07]. In [Răd06] the theorem is shown, for simplicity, only for $\mathbf{F}_2$ (in our analysis there is no saving in restricting to $\mathbf{F}_2$).

1.3 More related work and further motivation

Our work is inspired by that of the second author and Parzanchevski [PP15], where word measures on finite symmetric groups are considered. An element of a free group $\mathbf{F}$ is called *primitive* if it belongs to some free generating set of $\mathbf{F}$. The following estimate from [PP15, Theorem 1.8] is analogous to Theorem 1.2:

Theorem 1.15 (Puder-Parzanchevski)**.** *Let S_n be the symmetric group on n elements. For $w \in \mathbf{F}_r$ given as in (1.1), let w be the word map*

$$w : S_n^r \to S_n, \quad w(\sigma_1, \ldots, \sigma_r) \equiv \prod_{1 \le j \le |w|} \sigma_{i_j}^{\varepsilon_j},$$

just as in (1.2). Let $\sigma_1^{(n)}, \ldots, \sigma_r^{(n)}$ be r independent random permutations in S_n taken with respect to the uniform measure, viewed as 0-1 $n \times n$ matrices. Then

$$\mathbb{E}\left[\operatorname{tr}\left(w\left(\sigma_1^{(n)}, \ldots, \sigma_r^{(n)}\right)\right)\right] = 1 + \frac{|\mathrm{Crit}(w)|}{n^{\pi(w)-1}} + O\left(\frac{1}{n^{\pi(w)}}\right),$$

where $|\mathrm{Crit}(w)|$ and $\pi(w)$ are invariants of w. The primitivity rank $\pi(w)$ *is the minimal rank of a subgroup in*

$$\{ J \mid w \in J \le \mathbf{F}_r \text{ and } w \text{ is } \textbf{not} \text{ primitive in } J \}.$$

$\mathrm{Crit}(w)$ *is the set of subgroups attaining this minimum rank.*

The study leading to Theorem 1.15 had two main motivations, both of which are also relevant to the main result of the current paper. The first motivation is related to questions about word measures on finite, or more generally compact, groups. As mentioned above, the measure induced by $w \in \mathbf{F}_r$ on some compact group G is identical to the measure induced by $\theta(w)$ for any $\theta \in \mathrm{Aut}(\mathbf{F}_r)$. In particular, since the x_1-measure on G (the measure induced by the single letter word "x_1") is the Haar measure, or simply the uniform measure for finite groups, the same holds for the w-measure of every word w in the $\mathrm{Aut}(\mathbf{F}_r)$-orbit of x_1. This orbit consists precisely of the *primitive* words in $\mathbf{F}_r$. Several mathematicians have asked whether primitive words are the only words inducing the uniform (Haar) measure on every finite (compact, respectively) group (see [PP15] and the references therein). Theorem 1.15 answered this question to the positive, showing that every non-primitive word induces a non-uniform measure on S_n for n large enough. However, many conjectures revolving around word measures on groups remain open, and we see the current paper as a step towards their resolution. More details are given in Section 2.2.

The second motivation for Theorem 1.15 lies in the field of random graphs, and more precisely that of spectra of random graphs. A strengthened version of the asymptotic formula in Theorem 1.15 appears in [Pud15], where it is used in an approach to *Alon's second eigenvalue conjecture* from [Alo86] that says

'Almost all d-regular graphs are weakly Ramanujan.'

This conjecture was proved by Friedman in [Fri08] and a new proof has been given recently by Bordenave [Bor15]. While an approach using asymptotics of word maps has not yet proved the full strength of Alon's conjecture, the approach in [Pud15] comes very close (up to a small additive constant) while keeping the proof manageable. This approach has also given the best result to date regarding a natural generalization of Alon's conjecture to families of irregular graphs (see [Pud15]).

One can ask analogous questions about the spectrum of sums of Haar distributed unitary matrices in the large n limit. Consider, for example, the sum

$$\sum_{i=1}^{r} U_i^{(n)} + (U_i^{(n)})^*. \tag{1.10}$$

The connection to word measures on $\mathcal{U}(n)$ is that the N^{th} power of (1.10) is equal to the sum, over all not-necessarily-reduced words w of length N, of $w(U_1^{(n)}, \ldots, U_r^{(n)})$.

When one replaces unitaries in (1.10) with random permutation matrices, one gets the adjacency matrix of a graph sampled from the *permutation model* of random regular graphs. Hence the analogy with spectral graph theory. Heuristically, questions about the spectra of sums of unitary matrices should be much easier than the corresponding questions about sums of 0-1 permutation matrices[16], owing to the random unitary matrices being denser, and thus having more variables to average over.

Nevertheless, interesting analytic problems about random unitary matrices remain. In [HT05] Haagerup and Thorbjørnsen proved that a certain operator-theoretic semigroup $\mathrm{Ext}(\mathbf{F}_r)$ is not a group for $r \geq 2$, which had been an open problem for about 25 years. Their approach uses an observation of Voiculescu from [Voi93] that reduces the question to one about the existence of unitary representations of $\mathbf{F}_r$ with certain spectral features[17]. Building on the work of [HT05], Collins and Male [CM14] proved the *strong asymptotic freeness* of Haar unitary matrices from which they obtain:

Theorem 1.16 (Collins-Male)**.** *Almost surely*

$$\left\| \sum_{i=1}^{r} U_i^{(n)} + (U_i^{(n)})^* \right\| \xrightarrow{n\to\infty} 2\sqrt{2r-1}.$$

We expect that our Theorems 1.2 and 1.10, made suitably uniform in w or $w_1, \ldots, w_\ell$, should give an alternative approach to bounds such as in Theorem 1.16, as well as to the related questions of strong asymptotic freeness and properties of $\mathrm{Ext}(\mathbf{F}_r)$. Going further with these questions, one expects the following "folklore" conjecture:

> 'The largest eigenvalue of (1.10) should be governed by a suitably normalized Tracy-Widom law, in the limit $n \to \infty$.'

The set of solutions to (1.4) along with its $\mathrm{Aut}_\delta(\mathbf{F}_r)$-action is interesting even considered apart from the connection with Random Matrix Theory made in Theorem 1.2. In fact, it is the content of quite a few research papers.

Algorithms to compute commutator lengths of words in free groups were found independently by [Edm75], [GT79] and [Cul81]. The latter work, by Culler, is the most relevant to ours. His geometric approach to $\mathrm{cl}(w)$ which we mentioned above (and see Proposition 2.3 below), is further developed in the current paper and stands in the core of our methods. Culler also introduces an algorithm to obtain a representative of every equivalence class of solutions to (1.4), namely of every orbit of $\mathrm{Aut}_\delta(\mathbf{F}_{2g}) \backslash \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$ where $g = \mathrm{cl}(w)$. Although similar in spirit, our analysis yields a clearer description of the set of classes of solutions and, in particular, a more direct way to distinguish them from each other. See Remark 4.12 and Section 7 for comparison between Culler's approach and ours.

[16] We thank Peter Sarnak for an illuminating conversation about this subject.

[17] Voiculescu in [Voi93] also relates these questions to the existence of Ramanujan graphs, pleasantly completing a circle of ideas.

In addition, Culler proves that for every $w \in [\mathbf{F}_r, \mathbf{F}_r]$ there are only finitely many equivalence classes of solutions to (1.4). This extends an older result regarding words w with $\mathrm{cl}(w) = 1$ [Hme71], and we extend it further to equivalence classes of admissible incompressible maps of $w_1, \ldots, w_\ell$ – see Corollary 4.11. We remark that some researchers have looked at a larger group $\widehat{\mathrm{Aut}}_\delta(\mathbf{F}_{2g}) \supset \mathrm{Aut}_\delta(\mathbf{F}_{2g})$ acting on the solution space to (1.4). In geometric terms, one allows not only ordinary Dehn twists, but also "fractional" ones – see [BF05]. Bestvina and Feighn [BF05] study the problem of counting the number of $\widehat{\mathrm{Aut}}_\delta(\mathbf{F}_{2g})$-orbits of solutions to (1.4). They prove that for all $g \geq 1$ there is a word w with $\mathrm{cl}(w) = g$ which has at least 2^g distinct $\widehat{\mathrm{Aut}}_\delta(\mathbf{F}_{2g})$-orbits of solutions to (1.4). When $g = 1$, this is a result of Lyndon and Wicks [LW81]. The motivation for [BF05] comes for questions raised by Sela, who has introduced a very general framework for studying the solutions to systems of equations such as (1.4) in free groups (e.g. [Sel01]).

Before giving an overview of our proofs in Section 1.4 below, we trace the history of the ideas of this paper. A *ribbon graph*, also called a *fat graph*, is a graph where each vertex comes with a cyclic ordering of its incident edges. Ribbon graphs commonly serve as a combinatorial way to describe orientable surfaces with boundary: every vertex is magnified to a disc, and every edge widened to a strip. A standard reference is [Pen88, Section 1]. With some extra information ribbon graphs appear as the "dessins d'enfants" of Grothendieck [Gro]. The book of Lando and Zvonkin [LZ04] gives an encyclopedic overview of subjects related to ribbon graphs.

There are two central themes in the current paper:

A Certain integrals over random matrices can be computed by a sum of terms encoded by "ribbon graphs". Moreover, the order of contribution of each term corresponds to the *genus* or to the *degree sequence* of the corresponding ribbon graph.

B Certain contributions from the sum in **A** coincide with homotopy invariants of some topological spaces.

One early synthesis of these ideas is the following seminal result of Harer and Zagier [HZ86], independently discovered by Penner [Pen88].

Theorem 1.17 (Harer-Zagier, Penner)**.** *Assume $g \geq 1$. Let Σ_g^1 be the closed genus g surface with one point removed and let $\mathrm{MCG}\left(\Sigma_g^1\right)$ be the mapping class group of isotopy classes of orientation preserving homeomorphisms $\Sigma_g^1 \to \Sigma_g^1$. Then*

$$\chi\left(\mathrm{MCG}\left(\Sigma_{\mathrm{g}}^1\right)\right) = \zeta\left(1 - 2g\right), \tag{1.11}$$

where ζ is Riemann's zeta function.

Penner's approach in [Pen88] clarifies our discussion so we give a brief outline. Penner begins with the apriori unrelated[18] matrix integral

$$P_{v_3,\ldots,v_K}(n) = \frac{1}{\mu_n \prod_{j=1}^K v_j!} \int \prod_{j=1}^K \left(\frac{\mathrm{tr} H^j}{j}\right)^{v_j} \exp\left(\frac{-\mathrm{tr} H^2}{2}\right) dH, \tag{1.12}$$

where the integral is taken over the probability space of GUE $n \times n$ Hermitian matrices, v_k are non-negative integers and μ_n is a normalization factor. He proves that $P_{v_3,\ldots,v_K}$ is a polynomial

[18] As Penner puts it, "It is also noteworthy that the technique of perturbative series from particle physics so effectively captures the combinatorics of the bundle over Teichmüller space [...].".

in n that can be expressed as a sum over ribbon graphs with exactly v_j vertices of degree j for every $3 \leq j \leq K$ (and no vertices of degree 1, 2 or larger than K).

The general idea of equating matrix integrals with sum of terms encoded by diagrams goes back to the celebrated "Feynman diagrams" of [Fey48], and the first encoding by ribbon graphs seems to be due to by 't Hooft [tH74]. In [BIZ80], Bessis, Itzykson and Zuber consider a matrix integral roughly similar to (1.12), with an extra generating parameter λ, and show that in the sum they obtain over ribbon graphs, the exponent of λ in every term coincides with the genus of the corresponding ribbon graph.

As for Theme **B**, the key topological object related to Theorem 1.17 is the *fat graph complex* $\mathcal{G}_g^1$ of Penner, defined in [Pen88, Page 41][19]. An equivalent definition, and one more clearly related to our setting, is that $\mathcal{G}_g^1$ is a simplicial complex with one simplex of dimension k for each isotopy class of $k+1$ disjoint embedded arcs in Σ_g^1 with the following properties. The arcs begin and end at the puncture, must be pairwise non parallel, individually not homotopic into the puncture, and must cut Σ_g^1 into discs. Each of these discs must be bounded by at least 3 arcs. One simplex is a face of another if it can be obtained by deleting some arcs. Thus $\mathcal{G}_g^1$ carries the obvious action of the mapping class group by change of markings.

This $\mathcal{G}_g^1$ arises naturally from the Teichmüller space of Σ_g^1 and furthermore inherits its homotopy type[20]. By the well known work of Fenchel and Nielsen [FN03], the Teichmüller space of Σ_g^1 is contractible and thus so is $\mathcal{G}_g^1$. This is the fact that allows one to obtain an Euler characteristic in Theorem 1.17. Indeed this Euler characteristic can be obtained by counting MCG-orbits of simplices of $\mathcal{G}_g^1$, and after translation to fat/ribbon graphs this is exactly what shows up in the Feynman diagram expansion of Theme **A**.

A similar combinatorial model of the moduli space of curves was given by Kontsevich in [Kon92, Theorem 2.2] by means of Jenkins-Strebel quadratic differentials, and in Appendix D of *loc. cit.* Kontsevich gives a short proof of Theorem 1.17. These results appear in the context of the proof of a conjecture of Witten from [Wit91] asserting that two models of quantum gravity are equal.

1.4 Overview of the proof and paper organization

We now sketch the outline of the proofs of our main results.

Section 3: Formula for $\mathcal{T}r$ using pairs of matchings of letters

In the first stage of our analysis, a crucial role is played by a formula developed in [Xu97] and extended in [Col03] and [CŚ06] in the aim of giving a new proof to the asymptotic freeness of Haar Unitary matrices, namely, to (1.3). This is an integration formula for polynomials in the entries of a Haar unitary matrix and their conjugates, appearing as Theorem 3.6 below. For example, it allows one to compute

$$\int_{u\in\mathcal{U}(n)} u_{1,2}u_{3,4}\overline{u_{1,4}u_{3,2}}d\mu_n. \tag{1.13}$$

This formula is parallel to a moment formula for Gaussian variables that appears in the corresponding GUE analysis, a formula which usually goes under the name "Wick formula".

[19] Penner defines arc complexes $\mathcal{G}_g^s$ for surfaces of genus g with s punctures, and everything we say about Penner's work naturally extends to general g and s.

[20] Following [Pen88, Page 41], $\mathcal{G}_g^1$ is MCG-equivariantly homotopy equivalent to a MCG-invariant spine of some decorated Teichmüller space. This decorated version is homeomorphic to the Cartesian product of the usual Teichmüller space and $\mathbf{R}_+$.

As shown in [CŚ06], the evaluation of every such polynomial is a rational function in n. For example, the integral in (1.13) is equal to $\frac{-1}{n^3-n}$ for every $n \geq 4$. A key feature of this formula is that the leading term (exponent and coefficient) has combinatorial significance, and is related to the Möbius function of the poset (partially ordered set) of non-crossing partitions.

In the current paper, we fully expand out the product $\mathrm{tr}(w_1(U_1^{(n)}, \ldots, U_r^{(n)})) \cdots \mathrm{tr}(w_\ell(U_1^{(n)}, \ldots, U_r^{(n)}))$ as a sum over indices of rows and columns of the matrices $U_1^{(n)}, \ldots, U_r^{(n)}$, and, using the integration formula mentioned above, show its expected value $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ is indeed a rational function in n, which can be computed explicitly. This is the content of Theorem 3.7 below.

The formula we obtain for $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ can be viewed as a sum over pairs (σ,τ) of matchings of the letters of $w_1, \ldots, w_\ell$, where every letter x_i^{ε} is matched with some $x_i^{-\varepsilon}$. Indeed, by Claim 3.1 below, $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ vanishes unless the total number of instances of x_i^{-1} in $w_1, \ldots, w_\ell$ is equal to the total number of x_i^{+1}, for every $i \in [r]$. The latter holds if and only if $w_1 w_2 \cdots w_\ell \in [\mathbf{F}_r, \mathbf{F}_r]$, and we sometimes say that in this case $w_1, \ldots, w_\ell$ form a **balanced** set of words. The set of matchings associated with $w_1, \ldots, w_\ell$ is denoted $\mathrm{Match}(w_1, \ldots, w_\ell)$ and is formally described in Definition 4.1. balanced set of words

Section 4: Constructing surfaces from pairs of matchings and Theorems 1.1 and 1.8

In Section 4 we explain how to associate an orientable surface $\Sigma_{(\sigma,\tau)}$ with every pair of matchings $(\sigma,\tau) \in \mathrm{Match}(w_1, \ldots, w_\ell) \times \mathrm{Match}(w_1, \ldots, w_\ell)$. The surface $\Sigma_{(\sigma,\tau)}$, which is basically given in the form of a ribbon graph, has ℓ boundary components and its Euler characteristic is denoted $\chi(\sigma,\tau)$. This extends a construction of Culler [Cul81] that deals with a single matching $\sigma \in \mathrm{Match}(w)$ of a single word: his construction is the special case $\sigma = \tau$ in ours. The extension to pairs of matchings (and to multiple words) seems to be new here.

It so happens that in the formula for $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ given by a sum over pairs of matchings, the contribution of every pair (σ,τ) is of order $n^{\chi(\sigma,\tau)}$ (Proposition 4.6). Hence the contributions to $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ of largest order come from pairs (σ,τ) of largest Euler characteristic.

In addition, we associate with the pair (σ,τ) a map (defined up to homotopy)

$$f_{(\sigma,\tau)} \colon \Sigma_{(\sigma,\tau)} \to \bigvee^r S^1,$$

and show that $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$ is admissible for $w_1, \ldots, w_\ell$. Moreover, in Lemma 4.9 we explain that for every (Σ, f) admissible for $w_1, \ldots, w_\ell$ with f incompressible, there is a pair of matchings (σ,τ) such that $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right) \sim (\Sigma, f)$. This gives a procedure for computing $\mathrm{chi}(w_1, \ldots, w_\ell)$ which, again, generalizes a procedure suggested in [Cul81] to compute $\mathrm{cl}(w)$.

Since every admissible (Σ, f) with $\chi(\Sigma) = \mathrm{chi}(w_1, \ldots, w_\ell)$ is incompressible, we deduce in Corollary 4.13 the content of Theorems 1.8 and 1.1, namely, that $\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = O\left(n^{\mathrm{chi}(w_1,\ldots,w_\ell)}\right)$. This result roughly summarizes the role in the current work of Theme **A** from above (although we have not used the fine details of the ribbon graphs so far, only the Euler characteristics of the underlying surfaces).

Section 5: Poset of pairs of matchings for incompressible maps

Our next goal is to study the leading coefficient of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$, namely, the coefficient of $n^{\mathrm{chi}(w_1,\ldots,w_\ell)}$. For this sake, we gather all pairs of matchings (σ,τ) that are associated with the same class $[(\Sigma, f)]$ of admissible surfaces and maps, and denote the set $\mathcal{PMP}(\Sigma, f)$:

$$\mathcal{PMP}(\Sigma, f) \stackrel{\mathrm{def}}{=} \left\{ (\sigma,\tau) \in \mathrm{Match}(w_1, \ldots, w_\ell)^2 \,\middle|\, \left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right) \sim (\Sigma, f) \right\}.$$

We show that whenever f is incompressible, there is a natural partial order on the set $\mathcal{PMP}(\Sigma, f)$, which turns it into a poset we call the Pairs of Matchings Poset of (Σ, f) (Definition 5.1). This partial order is closely related to the aforementioned partial order on non-crossing partitions (e.g. Proposition 5.8).

The pairs of matchings poset is important mainly because of the role of its associated simplicial complex. This finite complex, the simplices of which corresponding to chains in $\mathcal{PMP}(\Sigma, f)$, is denoted $|\mathcal{PMP}(\Sigma, f)|$ – see Definition 5.10. Theorem 5.11 shows that the contributions to $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ of all pairs of matchings in $\mathcal{PMP}(\Sigma, f)$ sum to

$$\chi\left(|\mathcal{PMP}(\Sigma, f)|\right) \cdot n^{\chi(\Sigma)} + O\left(n^{\chi(\Sigma)-2}\right). \tag{1.14}$$

We remark again that the two instances of $\chi()$ in (1.14) are applied to very different topological objects: on the one hand an orientable compact surface Σ, and on the other hand a simplicial complex obtained from a poset whose elements are related to Σ.

$|\mathcal{PMP}(\Sigma, f)|$ as $K(G, 1)$-space

Finally, we prove that the simplicial complex $|\mathcal{PMP}(\Sigma, f)|$ is a $K(G, 1)$-space for $G = \mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$ whenever f is incompressible. This is the content of Theorem 5.12 below, and it obviously yields Theorems 1.12 and 1.4, and together with (1.14) implies our main result: Theorem 1.10 and its special case, Theorem 1.2.

Proving that $|\mathcal{PMP}(\Sigma, f)|$ is a $K(G, 1)$-space boils down to showing the following three facts (see Footnote 9):

(i) $|\mathcal{PMP}(\Sigma, f)|$ is path-connected.

(ii) The fundamental group of $|\mathcal{PMP}(\Sigma, f)|$ is isomorphic to $\mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$. And,

(iii) The universal cover of $|\mathcal{PMP}(\Sigma, f)|$ is contractible.

Establishing this result requires the most involved part of this work and the introduction of yet another poset: the arc poset of (Σ, f).

Section 6: The arc poset of (Σ, f)

Let Σ be a compact surface and $f\colon \Sigma \to \bigvee^r S^1$ incompressible so that (Σ, f) is admissible for $w_1, \ldots, w_\ell$. The arc poset of (Σ, f), denoted $\mathcal{AP}(\Sigma, f)$, is an infinite poset composed of "arc systems". An arc system consists of $|w_1| + \ldots + |w_\ell|$ disjoint arcs in Σ (defined up to isotopy). The boundary components of Σ are marked in a way that "spells out" $w_1, \ldots, w_\ell$ (via the functions $f_{w_i} \circ \partial_i^{-1}$ for $i \in [\ell]$), and the arcs represent a pair of matchings in $\mathrm{Match}(w_1, \ldots, w_\ell)$. Thus, every arc system in $\mathcal{AP}(\Sigma, f)$ is a specific geometric realization of a pair (σ, τ) in $\mathcal{PMP}(\Sigma, f)$. However, every pair (σ, τ) has (infinitely) many different geometric realizations. We endow the set of arc systems with a partial ordering, analogous to the one we defined on $\mathcal{PMP}(\Sigma, f)$. This order is too related to the order on non-crossing partitions. The construction of the arc poset $\mathcal{AP}(\Sigma, f)$ is detailed in Definition 6.4.

A major part of this work is devoted to the analysis of the arc poset. As in the case of $\mathcal{PMP}(\Sigma, f)$, we can associate a simplicial complex to $\mathcal{AP}(\Sigma, f)$, which we denote $|\mathcal{AP}(\Sigma, f)|$. It is clear that the mapping class group $\mathrm{MCG}(\Sigma)$ acts on arc systems, and we show it preserves the order we defined, so we obtain an action on the poset $\mathcal{AP}(\Sigma, f)$ (part of Theorem 6.8). To establish our results we show the following properties of $\mathcal{AP}(\Sigma, f)$ and of the action of $\mathrm{MCG}(\Sigma)$ on it:

1. Theorem 6.8: the infinite simplicial complex $|\mathcal{AP}(\Sigma, f)|$ is a topological covering space of $|\mathcal{PMP}(\Sigma, f)|$. Moreover, the action $\mathrm{MCG}(\Sigma) \curvearrowright \mathcal{AP}(\Sigma, f)$ extends to a covering space action $\mathrm{MCG}(\Sigma) \curvearrowright |\mathcal{AP}(\Sigma, f)|$ and

$$|\mathcal{AP}(\Sigma,f)|/\mathrm{MCG}(\Sigma) \cong |\mathcal{PMP}(\Sigma, f)|$$

is an isomorphism of simplicial complexes.

2. Theorem 6.12 (first part): there is a one-to-one correspondence between connected components in $|\mathcal{AP}(\Sigma, f)|$ and homotopy classes of functions in $[(\Sigma, f)]$.

3. Theorem 6.12 (second part): every connected component of $|\mathcal{AP}(\Sigma, f)|$ is contractible.

The first and last item show that every connected component of $|\mathcal{AP}(\Sigma, f)|$ is a universal covering space for $|\mathcal{PMP}(\Sigma, f)|$. The second item then shows that the fundamental group of $|\mathcal{PMP}(\Sigma, f)|$ is isomorphic to $\mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$.

The proof of contractability of the connected components of $\mathcal{AP}(\Sigma, f)$, the content of Theorem 6.12, requires the most technical proof of this paper, and we devote to it Section 6.3. The proof consists of a series of (countably many) deformation retracts which we define for each component of $|\mathcal{AP}(\Sigma, f)|$. This eventually shows that every component contracts to a point. Each step is described by a poset morphism which, by the content of Appendix A.1, corresponds to a deformation retract on the associated simplicial complex.

Remark 1.18. There is another *arc complex* that is similar to Penner's fat graph complex but with fewer constraints on the arcs: in particular, without the constraint that the arcs cut the surface into discs. In [Hat91], Hatcher extends earlier work of Harer [Har85] to prove under certain conditions that this arc complex is contractible by a direct combinatorial argument, in contrast to the proof of the contractability of the fat graph complex via Teichmüller theory. This direct argument, while less involved than our argument, is similar in flavor. We also point out that while at the level of objects our arc poset is related to Hatcher's arc complex from [Hat91], the topological claims we make are quite different. Indeed, a k-simplex in $|\mathcal{AP}(\Sigma, f)|$ is a chain of arc systems all with the same number of arcs, whereas in Hatcher's arc complex a k-simplex is a series of arc systems which are obtained by a series of arc deletions.

Remaining sections in the paper

Except for the sections mentioned above, the rest of the paper is organized as follows. In Section 2 we give some background for the ideas and tools in this paper: some basic facts about commutator length of words, Culler's construction and the correspondence between algebraic and geometric objects (Section 2.1); some comments and open questions regarding word measures on groups (Section 2.2), and some words about stable commutator length and the proof of Corollary 1.14 (Section 2.3).

After the core of the paper in Sections 3-6, Section 7 elaborates some further results derived from our analysis, especially regarding properties of the stabilizers from Theorems 1.2 and 1.10, and Section 8 contains some detailed examples. These are followed by some related open questions in Section 9 and a glossary of notation. The appendix contains some technical, mostly known, lemmas regarding posets and complexes. These are used in the proofs of Theorems 6.8 and 6.12.

1.5 Notations

For the convenience of the readers, there is a glossary on Page 63 listing most of the notations we use and where each one is defined. We also mention here some of the notation we will use. We use $\partial\Sigma$ to denote the boundary of the surface Σ. The word measures are coming from words in $\mathbf{F}_r$, and we denote the generators by $x_1, \ldots, x_r$. However, in examples we sometimes use x, y, z, t instead. We may use capital letters for inverses and occasionally enumerate the letters by their location in w. For example, we may write $w = [x, y]^2$ as $x_1 y_2 X_3 Y_4 x_5 y_6 X_7 Y_8$. We use $a_1, b_1, \ldots, a_g, b_g$ and their capital forms to write elements in the fundamental groups of surfaces.

Standard asymptotic notation is used to describe some of our results. This includes the big O notation "$f(n) = O(g(n))$" meaning that the functions f and g satisfy that for large enough n, $f(n) \leq C \cdot g(n)$ for some constant $C > 0$. Likewise, "$f(n) = o(g(n))$" means that for large enough n, $g(n) \neq 0$ and that $\frac{f(n)}{g(n)} \underset{n\to\infty}{\to} 0$. Finally, "$f(n) = \theta(g(n))$" means that for large enough n, $C_1 \cdot g(n) \leq f(n) \leq C_2 \cdot g(n)$ for some constants $C_1, C_2 > 0$.

2 Background

2.1 The geometric approach to commutator length

In this subsection we explain Culler's geometric interpretation of commutator length which yields that Theorem 1.1 is indeed a special case of Theorem 1.8. We also explain the other parallels mentioned in Section 1.2 and Table 1 between algebraic notions and geometric ones. In particular, we formulate the Dehn-Nielsen-Baer Theorem showing that $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ is isomorphic to the mapping class group of $\Sigma_{g,1}$, the genus g one boundary component orientable surface. This yields that Theorem 1.2 is a special case of Theorem 1.10.

We begin with an easy but useful characterization of maps from surfaces that are homotopic relative the boundary.

Lemma 2.1. *Let Σ be any orientable surface with ℓ boundary components and ℓ marked points as in Definition 1.5, and let $\gamma_1, \ldots, \gamma_t$ be a set of disjoint oriented arcs with endpoints in $v_1, \ldots, v_\ell$ which "fill Σ", i.e., which cut Σ into discs. Then two maps $f_1, f_2 \colon \Sigma \to \bigvee^r S^1$ which coincide on $\partial\Sigma$ and send all marked points $v_1, \ldots, v_\ell$ to the basepoint o are homotopic relative the boundary if and only if $[f_1(\gamma_j)] = [f_2(\gamma_j)]$ for all $j \in [t]$.*

This is also equivalent to f_1 and f_2 inducing the same map from the "fundamental groupoid" of Σ as a space with several marked points to $\pi_1\left(\bigvee^r S^1, o\right)$.

Proof. It f_1 and f_2 are homotopic and γ is any oriented arc from v_i to v_j, one can push forward this homotopy to show homotopy between $f_1(\gamma)$ and $f_2(\gamma)$, hence $[f_1(\gamma)] = [f_2(\gamma)]$. Conversely, assume that f_1 and f_2 satisfy the property with the arcs. We can then perturb f_2 so that it agrees with f_1 on these arcs (without changing the homotopy class of f_2). Then, on every disc D, f_1 and f_2 agree on the boundary, and it is enough to show that $f_1\Big|_D \simeq f_2\Big|_D$ are homotopic relative ∂D. Now f_1 and f_2 can be lifted to maps $\hat{f}_1, \hat{f}_2 \colon D \to \mathbb{T}_{2r}$ which coincide on ∂D, where the $2r$-regular tree $\mathbb{T}_{2r}$ is the universal covering space of $\bigvee^r S^1$. It is easy to see that $\hat{f}_1$ and $\hat{f}_2$ are homotopic: for every $x \in D$, let the homotopy move in a constant pace from $\hat{f}_1(x)$ to $\hat{f}_2(x)$ in $\mathbb{T}_{2r}$ along the sole geodesic between them. This homotopy can then be projected to a homotopy between f_1 and f_2. ◻

To give a precise formulation of the geometric analogue for $\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$, we first fix some more notation. Identify each circle in the wedge $\bigvee^r S^1$ with a distinct generator x_i of $\mathbf{F}_r$, orient each of the circles, and use these labeling and orientation to fix an isomorphism

$$\mathbf{F}_r \cong \pi_1\left(\bigvee^r S^1, o\right). \tag{2.1}$$

Recall that for every $w \in \mathbf{F}_r$, the map $f_w\colon \left(S^1, 1\right) \to \left(\bigvee^r S^1, o\right)$ is a fixed representative of w. More concretely,

Definition 2.2. For $1 \neq w \in \mathbf{F}_r$, let $f_w\colon \left(S^1, 1\right) \to \left(\bigvee^r S^1, o\right)$ be the sole non-backtracking closed path at o representing w (moving at arbitrary positive speed), so that $f_{w^{-1}}(z) = f_w(\overline{z})$. For $w = 1$ fix f_1 to be the constant map to o.

To fix the isomorphism of $\mathbf{F}_{2g}$ with $\pi_1(\Sigma_{g,1}, v_1)$, we need to fix $2g$ disjoint oriented arcs $\alpha_1, \beta_1, \ldots, \alpha_g, \beta_g$ in $\Sigma_{g,1}$ with endpoints in v_1 that serve as representatives for the basis $a_1, b_1, \ldots, a_g, b_g$ of $\mathbf{F}_{2g}$. We do this using the construction of $\Sigma_{g,1}$ from a $(4g+1)$-gon, as in Figure 2.1: we identify $\alpha_1, \ldots, \beta_g$ with the sides of this $(4g+1)$-gon that are being glued. Then there is an isomorphism

$$\mathbf{F}_{2g} \cong \pi_1(\Sigma_{g,1}, v_1) \tag{2.2}$$

mapping a_i to $[\alpha_i]$, b_i to $[\beta_i]$, and δ_g to $[\partial_1]$.

Recall from Section 1 that the commutator length of a word $w \in [\mathbf{F}_r, \mathbf{F}_r]$, denoted $\mathrm{cl}(w)$, is the smallest g such that there exist $u_1, v_1, \ldots, u_g, v_g \in \mathbf{F}_r$ with $[u_1, v_1] \ldots [u_g, v_g] = w$. Equivalently, $\mathrm{cl}(w)$ is the smallest g for which

$$\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r) = \{\phi \in \mathrm{Hom}(\mathbf{F}_{2g}, \mathbf{F}_r) \mid \phi(\delta_g) = w\}$$

is non-empty. The following proposition, basically due to [Cul81], explains why $\mathrm{cl}(w)$ is often called "the genus of w", and why Theorem 1.1 is a special case of Theorem 1.8.

Proposition 2.3. *Let $w \in [\mathbf{F}_r, \mathbf{F}_r]$ be a balanced word and $g \geq 0$ a non-negative integer. With the fixed isomorphisms (2.1) and (2.2), there is a one-to-one correspondence*

$$\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r) \quad \longleftrightarrow \quad \left\{ \begin{array}{c} \text{Homotopy classes (relative } \partial\Sigma_{g,1}) \text{ of maps } f\colon \Sigma_{g,1} \to \bigvee^r S^1 \\ \text{such that } \left(\Sigma_{g,1}, f\right) \text{ is admissible for } w \end{array} \right\}. \tag{2.3}$$

In particular, $\mathrm{cl}(w)$ is equal to the smallest genus g of a surface $\Sigma_{g,1}$ with an admissible map for w.

We note there are correspondences of the same spirit for maps admissible for several words.

Proof. It is clear that if $(\Sigma_{g,1}, f)$ is admissible for w, then $f_* \in \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$, and f_* only depends on the homotopy class of f. Conversely, given $\phi \in \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$, define $f : \Sigma_{g,1} \to \bigvee^r S^1$ as following. First, define $f\Big|_{\partial\Sigma_{g,1}}$ so that $f \circ \partial_1 = f_w$. For every $i \in [g]$ define $f\Big|_{\alpha_i}$ so that $f \circ \alpha_i = f_{\phi(a_i)}$ and $f\Big|_{\beta_i}$ so that $f \circ \beta_i = f_{\phi(b_i)}$. The arcs $\alpha_1, \ldots, \beta_g$ cut $\Sigma_{g,1}$ to a single polygon P, identical to the $(4g+1)$-gon used to construct $\Sigma_{g,1}$ - see Figure 2.1. It therefore remains to define f on the interior of P.

By the assumption on ϕ, the boundary ∂P is mapped by f to the trivial element of $\pi_1\left(\bigvee^r S^1, o\right)$.

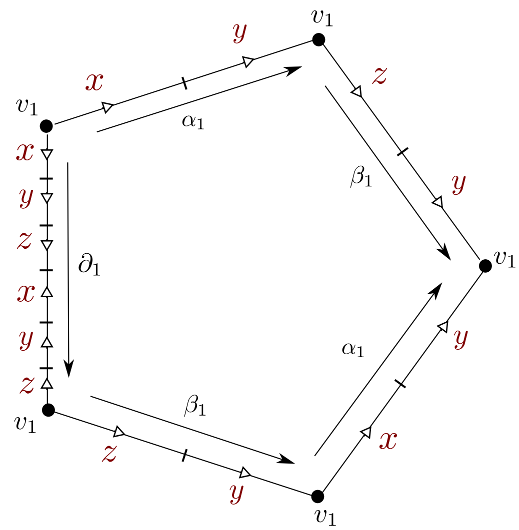

Figure 2.1: The word $w = xyzXYZ$ has commutator length 1 as shown by the solution $w = [xy, zy]$. To construct a corresponding map f from $\Sigma_{1,1}$, we first define f on the boundary of a 5-gon P described in the definition of the arcs $\alpha_1, \beta_1, \ldots, \alpha_g, \beta_g$. The letters x, y, z describe the image of $f\big|_{\partial P}$ in $\bigvee^r S^1$.

So there is a homotopy $T \colon \partial P \times [0,1] \to \bigvee^r S^1$ such that $T(x,0) \equiv f\big|_{\partial P}$ and $T(x,1)$ is constantly o. This map induces, therefore, a continuous map $\overline{T} \colon {}^{\partial P \times [0,1]}/_{(x,1) \sim (y,1)} \to \bigvee^r S^1$. Since ${}^{\partial P \times [0,1]}/_{(x,1) \sim (y,1)}$ is homeomorphic to P in a way that identifies $(x,0)$ with x, we can use $\overline{T}$ to get the required map f on all of P. Lemma 2.1 shows that the homotopy class of f is well defined. It is also clear that $f_* = \phi$. $\square$

Let us also mention a few facts about commutator length in free groups. As mentioned in Section 1.3, there are several algorithms for computing the commutator length of a given word $w \in [\mathbf{F}_r, \mathbf{F}_r]$. One of this algorithms, due to Culler, follows from our discussion in Section 4 below — see Remark 4.12. We also remark that the values taken by cl on $[\mathbf{F}_r, \mathbf{F}_r]$ $(r \geq 2)$ are all positive integers. An illuminating example is given in [Cul81, Section 2.6]:

$$\mathrm{cl}\left([x,y]^n\right) = \left\lfloor \frac{n}{2} \right\rfloor + 1.$$

For instance, $[x,y]^3 = \left[xyX, YxyX^2\right]\left[Yxy, y^2\right]$. Moreover, in the same paper Culler shows that for every $1 \neq w \in [\mathbf{F}_r, \mathbf{F}_r]$, $\mathrm{cl}\left(w^n\right) \underset{n \to \infty}{\to} \infty$. A tight lower bound $\mathrm{cl}\left(w^n\right) > \frac{n}{2}$ is given in [Cal09a, Theorem 4.111].

Finally, let us explain the last two lines of Table 1, showing that Theorems 1.2 and 1.4 are special cases of Theorems 1.10 and 1.12, respectively. Recall that $\mathrm{Aut}_\delta\left(\mathbf{F}_{2g}\right)$ is the subgroup of $\mathrm{Aut}\left(\mathbf{F}_{2g}\right)$ fixing $\delta_g = [a_1, b_1] \ldots [a_g, b_g]$. Via the isomorphism (2.2), we can view $\mathrm{Aut}_\delta\left(\mathbf{F}_{2g}\right)$ as the group of automorphisms of $\pi_1\left(\Sigma_{g,1}, v_1\right)$ fixing the element $[\partial_1]$.

Theorem 2.4 (Dehn-Nielsen-Baer)**.** *The map* $\theta \colon \mathrm{MCG}\left(\Sigma_{g,1}\right) \to \mathrm{Aut}_\delta\left(\mathbf{F}_{2g}\right)$ *defined by*

$$[\rho] \mapsto \rho_*$$

is an isomorphism.

A reference for the Dehn-Nielsen-Baer Theorem, including some historical notes, can be found in [FM12, Chapter 8]. However, the version that appears in [FM12] and usually found in the literature is slightly different and deals either with surfaces without boundary or with homeomorphisms of surfaces with boundary that do not necessarily fix the boundary. As we could not find any published reference for the exact version we need here, let us say a few words about the proof of Theorem 2.4.

That θ is a well-defined homomorphism of groups is trivial. The surjectivity of θ is a special case of [ZVC80, Theorem 5.7.1]. Finally, the injectivity of θ follows from the fact that $\Sigma_{g,1}$ is a $\mathrm{K}(\mathbf{F}_{2g},1)$-complex: indeed, $\Sigma_{g,1}$ is a $\mathrm{K}(\mathbf{F}_{2g},1)$-space (for example, because it deformation-retracts to a bouquet with $2g$ loops), which can be given a CW-complex structure. A basic feature of every $\mathrm{K}(G,1)$-complex Y is that any homomorphism $\pi_1(Y,y_0)\to\pi_1(Y,y_o)$ is induced by some map $(Y,y_0)\to(Y,y_0)$, which is unique up to homotopy fixing y_0 (e.g. [Hat02, Theorem 1B.9]). Since on surfaces homotopy of homeomorphisms is the same as isotopy ([FM12, Theorem 1.12]), we see that $\theta^{-1}(\mathrm{id})$ is precisely $\mathrm{Homeo}_0(\Sigma_{g,1})$.

Another remark worth mentioning is that the group $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ is torsion-free (e.g., [FM12, Corollary 7.3]), and thus so are the stabilizer subgroups in Theorem 1.4. This shows that a finite $\mathrm{K}(\mathrm{G},1)$-complex is plausible.

Finally, note that if $[\rho]\in\mathrm{MCG}(\Sigma_{g,1})$ and $f\colon(\Sigma_{g,1},v_1)\to\left(\bigvee^r S^1,o\right)$ is admissible for w, then the action of $[\rho]$ on $\tilde{f}$, the homotopy class of f, is given by $\widetilde{\rho\circ f}$. On the other hand, the action of $\rho_*\in\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ on $f_*\in\mathrm{Hom}_w(\mathbf{F}_{2g},\mathbf{F}_r)$ is given by $\rho_*\circ f_*=(\rho\circ f)_*$. This shows that the action of $\mathrm{MCG}(\Sigma_{g,1})$ on the homotopy classes in (2.3) is isomorphic to the action of $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ on $\mathrm{Hom}_w(\mathbf{F}_{2g},\mathbf{F}_r)$.

2.2 Word measures on compact groups

Let G be a compact group. As explained in Section 1.1, every word $w\in\mathbf{F}_r$ induces a measure on G, which we call the w-measure and denote in this subsection μ_G^w. This is the measure obtained by pushing forward the Haar measure on $G^r=\underbrace{G\times\ldots\times G}_{r\text{ times}}$ through the word map $w\colon G^r\to G$. μ_G^w

Namely, to sample an element from the w-measure on G, simply sample r independent elements $g_1,\ldots,g_r$ according to the Haar measure on G, and evaluate $w(g_1,\ldots,g_r)$. An important special case is when G is finite and then the Haar measure is simply the uniform distribution.

The following invariance of word measure motivates the theme that w-measures on groups encode algebraic information about w:

Fact 2.5. *Word measures are invariant under* $\mathrm{Aut}(\mathbf{F}_r)$. *Namely, if* $w\in\mathbf{F}_r$ *and* $\phi\in\mathrm{Aut}(\mathbf{F}_r)$, *then* w *and* $\phi(w)$ *induce the same measure on every compact group.*

Proof. Recall we denote the generators of $\mathbf{F}_r$ by $x_1,\ldots,x_r$. The automorphism group $\mathrm{Aut}(\mathbf{F}_r)$ is generated by the following "elementary Nielsen transformations" defined on the generators (e.g. [LS77, Section I.4]):

- The automorphism α_σ defined by a permutation $\sigma\in S_r$ on the generators
- The automorphism β defined by $x_1\mapsto x_1x_2$ and $x_i\mapsto x_i$ for $i\geq 2$
- The automorphism γ defined by $x_1\mapsto x_1^{-1}$ and $x_i\mapsto x_i$ for $i\geq 2$

Thus it is enough to show the word measures of a compact group G are invariant under these transformations. This is obvious for the automorphisms α_σ. For β, it is enough to show that if $g_1,g_2,\ldots,g_r\in G$ are r independent Haar random elements, then so are $g_1g_2,g_2,\ldots,g_r$. This is true by right-invariance of the Haar measure on compact groups: sample g_2 first. When sampling g_1, the measure on g_1g_2 is again the Haar measure. It also shows that g_1g_2 is independent of g_2. As for automorphism γ, given $g_1,\ldots,g_r$ as before, the independence of $g_1^{-1},g_2,\ldots,g_r$ is obvious. The transformation $g\mapsto g^{-1}$ turns a left Haar measure into a right one, but these two are the same in compact groups.

□

So two words in the same $\mathrm{Aut}(\mathbf{F}_r)$-orbit in $\mathbf{F}_r$ induce the same measure on every compact group. But is this the only reason for two words to have such a strong connection? A version of the following conjecture appears, for example, in [AV11, Question 2.2] and in [Sha13, Conjecture 4.2].

Conjecture 2.6. *If two words $w_1, w_2 \in \mathbf{F}_r$ induce the same measure on every compact group, then there exists $\phi \in \mathrm{Aut}(\mathbf{F}_r)$ with $w_2 = \phi(w_1)$.*

A special case of this conjecture, which attracted the attention of several researchers, deals with the $\mathrm{Aut}(\mathbf{F}_r)$-orbit of the single-letter word x_1, namely, with the set of primitive words. It was asked whether words inducing the Haar measure on every compact group are necessarily primitive. As mentioned in Section 1.3, this was settled in [PP15, Theorem 1.1] using word measures on symmetric groups:

Theorem 2.7 (Puder-Parzanchevski)**.** *A word inducing uniform measure on every finite group is necessarily primitive.*

Still, even in this special case, open problems remain: for example, can the symmetric groups be replaced in this result by, say, solvable groups? or compact Lie groups? Is there a single compact Lie group which suffices? We see our work here as a step towards answering these questions and, especially, Conjecture 2.6.

To the very least, we hope to be able to show that only primitive words induce the Haar measure on $\mathcal{U}(n)$ for every n. To date, we can use the current work to show that two words w_1 and w_2 with $\mathrm{scl}(w_1) \neq \mathrm{scl}(w_2)$ induce different measures on $\mathcal{U}(n)$ for every large enough n – see Section 2.3.

The first result in our paper deals with $\mathcal{T}r_w(n)$, the expected trace of a random matrix in $\mathcal{U}(n)$ sampled by the w-measure. Let us explain why this particular projection of the w-measure μ_G^w is a very natural first step.

Fact 2.8. *For any compact group G, the word measure μ_G^w is determined by the expected values of the irreducible characters $\left\{\int_{g\in G} \xi(g)\, d\mu_G^w(g)\right\}_{\xi\in\widehat{G}}$.*

Here $\widehat{G}$ marks the set of all irreducible characters of G.

Proof. The statement of the proposition holds for every conjugation-invariant measure. First we show why μ_G^w has this property, and then why this property yields the statement of the proposition. We ought to show that for every $g \in G$ and every measurable set $A \subseteq G$, we have $\mu_G^w(A) = \mu_G^w(gAg^{-1})$. This follows from the invariance of Haar measures under conjugation and the equality

$$w^{-1}\left(gAg^{-1}\right) = g\left(w^{-1}(A)\right)g^{-1},$$

the conjugation on the right hand side being the diagonal conjugation on G^r.

To see that a conjugation-invariant measure μ on a compact group G is completely determined by the expectation of irreducible characters[21], consider any μ-measurable function $f\colon G \to \mathbb{C}$ with finite expectation. Then, by conjugation-invariance, for every $h \in G$,

$$\int_G f(g)\, d\mu(g) = \int_G f\left(hgh^{-1}\right) d\mu(g).$$

[21] For finite groups, this follows by viewing the measure as a function and the fact that the irreducible characters form a basis for class functions.

Thus,

$$\int_{g\in G} f(g)\,d\mu(g) = \int_{h\in G}\left[\int_{g\in G} f\left(hgh^{-1}\right)d\mu(g)\right]d\mu(h) = \int_{g\in G}\left[\int_{h\in G} f\left(hgh^{-1}\right)d\mu(h)\right]d\mu(g),$$

where we used Fubini's theorem. Defining the class function $\overline{f}(g) = \int_{h\in G} f\left(hgh^{-1}\right)d\mu(h)$, we obtain, as $\overline{f} = \sum_{\xi\in\hat{G}}\left\langle \overline{f},\xi\right\rangle \xi$, that

$$\int_{g\in G} f(g)\,d\mu(g) = \int_{g\in G}\overline{f}(g)\,d\mu(g) = \sum_{\xi\in\hat{G}}\left\langle\overline{f},\xi\right\rangle\cdot\int_{g\in G}\xi(g)\,d\mu(g).$$

□

Thus it makes sense to study word measures via the expectation of irreducible characters. In this language, for example, Conjecture 2.6 says that if w_1 and w_2 do not belong to the same $\mathrm{Aut}(\mathbf{F}_r)$-orbit, then there is some compact group G and some non-trivial character $1\neq\xi\in\widehat{G}$ so that ξ has different expectations under $\mu_G^{w_1}$ and $\mu_G^{w_2}$. In the case of $\mathcal{U}(n)$, it is fair to say the simplest irreducible character is the trace of the standard representation, and its expected value under the w-measure $\mu^w_{\mathcal{U}(n)}$ is, by definition, $\mathcal{T}r_w(n)$.

Remark 2.9. As hinted in Section 1.2, our more general results regarding $\mathcal{T}r_{w_1,\ldots,w_\ell}$ and finite sets of words give much more information about word measures in $\mathcal{U}(n)$. In particular, they give similar kind of control we get over $\mathcal{T}r_w(n)$ for many other irreducible characters of $\mathcal{U}(n)$. For example, consider the irreducible character of $\mathcal{U}(n)$ which corresponds to the highest weight vector $(2,0,\ldots,0,-1)$. It is given by

$$\xi(A) = \frac{\mathrm{tr}\left(A^2\right)+\mathrm{tr}^2(A)}{2}\cdot\mathrm{tr}\left(A^{-1}\right)-\mathrm{tr}(A).$$

So the expected value of ξ in the measure $\mu^w_{\mathcal{U}(n)}$ is

$$\mathbb{E}_{\mu^w_{\mathcal{U}(n)}}[\xi] = \frac{1}{2}\mathcal{T}r_{w^2,w^{-1}}(n)+\frac{1}{2}\mathcal{T}r_{w,w,w^{-1}}(n)-\mathcal{T}r_w(n),$$

and Theorem 1.10 gives information about the leading term of this expression. The same is true for any "non-balanced" irreducible character: a character the corresponding highest weight vector of which sums to zero, or equivalently, a character which is not invariant under multiplication by central elements of $\mathcal{U}(n)$. In contrast, the character corresponding to $(1,0,\ldots,0,-1)$, which is given by

$$|\mathrm{tr}(A)|^2-1,$$

is balanced, and its expected value under $\mu^w_{\mathcal{U}(n)}$ is

$$\mathcal{T}r_{w,w^{-1}}(n)-1. \tag{2.4}$$

Because of the free term "-1" in (2.4), Theorem 1.10 gives weaker information about the leading coefficient of (2.4), and only determines the limit of the character as $n\to\infty$, rather than its leading term.

Finally, let us remark that many works in the area of word measures focus on questions of slightly different flavor: the word measures induced by a fixed word across all finite/compact groups; the support of word measures; the probability, in word measures on finite groups, of the identity, etc. A survey containing many references is [Sha13].

2.3 Stable commutator length

Recall that Corollary 1.14 states that the w-measures on $\{\mathcal{U}(n)\}_{n\in\mathbb{N}}$ determine $\mathrm{scl}(w)$, the stable commutator length of $w \in \mathbf{F}_r$ (see (1.8)). In this subsection we explain how this result follows from Theorem 1.10 and from Calegari's rationality theorem.

Calegari's theorem, which is the main result of [Cal09b], says that $\mathrm{scl}(w)$ is rational for every $w \in [\mathbf{F}_r, \mathbf{F}_r]$. The proof goes through showing the existence of "extremal surfaces" for w: an extremal surface for w is an admissible (Σ, f) for some set of powers of w, say $w^{j_1}, \ldots, w^{j_\ell}$ with $j_1, \ldots, j_\ell \in \mathbb{Z}$, so that $\frac{-\chi(\Sigma)}{2(j_1+\ldots+j_\ell)}$ achieves the infimum of the values of its kind. This infimum is $\mathrm{scl}(w)$, the stable commutator length of w [Cal09b, Lemma 2.6].

The main theorem of [Cal09b] states that if $w \in [\mathbf{F}_r, \mathbf{F}_r]$ then w admits an extremal surface (Σ, f). Moreover, by [Cal09b, Lemma 2.7], this extremal surface can be taken to be admissible for $w^{j_1}, \ldots, w^{j_\ell}$ with $j_1, \ldots, j_\ell > 0$. By definition of extremal surface, Σ has maximal Euler characteristic for $w^{j_1}, \ldots, w^{j_\ell}$, namely, $\chi(\Sigma) = \mathrm{chi}\left(w^{j_1}, \ldots, w^{j_\ell}\right)$. Moreover, every surface which is admissible for $w^{j_1}, \ldots, w^{j_\ell}$ with Euler characteristic $\mathrm{chi}\left(w^{j_1}, \ldots, w^{j_\ell}\right)$ is extremal. By [Cal09b, Lemma 2.9], the maps associated with extremal surfaces are π_1-injective, namely, if γ is a non-nullhomotopic closed curve, then $f(\gamma)$ is not nullhomotopic. Note that this condition is stronger than incompressibility, which only deals with *simple* closed curves. The crux of the matter is the following lemma, a special case of which is discussed in Remark 1.3:

Lemma 2.10. *If (Σ, f) is π_1-injective, then $\mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$ is trivial.*

Proof. Let $[\rho] \in \mathrm{MCG}(\Sigma)$ fix $\tilde{f}$, so $f \circ \rho \simeq f$ are homotopic. Let $\gamma_2, \ldots, \gamma_\ell$ be a set of $\ell - 1$ disjoint arcs in Σ, where γ_i leads from v_1 to v_i. The arc $\rho(\gamma_2)$ is homotopic to the concatenation $\beta * \gamma_2$ where β is a closed loop at v_1, but

$$[f(\gamma_2)] = [f(\rho(\gamma_2))] = [f(\beta * \gamma_2)] = [f(\beta)] \cdot [f(\gamma_2)]$$

and so $[f(\beta)] = 1$ and by π_1-injectivity, $[\beta] = 1 \in \pi_1(\Sigma, v_1)$. Hence $\rho(\gamma_2) \simeq \gamma_2$ and we may perturb ρ so that it fixes γ_2. We can do the same for γ_3 without modifying $\rho|_{\gamma_2}$ and so on, until ρ fixes $\gamma_2 \cup \ldots \cup \gamma_\ell$ pointwise. Now we can cut Σ along $\gamma_2, \ldots, \gamma_\ell$ and get a surface Σ' with one boundary component, a map $f'\colon \Sigma' \to \bigvee^r S^1$ and an induced homeomorphism ρ' which fixes $\partial\Sigma'$ pointwise and such that $f' \circ \rho' \simeq \rho'$. By Theorem 2.4, $[\rho']$ corresponds to some $\phi \in \mathrm{Aut}_\delta(\pi_1(\Sigma', v_1))$. As f' is still π_1-injective, we see that $(f')_*$ cannot be fixed by any non-trivial element of $\mathrm{Aut}(\pi_1(\Sigma', v_1))$, let alone of $\mathrm{Aut}_\delta(\pi_1(\Sigma', v_1))$, hence $\phi = 1$ and $[\rho'] = [\mathrm{id}]$. Thus $[\rho] = 1$. □

We infer that if one of the extremal surfaces of w is admissible for $w^{j_1}, \ldots, w^{j_\ell}$ with $j_1, \ldots, j_\ell > 0$, then Theorem 1.10 translates in this case to

$$\mathcal{T}r_{w^{j_1},\ldots,w^{j_\ell}}(n) = n^{\mathrm{chi}\left(w^{j_1},\ldots,w^{j_\ell}\right)} \cdot \left|\mathrm{Solu}\left(w^{j_1}, \ldots, w^{j_\ell}\right)\right| + O\left(n^{\mathrm{chi}\left(w^{j_1},\ldots,w^{j_\ell}\right)-2}\right), \tag{2.5}$$

which is strictly positive for large enough n. Hence,

$$\frac{-\lim_{n\to\infty} \log_n \mathcal{T}r_{w^{j_1},\ldots,w^{j_\ell}}(n)}{2(j_1 + \ldots + j_\ell)} = \frac{-\mathrm{chi}\left(w^{j_1}, \ldots, w^{j_\ell}\right)}{2(j_1 + \ldots + j_\ell)} = \mathrm{scl}(w).$$

On the other hand, for an arbitrary $\ell > 0$ and $j_1, \ldots, j_\ell > 0$ we have

$$\frac{-\lim_{n\to\infty} \log_n \mathcal{T}r_{w^{j_1},\ldots,w^{j_\ell}}(n)}{2(j_1 + \ldots + j_\ell)} \geq \frac{-\mathrm{chi}\left(w^{j_1}, \ldots, w^{j_\ell}\right)}{2(j_1 + \ldots + j_\ell)} \geq \mathrm{scl}(w).$$

This proves (1.9) and Corollary 1.14. □

Corollary 2.11. *If* $\mathrm{scl}(w_1) \neq \mathrm{scl}(w_2)$ *then for every large enough* n*, the* w_1*-measure on* $\mathcal{U}(n)$ *is different from the* w_2*-measure on* $\mathcal{U}(n)$*. In particular, if* $w_1 \in [\mathbf{F}_r, \mathbf{F}_r]$ *and* $w_2 \notin [\mathbf{F}_r, \mathbf{F}_r]$ *then they induce different measures on* $\mathcal{U}(n)$ *for almost all* n*.*

Proof. Assume without loss of generality that $\mathrm{scl}(w_1) < \mathrm{scl}(w_2)$, and let $j_1, \ldots, j_\ell > 0$ be so that $w_1^{j_1}, \ldots, w_1^{j_\ell}$ admit an extremal surface. Then by the above discussion, $\mathcal{T}r_{w_1^{j_1},\ldots,w_1^{j_\ell}}(n)$ is strictly larger than $\mathcal{T}r_{w_2^{j_1},\ldots,w_2^{j_\ell}}(n)$ for any large enough n. In particular, if w_2 is not balanced, i.e. $w_2 \notin [\mathbf{F}_r, \mathbf{F}_r]$ and $\mathrm{scl}(w_2) = \infty$, then nor is the set $w_2^{j_1}, \ldots, w_2^{j_\ell}$ balanced as we assume $j_1, \ldots, j_\ell > 0$. By Claim 3.1, $\mathcal{T}r_{w_2^{j_1},\ldots,w_2^{j_\ell}}(n) \equiv 0$ for every n. □

3 A Rational Expression for $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$

In this section we prove that $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ is a rational function in n (Theorem 3.7). First, we prove the observation mentioned above regarding non-balanced sets of words:

Claim 3.1. *If* $w_1 w_2 \cdots w_\ell \in \mathbf{F}_r \setminus [\mathbf{F}_r, \mathbf{F}_r]$ *then* $\mathcal{T}r_{w_1,\ldots,w_\ell}(n) \equiv 0$*.*

Proof. By the assumption, there is some $j \in [r]$ so that α_j, the sum of exponents of the letter x_j in $w_1, \ldots, w_\ell$, satisfies $\alpha_j \neq 0$. Recall that the Haar measure of a compact group is invariant under left multiplication by any element. Since for $\theta \in [0, 2\pi]$, the diagonal central matrix $e^{i\theta} I_n$ is in $\mathcal{U}(n)$, we obtain

$$\begin{aligned}
&\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = \\
&= \mathbb{E}_{\mathcal{U}(n)^r}\left[tr\left(w_1\left(U_1^{(n)},\ldots,U_j^{(n)},\ldots,U_r^{(n)}\right)\right)\cdots tr\left(w_\ell\left(U_1^{(n)},\ldots,U_j^{(n)},\ldots,U_r^{(n)}\right)\right)\right] \\
&= \mathbb{E}_{\mathcal{U}(n)^r}\left[tr\left(w_1\left(U_1^{(n)},\ldots,e^{i\theta}U_j^{(n)},\ldots,U_r^{(n)}\right)\right)\cdots tr\left(w_\ell\left(U_1^{(n)},\ldots,e^{i\theta}U_j^{(n)},\ldots,U_r^{(n)}\right)\right)\right] \\
&= e^{i\theta\alpha_j} \cdot \mathcal{T}r_{w_1,\ldots,w_\ell}(n).
\end{aligned}$$

The claim follows as this equality holds for every $\theta \in [0, 2\pi]$. □

3.1 Weingarten function and integrals over $\mathcal{U}(n)$

The main tool used in this section is a formula, basically due to Xu [Xu97] and, more neatly, to Collins and Śniady [CŚ06], which expresses integrals with respect to $(\mathcal{U}(n), \mu_n)$. These integrals are expressed in terms of the *Weingarten* function, first studied in [Wei78] and formally defined and named in [Col03]. Let $\mathbb{Q}(n)$ denote the field of rational functions with rational coefficients in the variable n. Let S_L denote the symmetric group on L elements. The Weingarten function maps[22] S_L to $\mathbb{Q}(n)$ (for every L). We think of such functions as elements of the group ring $\mathbb{Q}(n)[S_L]$. S_L

Definition 3.2. The **Weingarten function** $\mathrm{Wg}: S_L \to \mathbb{Q}(n)$ is the inverse, in the group ring $\mathbb{Q}(n)[S_L]$, of the function $\sigma \mapsto n^{\#\mathrm{cycles}(\sigma)}$. Wg

That the function $\sigma \mapsto n^{\#\mathrm{cycles}(\sigma)}$ is invertible for every L follows from [CŚ06, Proposition 2.3] and the discussion following it. Clearly, Wg is constant on conjugacy classes. For example, for $L = 2$, the inverse of $\left(n^2 \cdot (1)(2) + n \cdot (12)\right) \in \mathbb{Q}(n)[S_2]$ is $\left(\frac{1}{n^2-1} \cdot (1)(2) - \frac{1}{n(n^2-1)} \cdot (12)\right)$,

[22] More precisely, it is a function from the disjoint union $\bigcup_{L=1}^{\infty} S_L$ to $\mathbb{Q}(n)$.

so $\mathrm{Wg}((1)(2)) = \frac{1}{n^2-1}$ while $\mathrm{Wg}((12)) = \frac{-1}{n(n^2-1)}$. For $L = 3$ the values of the Weingarten function are

$$\mathrm{Wg}((1)(2)(3)) = \frac{n^2-2}{n(n^2-1)(n^2-4)} \quad \mathrm{Wg}((12)(3)) = \frac{-1}{(n^2-1)(n^2-4)}$$
$$\mathrm{Wg}((123)) = \frac{2}{n(n^2-1)(n^2-4)}.$$

(We use here a non-standard cycle notation for permutations where we write fixed points as well. This is to stress the dependency of $\mathrm{Wg}(\sigma)$, for $\sigma \in S_L$, on L. E.g., $\mathrm{Wg}((12)) \neq \mathrm{Wg}((12)(3))$.)

Collins and Śniady also provide an explicit formula for Wg in terms of the irreducible characters of S_L and Schur polynomials [CŚ06, Equation (13)]: for $\sigma \in S_L$,

$$\mathrm{Wg}(\sigma) = \frac{1}{(L!)^2} \sum_{\lambda \vdash L} \frac{\chi_\lambda(e)^2}{d_\lambda(n)} \chi_\lambda(\sigma),$$

where λ runs over all partitions of L, χ_λ is the character of S_L corresponding to λ, and $d_\lambda(n)$ is the number of semistandard Young tableaux with shape λ, filled with numbers from $[n]$. A well known formula for $d_\lambda(n)$ states $d_\lambda(n) = \frac{\chi_\lambda(e)}{L!} \prod_{(i,j)\in\lambda} (n+j-i)$, where (i,j) are the coordinates of cells in the Young diagram with shape λ (e.g. [Ful97, Section 4.3, Equation (9)]). Thus,

Corollary 3.3. *For $\sigma \in S_L$, $\mathrm{Wg}(\sigma)$ may have poles only at integers n with $-L < n < L$.*

The key feature of Wg that we need is the value of its leading term. This is expressed in terms of a certain Möbius function which we now define. For every permutation $\sigma \in S_L$ denote by $\|\sigma\|$ its norm, defined as the length of the shortest product of transpositions giving σ. Equivalently, $\|\sigma\| = L - \#\mathrm{cycles}(\sigma)$. This norm can be used to define a poset structure on S_L: say that $\sigma \preceq \tau$ if and only if $\|\tau\| = \|\sigma\| + \|\sigma^{-1}\tau\|$. That is, $\sigma \preceq \tau$ if and only if there is a product of transpositions of minimal length giving τ, such that some prefix of this product is equal to σ. This poset is closely related to that of non-crossing partitions — see [NS06, Lecture 23]. $\|\sigma\|$ $\sigma \preceq \tau$

Every locally finite poset[23] gives rise to a Möbius function defined on comparable pairs of elements. This is defined to be the only function $\mu : \{(x,y) \mid x \preceq y\} \to \mathbb{Z}$ that satisfies

$$\sum_{z: x \leq z \leq y} \mu(x,z) = \delta_{x,y} \tag{3.1}$$

for every x, y in the poset with $x \preceq y$ (see [Sta12, Section 3.7]).

In the case of the poset $(S_L, \preceq)$, the corresponding Möbius function has a nice combinatorial description:

Proposition 3.4. *[CŚ06, Section 2.3] The Möbius function of the poset $(S_L, \preceq)$ is given by $\mu(\sigma,\tau) = \mathrm{M\ddot{o}b}(\sigma^{-1}\tau)$, where* $\mathrm{M\ddot{o}b}(\sigma)$

$$\mathrm{M\ddot{o}b}(\sigma) = \mathrm{sgn}(\sigma) \prod_{i=1}^{k} c_{|C_i|-1}, \tag{3.2}$$

with $C_1, \ldots, C_k$ the cycles composing σ, and

$$c_m = \frac{(2m)!}{m!(m+1)!}$$

the m-th Catalan number.

[23] A poset $(P, \leq)$ is said to be locally finite if for every $x \leq y$ in P, the interval $[x,y] = \{z \mid x \leq z \leq y\}$ is finite.

The content of Proposition 3.4 is that if $\sigma \preceq \tau$ in S_L, then

$$\sum_{\pi \in S_L \text{ s.t. } \sigma \preceq \pi \preceq \tau} \text{Möb}\left(\sigma^{-1}\pi\right) = \delta_{\sigma,\tau}.$$

Proposition 3.5. *[CŚ06, Corollary 2.7] Let $\sigma \in S_L$. The Weingarten function satisfies*

$$\text{Wg}\left(\sigma\right) = \frac{\text{Möb}\left(\sigma\right)}{n^{L+\|\sigma\|}} + O\left(\frac{1}{n^{L+\|\sigma\|+2}}\right).$$

Note the jump of 2 in the exponent after the subtraction of the leading term. In fact, this is shown to go on: in the Taylor expansion of $\text{Wg}\left(\sigma\right)$ in $\frac{1}{n}$, every other term vanishes [CŚ06, Proposition 2.6].

The formula of Collins and Śniady evaluates integrals of monomials in the entries $u_{i,j}$ and their conjugates $\overline{u_{i,j}}$ of a Haar distributed unitary matrix $u \in \mathcal{U}\left(n\right)$. The simple argument in the proof of Claim 3.1 shows that such an integral vanishes whenever the monomial is not balanced, namely whenever the number of $u_{i,j}$'s is different from the number of $\overline{u_{i,j}}$'s. The following formula deals with the interesting case, where the monomial is balanced:

Theorem 3.6. *[CŚ06, Proposition 2.5] Let m and n_0 be positive integers and $(i_1, \ldots, i_m)$, $(j_1, \ldots, j_m)$, $(i'_1, \ldots, i'_m)$ and $(j'_1, \ldots, j'_m)$ be m-tuples of indices in $[n_0]$. Then*

$$\int_{\mathcal{U}(n)} u_{i_1,j_1} u_{i_2,j_2} \ldots u_{i_m,j_m} \overline{u_{i'_1,j'_1} u_{i'_2,j'_2}} \ldots \overline{u_{i'_m,j'_m}} d\mu_n$$

is a rational function in n (valid for $n \geq n_0$), which is equal to

$$\sum_{\sigma,\tau \in S_m} \delta_{i_1 i'_{\sigma(1)}} \ldots \delta_{i_m i'_{\sigma(m)}} \delta_{j_1 j'_{\tau(1)}} \ldots \delta_{j_m j'_{\tau(m)}} \text{Wg}\left(\sigma^{-1}\tau\right). \tag{3.3}$$

Put differently, the rational function is given by $\sum_{\sigma,\tau} \text{Wg}\left(\sigma^{-1}\tau\right)$, where σ runs over all rearrangements of $(i'_1, \ldots, i'_m)$ which make it identical to $(i_1, \ldots, i_m)$, and τ runs over all rearrangements of $(j'_1, \ldots, j'_m)$ which make it identical to $(j_1, \ldots, j_m)$. In particular, the possible poles of the Weingarten function at n, for every $n \geq n_0$, are guaranteed to cancel out in this summation (see the example following Proposition 2.5 in [CŚ06]). We mention that a result of the type of Theorem 3.6, where integrals over $\mathcal{U}\left(n\right)$ are expressed as combinatorial formulas involving permutations, is possible thanks to the Schur-Weyl duality.

3.2 Word integrals over $\mathcal{U}\left(n\right)$

We use (3.3) to analyze

$$\mathcal{T}r_{w_1,\ldots,w_\ell}\left(n\right) = \int_{\mathcal{U}(n)\times\mathcal{U}(n)\times\ldots\times\mathcal{U}(n)} \text{tr}\left(w_1\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right) \cdots \text{tr}\left(w_\ell\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right) d\mu_n^r.$$

We explain our approach by way of an example. Let $w = [x,y]^2 = xyXYxyXY \in \mathbf{F}_2$. Then,

$$
\begin{aligned}
\mathcal{T}r_w(n) &= \int_{(A,B)\in\mathcal{U}(n)\times\mathcal{U}(n)} \operatorname{tr}\left(ABA^{-1}B^{-1}ABA^{-1}B^{-1}\right) d\mu_n^2 \\
&= \int_{(A,B)\in\mathcal{U}(n)\times\mathcal{U}(n)} \sum_{i,j,k,\ell,I,J,K,L\in[n]} A_{i,j}B_{j,k}A^{-1}_{k,\ell}B^{-1}_{\ell,I}A_{I,J}B_{J,K}A^{-1}_{K,L}B^{-1}_{L,i}d\mu_n^2 \qquad (3.4)\\
&= \sum_{i,j,k,\ell,I,J,K,L\in[n]} \int_{(A,B)\in\mathcal{U}(n)\times\mathcal{U}(n)} A_{i,j}B_{j,k}\overline{A_{\ell,k}B_{I,\ell}}A_{I,J}B_{J,K}\overline{A_{L,K}B_{i,L}}d\mu_n^2 \\
&= \sum_{i,j,k,\ell,I,J,K,L\in[n]} \left[\int_{A\in\mathcal{U}(n)} A_{i,j}A_{I,J}\overline{A_{\ell,k}A_{L,K}}d\mu_n\right]\cdot\left[\int_{B\in\mathcal{U}(n)} B_{j,k}B_{J,K}\overline{B_{I,\ell}B_{i,L}}d\mu_n\right].
\end{aligned}
$$

Now we use Theorem 3.6 to replace each of the two integrals inside the sum by a summation over pairs of permutations in S_2. For the first integral we go over all bijections $\sigma_a\colon \{i,I\} \xrightarrow{\sim} \{\ell,L\}$ and $\tau_a\colon \{j,J\} \xrightarrow{\sim} \{k,K\}$, and similarly over bijections σ_b and τ_b for the second integral. We think of these sets as ordered, so $\sigma_a = (12)$ means it maps $i \mapsto L$, $I \mapsto \ell$. We change the order of summation, and sum first over σ_a, τ_a, σ_b and τ_b, and only then over the indices $i,j,\ldots,L$. In fact, for every set of permutations, we only need to count the number of evaluations of $i,j,\ldots,L$ which "agree" with the permutations. For example, consider the case where

$$
\begin{array}{cccc}
\underline{\sigma_a = \mathrm{id}} & \underline{\tau_a = (12)} & \underline{\sigma_b = (12)} & \underline{\tau_b = (12)} \\
i \mapsto \ell & j \mapsto K & j \mapsto i & k \mapsto L \\
I \mapsto L & J \mapsto k & J \mapsto I & K \mapsto \ell
\end{array}.
$$

The summand corresponding to these permutations is

$$
\mathrm{Wg}\left((12)\right)\cdot \mathrm{Wg}\left((1)(2)\right)\cdot \sum_{i,j,k,\ell,I,J,K,L\in[n]} \delta_{i\ell}\delta_{IL}\delta_{jK}\delta_{Jk}\delta_{ji}\delta_{JI}\delta_{kL}\delta_{K\ell},
$$

and the product inside the last sum is 1 (and not 0) if and only if $i=\ell=K=j$ and $I=L=k=J$. So there are exactly n^2 such sets of indices and the total contribution of these particular 4 permutations is

$$
\mathrm{Wg}\left((12)\right)\cdot \mathrm{Wg}\left((1)(2)\right)\cdot n^2 = \frac{-1}{n\left(n^2-1\right)}\cdot\frac{1}{n^2-1}\cdot n^2 = \frac{-n}{\left(n^2-1\right)^2}.
$$

If we perform the same calculation for all 16 possible sets of permutations and sum the contributions, we obtain that

$$
\mathcal{T}r_{[x,y]^2}(n) = \frac{-4}{n^3-n}. \qquad (3.5)
$$

Of course, similar analysis works for any word $w \in [\mathbf{F}_r,\mathbf{F}_r]$ and any (balanced) finite sets $w_1,\ldots,w_\ell \in \mathbf{F}_r$. As $\{w_1,\ldots,w_\ell\}$ is balanced, the total length of the words is even, and we denote it by $2L \overset{\text{def}}{=} |w_1|+\ldots+|w_\ell|$. Let L_i denote the number of appearances of x_i in $w_1,\ldots,w_\ell$ (appearances with positive exponent $+1$), so $\sum_{i=1}^r L_i = L$. Let $\mathrm{BIJ}_i(w_1,\ldots,w_\ell)$ denote the set of bijections from the appearances of x_i^{+1} to those of x_i^{-1}, so $|\mathrm{BIJ}_i(w_1,\ldots,w_\ell)| = L_i!$. To compute $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$, we go over all $(2r)$-tuples of bijections $(\sigma_1,\tau_1,\ldots,\sigma_r,\tau_r)$, with $\sigma_i,\tau_i \in \mathrm{BIJ}_i(w_1,\ldots,w_\ell)$. Note that $\sigma_i^{-1}\tau_i$ can be thought of as a permutation of the appearances of x_i^{+1}, and so $\sigma_i^{-1}\tau_i$ belongs to a well-defined conjugacy class in S_{L_i}.

L, L_i

$\mathrm{BIJ}_{i(w_1,\ldots,w_\ell)}$

As in the example, each tuple induces a partition on a set of $|w_1| + \ldots + |w_\ell| = 2L$ indices, and we denote the number of blocks in this partition by $B(\sigma_1, \tau_1, \ldots, \sigma_r, \tau_r)$. The number of evaluations of the indices which agree with these bijections is $n^{B(\sigma_1,\ldots,\tau_r)}$. Hence, $B(\sigma_1,\ldots,\tau_r)$

Theorem 3.7. *In the notations of the previous paragraph, for every $n \geq \max_i L_i$,*

$$\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = \sum_{\sigma_1,\tau_1 \in \mathrm{BIJ}_1(w_1,\ldots,w_\ell),\ \ldots\ ,\sigma_r,\tau_r \in \mathrm{BIJ}_r(w_1,\ldots,w_\ell)} \mathrm{Wg}\left(\sigma_1^{-1}\tau_1\right) \ldots \mathrm{Wg}\left(\sigma_r^{-1}\tau_r\right) \cdot n^{B(\sigma_1,\tau_1,\ldots,\sigma_r,\tau_r)}. \tag{3.6}$$

In particular, for $n \geq \max_i L_i$, $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ is given by a rational function in n.

We have to restrict to $n \geq \max_i L_i$ because of possible poles of the Weingarten function[24] (Corollary 3.3). When this function has no poles, Theorem 3.6 guarantees that the expression we get gives the right answer.

4 Constructing Surfaces from Pairs of Matchings

In this section we associate a surface for every $2r$-tuple of bijections $(\sigma_1, \ldots, \tau_r)$ appearing in Theorem 3.7. This allows a better understanding of the summation (3.6) and the order of its terms, and leads to Theorem 1.1 about the leading exponent of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ (Corollary 4.13 below). Together with a suitable map, the surface we construct will be admissible for $w_1, \ldots, w_\ell$. As shown in Proposition 4.6 below, the order of the contribution of a $2r$-tuple of bijections in (3.6) is given by the Euler characteristic of its associated surface.

Notation-wise, instead of keeping track of $2r$ different bijections, it is more convenient to regard them as a pair of matchings between the letter with positive exponent in $w_1, \ldots, w_\ell$ and the letters with negative exponents. To formalize this, let L and L_i ($i \in [r]$) be as in Section 3 above (we keep restricting to the interesting case where $w_1, \ldots, w_\ell$ is a balanced set). Let E_i^+ be the set of appearances of x_i^{+1} and E_i^- be the set of appearances of x_i^{-1}, so that $\left|E_i^+\right| = \left|E_i^-\right| = L_i$. We also let $E^+ = \bigcup_i E_i^+$ and $E^- = \bigcup_i E_i^-$, so $|E^+| = |E^-| = L$. We then consider the set $\{\sigma_i : E_i^+ \xrightarrow{\sim} E_i^-\}_{i\in[r]}$ encoded in a single bijection $\sigma : E^+ \xrightarrow{\sim} E^-$. Likewise, we encode $\{\tau_i : E_i^+ \xrightarrow{\sim} E_i^-\}_{i\in[r]}$ in a single $\tau : E^+ \xrightarrow{\sim} E^-$. E_i^+, E_i^- E^+, E^-

Definition 4.1. Denote by $\mathrm{Match}(w_1, \ldots, w_\ell)$ the set of bijections $\sigma : E^+ \xrightarrow{\sim} E^-$ which are compatible with the colors of the edges. Namely, $\mathrm{Match}(w_1,\ldots,w_\ell)$

$$\mathrm{Match}(w_1, \ldots, w_\ell) = \left\{\sigma \colon E^+ \xrightarrow{\sim} E^- \;\middle|\; \sigma\left(E_i^+\right) = E_i^- \ \forall i \in [r]\right\}.$$

For a pair of matchings $(\sigma, \tau) \in \mathrm{Match}(w_1, \ldots, w_\ell)^2$ we let $B_{(\sigma,\tau)}$

$$B_{(\sigma,\tau)} = B(\sigma\Big|_{E_1^+}, \tau\Big|_{E_1^+}, \ldots, \sigma\Big|_{E_r^+}, \tau\Big|_{E_r^+})$$

denote the number of blocks in the partition of $2L$ indices induced by σ and τ.

Clearly, for $\sigma, \tau \in \mathrm{Match}(w_1, \ldots, w_\ell)$, $\sigma^{-1}\tau$ is a permutation of E^+ which only mixes edges with the same color, and belongs to a well-defined conjugacy class in S_L.

For every pair of matchings (σ, τ) we construct a surface as a CW-complex. We begin by the ℓ boundary components of the surface. These are, of course, merely ℓ pointed 1-spheres,

[24] Interestingly, very similar constraints on n appear in a formula for the trace of w in r uniform *permutation* matrices — see [Pud14, Section 5].

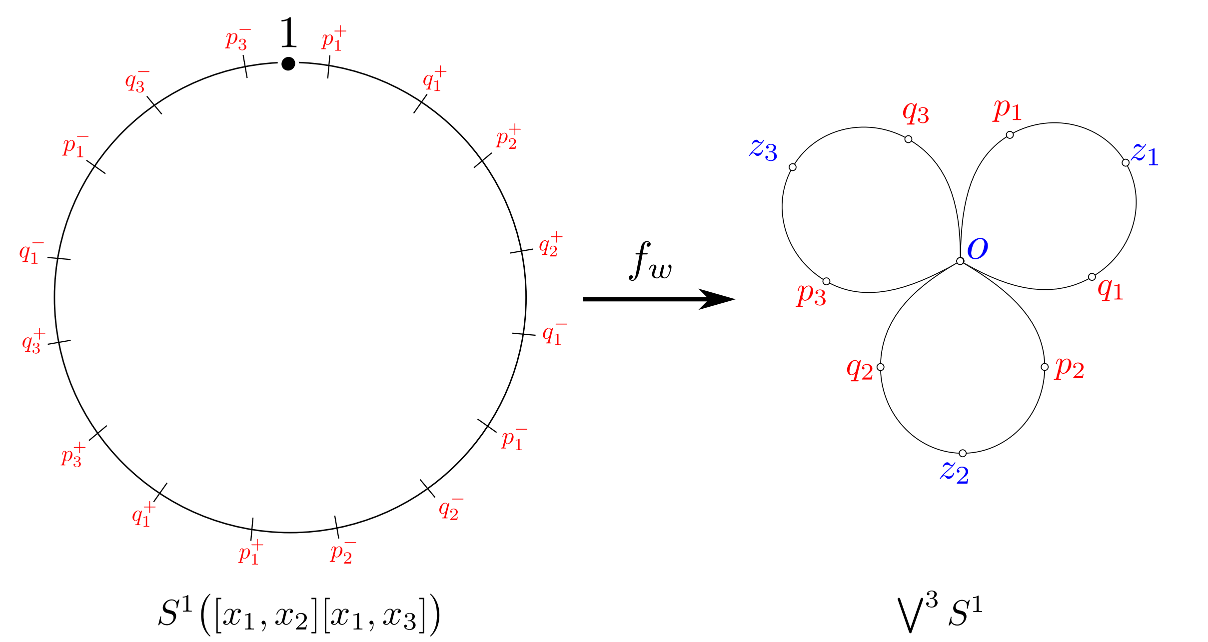

Figure 4.1: The marked 1-sphere $S^1(w)$ for $w=[x_1,x_2][x_1,x_3]\in\mathbf{F}_3$ together with the marked wedge $\bigvee^3 S^1$.

but we want to mark some additional points on each of them. For this sake, we first mark points on the wedge $\bigvee^r S^1$ (in addition to the basepoint o): on the circle corresponding to the generator x_i, we mark, in the order of the circle's orientation, distinct points p_i, z_i and q_i that[25] are also distinct from o – this is illustrated in the right hand side of Figure 4.1. p_i, q_i, z_i

Now, for every $w\in\mathbf{F}_r$, define $S^1(w)$ to be the pointed 1-sphere $(S^1,1)$ with additional $2|w|$ marked points with set of colors $\{p_i^+,p_i^-,q_i^+,q_i^-\mid i\in[r]\}$. The marking is induced by the maps $f_w\colon(S^1,1)\to(\bigvee^r S^1,o)$ from Definition 2.2: the marked points are $f_w^{-1}(\{p_1,q_1,\ldots,p_r,q_r\})$. The color of a marked point p is determined by $f_w(p)$ and the orientation. For example, if $f_w(p)=q_i$ and f_w advances at p against the orientation of the circle corresponding to x_i, then p gets the color q_i^-. This is illustrated in Figure 4.1. $S^1(w)$ p_i^+,p_i^-,q_i^+,q_i^-

We think of the points $p_i^\pm$ and $q_i^\pm$ in $S^1(w_1),\ldots,S^1(w_\ell)$ as representing the $2L$ indices associated with the different letters of $w_1,\ldots,w_\ell$ in the computation of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$, as in (3.4). By definition, the second index of every letter of w_i must be identical to the first index of the cyclically subsequent letter of w_i. The other type of identifications of indices comes from the fixed bijections $\sigma,\tau\in\mathrm{Match}(w_1,\ldots,w_\ell)$. Every p_i^+-point is matched by σ to a p_i^--point. Similarly, every q_i^+-point is matched by τ with a q_i^--point.

Definition 4.2. Let $w_1,\ldots,w_\ell$ be a balanced set of words and let $\sigma,\tau\in\mathrm{Match}(w_1,\ldots,w_\ell)$. We associate with the pair (σ,τ) a 2-dimensional CW-complex, denoted $\Sigma_{(\sigma,\tau)}$. Its 1-dimensional skeleton consists of $S^1(w_1),\ldots,S^1(w_\ell)$ together with edges (1-dimensional cells) depicting the matchings σ and τ as above. Namely, for every $i\in[r]$, there is an edge connecting every p_i^+-point with its σ-image, and an edge connecting every q_i^+-point with its τ-image. We call these edges **matching-edges**. $\Sigma_{(\sigma,\tau)}$ matching-edges

To define the 2-dimensional cells, consider cycles in the 1-skeleton which are obtained by starting in some marked point on $S^1(w_i)$, moving orientably along $S^1(w_i)$ until the next marked point, then following the matching-edge emanating from this point and arriving at some marked

[25] Our immediate aim requires only the points p_i and q_i. The role of z_i is explained in Claim 4.4 below.

point in $S^1(w_j)$, then moving orientably along $S^1(w_j)$ to the next marked point, following a matching-edge and so forth, until a cycle has been completed. A 2-cell (a disc) is glued along every such cycle.

Finally, we denote by $v_1, \ldots, v_\ell$ the basepoints of $S^1(w_1), \ldots, S^1(w_\ell)$, respectively, and for $i \in [\ell]$ define $\partial_i : (S^1, 1) \to (\Sigma_{(\sigma,\tau)}, v_i)$ by the identification of $(S^1, 1)$ with $(S^1(w_i), 1) \subset \partial\Sigma_{(\sigma,\tau)}$.

We think of the $4L$ marked points as the vertices, or 0-skeleton of $\Sigma_{(\sigma,\tau)}$. Note the description of cycles we gave in the definition does indeed yield cycles because the walks on the 1-skeleton are invertible: to get the inverse walks use the same instructions only with reversed orientation on $S^1(w_1), \ldots, S^1(w_\ell)$. In Figures 4.2 and 4.3 we illustrate the 1-skeleton and surface associated with a particular pair of matchings for the word $w = [x,y][x,z]$.

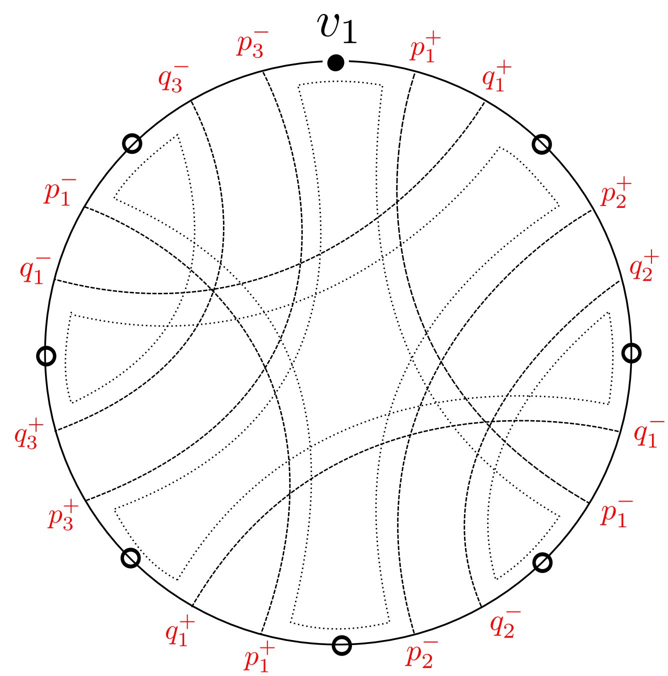

Figure 4.2: The 1-skeleton of $\Sigma_{(\sigma,\tau)}$ for $w = [x_1,x_2][x_1,x_3] = [x,y][x,z] = x_1y_2X_3Y_4x_5z_6X_7Z_8$ and the matchings $\sigma = \begin{pmatrix} x_1 & y_2 & x_5 & z_6 \\ X_3 & Y_4 & X_7 & Z_8 \end{pmatrix}$ and $\tau = \begin{pmatrix} x_1 & y_2 & x_5 & z_6 \\ X_7 & Y_4 & X_3 & Z_8 \end{pmatrix}$. Dashed lines are matching-edges. The dotted lines trace the boundaries of the two type-o disc to be glued in (see Claim 4.4). Three additional discs, one of type-z_1, one of type-z_2 and one of type-z_3, are glued in inside the other types of cycles one can follow (unmarked). For convenience, we also mark here the additional seven points of $f_w^{-1}(o)$ in $S^1(w)$, along v_1, by black circles.

Remark 4.3. For completeness we need also describe what happens when some of $S^1(w_1), \ldots, S^1(w_\ell)$ have no marked points, namely, when some of $w_1, \ldots, w_\ell$ are the empty word 1. In this case, whenever $w_i = 1$, we simply glue a disc along $S^1(w_i)$. To formally make it a CW-complex we also need to specify a vertex at the boundary of such disc, say, the basepoint 1 of $S^1(w_i)$. All the results below work just as well with this extension to trivial words, and the adjustments required in the proofs are trivial. However, to keep the writing slightly simpler, we ignore this case in what follows.

Claim 4.4. *The CW-complex $\Sigma_{(\sigma,\tau)}$ has the following properties:*

1. *Topologically, it is an orientable surface with ℓ boundary components.*

2. *Each 2-cell D is of one of two types:*

   (a) *Either $\partial D \cap (S^1(w_1) \cup \ldots \cup S^1(w_\ell))$ contains o-points (points from $f_{w_1}^{-1}(o) \cup \ldots \cup f_{w_\ell}^{-1}(o)$), in which case we call it a **type-o disc**,* type-o disc

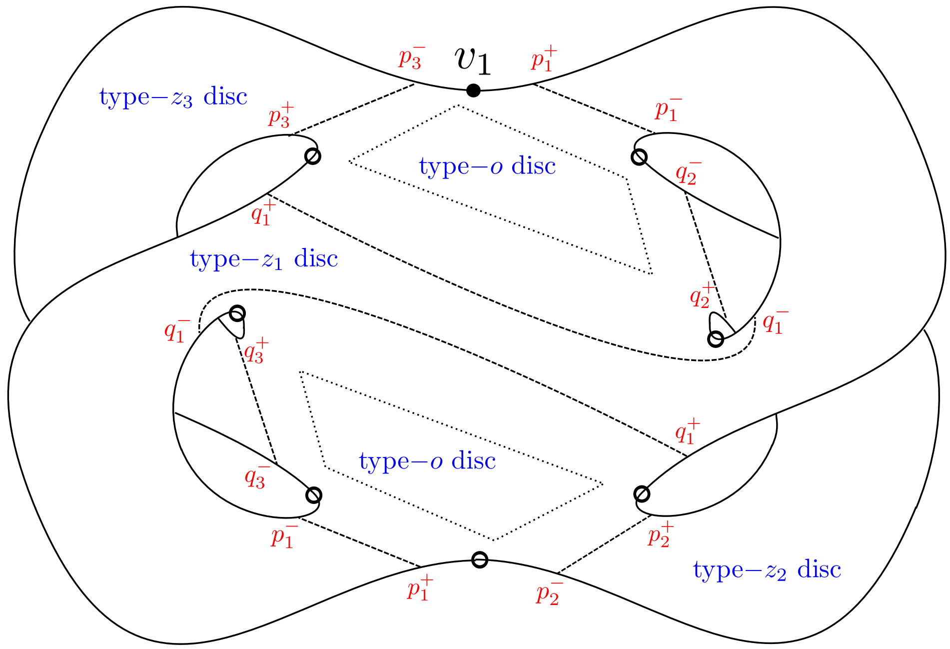

Figure 4.3: The CW-complex $\Sigma_{(\sigma,\tau)}$ corresponding to the word and matchings from Figure 4.2. Dashed and dotted lines correspond to those of Figure 4.2

(b) *Or $\partial D \cap \left(S^1(w_1) \cup \ldots \cup S^1(w_\ell)\right)$ contains z_i-points (points from $f_{w_1}^{-1}(z_i) \cup \ldots \cup f_{w_\ell}^{-1}(z_i)$) for some unique i, in which case we call it a **type-z_i disc**.* type-z_i disc

3. *Every type-o disc corresponds to a block of indices in the partition induced by σ and τ, so that $B_{(\sigma,\tau)}$ is the number of type-o discs.*

4. *Every type-z_i disc corresponds to a cycle of the permutation $\left(\sigma^{-1}\tau\right)\Big|_{E_i^+}$.*

5. *Every matching-edge is contained in the boundaries of exactly one type-o disc and exactly one type-z_i disc.*

Proof. Every segment in $S^1(w_{j_1})$ between two marked points contains either an o-point or a z_i-point for some unique i. If the boundary ∂D of a 2-cell D follows a segment containing an o-point, then ∂D goes on to follow a matching-edge emanating at the first marked point of a letter in $E^+ \cup E^-$, which, by construction, arrives at a second marked point of some other letter in $E^- \cup E^+$. So it then follows, again, a segment of $S^1(w_{j_2})$ containing an o-point. A similar argument shows that if ∂D contains a segment of $S^1(w_{j_1})$ with a z_i-point, then all the segments of $S^1(w_1) \cup \ldots \cup S^1(w_\ell)$ it contains have the same property. This shows item (2).

Items (3), (4) and (5) are evident from the construction. Every segment of $S^1(w_1) \cup \ldots \cup S^1(w_\ell)$ between two adjacent marked points is contained in the boundary of exactly one disc. This and item (5) show that $\Sigma_{(\sigma,\tau)}$ is a surface with $S^1(w_1) \cup \ldots \cup S^1(w_\ell)$ its boundary, hence ℓ boundary components. We can orient every disc according to the orientation of the $\left(S^1(w_1) \cup \ldots \cup S^1(w_\ell)\right)$-segments at its boundary, which shows the global orientability and item (1). □

We can now rewrite (3.6) as

$$\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = \sum_{\substack{(\sigma,\tau)\in\\ \mathrm{Match}(w_1,\ldots,w_\ell)^2}} \mathrm{Wg}\left(\left(\sigma^{-1}\tau\right)\Big|_{E_1^+}\right)\cdot\ldots\cdot\mathrm{Wg}\left(\left(\sigma^{-1}\tau\right)\Big|_{E_r^+}\right)\cdot n^{\#\left\{\text{type}-o\text{ discs in }\Sigma_{(\sigma,\tau)}\right\}}. \tag{4.1}$$

Definition 4.5. For $(\sigma,\tau)\in\mathrm{Match}(w_1,\ldots,w_\ell)^2$ denote by $\chi(\sigma,\tau)$ the Euler characteristic of $\Sigma_{(\sigma,\tau)}$. $\chi(\sigma,\tau)$

Proposition 4.6. *The contribution of* $(\sigma,\tau)\in\mathrm{Match}(w_1,\ldots,w_\ell)^2$ *to the summation* (4.1) *giving* $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ *is*

$$\text{Möb}\left(\sigma^{-1}\tau\right)\cdot n^{\chi(\sigma,\tau)}+O\left(n^{\chi(\sigma,\tau)-2}\right).$$

Proof. Although the Weingarten function of a permutation is *not* the product of the Weingarten functions of its disjoint cycles, the leading term does have this property. More generally, if

$$\pi=(\pi_1,\ldots,\pi_r)\in S_{L_1}\times\ldots\times S_{L_r}\le S_L,$$

then $\|\pi\|=\|\pi_1\|+\ldots+\|\pi_r\|$ and, by (3.2), $\text{Möb}(\pi)=\text{Möb}(\pi_1)\cdot\ldots\cdot\text{Möb}(\pi_r)$. Proposition 3.5 therefore yields that

$$\begin{aligned}\mathrm{Wg}(\pi_1)\cdot\ldots\cdot\mathrm{Wg}(\pi_r) &= \left(\frac{\text{Möb}(\pi_1)}{n^{L_1+\|\pi_1\|}}+O\left(\frac{1}{n^{L_1+\|\pi_1\|+2}}\right)\right)\cdot\ldots\cdot\left(\frac{\text{Möb}(\pi_r)}{n^{L_r+\|\pi_r\|}}+O\left(\frac{1}{n^{L_r+\|\pi_r\|+2}}\right)\right)\\ &= \frac{\text{Möb}(\pi)}{n^{L+\|\pi\|}}+O\left(\frac{1}{n^{L+\|\pi\|+2}}\right).\end{aligned}$$

Since $\|\pi_i\|=L_i-\#\text{cycles}(\pi_i)$, Claim 4.4(4) yields that

$$\left\|\sigma^{-1}\tau\right\|=L-\sum_i\#\left\{\text{type}-z_i\text{ discs in }\Sigma_{(\sigma,\tau)}\right\},$$

so the term corresponding to (σ,τ) in (4.1) is

$$\begin{aligned}&\frac{\text{Möb}\left(\sigma^{-1}\tau\right)}{n^{2L-\sum_i\#\left\{\text{type}-z_i\text{ discs in }\Sigma_{(\sigma,\tau)}\right\}}}\cdot n^{\#\left\{\text{type}-o\text{ discs in }\Sigma_{(\sigma,\tau)}\right\}}\cdot\left(1+O\left(\frac{1}{n^2}\right)\right)\\ &=\text{Möb}\left(\sigma^{-1}\tau\right)\cdot n^{\#\left\{\text{discs in }\Sigma_{(\sigma,\tau)}\right\}-2L}\cdot\left(1+O\left(\frac{1}{n^2}\right)\right).\end{aligned}$$

The statement of the proposition follows by noting that the 1-skeleton of $\Sigma_{(\sigma,\tau)}$ has $4L$ 0-cells (2 marked points associated with every letter of $w_1,\ldots,w_\ell$), and $6L$ 1-cells ($4L$ of them as segments of $S^1(w_1)\cup\ldots\cup S^1(w_\ell)$ and $2L$ matching-edges), so

$$\#\left\{\text{discs in }\Sigma_{(\sigma,\tau)}\right\}-2L=4L-6L+\#\left\{\text{discs in }\Sigma_{(\sigma,\tau)}\right\}=\chi\left(\Sigma_{(\sigma,\tau)}\right)=\chi(\sigma,\tau).$$

□

Next, we define (the homotopy class of) a function $f_{(\sigma,\tau)}\colon\Sigma_{(\sigma,\tau)}\to\bigvee^r S^1$ which makes $\left(\Sigma_{(\sigma,\tau)},f_{(\sigma,\tau)}\right)$ admissible for $w_1,\ldots,w_\ell$.

Definition 4.7. Given $\sigma, \tau \in \mathrm{Match}\,(w_1, \ldots, w_\ell)$, define the homotopy class (relative $\partial\Sigma_{(\sigma,\tau)}$) of a map $f_{(\sigma,\tau)} \colon \Sigma_{(\sigma,\tau)} \to \bigvee^r S^1$ as follows: $f_{(\sigma,\tau)}$

- Define $f_{(\sigma,\tau)}$ on $\partial\Sigma_{(\sigma,\tau)}$ by setting $f_{(\sigma,\tau)}\Big|_{S^1(w_i)} \equiv f_{w_i} \circ \partial_i^{-1}$ for every $i \in [\ell]$.
- Extend $f_{(\sigma,\tau)}$ to the entire 1-skeleton of $\Sigma_{(\sigma,\tau)}$ by setting $f_{(\sigma,\tau)}$ to be constant on every matching-edge, namely, $f_{(\sigma,\tau)}\Big|_e \equiv p_i$ for every p_i-matching-edge e etc.
- On every disc (2-cell) D, $f_{(\sigma,\tau)}$ now maps its boundary to a nullhomotopic loop in $\bigvee^r S^1$, so there exists a unique way, up to homotopy, to extend $f_{(\sigma,\tau)}$ to the interior of D (as in Lemma 2.1).

From Definitions 4.2 and 4.7 and Claim 4.4 we conclude:

Corollary 4.8. *For every $\sigma, \tau \in \mathrm{Match}\,(w_1, \ldots, w_\ell)$, the pair $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$ is admissible for $w_1, \ldots, w_\ell$.*

It turns out that all admissible maps (Σ, f) for $w_1, \ldots, w_\ell$ can be basically obtained this way, as long as f is incompressible.

Lemma 4.9. *If (Σ, f) is admissible for $w_1, \ldots, w_\ell$ and f is incompressible, then there is a pair of matchings $(\sigma, \tau) \in \mathrm{Match}\,(w_1, \ldots, w_\ell)$ so that $(\Sigma, f) \sim \left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$.*

Proof. Let[26] (Σ, f) be admissible for $w_1, \ldots, w_\ell$ and f incompressible. As in Lemma 2.1, we find a finite set of oriented disjoint arcs $\gamma_1, \ldots, \gamma_m \colon [0,1] \to \Sigma$ with endpoints in $\{v_1, \ldots, v_\ell\}$ which cut Σ into discs. For every $j \in [m]$, denote $u_j = [f(\gamma_j)] \in \mathbf{F}_r$. We now want to mark the arc γ_j with $2\,|u_j|$ points colored with $\{p_i, q_i \,|\, i \in [r]\}$ as we did in $S^1(w)$ in the beginning of this section, only, for now, without the $\pm$ sign. Namely, using the function $g \colon [0,1] \to S^1$ defined by $t \mapsto e^{2\pi i t}$, use $S^1(u_j)$ to mark and color points on γ_j.

Now, in every disc D which is cut from Σ by the arcs $\gamma_1, \ldots, \gamma_m$, use the orientation on D (induced from the one on Σ) to orient each arc γ_j at the boundary of D, and add accordingly $\pm$ signs to the colors of the marked points on this arc. In particular, every marked point on a γ_j is signed "+" for one of the two discs it borders and signed "−" for the other.

Since the image of ∂D through f is nullhomotopic, the sequence of marked points one reads along ∂D can be reduced to an empty sequence by successive deletions of pairs of the form $p_i^+ p_i^-$, $p_i^- p_i^+$, $q_i^+ q_i^-$ or $q_i^+ q_i^-$. We use one of these reduction processes and, at each step, draw an arc inside D between the two marked points we delete at that step. A simple inductive argument shows that at each step, the remaining unpaired points are all in the boundary of the same disc bounded by parts of ∂D and the existing arcs (with no arcs inside the disc), so one can draw in its interior a new arc connecting the next pair of points.

Next, use the new "reduction" arcs to determine σ and τ: for every marked point t on $\partial\Sigma$, t belongs to some disc D, and follow the arc emanating from t it to some $t' \in \partial D$. If t' is not in $\partial\Sigma$, but, say, in γ_j, follow the arc from t' inside the other disc bordering γ_j. Continue in the same way until a point from $\partial\Sigma$ is reached. It is easy to see that this induces matchings $\sigma, \tau \in \mathrm{Match}\,(w_1, \ldots, w_\ell)$: for example, a q_i^+-point in $\partial\Sigma$ is connected by an arc to a q_i^--point. If the latter is not on $\partial\Sigma$, it is identified with a q_i^+-point on a neighboring disc, which is then connected to another q_i^--point, and so forth. Note that some of the arcs may form cycles in the interior of Σ, and simply disregard or delete these one. Let A be the set of arcs we used for determining σ and τ. This is illustrated in Figure 4.4.

[26] A straight-forward argument is available when f is smooth outside $f^{-1}(o)$ and p_i and q_i are regular points for each $i \in [r]$. In this case, the desired matchings are obtained by the arc parts of $f^{-1}(p_i)$ and $f^{-1}(q_i)$.

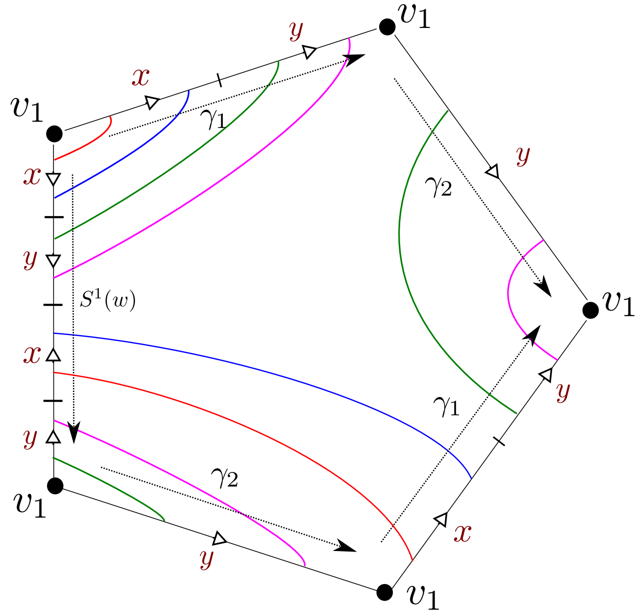

Figure 4.4: Let $w = [x, y] = x_1y_2X_3Y_4$ and $(\Sigma_{1,1}, f)$ be admissible for w and representing the solution $w = [xy, y]$. Namely, if γ_1 and γ_2 are arcs in $\Sigma_{1,1}$ representing two basis elements in $\mathbf{F}_2$, then $[f \circ \gamma_1] = xy$ and $[f \circ \gamma_2] = y$. The arcs γ_1 and γ_2 cut $\Sigma_{1,1}$ to a sole disc D, and the colored arcs in the figure correspond to a particular reduction process of the word read along ∂D, as explained in the proof of Lemma 4.9. The matchings we get here are $\sigma = \tau = \begin{pmatrix} x_1 & y_2 \\ X_3 & Y_4 \end{pmatrix}$ (this is the only possible matching for this particular word).

We claim that $\Sigma \setminus \bigcup_{\alpha \in A} \alpha$ is a union of discs. To see this, we first perturb f so that it agrees with f_{u_j} on γ_j for every $j \in [m]$. We then perturb it so that it is constant on every arc drawn in the reduction process: this only requires to change f in the interior of every disc D which is cut from Σ by the arcs γ_j. Now regard the arcs $\alpha \in A$ as the matching-edges in Definition 4.2, and follow the cycles along these arcs and $\partial\Sigma$ described in the same definition. These are precisely the boundaries of the connected components of $\Sigma \setminus \bigcup_{\alpha \in A} \alpha$. As in Definition 4.7, the image of f through each such cycle is easily seen to be nullhomotopic. But f is incompressible, hence each such circle must bound a disc.

This shows that Σ is homeomorphic to $\Sigma_{(\sigma,\tau)}$ with the arcs $\alpha \in A$ mapped to the matching-edges in $\Sigma_{(\sigma,\tau)}$. Since f and $f_{(\sigma,\tau)}$ agree on the 1-skeleton, they are homotopic (using, again, Lemma 2.1). Hence $(\Sigma, f) \sim \left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$. □

Since in every admissible (Σ, f) for $w_1, \ldots, w_\ell$ with maximal Euler characteristic f is incompressible, we deduce from Corollary 4.8 and Lemma 4.9 that,

Corollary 4.10. *The highest Euler characteristic of a pair of matchings is* $\operatorname{chi}(w_1, \ldots, w_\ell)$, *namely,*

$$\max_{(\sigma,\tau) \in \mathrm{Match}(w_1,\ldots,w_\ell)^2} \chi(\sigma, \tau) = \operatorname{chi}(w_1, \ldots, w_\ell).$$

Moreover, we get an extension to a Theorem of Culler [Cul81, Thm 4.1], stating that the number of equivalence classes of solutions to $[u_1, v_1] \cdots [u_g, v_g] = w$ with $g = \operatorname{cl}(w)$ is finite:

Corollary 4.11. *For every $w_1, \ldots, w_\ell \in \mathbf{F}_r$, there are at most finitely many equivalence classes of (Σ, f) which are admissible for $w_1, \ldots, w_\ell$ and incompressible. In particular, the set* $\mathrm{Solu}(w_1, \ldots, w_\ell)$ *is finite.*

Remark 4.12. In the proof of Lemma 4.9, we could choose in each disc D a reduction process that comes from a reduction of the word we read along ∂D. This would mean that whenever we pair two p_i-points, we also match their associated two q_i-points. In other words, the bijections we obtain satisfy $\sigma = \tau$. Thus,

$$\max_{\sigma \in \mathrm{Match}(w_1,\ldots,w_\ell)} \chi(\sigma, \sigma) = \operatorname{chi}(w_1, \ldots, w_\ell).$$

This fact, in a slightly different language and for a single word, appears already in Culler's work, where it is used as an algorithm to compute $\mathrm{cl}(w)$ [Cul81, Theorem 2.1]. More generally, Corollary 4.10 provides an algorithm to compute $\mathrm{chi}(w_1,\ldots,w_\ell)$ for every $w_1,\ldots,w_\ell \in \mathbf{F}_r$. Furthermore, Lemma 4.9 shows that by going over all matchings, we can find representatives for all admissible (Σ, f) for $w_1,\ldots,w_\ell$ with f incompressible. It is still not clear at this point how to tell apart the different equivalence classes of admissible maps, but we face this challenge in Section 7 below.

Finally, using (4.1) and Proposition 4.6, we can now deduce Theorem 1.8:

Corollary 4.13. *For any $w_1,\ldots,w_\ell \in \mathbf{F}_r$,*

$$\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = n^{\mathrm{chi}(w_1,\ldots,w_\ell)} \left[\sum_{\substack{(\sigma,\tau)\in \mathrm{Match}(w_1,\ldots,w_\ell)^2 \\ \text{with } \chi(\sigma,\tau) = \mathrm{chi}(w_1,\ldots,w_\ell)}} \text{Möb}\left(\sigma^{-1}\tau\right) \right] + O\left(n^{\mathrm{chi}(w_1,\ldots,w_\ell)-2}\right). \tag{4.2}$$

In particular, $\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = O\left(n^{\mathrm{chi}(w_1,\ldots,w_\ell)}\right)$.

Example 4.14. As an example, consider the word $w = [x,y]\,[x,z] = x_1 y_2 X_3 Y_4 x_5 z_6 X_7 Z_8$. The two possible matchings of E^+ and E^- which preserve the alphabet are $\begin{pmatrix} x_1 & y_2 & x_5 & z_6 \\ X_3 & Y_4 & X_7 & Z_8 \end{pmatrix}$ and $\begin{pmatrix} x_1 & y_2 & x_5 & z_6 \\ X_7 & Y_4 & X_3 & Z_8 \end{pmatrix}$, so there are exactly 4 pairs of matchings in this case. A simple computation shows all of them have Euler characteristic -3, which shows that $\mathrm{chi}(w) = -3$ (and $\mathrm{cl}(w) = 2$). For two of the pairs, $\text{Möb}\left(\sigma^{-1}\tau\right) = 1$ and for the other two $\text{Möb}\left(\sigma^{-1}\tau\right) = -1$. Hence, by Corollary 4.13, $\mathcal{T}r_{[x,y][x,z]}(n) = n^{-3}\cdot 0 + O\left(n^{-5}\right)$. In fact, the full computation in this case (by Theorem 3.7) shows that $\mathcal{T}r_{[x,y][x,z]}(n)$ is identically zero for every $n \geq 2$. In particular, this example shows that it is not true in general that $\mathcal{T}r_w(n) = \theta\left(\frac{1}{n^{2\cdot \mathrm{cl}(w)-1}}\right)$, nor that $\mathcal{T}r_w(n) \not\equiv 0$ for $w \in [\mathbf{F}_r, \mathbf{F}_r]$.

Example 4.15. As another example, consider $w = [x,y]^2$. There are four matchings in $\mathrm{Match}(w)$, hence 16 pairs. Among them, twelve have $\chi = -3$ and four have $\chi = -5$. Of the twelve with $\chi = -3$, four have $\text{Möb}\left(\sigma^{-1}\tau\right) = 1$ and eight have $\text{Möb}\left(\sigma^{-1}\tau\right) = -1$. Corollary 4.13 thus gives $\mathcal{T}r_{[x,y]^2} = \frac{-4}{n^3} + O\left(\frac{1}{n^5}\right)$. (Compare with the exact rational expression in (3.5)).

We end this section with one more interesting property of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$.

Corollary 4.16. *In the Laurent series in n expressing $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$, the coefficient of* every other exponent *vanishes. If ℓ is odd, only terms with odd exponents may not vanish, and if ℓ is even, only terms with even exponents may not vanish.*

Proof. Actually, this is true for the contribution of every $(\sigma,\tau)\in \mathrm{Match}(w_1,\ldots,w_\ell)^2$ separately. That the leading exponent of every contribution has the same parity as ℓ follows from the orientability of the surface $\Sigma_{(\sigma,\tau)}$: we saw that this leading exponent is $\chi(\sigma,\tau) = 2 - 2\cdot \mathrm{genus}\left(\Sigma_{(\sigma,\tau)}\right) - \ell$. The statement now follows from the property of the Weingarten function that the coefficient of every other exponent vanishes (see the paragraph right after Proposition 3.5). ∎

5 The Pairs of Matchings Poset

Corollary 4.13 shows that in order to prove Theorem 1.10, it is enough to restrict attention to pairs of matchings $(\sigma,\tau)\in \text{Match}\,(w_1,\ldots,w_\ell)^2$ with $\chi(\sigma,\tau)=\text{chi}\,(w_1,\ldots,w_\ell)$, namely, with (σ,τ) so that $\left[\left(\Sigma_{(\sigma,\tau)},f_{(\sigma,\tau)}\right)\right]\in \text{Solu}\,(w_1,\ldots,w_\ell)$. However, all our proofs below regarding these matchings only use the fact that $f_{(\sigma,\tau)}$ is incompressible. Therefore, we continue analyzing pairs (σ,τ) with $\left(\Sigma_{(\sigma,\tau)},f_{(\sigma,\tau)}\right)$ incompressible. This also allows us to prove Theorem 1.12 in its full generality.

To continue our analysis, we gather all pairs (σ,τ) which correspond to the same equivalence class of an admissible, incompressible (Σ,f). The main result of this section is that there is a natural poset structure on every such set of pairs, and that the leading coefficient of the contribution of this set to $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ is the Euler characteristic of (the simplicial complex associated with) this poset.

First, we introduce an order on pairs of permutations which is related to the partial order on S_L defined in Section 3.1: for $\sigma,\tau,\sigma',\tau'\in S_L$, we write[27] $(\sigma',\tau')\preceq(\sigma,\tau)$ if $(\sigma',\tau')\preceq(\sigma,\tau)$

$$\left\|\sigma^{-1}\tau\right\|=\left\|\sigma^{-1}\sigma'\right\|+\left\|\left(\sigma'\right)^{-1}\tau'\right\|+\left\|\left(\tau'\right)^{-1}\tau\right\|.$$

In other words, consider the Cayley graph of S_L with respect to all transpositions. We say that $(\sigma',\tau')\preceq(\sigma,\tau)$ if and only if there is a geodesic in this Cayley graph from σ to τ which goes through σ' and then through τ'.

$$\sigma---\sigma'---\tau'---\tau$$

Clearly, this order, with the same definition, can be applied just as well to pairs of bijections $\sigma,\tau,\sigma',\tau'\colon E^+\xrightarrow{\sim}E^-$. In fact, we can identify the set of bijections $E^+\xrightarrow{\sim}E^-$ with S_L by declaring an arbitrary bijection as the identity element. We can then think of $\text{Match}\,(w_1,\ldots,w_\ell)$ as a set of permutations in S_L. We shall use both points of views interchangeably.

Definition 5.1. Let (Σ,f) be admissible for $w_1,\ldots,w_\ell$ and incompressible. The **pairs of matchings poset** of (Σ,f), denoted $\mathcal{PMP}(\Sigma,f)$, consists of pairs of matchings in $\text{Match}\,(w_1,\ldots,w_\ell)^2$ which are associated, up to equivalence, with (Σ,f). Namely, $\mathcal{PMP}(\Sigma,f)$

$$\mathcal{PMP}(\Sigma,f)\overset{\text{def}}{=}\left\{(\sigma,\tau)\in\text{Match}\,(w_1,\ldots,w_\ell)^2\,\middle|\,\left(\Sigma_{(\sigma,\tau)},f_{(\sigma,\tau)}\right)\sim(\Sigma,f)\right\}.$$

The partial order $\preceq$ on $\mathcal{PMP}(\Sigma,f)$ is induced from the partial order on pairs of bijections $E^+\xrightarrow{\sim}E^-$.

The following property of pairs of matching associated with an incompressible map is important in what follows.

Lemma 5.2. *If $f_{(\sigma,\tau)}$ is incompressible for some $(\sigma,\tau)\in\text{Match}\,(w_1,\ldots,w_\ell)^2$, then any two neighboring discs in $\Sigma_{(\sigma,\tau)}$, which are necessarily of type-o and of type-z_i for some $i\in[r]$, have at most two common matching-edges at their boundaries: at most one p_i-edge and at most one q_i-edge.*

[27] This paper uses the same symbol $\preceq$ to denote different partial orders. However, two different partial orders are always defined on different types of elements, so it should be easy to realize which partial order is referred to at any point in the text.

Proof. Assume, to the contrary, that there are discs D_1 of type-o and D_2 of type-z_i so that $\partial D_1 \cap \partial D_2$ contains two distinct matching-edges e_1 and e_2 of the same color, say q_i. Let γ be a simple closed curve that traverses exactly two matching-edges – e_1 and e_2 – and each one exactly once. It is easy to see that $f_{(\sigma,\tau)}(\gamma)$ is then nullhomotopic in $\bigvee^r S^1$, and so γ bounds a disc by the assumption. But this is impossible as there are points from $\partial\Sigma_{(\sigma,\tau)}$ at both sides of γ (e.g. the points at the endpoints of e_1 and e_2). $\square$

The poset $\mathcal{PMP}(\Sigma, f)$ is a downward-closed sub-poset of the poset of pairs of bijections. Namely,

Lemma 5.3. *Assume that $(\sigma,\tau) \in \mathcal{PMP}(\Sigma, f)$, that $\sigma', \tau' \colon E^+ \xrightarrow{\sim} E^-$ are bijections and that $(\sigma',\tau') \preceq (\sigma,\tau)$. Then $(\sigma',\tau') \in \mathcal{PMP}(\Sigma, f)$.*

Proof. First note the following observation: let $\pi \in S_L$ satisfy $\|\pi\| = k$ and let $t_1 t_2 \dots t_k$ be a product of transpositions giving π. Then for every j, the two elements $x, y \in [L]$ swapped by t_j must be two elements which sit in two different cycles in $t_1 t_2 \dots t_{j-1}$ but which belong to the same cycle in π. This follows from the identity $\|\pi\| = L - \#\text{cycles}(\pi)$ and from the fact that when a permutation is multiplied by a transposition either two of its cycles are merged together or one of its cycles is split into two.

We claim that from this simple observation it follows that $\sigma', \tau' \in \text{Match}(w_1, \dots, w_\ell)$, i.e. that σ' and τ' map E_i^+ to E_i^- for every $i \in [r]$. Indeed, this is certainly true for σ and τ and thus $\sigma^{-1}\tau$ maps E_i^+ to E_i^+ for every i. By assumption, there is a product of transpositions in $\text{Sym}(E^+)$ of minimal length which gives $\sigma^{-1}\tau$ such that two of its prefixes equal $\sigma^{-1}\sigma'$ and $\sigma^{-1}\tau'$. By the observation, no transposition in the product can mix elements of E_i^+ and E_j^+ with $i \neq j$, and thus this is also true for $\sigma^{-1}\sigma'$ and $\sigma^{-1}\tau'$, and indeed $\sigma', \tau' \in \text{Match}(w_1, \dots, w_\ell)$.

It is left to show that $\left(\Sigma_{(\sigma',\tau')}, f_{(\sigma',\tau')}\right) \sim \left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$. It is enough to show this in the case when (σ,τ) covers (σ',τ') (see Footnote 28). In this case, either $\sigma' = \sigma$ and $\tau^{-1}\tau'$ is a transposition, or $\tau' = \tau$ and $\sigma^{-1}\sigma'$ is a transposition. Assume the former case, the latter having the exact same proof. So $(\sigma',\tau') = (\sigma,\tau')$ is the same as (σ,τ), except for two q_i^+-points j and k, for some i, with $\tau'(j) = \tau(k)$ and $\tau'(k) = \tau(j)$. If we abuse notation and let j and k denote also the corresponding letters in E^+, then $\sigma^{-1}\tau' \cdot (j\,k) = \sigma^{-1}\tau \in \text{Sym}(E^+)$. Because of the equality $\left\|\sigma^{-1}\tau'\right\| = \left\|\sigma^{-1}\tau\right\| - 1$, j and k must belong to different cycles of $\sigma^{-1}\tau'$ and to the same cycle of $\sigma^{-1}\tau$. Namely, the q_i^+-points j and k are at the boundary of the same type-z_i disc of $\Sigma_{(\sigma,\tau)}$.

Consider $\Sigma_{(\sigma,\tau)}$ and the two matching-edges e_j and e_k emanating from j and k, respectively, and let D denote the type-z_i disc they both belong to. By Lemma 5.2, they belong to two different type-o discs. The change in these two edges is the only change in the 1-skeleton of the CW-complex when moving from $\Sigma_{(\sigma,\tau)}$ to $\Sigma_{(\sigma,\tau')}$. In fact, to obtain $\Sigma_{(\sigma,\tau')}$ from $\Sigma_{(\sigma,\tau)}$ we can do the following: (i) draw two new disjoint edges (arcs) inside D: e'_j from j to $\tau(k)$ and e'_k from k to $\tau(j)$ – this is always possible because all q_i-matching-edges at the boundary of a type-z_i disc are oriented. (ii) Replace e_j and e_k by e'_j and e'_k. The change results in splitting the joint type-z_i into two discs and merging the two type-o discs into one. We illustrate this in figure 5.1.

From this description of $\Sigma_{(\sigma,\tau')}$ there is a natural homeomorphism $\Sigma_{(\sigma,\tau)} \cong \Sigma_{(\sigma,\tau')}$, and $f_{(\sigma,\tau)}$ and $f_{(\sigma,\tau')}$ agree on the entire common parts of the 1-skeletons, i.e. on all boundary and matching-edges surfaces, except for, possibly, on e_j, e_k and e'_j, e'_k. But within the freedom left in the definition of these functions (Definition 4.7), we can assume that both are constant functions along all of e_j, e_k, e'_j and e'_k, mapping all four matching edges to $q_i \in \bigvee^r S^1$. Then, by Lemma 2.1, they are homotopic to each other. Thus $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right) \sim \left(\Sigma_{(\sigma,\tau')}, f_{(\sigma,\tau')}\right)$. $\square$

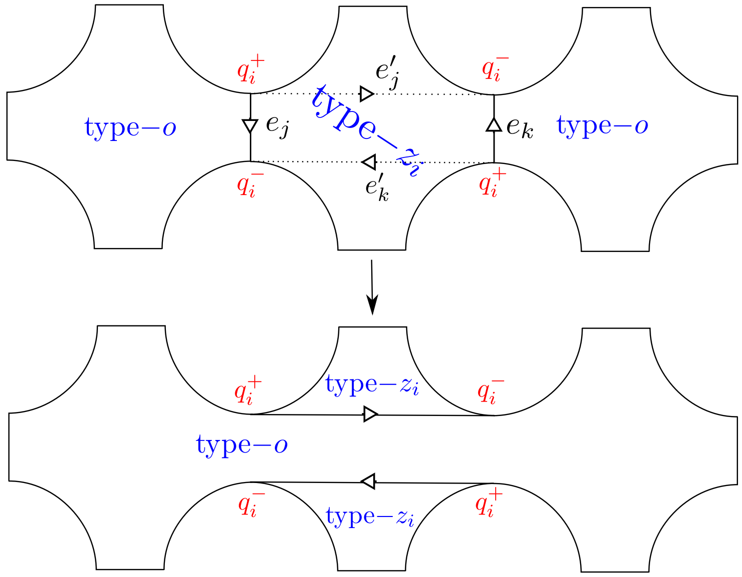

Figure 5.1: Swapping two q_i-matching-edges in the boundary of the same type-z_i disc in $\Sigma_{(\sigma,\tau)}$ for some $(\sigma,\tau) \in \mathcal{PMP}(\Sigma, f)$ results in $\Sigma_{(\sigma,\tau')}$ for some other $(\sigma,\tau') \in \mathcal{PMP}(\Sigma, f)$. The number of type-$z_i$ discs increases by one, while the number of type-o discs decreases by one. This corresponds to moving one step down, namely, to a covered element, in the poset $\mathcal{PMP}(\Sigma, f)$.

As an example, let $w = [x, y][x, z]$. We already mentioned in Example 4.14 above that there are four pairs of matchings, all of which with $\chi = -3$. An easy application of Lemma 5.3 shows that all four belong to same class of admissible incompressible $[(\Sigma, f)]$. Two of the four pairs satisfy $\sigma = \tau$, and both are smaller ($\prec$) than the other two pairs in which $\sigma^{-1}\tau$ is a transposition.

From the last lemma we can deduce that the poset $\mathcal{PMP}(\Sigma, f)$ is a *graded poset*[28], with rank function $\mathcal{PMP}(\Sigma, f) \to \mathbb{Z}_{\geq 0}$ given by $(\sigma, \tau) \mapsto \left\|\sigma^{-1}\tau\right\|$. Moreover, recall from the proof of Proposition 4.6 that $\chi(\sigma,\tau) = \#\left\{\text{discs in}\,\Sigma_{(\sigma,\tau)}\right\} - 2L$. Among the pairs in $\mathcal{PMP}(\Sigma, f)$ the Euler characteristic $\chi(\sigma,\tau)$ is constant, and thus so is the total number of discs. The total number of type-z_i discs is equal to $L - \left\|\sigma^{-1}\tau\right\|$, hence we obtain:

Claim 5.4. *The poset $\mathcal{PMP}(\Sigma, f)$ is graded, with two possible, natural rank functions: either $\left\|\sigma^{-1}\tau\right\|$, or the number of type-$o$ discs in $\Sigma_{(\sigma,\tau)}$.*

Remark 5.5. More generally, a similar argument as in the proof of Lemma 5.3 shows that if $(\sigma,\tau), (\sigma',\tau') \in \text{Match}(w_1, \ldots, w_\ell)^2$ and $(\sigma',\tau') \preceq (\sigma,\tau)$, then $\chi(\sigma',\tau') \geq \chi(\sigma,\tau)$. If, moreover, (σ,τ) covers (σ',τ'), then $\chi(\sigma',\tau') - \chi(\sigma,\tau) \in \{0, 2\}$.

The last argument in the proof of Lemma 5.3, where we made changes to matching-edges in $\Sigma_{(\sigma,\tau)}$, can be generalized to the following definition which allows a more geometric definition of the order $\preceq$ on $\mathcal{PMP}(\Sigma, f)$. This equivalent definition will be of great importance in Section 6.

Definition 5.6. A partition P of the matching-edges at the boundary of a disc of $\Sigma_{(\sigma,\tau)}$ is called a **colored non-crossing partition**, if

[28] A graded poset is a poset $(P, \leq)$ together with a rank function $\text{rk} : P \to \mathbb{Z}_{\geq 0}$, such that if $x < y$ then $\text{rk}(x) < \text{rk}(y)$, and if y covers x (that is, $x < y$ and there is no z with $x < z < y$) then $\text{rk}(y) = \text{rk}(x) + 1$. We note that the definition in [Sta12, Section 3.1] is slightly less general.

- it is colored: every block of P is monochromatic (contains matching-edges of the same color), and
- it is non-crossing: there are no four matching-edges which in cyclic order are e_1, e_2, e_3, e_4 and such that e_1 and e_3 belong to one block and e_2 and e_4 to another.

This is the same as the usual notion of non-crossing partitions (see [NS06, Lecture 9]), only with the additional constraint of monochromatic blocks.

Lemma 5.7. *Given $(\sigma, \tau) \in \mathcal{PMP}(\Sigma, f)$ and a colored non-crossing partition P of a disc (2-cell) D of $\Sigma_{(\sigma,\tau)}$, we can obtain a new pair of matchings $(\sigma', \tau') \in \mathcal{PMP}(\Sigma, f)$ by the following procedure: using the orientation on ∂D, match the second endpoint of a matching-edge with the first endpoint of the following edge in the same block of P. Now replace the old matching-edges along ∂D with the new ones.*

Proof. First, all matching-edges of a fixed color at the boundary of D have the same orientation, so the instructions in the claim indeed match marked points on E^+ with marked points on E^-, and lead to a new pair $(\sigma', \tau') \in \mathrm{Match}(w_1, \ldots, w_\ell)^2$. It remains to show that $(\sigma', \tau') \in \mathcal{PMP}(\Sigma, f)$, and we now show this basically follows from the same argument as in the proof of Lemma 5.3.

Note that the new matching-edges can be drawn as disjoint arcs inside D: the disjointness can be achieved thanks to P being non-crossing. By Lemma 5.2, the discs on the other side of the matching-edges in the same block $B \in P$ are distinct. Thus, after replacing the matching-edges along ∂D with the new ones, the surface is still cut to discs, and so the CW-complex obtained that way from $\Sigma_{(\sigma,\tau)}$ is exactly $\Sigma_{(\sigma',\tau')}$. Finally, we can choose $f_{(\sigma,\tau)}$ so that it is constant not only on all matching-edges of $\Sigma_{(\sigma,\tau)}$ but also on the new matching-edges in D. Then, with Lemma 2.1, we get that $f_{(\sigma,\tau)}$ and $f_{(\sigma',\tau')}$ are homotopic. We illustrate this in Figure 5.2. ☐

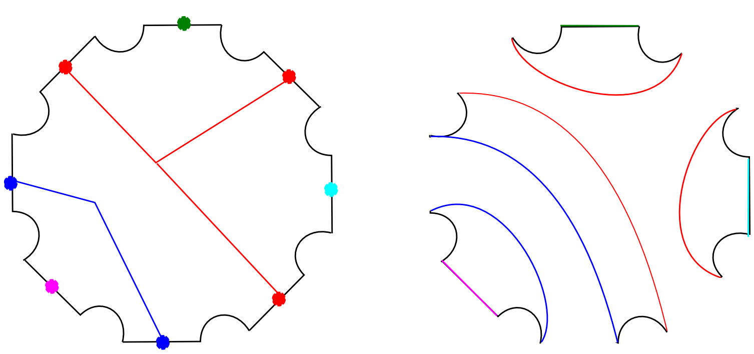

Figure 5.2: The figure on the left shows a non-crossing partition of the eight matching-edges along the boundary of a disc D: every block is marked by a different color. (The matching-edges in every block need be of the same color of q_i or p_i, but this is not shown in the figure.) Rewiring the matching-edges according to this partition results in the figure on the right: the disc D is split into four smaller discs, and some of its area serves as "corridors" which merge neighboring discs.

Proposition 5.8. *Assume that (σ', τ') and (σ, τ) are both in $\mathcal{PMP}(\Sigma, f)$. Then the following are equivalent:*

1. $(\sigma',\tau') \preceq (\sigma,\tau)$

2. *$\Sigma_{(\sigma,\tau)}$ can be obtained from $\Sigma_{(\sigma',\tau')}$ by a rewiring of matching-edges according to colored non-crossing partitions in type-o discs.*

3. *$\Sigma_{(\sigma',\tau')}$ can be obtained from $\Sigma_{(\sigma,\tau)}$ by a rewiring of matching-edges according to colored non-crossing partitions in type-z_i discs (for all i together).*

Moreover, if indeed $(\sigma',\tau') \preceq (\sigma,\tau)$, then the set of colored non-crossing partitions in item 2 *(item* 3*) is unique.*

Proof. The uniqueness of the partitions is obvious. For example, in item (2) the partition in every type-o disc can be read from the pair (σ,τ), which is given. We now prove (1) $\iff$ (2), the equivalence (1) $\iff$ (3) being completely analogous.

(1) $\Longrightarrow$ (2)**:** We show that if $(\sigma',\tau') \preceq (\sigma,\tau)$ then there is a rewiring of matching-edges inside type-o discs of $\Sigma_{(\sigma',\tau')}$ which results in $\Sigma_{(\sigma,\tau)}$. It is then obvious that the rewiring in every type-o disc corresponds to a colored non-crossing partition of its matching-edges. We prove there is such rewiring by induction on the difference in ranks $t = \left\|\sigma^{-1}\tau\right\| - \left\|(\sigma')^{-1}\tau'\right\|$.

If $t=1$, namely, if (σ,τ) covers (σ',τ'), we repeat the argument in the proof of Lemma 5.3: the difference in the 1-skeletons is exactly in two matching-edges. These two matching-edges in $\Sigma_{(\sigma,\tau)}$ must belong to the same type-z_i disc. Hence, by the proof of Lemma 5.3 and Figure 5.1, these the two matching-edges in $\Sigma_{(\sigma',\tau')}$ must belong to the same type-o disc, and we can rewire both of them inside this disc to obtain $\Sigma_{(\sigma,\tau)}$.

If $t \geq 2$, let (σ'',τ'') be an intermediate pair which is covered by (σ,τ). Use the induction hypothesis to find a rewiring inside type-o discs of $\Sigma_{(\sigma',\tau')}$ which gives $\Sigma_{(\sigma'',\tau'')}$. Of course, we can now find a rewiring of two matching-edges inside a type-o disc of $\Sigma_{(\sigma'',\tau'')}$ which gives $\Sigma_{(\sigma,\tau)}$. The crux of the argument is that type-o discs of $\Sigma_{(\sigma'',\tau'')}$ are *completely contained* inside type-o discs of $\Sigma_{(\sigma',\tau')}$, so the whole rewiring takes places inside type-o discs of $\Sigma_{(\sigma',\tau')}$.

(2) $\Longrightarrow$ (1): By Lemma 5.7, we can perform the rewiring at one type-o disc at a time and obtain a surface corresponding to some pair in $\mathcal{PMP}(\Sigma,f)$ at each step. Thus it is enough to show this implication if the rewiring is in a single type-o disc D, and by the colored non-crossing partition P.

Let $P_0, P_1, \ldots, P_m = P$ be a sequence of partitions of the matching-edges in D, each obtained from the former by merging together two blocks, so that P_0 consists entirely of singletons. Denote by (σ_j,τ_j) the pair of matchings in $\mathcal{PMP}(\Sigma,f)$ corresponding to the rewiring by P_j. Now, $\Sigma_{(\sigma_j,\tau_j)}$ can be obtained from $\Sigma_{(\sigma_{j-1},\tau_{j-1})}$ by rewiring a single pair of matching-edges inside a type-o disc. Thus, it suffices to show that in this case we go up in the poset $\mathcal{PMP}(\Sigma,f)$. Without loss of generality, assume that this single pair of matching-edges is of color q_i. Thus, $\sigma_j = \sigma_{j-1}$ and $\tau_j^{-1}\tau_{j-1}$ is a transposition. So the pairs $(\sigma_{j-1},\tau_{j-1})$ and (σ_j,τ_j) are necessarily comparable, and indeed $(\sigma_{j-1},\tau_{j-1}) \prec (\sigma_j,\tau_j)$ because the number of type-o discs increases in this rewiring. □

Before stating the main theorem of this section we need one more simple lemma:

Lemma 5.9. *Let $\sigma_0,\tau_0 \in S_L$. Then*

$$\sum_{(\sigma,\tau)\preceq(\sigma_0,\tau_0)} \text{Möb}\left(\sigma^{-1}\tau\right) = 1.$$

Proof. By the definition of the order $\preceq$ on pairs, $(\sigma,\tau)\preceq(\sigma_0,\tau_0)$ if and only if $\mathrm{id}\preceq\sigma_0^{-1}\sigma\preceq\sigma_0^{-1}\tau\preceq\sigma_0^{-1}\tau_0$ in S_L. By Proposition 3.4 and the definition (3.1) of the Möbius function μ of the poset $(S_L,\preceq)$,

$$\begin{aligned}\sum_{(\sigma,\tau)\preceq(\sigma_0,\tau_0)}\text{Möb}\left(\sigma^{-1}\tau\right) &= \sum_{\sigma,\tau:\,\mathrm{id}\preceq\sigma_0^{-1}\sigma\preceq\sigma_0^{-1}\tau\preceq\sigma_0^{-1}\tau_0}\text{Möb}\left(\sigma^{-1}\tau\right)\\ &= \sum_{\sigma,\tau:\,\mathrm{id}\preceq\sigma\preceq\tau\preceq\sigma_0^{-1}\tau_0}\text{Möb}\left(\sigma^{-1}\tau\right) = \sum_{\sigma:\,\mathrm{id}\preceq\sigma\preceq\sigma_0^{-1}\tau_0}\left(\sum_{\tau:\,\sigma\preceq\tau\preceq\sigma_0^{-1}\tau_0}\text{Möb}\left(\sigma^{-1}\tau\right)\right)\\ &= \sum_{\sigma:\,\mathrm{id}\preceq\sigma\preceq\sigma_0^{-1}\tau_0}\delta_{\sigma,\sigma_0^{-1}\tau_0} = 1.\end{aligned}$$

□

Definition 5.10. [Sta12, Section 3.8] For every locally finite poset[29] $(P,\leq)$ there is an associated simplicial complex, the vertices of which are the elements of P and the simplices are the chains. That is, $x_1,\ldots,x_k\in P$ form a simplex if and only if, after possible rearrangement, $x_1<x_2<\ldots<x_k$. We let $|P|$ denote the *geometric realization* of this simplicial complex[30]. $|P|$

The following theorem shows that the Euler characteristic of the simplicial complex $|\mathcal{PMP}(\Sigma,f)|$ captures the leading coefficient of the contribution of the pairs of matchings in $\mathcal{PMP}(\Sigma,f)$ to $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ from Corollary 4.13. Recall that $\chi()$ marks Euler characteristic.

Theorem 5.11. *If (Σ,f) is admissible for $w_1,\ldots,w_\ell$ and incompressible, then*

$$\sum_{(\sigma,\tau)\in\mathcal{PMP}(\Sigma,f)}\text{Möb}\left(\sigma^{-1}\tau\right)=\chi\left(|\mathcal{PMP}(\Sigma,f)|\right).$$

In particular,

$$\mathcal{T}r_{w_1,\ldots,w_\ell}(n)=n^{\mathrm{chi}(w_1,\ldots,w_\ell)}\left[\sum_{[(\Sigma,f)]\in\mathrm{Solu}(w_1,\ldots,w_\ell)}\chi\left(|\mathcal{PMP}(\Sigma,f)|\right)\right]+O\left(n^{\mathrm{chi}(w_1,\ldots,w_\ell)-2}\right). \tag{5.1}$$

Proof. Recall that for a simplicial complex Δ, the Euler characteristic is

$$\chi(\Delta)=\sum_{\emptyset\neq s}(-1)^{\dim s},$$

the sum being over all non-empty simplices in Δ, and $\dim s=|s|-1$. We prove the statement for any poset P of pairs of bijections with the downward-closure property elaborated in Lemma 5.3. It is enough to show that for every pair (σ_0,τ_0) we have

$$\text{Möb}\left(\sigma_0^{-1}\tau_0\right)=\sum_{s\subseteq P:\,\max s=(\sigma_0,\tau_0)}(-1)^{\dim s}, \tag{5.2}$$

[29] See footnote on Page 25.

[30] The space $|P|$ is a topological space with the following topology: every simplex s has the Euclidean topology. A general set $A\subseteq|P|$ is closed if and only if $A\cap s$ is closed in s for every simplex s.

the sum being over all chains in P with maximal element (σ_0, τ_0). Indeed, if (5.2) holds, then

$$\sum_{(\sigma_0,\tau_0)\in P} \text{Möb}\left(\sigma_0^{-1}\tau_0\right) = \sum_{(\sigma_0,\tau_0)\in P} \left[\sum_{s\subseteq P\colon \max s=(\sigma_0,\tau_0)} (-1)^{\dim s}\right] = \sum_{\emptyset\neq s\subseteq P} (-1)^{\dim s} = \chi\left(|P|\right).$$

So we only need to prove (5.2). Denote by $(-\infty, (\sigma_0, \tau_0)]_{\preceq}$ all pairs below (or equal to) (σ_0, τ_0) according to $\preceq$. We prove (5.2) by induction on the size t of $(-\infty, (\sigma_0, \tau_0)]_{\preceq}$. It clearly holds for $t = 1$, in which case necessarily $\sigma_0 = \tau_0$ by the downward-closeness property. For $t \geq 2$, note the one-to-one correspondence among the chains in $\big(-\infty, (\sigma_0, \tau_0)\big]_{\preceq}$ between those containing (σ_0, τ_0) and those not containing it. This correspondence is given by $s \mapsto s \setminus \{(\sigma_0, \tau_0)\}$. Now,

$$\begin{aligned}
\sum_{s\subseteq P\colon \max s=(\sigma_0,\tau_0)} (-1)^{\dim s} &= \left(\sum_{s\subseteq P\colon \max s=(\sigma_0,\tau_0)} \left[(-1)^{\dim s} + (-1)^{\dim(s\setminus\{(\sigma_0,\tau_0)\})}\right]\right) \\
&\quad - \left((-1)^{\dim\emptyset} + \sum_{\emptyset\neq s\subseteq P\colon \max s\prec(\sigma_0,\tau_0)} (-1)^{\dim s}\right) \\
&= 0 - \left(-1 + \sum_{(\sigma,\tau)\prec(\sigma_0,\tau_0)} \sum_{s\subseteq P\colon \max s=(\sigma,\tau)} (-1)^{\dim s}\right) \\
&\overset{(1)}{=} 1 - \sum_{(\sigma,\tau)\prec(\sigma_0,\tau_0)} \text{Möb}\left(\sigma^{-1}\tau\right) \overset{(2)}{=} \text{Möb}\left(\sigma_0^{-1}\tau_0\right),
\end{aligned}$$

where in $\overset{(1)}{=}$ we used the induction hypothesis for smaller values of t, and in $\overset{(2)}{=}$ we used Lemma 5.9. ∎

As an example, consider again $w = [x, y]\,[x, z]$. We already described above (in Page 38) the poset $\mathcal{PMP}(\Sigma, f)$ of the only equivalence class in this case. The associated simplicial complex is one dimensional with the shape of a 4-cycle. Topologically, this is simply S^1, and the Euler characteristic is 0. This agrees, of course, with the direct computation carried out in Example 4.14.

In the next section we shall prove the following:

Theorem 5.12. *Let (Σ, f) be admissible for $w_1, \ldots, w_\ell$ and incompressible. As above, denote by $\tilde{f}$ the homotopy class of f, relative $\partial\Sigma$. Then $|\mathcal{PMP}(\Sigma, f)|$ is a $\mathrm{K}(G, 1)$-space for $G = \mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$.*

As explained in Section 1, in order to prove this theorem we show in the next section that (i) $|\mathcal{PMP}(\Sigma, f)|$ is (path) connected, (ii) its fundamental group is isomorphic to $\mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$, and (iii) its universal cover in contractible.

Our main theorems now follow immediately from Theorem 5.12: Theorem 1.12 follows as $|\mathcal{PMP}(\Sigma, f)|$ is a finite simplicial complex, and Theorem 1.10 follows using (5.1).

6 The Arc Poset

In this section we construct yet another poset related to some (Σ, f) (we assume (Σ, f) is admissible for $w_1, \ldots, w_\ell$ and incompressible throughout this section). This poset is named the "arc poset" of (Σ, f), and its elements consist of sets of arcs on the surface Σ. Each one of them looks like a specific geometric realization of the matching-edges in $\Sigma_{(\sigma,\tau)}$ for some $(\sigma, \tau) \in \mathcal{PMP}(\Sigma, f)$. However, in the arc poset we let $\mathrm{MCG}(\Sigma)$ act freely. Namely, different sets of arcs representing the same pair of matchings will constitute different elements in the arc poset as long as they differ by the action of a non-trivial element of $\mathrm{MCG}(\Sigma)$. As we show below, the connected components of the arc poset shall serve as universal cover of $|\mathcal{PMP}(\Sigma, f)|$ and enable us to prove Theorem 5.12.

6.1 Arc systems

Recall that if (Σ, f) is admissible for $w_1, \ldots, w_\ell$, then the ℓ boundary components of Σ are identified with $S^1(w_1), \ldots, S^1(w_\ell)$ and have $4L$ marked points on them which spell out $w_1, \ldots, w_\ell$ (consult also the glossary on Page 63).

Definition 6.1. Let (Σ, f) be admissible for $w_1, \ldots, w_\ell$ and incompressible. An **arc system** for (Σ, f) is an ambient isotopy (relative to the boundary $\partial\Sigma$) class of sets of $2L$ disjoint arcs embedded in Σ, which meet $\partial\Sigma$ only at their endpoints and so that the matching they induce on the $4L$ marked points in $\partial\Sigma$ is identical to the one induced by some $(\sigma, \tau) \in \mathcal{PMP}(\Sigma, f)$.

We denote by $[\{\alpha_1, \ldots, \alpha_{2L}\}]$ the arc system with representative $\{\alpha_1, \ldots, \alpha_{2L}\}$. We also denote by $(\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}})$ the pair of matchings in $\mathcal{PMP}(\Sigma, f)$ associated with the arc system $\vec{\alpha} = [\{\alpha_1, \ldots, \alpha_{2L}\}]$.

$[\{\alpha_1, \ldots, \alpha_{2L}\}]$ $\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}}$

Note, in particular, that an arc system for (Σ, f) must connect p_i^+-points in $\partial\Sigma$ to p_i^--points, and q_i^+-points to q_i^--points, for every $i \in [r]$. We call an arc a **p_i-arc** (a **q_i-arc**, respectively) if it connects a p_i^+-point with a p_i^--point (a q_i^+-point with a q_i^--point, respectively). We think of the arcs as colored by $\{p_i, q_i \mid i \in [r]\}$.

Claim 6.2. *An arc system $\vec{\alpha}$ for (Σ, f) cuts Σ into discs.*

Proof. By definition, the matching-edges in $\Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}$ cut $\Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}$ into discs. Since $\Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})} \cong \Sigma$, a simple Euler characteristic argument shows the arcs in $\vec{\alpha}$ must also cut Σ into discs: otherwise, the Euler characteristic is too small. □

We can therefore think of Σ with the arc system $\vec{\alpha}$ as a CW-complex which is isomorphic to the CW-complex $\Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}$. We let $\Sigma_{\vec{\alpha}}$ denote this CW-complex. We extend some of the notions we had for $\Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}$ to $\Sigma_{\vec{\alpha}}$: As in Claim 4.4, every disc D in $\Sigma_{\vec{\alpha}}$ is either a **type-o disc** (if ∂D contains o-points, i.e. points from $f_{w_1}^{-1}(o) \cup \ldots \cup f_{w_\ell}^{-1}(o)$) or a **type-$z_i$ disc** (if ∂D contains z_i-points for some i, i.e. points from $f_{w_1}^{-1}(z_i) \cup \ldots \cup f_{w_\ell}^{-1}(z_i)$). This is illustrated in Figure 6.1. $\Sigma_{\vec{\alpha}}$

We also define a (homotopy class of a) map $f_{\vec{\alpha}} \colon \Sigma \to \bigvee^r S^1$ as in Definition 4.7: we let $f_{\vec{\alpha}}$ extend $f_{w_1}, \ldots, f_{w_\ell}$ on $\partial\Sigma$, and be constant on the arcs. There is then a unique way (up to homotopy) to extend $f_{\vec{\alpha}}$ in the discs of $\Sigma_{\vec{\alpha}}$. Evidently, $(\Sigma, f_{\vec{\alpha}}) \sim \left(\Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}, f_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}\right) \sim (\Sigma, f)$ and, in particular, $(\Sigma, f_{\vec{\alpha}})$ is admissible for $w_1, \ldots, w_\ell$. Finally, as in Lemma 5.2, two bordering discs of $\Sigma_{\vec{\alpha}}$, which must be one of type-o and the other of type-z_i, have at most 2 common arcs at their boundaries: at most one p_i-arc and at most q_i-arc. $f_{\vec{\alpha}}$

The following useful claim is evident from the definition of $f_{\vec{\alpha}}$:

Claim 6.3. *Let $\vec{\alpha}$ be an arc system for (Σ, f), and let γ be an oriented arc in Σ with endpoints in $v_1, \ldots, v_\ell$. Then the word $[f_{\vec{\alpha}}(\gamma)] \in \mathbf{F}_r$ can be computed as follows: fix a representative $\{\alpha_1, \ldots, \alpha_{2L}\}$ of $\vec{\alpha}$ which meets γ transversely. Now follow the intersections of γ with the α_i's:*

- *Whenever γ enters a type-z_i disc through a p_i-arc and leaves through a q_i-arc, write x_i.*

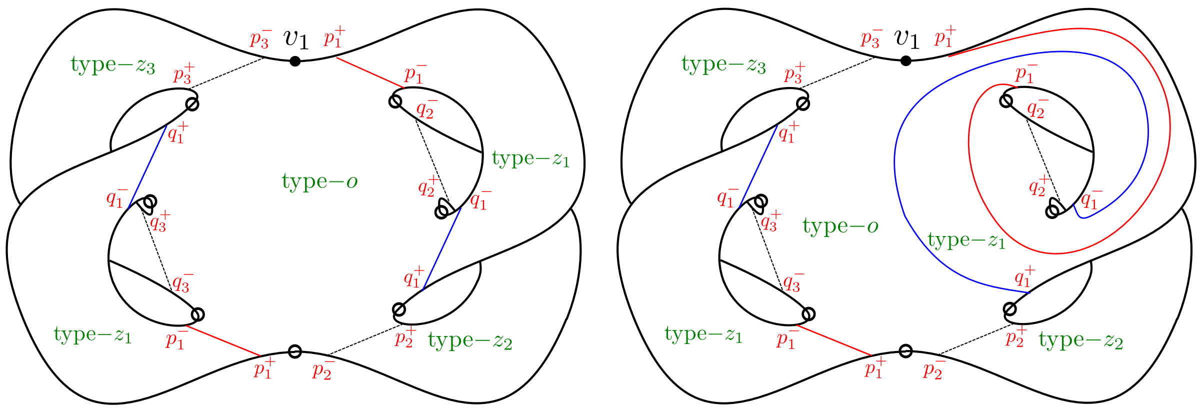

Figure 6.1: Two arc systems drawn on Σ of genus 2 and with one boundary component identified with $S^1(w)$ for the word $w = [x,y][x,z] = [x_1,x_2][x_1,x_3]$. The p_1-arcs are red, the q_1-arcs are blue and all the others are drawn in black. These two arc systems are distinct yet induce the same pair of matchings. Each of the associated CW-complexes has five discs: one of type-o, two of type-z_1, one of type-z_2 and one of type-z_3.

- *Whenever γ enters a type-z_i disc through a q_i-arc and leaves through a p_i-arc, write x_i^{-1}.*
- *Whenever γ enters and leaves a type-z_i disc through p_i-arcs, or enter and leaves through q_i-arcs, write nothing.*

The final result is $[f_{\vec{\alpha}}(\gamma)]$, albeit not necessarily in reduced form.

6.2 The Arc Poset of (Σ, f)

Definition 6.4. Let (Σ, f) be admissible for $w_1, \ldots, w_\ell$ and incompressible. The **arc poset of $(\mathbf{\Sigma}, \mathbf{f})$**, denoted $\mathcal{AP}(\Sigma, f)$, consists of the set of all arc systems for (Σ, f) together with the partial order $\preceq$ defined by $\mathcal{AP}(\Sigma, f)$ $\preceq$

$$\vec{\alpha} \preceq \vec{\beta}$$

whenever, for some representatives of $\vec{\alpha}$ and $\vec{\beta}$, the arcs of $\vec{\beta}$ are embedded entirely inside type-o discs of $\vec{\alpha}$.

Remark 6.5.

1. The type-o discs in the definition can be taken to be either open or closed (although the endpoints of the arcs, of course, are always contained in their boundaries). However, using closed discs is more convenient: some of the arcs can be left unchanged when moving from $\vec{\alpha}$ to $\vec{\beta}$.

2. Of course, if $\vec{\alpha} \preceq \vec{\beta}$ then for every representative $\{\alpha_1, \ldots, \alpha_{2L}\}$ of $\vec{\alpha}$ there is a representative $\{\beta_1, \ldots, \beta_{2L}\}$ of $\vec{\beta}$ with arcs embedded inside the type-o discs defined by $\{\alpha_1, \ldots, \alpha_{2L}\}$.

3. This rewiring of arcs is completely analogous to the one in Proposition 5.8. As we explained there, if $\vec{\alpha} \preceq \vec{\beta}$ then this rewiring corresponds to a unique set of colored non-crossing partitions of the arcs of $\vec{\alpha}$ inside its type-o discs.

4. An equivalent definition for the order $\preceq$ in $\mathcal{AP}(\Sigma, f)$ is the following: $\vec{\alpha} \preceq \vec{\beta}$ if and only if for some representatives of $\vec{\alpha}$ and $\vec{\beta}$, the arcs of $\vec{\alpha}$ are embedded entirely inside type-z_i discs of $\vec{\beta}$ (union of type-z_i discs for all i).

The following claim says, in particular, that the partial order we just defined is indeed an order:

Claim 6.6. *1. If $\vec{\alpha} \preceq \vec{\beta}$ and $\vec{\alpha} \neq \vec{\beta}$ then the number of type-o discs in $\vec{\beta}$ is strictly larger.*

2. *Moreover, the number of type-o discs can serve as a rank for the poset $\mathcal{AP}(\Sigma, f)$, which turns it into a* graded poset[31].

3. *If $\vec{\alpha} \preceq \vec{\beta}$ and $\vec{\beta} \preceq \vec{\gamma}$ then $\vec{\alpha} \preceq \vec{\gamma}$.*

Proof. (1) Let D be a type-o disc of $\vec{\alpha}$ where new arcs of $\vec{\beta}$ are introduced (namely, where the non-crossing partition is non-trivial). With the new arcs instead of the old ones, at least two of the regions of D are now disjoint type-o discs of $\vec{\beta}$, thus strictly increasing the total number of type-o discs. (The other effect is that the other areas in D now serve as "corridors", merging together several neighboring type-z_i discs, as in Figure 5.2.)

(2) One needs to show that if $\vec{\alpha}$ is covered by $\vec{\beta}$ (see footnote on Page 28), then $\vec{\beta}$ has exactly one more type-o disc than $\vec{\alpha}$. Let $\{\alpha_1, \ldots, \alpha_{2L}\}$ and $\{\beta_1, \ldots, \beta_{2L}\}$ be representatives with the β_i's contained in the type-o discs defined by the α_i's. Assume without loss of generality that β_1 is a genuine new arc (does not share the same two endpoints as any of the α_i's), which is contained inside the type-o disc D and meets at its two endpoints α_1 and α_2. It is evident that we can draw an arc β' embedded in D and disjoint from all the (interiors of) $\beta_1, \ldots, \beta_{2L}$, which connects the other endpoints of α_1 and α_2. Then $\vec{\gamma} = [\{\beta_1, \beta', \alpha_3, \ldots, \alpha_{2L}\}]$ clearly satisfies $\vec{\alpha} \prec \vec{\gamma} \preceq \vec{\beta}$, and by the covering assumption, $\vec{\gamma} = \vec{\beta}$. The number of type-$o$ discs in $\vec{\gamma}$ is clearly one larger than in $\vec{\alpha}$.

(3) This is true by an argument similar to the one in the proof of Proposition 5.8: if $\beta_1, \ldots, \beta_{2L}$ are contained inside type-o discs defined by $\{\alpha_1, \ldots, \alpha_{2L}\}$, then the union of type-o discs associated with $\{\beta_1, \ldots, \beta_{2L}\}$ is contained in the union of type-o discs associated with $\{\alpha_1, \ldots, \alpha_{2L}\}$. Thus, if $\gamma_1, \ldots, \gamma_{2L}$ are contained inside type-o discs defined by $\{\beta_1, \ldots, \beta_{2L}\}$, they are also contained inside type-o discs defined by $\{\alpha_1, \ldots, \alpha_{2L}\}$. □

Proposition 6.7. *The map $\Psi : \mathcal{AP}(\Sigma, f) \to \mathcal{PMP}(\Sigma, f)$ defined by $\vec{\alpha} \mapsto (\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}})$ is a graded poset surjective morphism*[32].

Proof. Let $\vec{\alpha} \preceq \vec{\beta}$ in $\mathcal{AP}(\Sigma, f)$ and let $\{\alpha_1, \ldots, \alpha_{2L}\}$ and $\{\beta_1, \ldots, \beta_{2L}\}$ be representatives so that $\beta_1, \ldots, \beta_{2L}$ are embedded inside the type-o discs defined by $\{\alpha_1, \ldots, \alpha_{2L}\}$. Using the isomorphism of CW-complexes $\Sigma_{\vec{\alpha}} \cong \Sigma_{(\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}})}$ we can use the same rewiring of the arcs inside type-o discs in $\Sigma_{\vec{\alpha}}$, to get a rewiring of matching-edges inside type-o discs of $\Sigma_{(\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}})}$. The resulting CW-complex is $\Sigma_{(\sigma_{\vec{\beta}}, \tau_{\vec{\beta}})}$. By Proposition 5.8, this means that $(\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}}) \preceq (\sigma_{\vec{\beta}}, \tau_{\vec{\beta}})$, hence Ψ is order preserving. Since the number of type-o discs can serve as a rank for both posets (Claims 5.4 and 6.6), Ψ is a graded-poset morphism. It is surjective because given $(\sigma, \tau) \in \mathcal{PMP}(\Sigma, f)$, the homeomorphism $\Sigma_{(\sigma,\tau)} \cong \Sigma$ which yields the equivalence $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right) \sim (\Sigma, f)$ can map the matching-edges in $\Sigma_{(\sigma,\tau)}$ to a valid arc system for (Σ, f), and this system is mapped by Ψ to (σ, τ). □

As before, we denote by $|\mathcal{AP}(\Sigma, f)|$ the (geometric realization of the) simplicial complex associated with $\mathcal{AP}(\Sigma, f)$ (see Definition 5.10). $|\mathcal{AP}(\Sigma, f)|$

[31] See footnote on Page 38.

[32] For our cause, a map $\varphi\colon (P_1, \leq) \to (P_2, \leq)$ between two graded posets is a graded-poset morphism if it preserves the order ($x \leq y \Rightarrow \varphi(x) \leq \varphi(y)$) and preserves the rank up to a constant shift: $\mathrm{rank}(\varphi(x)) = \mathrm{rank}(x) + c_0$.

Recall $\mathrm{MCG}(\Sigma)$, the mapping class group of Σ defined on Page 8. Clearly, the action of homeomorphisms of Σ relative $\partial\Sigma$ on sets of arcs $\{\alpha_1, \ldots, \alpha_{2L}\}$ as in Definition 6.1 descends to an action of $\mathrm{MCG}(\Sigma)$ on their isotopy classes, namely, on arc systems. In the following theorem we analyze this action:

Theorem 6.8. *1. The action $\mathrm{MCG}(\Sigma) \curvearrowright \mathcal{AP}(\Sigma, f)$ is a graded-poset free action*[33]*. The quotient is isomorphic to $\mathcal{PMP}(\Sigma, f)$ as a graded poset.*

2. The action $\mathrm{MCG}(\Sigma) \curvearrowright |\mathcal{AP}(\Sigma, f)|$ is a covering space action[34]*. The quotient is isomorphic to $|\mathcal{PMP}(\Sigma, f)|$ as a simplicial complex.*

Remark 6.9. Item 2 of Theorem 6.8 does not automatically follow from item 1. Consider, for example, the poset $P = \{x_1, x_2, y_1, y_2\}$ with order $x_i \prec y_j$ for every i and j, and the action of $G = \mathbb{Z}/2\mathbb{Z}$ on P by swapping x_1 with x_2 and y_1 with y_2. Whereas P/G is the poset $\{x \prec y\}$ and $|P/G|$ consists of two vertices and an edge connecting them, the quotient $|P|/G$ consists of two vertices with *two* edges connecting them, and is not even a simplicial complex. See Appendix A.2 for more details.

Proof. **Item 1:** It is clear that the action of $\mathrm{MCG}(\Sigma)$ on $\mathcal{AP}(\Sigma, f)$ preserves the number of discs of each type, which shows it preserves the rank of the elements. It is also clear that the action commutes with rewiring of arcs inside type-o discs, which shows it is order-preserving. Assume that $[\varphi] \in \mathrm{MCG}(\Sigma)$ fixes $\vec{\alpha} = [\{\alpha_1, \ldots, \alpha_{2L}\}] \in \mathcal{AP}(\Sigma, f)$. Since $[\varphi]$ and $\vec{\alpha}$ are defined up to $\mathrm{Homeo}_0(\Sigma)$, we can assume $\varphi \in \mathrm{Homeo}_\delta(\Sigma)$ fixes $\partial\Sigma \cup \alpha_1 \cup \ldots \cup \alpha_{2L}$ pointwise. Because the boundary of every disc D in Σ contains segments from $\partial\Sigma$, the homeomorphism φ maps D to itself, and is the identity on ∂D. But $\mathrm{MCG}(D)$ is trivial (by the Alexander Lemma, e.g. [FM12, Lemma 2.1]), and so $\varphi\big|_D$ is isotopic (inside D, relative to ∂D) to $\mathrm{id}\big|_D$. Thus φ is isotopic to the identity in the whole of Σ, and so $[\varphi]$ is trivial. This proves the action $\mathrm{MCG}(\Sigma) \curvearrowright \mathcal{AP}(\Sigma, f)$ is free.

To see the quotient is $\mathcal{PMP}(\Sigma, f)$, we need to show a correspondence between the orbits of the action and the elements of $\mathcal{PMP}(\Sigma, f)$. Note first that $\Psi(\vec{\alpha}) = (\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}})$ only depends on the endpoints of the arcs which sit at the boundary of Σ, and the elements of $\mathrm{MCG}(\Sigma)$ fix the boundary pointwise. Thus the action commutes with Ψ. On the other hand, if $\Psi(\vec{\alpha}) = \Psi(\vec{\beta})$, then the isomorphisms of CW-complexes $\varphi_{\vec{\alpha}} \colon \Sigma_{\vec{\alpha}} \xrightarrow{\cong} \Sigma_{(\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}})}$ and $\varphi_{\vec{\beta}} \colon \Sigma_{\vec{\beta}} \xrightarrow{\cong} \Sigma_{(\sigma_{\vec{\beta}}, \tau_{\vec{\beta}})} = \Sigma_{(\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}})}$ satisfy that $[\varphi_{\vec{\beta}}^{-1} \circ \varphi_{\vec{\alpha}}] \in \mathrm{MCG}(\Sigma)$ maps $\vec{\alpha}$ to $\vec{\beta}$. So, indeed, the orbits of the action $\mathrm{MCG}(\Sigma) \curvearrowright \mathcal{AP}(\Sigma, f)$ correspond to the elements of $\mathcal{PMP}(\Sigma, f)$. That $\mathcal{AP}(\Sigma,f)/\mathrm{MCG}(\Sigma) \cong \mathcal{PMP}(\Sigma, f)$ is an isomorphism of graded-posets now follows from the fact that Ψ is a graded-poset morphism (which is the content of Proposition 6.7).

Item 2: A simplicial action of a group G on (the geometric realization of) a simplicial complex K is a covering space action if and only if the action is free: there is clearly a neighborhood U_x for every point x such that if $g.x \neq x$ then $g.U_x \cap U_x = \emptyset$ (take U_x that does not intersect any closed simplices in the barycentric subdivision of K which do not contain x). In our case, the freeness of the action $\mathrm{MCG}(\Sigma) \curvearrowright |\mathcal{AP}(\Sigma, f)|$ on the vertices is proved in item 1. Since the action preserves ranks, it cannot mix different vertices of the same simplex, so if $g.s = s$ for some simplex s and $g \in \mathrm{MCG}(\Sigma)$, then necessarily g fixes the vertices of s, hence $g = \mathrm{id}$. So the action is free on all points.

[33] A group action is said to be a graded-poset action if is order-preserving and rank-preserving.

[34] Namely, every point in $|\mathcal{AP}(\Sigma, f)|$ has a neighborhood U so that $g.U \cap U = \emptyset$ for every $\mathrm{id} \neq g \in \mathrm{MCG}(\Sigma)$.

To see that $|\mathcal{AP}(\Sigma,f)|/\mathrm{MCG}(\Sigma) \cong |\mathcal{AP}(\Sigma,f)/\mathrm{MCG}(\Sigma)|$, we use Corollary A.7 from the Appendix. According to this corollary, it is enough to check that if $\vec{\alpha}_0 \prec \ldots \prec \vec{\alpha}_r$ in $\mathcal{AP}(\Sigma,f)$ and $g_0.\vec{\alpha}_0 \prec \ldots \prec g_r.\vec{\alpha}_r$ for some $g_0,\ldots,g_r \in \mathrm{MCG}(\Sigma)$, then there is a $g \in \mathrm{MCG}(\Sigma)$ with $g.\vec{\alpha}_i = g_i.\vec{\alpha}_i$ for every i. In fact, we show more: we show that in this case, necessarily $g_0 = g_1 = \ldots = g_r$. To prove this stronger property, it is enough to show it for a pair of elements, namely, that if $\vec{\alpha} \prec \vec{\beta}$ and $g.\vec{\alpha} \prec g'.\vec{\beta}$, then $g = g'$. By acting on the latter pair by g^{-1}, we get that $\vec{\alpha} \prec \left(g^{-1}g'\right).\vec{\beta}$. So, replacing $g^{-1}g'$ with g, we reduce to showing that if $\vec{\alpha} \prec \vec{\beta}$ and $\vec{\alpha} \prec g.\vec{\beta}$ then $g = \mathrm{id}$.

Consider again the isomorphism of CW-complexes $\varphi_{\vec{\alpha}}\colon \Sigma_{\vec{\alpha}} \xrightarrow{\cong} \Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}$. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be the unique sets of colored non-crossing partitions of the arcs in type-o discs of $\Sigma_{\vec{\alpha}}$ which yield $\vec{\beta}$ and $g.\vec{\beta}$, respectively. They both pass through the homeomorphism induced by $\varphi_{\vec{\alpha}}$ to the unique set of colored non-crossing partitions of type-o discs in $\Sigma_{(\sigma_{\vec{\alpha}},\tau_{\vec{\alpha}})}$ yielding $(\sigma_{\vec{\beta}},\tau_{\vec{\beta}}) = \Psi(\vec{\beta}) = \Psi(g.\vec{\beta})$. Thus, $\mathcal{P}_1 = \mathcal{P}_2$ and $\vec{\beta} = g.\vec{\beta}$. Using the freeness from item **1**, we obtain that $g = \mathrm{id}$. □

Example 6.10. We already analyzed above the pairs of matchings poset $\mathcal{PMP}(\Sigma,f)$ and the simplicial complex $|\mathcal{PMP}(\Sigma,f)|$ of the sole incompressible (Σ,f) which is admissible for $w = [x,y][x,z]$ (see example 4.14 as well as Pages 38 and 42). We saw that $|\mathcal{PMP}(\Sigma,f)|$ was a cycle (composed of 4 vertices and 4 edges). We already know that $|\mathcal{AP}(\Sigma,f)|$ is a covering space of $|\mathcal{PMP}(\Sigma,f)|$, so every connected component of it is either a cycle or an infinite line. In Figure 6.2 we show a piece of a connected component of $|\mathcal{AP}(\Sigma,f)|$ made of three elements of smallest rank together with two elements of one rank higher, forming together a path of four edges. By carefully analyzing this component, it is possible to see that it is actually homeomorphic to an infinite line, and by Theorem 6.8 it follows that all components are of the same form. The fact it is a line is an instance of Theorem 6.12 below.

The middle element in Figure 6.2 is the same as the left element in Figure 6.1. The right element in Figure 6.1 is yet another element of the same poset $\mathcal{AP}(\Sigma,f)$. It is easy to see (by, e.g., Claim 6.3) that this element induces a different homotopy class of maps to $\bigvee^r S^1$. By Theorem 6.12 below this means it belongs to a different connected component of $|\mathcal{AP}(\Sigma,f)|$. However, this element induces the same bijections as the middle element in Figure 6.2 and thus can be mapped to it by some mapping class in $\mathrm{MCG}(\Sigma)$ (a Dehn twist in this case).

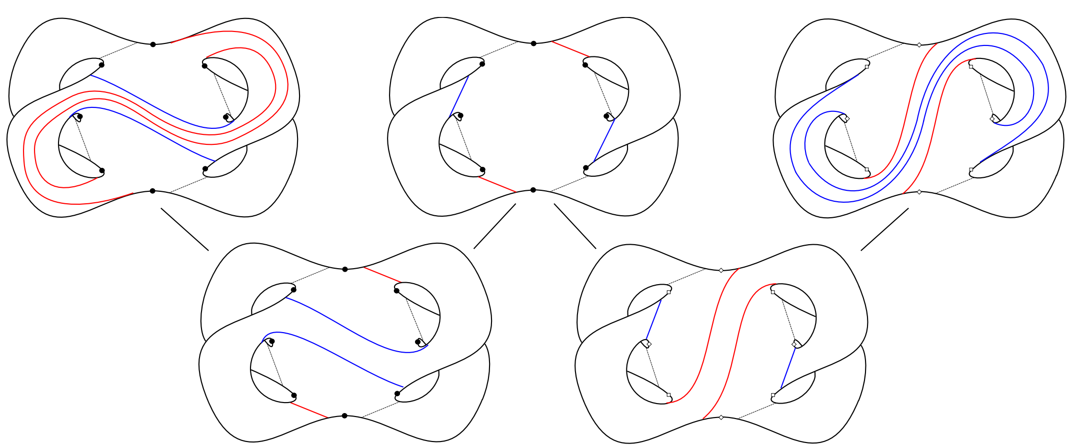

Figure 6.2: A series of five elements in the same connected component of the arc poset $\mathcal{AP}(\Sigma,f)$ of the sole equivalence class $[(\Sigma,f)]$ of incompressible map admissible for the word $[x,y][x,z]$. The red lines are p_1-arcs and the blue lines are q_1-arcs. The first and last elements differ by an element of $\mathrm{MCG}(\Sigma)$ and induce the same matchings $E^+ \xrightarrow{\sim} E^-$.

Example 6.11. Now consider $w = \left[x^2, y\right] = x_1 x_2 y_3 X_4 X_5 Y_6$. An easy computation yields that $\mathrm{Solu}\,(w)$ consists of exactly two equivalence classes. One $[(\Sigma_{1,1}, f)]$ is represented by the pair of matchings $\sigma = \tau = \begin{pmatrix} x_1 & x_2 & y_3 \\ X_5 & X_4 & Y_6 \end{pmatrix}$ and corresponds to the presentation of w as the commutator $\left[x^2, y\right]$; the other equivalence class $[(\Sigma_{1,1}, f')]$ is represented by the pair of matchings $\sigma = \tau = \begin{pmatrix} x_1 & x_2 & y_3 \\ X_4 & X_5 & Y_6 \end{pmatrix}$ and corresponds to the non-equivalent (under $\mathrm{Aut}_\delta\,(\mathbf{F}_2)$) presentation as $\left[x^2, yx\right]$. Both $|\mathcal{PMP}\,(\Sigma_{1,1}, f)|$ and $|\mathcal{PMP}\,(\Sigma_{1,1}, f')|$ are each an isolated point. It follows from Theorem 6.8 that $|\mathcal{AP}\,(\Sigma_{1,1}, f)|$ and $|\mathcal{AP}\,(\Sigma_{1,1}, f')|$ are also composed of isolated points. In fact, there are infinitely countably many of them in each of the two (this follows from Theorem 6.12 below). In Figure 6.3 we draw three elements from these two arc posets.

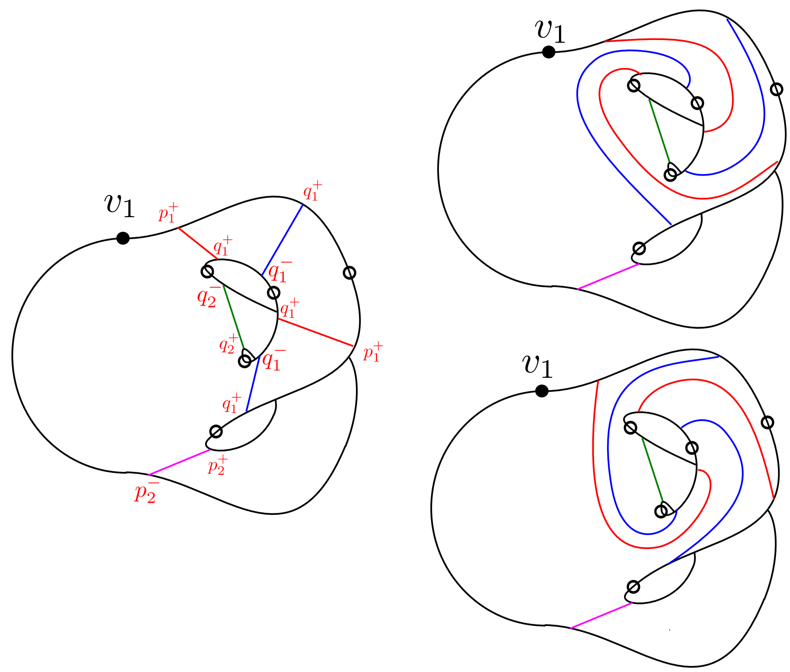

Figure 6.3: on the left: an element from $\mathcal{AP}\,(\Sigma_{1,1}, f)$, where $(\Sigma_{1,1}, f)$ is admissible for $w = \left[x^2, y\right]$ and corresponds to the solution $w = \left[x^2, y\right]$. On the right: two elements from $\mathcal{AP}\,(\Sigma_{1,1}, f')$, where $(\Sigma_{1,1}, f')$ is admissible for $w = \left[x^2, y\right]$ and corresponds to the solution $w = \left[x^2, yx\right]$. Clearly, the two arc systems on the right are in the same orbit of the action of $\mathrm{MCG}\,(\Sigma_{1,1})$.

In both examples the connected components of $|\mathcal{AP}\,(\Sigma, f)|$ are contractible: infinite lines in Example 6.10 and isolated points in Example 6.11. In particular, in both examples, every connected component is the universal covering space of the corresponding connected component of $|\mathcal{PMP}\,(\Sigma, f)|$. This turns out to be the general case:

Theorem 6.12. *The map $\mathcal{AP}\,(\Sigma, f) \to \left\{\Sigma \to \bigvee^r S^1\right\}$ given by $\vec{\alpha} \mapsto f_{\vec{\alpha}}$ induces a one-to-one correspondence between the connected components of $|\mathcal{AP}\,(\Sigma, f)|$ and the homotopy classes (relative $\partial\Sigma$) of maps $\Sigma \to \bigvee^r S^1$ which are equivalent to f:*

$$\pi_0\left(|\mathcal{AP}\,(\Sigma, f)|\right) \xrightarrow{\sim} \left\{ \begin{array}{c} \text{homotopy classes relative } \partial\Sigma \text{ of} \\ f' \colon \Sigma \to \bigvee^r S^1 \end{array} \middle| \, (\Sigma, f') \sim (\Sigma, f) \right\}.$$

Moreover, every connected component of $|\mathcal{AP}\,(\Sigma, f)|$ is contractible.

Recall that homotopy classes relative $\partial\Sigma$ of maps $\Sigma \to \bigvee^r S^1$ are in one-to-one correspondence with the homomorphisms of "fundamental groupoid" (Lemma 2.1). In particular, when

$\ell = 1$, if $g = \mathrm{cl}(w_1)$, the correspondence in Theorem 6.12 can be interpreted as a one-to-one correspondence between $\pi_0(|\mathcal{AP}(\Sigma, f)|)$ and the elements in the $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$-orbit of f_* in $\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$.

The proof of Theorem 6.12 is the most technical in the paper, and we postpone it to Section 6.3. We first explain how it readily yields Theorem 5.12 and thus our main results.

Proof of Theorem 5.12 given Theorem 6.12: Recall that to show that $|\mathcal{PMP}(\Sigma, f)|$ is a $\mathrm{K}(G,1)$-space for $G = \mathrm{Stab}_{\mathrm{MCG}(\Sigma)}(\tilde{f})$, one needs to establish that $|\mathcal{PMP}(\Sigma, f)|$ is path-connected, that its fundamental group is isomorphic to G and that its universal covering is contractible.

By Theorem 6.8, $|\mathcal{AP}(\Sigma, f)|$ is a covering space of $|\mathcal{PMP}(\Sigma, f)|$. In particular, so is every connected component of $|\mathcal{AP}(\Sigma, f)|$. For instance, by Theorem 6.12, we can take the connected component of $|\mathcal{AP}(\Sigma, f)|$ corresponding to $\tilde{f}$, the homotopy class of f. Denote this component by C. By Theorem 6.12 again, C is contractible and therefore the universal covering of $|\mathcal{PMP}(\Sigma, f)|$.

The subgroup of $\mathrm{MCG}(\Sigma)$ of elements mapping C to itself are precisely those preserving $\tilde{f}$, namely, precisely $G = \mathrm{Stab}_{\mathrm{MCG}(\Sigma)}(\tilde{f})$. Thus, the action of $\mathrm{MCG}(\Sigma)$ on $|\mathcal{AP}(\Sigma, f)|$ restricts to the action of G on C. This action is precisely the covering action, hence $G \cong \pi_1(|\mathcal{PMP}(\Sigma, f)|)$.

Finally, to show $|\mathcal{PMP}(\Sigma, f)|$ is path-connected, it is enough to show there is a path between any two of its vertices. Let $(\sigma, \tau) \in \mathcal{PMP}(\Sigma, f)$ be a pair of matchings. By definition, since $(\Sigma, f) \sim \left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$, there is a homeomorphism $\rho\colon \Sigma \to \Sigma_{(\sigma,\tau)}$ with $f \simeq f_{(\sigma,\tau)} \circ \rho$ homotopic. We can use the image through ρ^{-1} of the matching-edges in $\Sigma_{(\sigma,\tau)}$ to get an arc system $\vec{\alpha} \in \mathcal{AP}(\Sigma, f)$ with $f_{\vec{\alpha}} \simeq f$ homotopic, and so that $\Psi(\vec{\alpha}) = (\sigma, \tau)$ (see the notation from Proposition 6.7). But $(\sigma, \tau) \in \mathcal{PMP}(\Sigma, f)$ was arbitrary, and we can, likewise, obtain $\vec{\beta} \in \mathcal{AP}(\Sigma, f)$ with $f_{\vec{\beta}} \simeq f$ and $\Psi(\vec{\beta}) = (\sigma', \tau')$ for any $(\sigma', \tau') \in \mathcal{PMP}(\Sigma, f)$. By Theorem 6.12, $\vec{\alpha}$ and $\vec{\beta}$ belong to the same connected component of $|\mathcal{AP}(\Sigma, f)|$ (specifically, to C, the one corresponding to $\tilde{f}$). We can now take any path between them in C and project it to a path between (σ, τ) and (σ', τ') in $|\mathcal{PMP}(\Sigma, f)|$. □

6.3 Contractability of connected components

We now come to prove Theorem 6.12, regarding the connected components of $\mathcal{AP}(\Sigma, f)$. Let

$$\Upsilon\colon \pi_0(|\mathcal{AP}(\Sigma, f)|) \to \left\{ \begin{matrix} \text{homotopy classes relative } \partial\Sigma \text{ of} \\ f'\colon \Sigma \to \bigvee^r S^1 \end{matrix} \;\middle|\; (\Sigma, f') \sim (\Sigma, f) \right\}$$

be the map defined on every connected component C by taking an arbitrary vertex $\vec{\alpha} \in C$ and mapping C to $f_{\vec{\alpha}}$. We need to show that Υ is a well-defined bijection, and that every such C is contractible.

Lemma 6.13. Υ *is well-defined.*

Proof. To see that Υ is well-defined, it is enough to show that if $\vec{\beta}$ covers $\vec{\alpha}$ in $\mathcal{AP}(\Sigma, f)$, then $f_{\vec{\beta}} \simeq f_{\vec{\alpha}}$ are homotopic. This is shown by an argument we already used in Section 5: in this case, there is a particular type-z_i disc D defined by $\vec{\beta}$, and two equally-colored arcs at its boundary, say β_1 and β_2, which are replaced by α_1 and α_2 to obtain $\vec{\alpha} = [\{\alpha_1, \alpha_2, \beta_3, \ldots, \beta_{2L}\}]$. We can take both $f_{\vec{\alpha}}$ and $f_{\vec{\beta}}$ to be constant (and identical) on all arcs $\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3, \ldots, \beta_{2L}$. Since these arcs cut Σ to discs, Lemma 2.1 shows $f_{\vec{\alpha}} \simeq f_{\vec{\beta}}$. □

Lemma 6.14. Υ *is onto.*

Proof. Let $f' \colon \Sigma \to \bigvee^r S^1$ satisfy $(\Sigma, f') \sim (\Sigma, f)$. We want to show there is an arc system $\vec{\beta} \in \mathcal{AP}(\Sigma, f)$ with $f_{\vec{\beta}} \simeq f'$. First, note we have already seen that $\mathcal{PMP}(\Sigma, f)$ is non-empty (Lemma 4.9), and thus nor is $\mathcal{AP}(\Sigma, f)$ (Proposition 6.7). So there is some $\vec{\alpha} \in \mathcal{AP}(\Sigma, f)$. Now we can repeat an argument we used in the very end of Section 6.2: by definition, $(\Sigma, f') \sim (\Sigma, f_{\vec{\alpha}})$, so there is a homeomorphism $\rho \colon \Sigma \to \Sigma$ with $f' \simeq f_{\vec{\alpha}} \circ \rho$ homotopic. The arc system $\vec{\beta} = \rho^{-1}(\vec{\alpha})$ now satisfies $f_{\vec{\beta}} \simeq f'$. □

We are left to show that Υ is injective and that every connected component of $|\mathcal{AP}(\Sigma, f)|$ is contractible. Although the former is easier than the latter, we prove both at once. Consider the subposet $\mathcal{P}(f)$

$$\mathcal{P}(f) \stackrel{\text{def}}{=} \{\vec{\alpha} \in \mathcal{AP}(\Sigma, f) \mid f_{\vec{\alpha}} \simeq f\} \subseteq \mathcal{AP}(\Sigma, f).$$

We show that $|\mathcal{P}(f)|$ is connected and, moreover, contractible. Since f is arbitrary (if $(\Sigma, f') \sim (\Sigma, f)$, then $\mathcal{AP}(\Sigma, f) = \mathcal{AP}(\Sigma, f')$ and we could work just as well with f'), this yields that the same is true for any $f' \colon \Sigma \to \bigvee^r S^1$ with $(\Sigma, f') \sim (\Sigma, f)$, and thus proves Theorem 6.12.

It already follows from Lemmas 6.13 and 6.14 that $|\mathcal{P}(f)|$ is a non-empty collection of connected components of $|\mathcal{AP}(\Sigma, f)|$. It is left to show it consists of a single component, and that this component is contractible.

Guide-arcs

Fix $\vec{\alpha}_0 \in \mathcal{P}(f)$ (so $f_{\vec{\alpha}_0} \simeq f$). Let $\{\alpha_1, \ldots, \alpha_{2L}\}$ be a representative of $\vec{\alpha}_0$.

Definition 6.15. A finite set of arcs $\gamma_1, \ldots, \gamma_M$ embedded in Σ is said to be **a set of guide-arcs** for $\{\alpha_1, \ldots, \alpha_{2L}\}$ if

- the γ_m's are disjoint from each other and from the α_i's, and
- the only arc system in $\mathcal{AP}(\Sigma, f)$ with a representative which is disjoint from $\gamma_1 \cup \ldots \cup \gamma_M$ is $\vec{\alpha}$.

Every (representative of an) arc system has a set of guide-arcs: for example, for every arc α in the system take two guide arcs which follow α very closely, one from each side, in a parallel fashion. Figure 6.4 illustrates a set of guide-arcs of size five for an element of $\mathcal{AP}(\Sigma, f)$ where $[(\Sigma, f)] \in \mathrm{Solu}([x, y][x, z])$.

Given a set of guide-arcs $\gamma_1, \ldots, \gamma_M$, let $\mathcal{P}_m$, $0 \le m \le M$, denote the subposet of $\mathcal{P}(f)$ consisting of arc systems which have a representative which does not cross $\gamma_{m+1}, \ldots, \gamma_M$ (but may cross $\gamma_1, \ldots, \gamma_m$). So $\mathcal{P}_m$

$$\{\vec{\alpha}_0\} = \mathcal{P}_0 \subseteq \mathcal{P}_1 \subseteq \ldots \subseteq \mathcal{P}_M = \mathcal{P}(f)$$

is an increasing sequence of posets. Consider, for example, the set of guide arcs given in Figure 6.4, and denote the five elements in Figure 6.2, from left to right, by $\vec{\alpha}_{-2}$, $\vec{\alpha}_{-1}$, $\vec{\alpha}_0$, $\vec{\alpha}_1$ and $\vec{\alpha}_2$. Then, $\mathcal{P}_0 = \{\vec{\alpha}_0\}$, $\mathcal{P}_1 = \mathcal{P}_2 = \{\vec{\alpha}_0, \vec{\alpha}_1\}$, $\mathcal{P}_3 = \{\vec{\alpha}_0, \vec{\alpha}_1, \vec{\alpha}_2\}$ and $\mathcal{P}_4 = \mathcal{P}_5 = \mathcal{P}(f)$ contain the entire connected component of $|\mathcal{AP}(\Sigma, f)|$ (the component a piece of which is given in Figure 6.2). We stress that there may be many more elements in $\mathcal{AP}(\Sigma, f)$ with representatives which do not cross subsets of the guide-arcs (for instance, the arc system in the right hand side of Figure 6.1 does not cross γ_2, γ_3 nor γ_5), but they do not belong to $\mathcal{P}(f)$, and thus nor to the $\mathcal{P}_m$'s.

We shall prove the contractability of $|\mathcal{P}(f)|$ by showing that each $|\mathcal{P}_m|$ deformation retracts to $|\mathcal{P}_{m-1}|$.

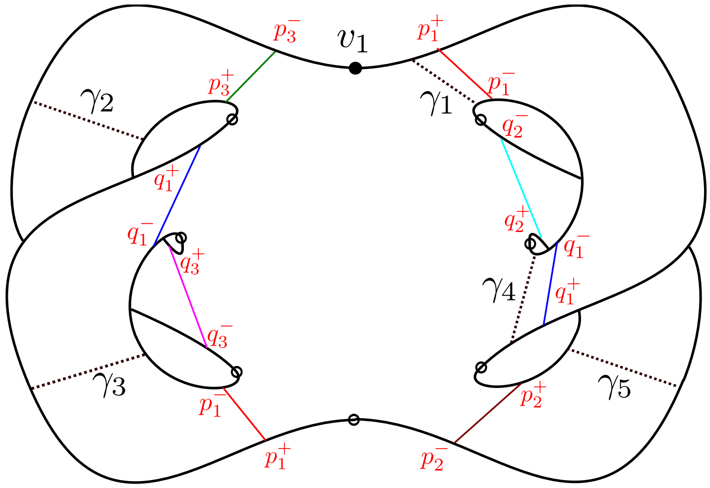

Figure 6.4: A set of guide-arcs (marked in dotted lines) for an element of $\mathcal{AP}(\Sigma, f)$ with $[(\Sigma, f)] \in \mathrm{Solu}([x, y][x, z])$. This element is the central one in Figure 6.2.

Depth of words along guide-arcs

Fix an arbitrary orientation for each guide-arc γ_m. For every $\vec{\alpha} \in \mathcal{P}(f)$ find a representative which meets the guide-arcs transversely and in minimal position (so that no arc of $\vec{\alpha}$ crosses twice in a row the same guide-arc). Define $u_m(\vec{\alpha})$ to be a word in the alphabet $\{p_1, q_1, \ldots, p_r, q_r\}$ $u_m(\vec{\alpha})$
which describes the sequence of crossings between γ_m and $\vec{\alpha}$: simply follow γ_m according to the given orientation, whenever it crosses a p_i-arc write p_i, and whenever it crosses a q_i-arc, write q_i. In this language,

$$\mathcal{P}_m = \{\vec{\alpha} \in \mathcal{P}(f) \,|\, u_{m+1}(\vec{\alpha}) = u_{m+2}(\vec{\alpha}) = \ldots = u_M(\vec{\alpha}) \text{ are all the empty word}\}.$$

Lemma 6.16. *For every $\vec{\alpha} \in \mathcal{P}(f)$ and every $1 \leq m \leq M$, the formal word $u_m(\vec{\alpha})$ can be reduced to the empty word by a series of deletions of subwords $p_i p_i$ and $q_i q_i$.*

Proof. Recall that for any path $\overline{\gamma}$ in Σ from v_i to v_j $(i, j \in [\ell])$ which meets the arcs of $\vec{\alpha}$ transversely, the value of $[f_{\vec{\alpha}}(\overline{\gamma})]$ is determined by the sequence of crossings between $\overline{\gamma}$ and the arcs, as detailed in Claim 6.3. It is easy to see that an equivalent way to define $[f_{\vec{\alpha}}(\overline{\gamma})]$ is the following: write a word in $\{p_1, q_1, \ldots, p_r, q_r\}$ which depicts the sequence of crossings of $\overline{\gamma}$ with the arcs of $\vec{\alpha}$ (as in the definition of $u_m(\vec{\alpha})$), then reduce this word by deleting subwords of the form $p_i p_i$ or $q_i q_i$, and eventually scan the word from beginning to end and replace every $p_i q_i$ with x_i and every $q_i p_i$ with x_i^{-1}. (It is standard that the order of reductions does not effect the final result.)

Now let $\gamma = \gamma_m$ for some m starting at the boundary component i and arriving at the boundary component j. Let $\overline{\gamma}$ be a path in Σ which begins at v_i, then goes along $\partial\Sigma$ from v_i to the beginning of γ, then goes along γ, and then arrives to v_j through $\partial\Sigma$ (the parts trough $\partial\Sigma$ can be chosen arbitrarily). Since the arcs of an arc system always meet the boundary only at their endpoints, the sequence of crossings along the pieces of $\overline{\gamma}$ at the boundary are the same for all elements of $\mathcal{AP}(\Sigma, f)$. Since $[f_{\vec{\alpha}}(\overline{\gamma})] = [f_{\vec{\alpha}_0}(\overline{\gamma})]$, the words $u_m(\vec{\alpha})$ and $u_m(\vec{\alpha}_0)$ must be equivalent (through reductions). We are done as $u_m(\vec{\alpha}_0)$ is empty by the definition of guide-arcs. □

For example, for $\vec{\alpha} \in \mathcal{AP}(\Sigma, f)$ the element in Figure 6.1 on the right and the element $\vec{\alpha}_0$ and guide arcs in Figure 6.4, $u_1(\vec{\alpha}) = q_1 p_1$ and thus $\vec{\alpha} \notin \mathcal{P}(f_{\vec{\alpha}_0})$.

Next, we define the depth of $u_m(\vec{\alpha})$. Let $\mathbb{T}_{2r,2}$ be the infinite $(2r,2)$-biregular tree[35]. We think of it as the universal cover of the graph $\bigvee^r S^1$, where the point o and the points $\{z_i \mid i \in [r]\}$ are vertices. We also label every vertex of $\mathbb{T}_{2r,2}$ by o or z_i according to the vertex it covers, and every edge of $\mathbb{T}_{2r,2}$ by p_i or q_i, according to the marked point contained in the edge of $\bigvee^r S^1$ it covers. $\mathbb{T}_{2r,2}$

Since γ_m is disjoint from the arcs of $\vec{\alpha}_0$, it is completely embedded in a (closed, type-o or type-z_i) disc of $\vec{\alpha}_0$. If this disc is type-o (type-z_i), then γ_m begins and ends in a type-o (type-z_i, respectively) disc in any arc system in $\mathcal{AP}(\Sigma, f)$. If it begins and ends in a type-o (type-z_i) disc, we choose a basepoint $\otimes_m$ for $\mathbb{T}_{2r,2}$ in some o-vertex (z_i-vertex, respectively). We can think of $u_m(\vec{\alpha})$ as a path in the tree: we begin at the basepoint $\otimes_m$, whenever we write p_i, we traverse a p_i-edge, and whenever we write q_i we traverse a q_i-edge. It it easy to verify that we never get stuck (if our walk reaches a z_i-vertex, the following step will necessarily be a p_i or a q_i with the same i). Moreover, $u_m(\vec{\alpha})$ reduces to the empty word if and only if the associated walk in the tree is closed. $\otimes_m$

We define **the depth of** $u_m(\vec{\alpha})$, denoted $\operatorname{depth}(u_m(\vec{\alpha}))$, to be the largest distance from the basepoint $\otimes_m$ of a vertex in $\mathbb{T}_{2r,2}$ visited in the walk of $u_m(\vec{\alpha})$. For example, in the following word we write the distance from the basepoint to the vertex visited after every step: $\operatorname{depth}(u_m(\vec{\alpha}))$

$$^0p_1{}^1q_1{}^2p_2{}^3p_2{}^2q_3{}^3q_3{}^2q_4{}^3p_4{}^4q_4{}^5q_4{}^4p_4{}^3q_4{}^2q_1{}^1q_1{}^2q_1{}^1p_1{}^0$$

hence the depth of this word is 5.

This notion of depth allows us to define a finer sequence of nested subposets $\mathcal{P}_{m,n}$ ($1 \le m \le M$ and $n \in \mathbb{Z}_{\ge 0}$) as follows: $\mathcal{P}_{m,n}$

$$\mathcal{P}_{m,n} \overset{\text{def}}{=} \left\{ \vec{\alpha} \in \mathcal{P}(f) \;\middle|\; \begin{array}{c} \operatorname{depth}(u_m(\vec{\alpha})) \le n, \text{ and} \\ u_{m+1}(\vec{\alpha}) = u_{m+2}(\vec{\alpha}) = \ldots = u_M(\vec{\alpha}) \text{ are all the empty word} \end{array} \right\}.$$

So

$$\mathcal{P}_{m-1} = \mathcal{P}_{m,0} \subseteq \mathcal{P}_{m,1} \subseteq \ldots \subseteq \mathcal{P}_{m,n} \subseteq \ldots \subseteq \mathcal{P}_m,$$

and

$$\bigcup_{n=0}^{\infty} \mathcal{P}_{m,n} = \mathcal{P}_m.$$

For instance, if we continue with the example of $w = [x,y][x,z]$, the five guide-arcs drawn in Figure 6.4 and the five elements $\vec{\alpha}_{-2}, \ldots, \vec{\alpha}_2$ in Figure 6.2, then $\mathcal{P}_0 = \mathcal{P}_{1,0} = \{\vec{\alpha}_0\}$ and $\mathcal{P}_{1,1} = \mathcal{P}_{1,2} = \ldots = \mathcal{P}_1 = \mathcal{P}_{2,0} = \{\vec{\alpha}_0, \vec{\alpha}_1\}$. "Opening" γ_2 does not add elements so $\mathcal{P}_{2,n} = \mathcal{P}_2 = \mathcal{P}_{3,0} = \{\vec{\alpha}_0, \vec{\alpha}_1\}$ for every n. When we allow words of depth 1 on γ_3 we get $\mathcal{P}_{3,1} = \{\vec{\alpha}_0, \vec{\alpha}_1, \vec{\alpha}_2\}$, but allowing bigger depth there without "opening" γ_4 does not add any elements, so $\mathcal{P}_{3,n} = \mathcal{P}_3 = \mathcal{P}_{4,0}$ for every $n \ge 1$. The subposet $\mathcal{P}_{4,1}$ already contains, in addition, $\vec{\alpha}_{-1}$ as well as the element to the right of $\vec{\alpha}_2$ which we may denote by $\vec{\alpha}_3$. The leftmost element in Figure 6.2, $\vec{\alpha}_{-2}$, is contained only in $\mathcal{P}_{4,2}$, and so does "$\vec{\alpha}_4$". This goes on: $\mathcal{P}_{4,n}$ consists of $\mathcal{P}_{4,n-1}$ together with one more element to the right and one more element to the left in the component a piece of which is given in Figure 6.2. Finally, $\mathcal{P}_4 = \mathcal{P}_{5,n} = \mathcal{P}_5 = \mathcal{P}(f)$ for every n.

Using Corollary A.4, we now show that $|\mathcal{P}_{m,n}|$ deformation retracts to $|\mathcal{P}_{m,n-1}|$. Namely, we show there is a map $|\mathcal{P}_{m,n}| \to |\mathcal{P}_{m,n-1}|$ which restricts to the identity in $|\mathcal{P}_{m,n-1}|$ and is homotopic to the identity in $|\mathcal{P}_{m,n}|$, through an homotopy that fixes $|\mathcal{P}_{m,n-1}|$ pointwise. Showing this means that $\mathcal{P}_m$ deformation retracts to $\mathcal{P}_{m-1}$, and thus completes the proof. (To

[35] A $(2r,2)$-biregular tree has vertices of degrees $2r$ and 2. Every vertex of degree $2r$ is connected only with vertices of degree 2, and vice-versa.

be sure: we can let the deformation retract $|\mathcal{P}_{m,n}| \to |\mathcal{P}_{m,n-1}|$ take place at time $\left[\frac{1}{2^n}, \frac{1}{2^{n-1}}\right]$. This is a well-defined deformation retract $|\mathcal{P}_m| \to |\mathcal{P}_{m-1}|$ since every point in $\mathcal{P}_m$ belongs to some $\mathcal{P}_{m,n}$, and the retracts of $|\mathcal{P}_{m,n+1}|, |\mathcal{P}_{m,n+2}|, \ldots$ leave $|\mathcal{P}_{m,n}|$ fixed pointwise.)

A deformation retract $|\mathcal{P}_{m,n}| \to |\mathcal{P}_{m,n-1}|$

The retract $|\mathcal{P}_{m,n}| \to |\mathcal{P}_{m,n-1}|$ is defined by a map $h_{m,n} : \mathcal{P}_{m,n} \to \mathcal{P}_{m,n-1}$ which prunes all leaves $h_{m,n}$
of depth n in the walk $u_m(\vec{\alpha})$ for every $\vec{\alpha} \in \mathcal{P}_{m,n}$. The basic idea is that if $\vec{\alpha} \in \mathcal{P}_{m,n} \setminus \mathcal{P}_{m,n-1}$ then $u_m(\vec{\alpha})$ has at least one leaf of depth n. Every such leaf means that γ_m crosses two equally-colored arcs of $\vec{\alpha}$ in a row, and we can "rewire" these two arcs locally to prune the leaf, as in Figure 6.5. We remark that in every such step, $\vec{\alpha}$ is modified to some *comparable* $\vec{\beta}$, so $\vec{\beta}$ is in the same connected component of $|\mathcal{AP}(\Sigma, f)|$ as $\vec{\alpha}$. By successive steps of this kind we can decrease the depth of all $u_m(\vec{\alpha})$ until they are all empty and we arrive at $\vec{\alpha}_0$. This alone suffices to show the connectivity of $|\mathcal{P}(f)|$.

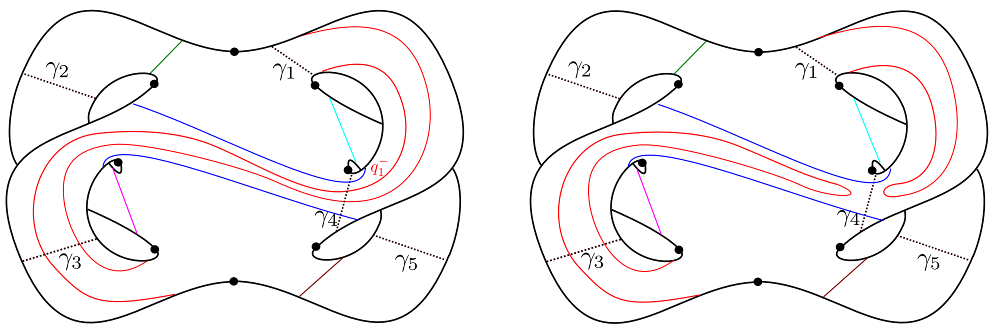

Figure 6.5: Pruning (from left to right) a leaf node of depth 2 in $u_4(\vec{\alpha}_{-2})$, where $\vec{\alpha}_{-2}$ is the left most arc system in Figure 6.2. The result is $\vec{\alpha}_{-1}$, the left of center arc system in Figure 6.2. We use the guide-arcs from Figure 6.4. This pruning is the resulting of applying $h_{4,2}$ on $\vec{\alpha}_{-2}$.

More formally, fix m and n and consider all leaves of depth n in $u_m(\vec{\alpha})$ (every visit of the walk to a vertex of distance n from $\otimes_m$ is considered a leaf). Every such leaf corresponds to some backtracking move $p_i p_i$ or $q_i q_i$, and we consider the segment of γ_m which lies between these two crossings (between the two crossings with p_i-arcs of $\vec{\alpha}$, or two crossings with q_i-arcs of $\vec{\alpha}$). From the point of view of the arc system $\vec{\alpha}$, these segments of γ_m correspond to disjoint arcs, which we call **γ-arcs**, inside the discs of $\vec{\alpha}$. Each γ-arc meets the boundary of the disc γ-arcs
only at its endpoints, and at two equally-colored $\vec{\alpha}$-arcs. Moreover, the γ-arcs never cross each other as γ_m is embedded in Σ (and does not self-intersect). In addition, all vertices at distance n from the basepoint $\otimes_m$ in $\mathbb{T}_{2r,2}$ are of type-o, or all are of type-z_i (not necessarily the same i for all vertices), depending solely on the parity of n. In the former case, all γ-arcs are contained in type-o discs; in the latter in type-z_i discs. From now on we assume that n is such that the γ-arcs are all contained in type-o discs, the other case being completely analogous.

For every type-o disc D of $\vec{\alpha}$ ($\vec{\alpha} \in \mathcal{P}_{m,n}$), the γ-arcs determine a partition P_D of the arcs in (the boundary of) D: this is the finest partition such that any two arcs connected by a γ-arc belong to the same block. We claim that P_D is colored and non-crossing. The monochromaticity of blocks stems from the fact that the γ-arcs correspond to subwords of the form $p_i p_i$ or $q_i q_i$ for some i. The partition P_D is non-crossing because the γ-arcs are disjoint. We define $h_{m,n}(\vec{\alpha})$ to $h_{m,n}(\vec{\alpha})$
be the arc system obtained from $\vec{\alpha}$ by the set of partitions P_D of its type-o discs (see Definition 6.4 and Remark 6.5).

It is evident that $h_{m,n}\Big|_{\mathcal{P}_{m,n-1}}$ is the identity, and that $\vec{\alpha} \preceq h_{m,n}(\vec{\alpha})$ for every $\vec{\alpha} \in \mathcal{P}_{m,n}$. Moreover, we claim that indeed $h_{m,n}(\mathcal{P}_{m,n}) \subseteq \mathcal{P}_{m,n-1}$: to see this, we show that the modification we made to obtain $h_{m,n}(\vec{\alpha})$ from $\vec{\alpha}$ prunes all backtracking steps of $u_m(\vec{\alpha})$ which correspond to leaves at depth n and does not introduce any new steps in $u_m(\vec{\alpha})$ or in $u_{m'}(\vec{\alpha})$ for any m'. (In contrast, $h_{m,n}$ may prune backtracking steps at depth smaller than n in $u_m(\vec{\alpha})$ or at any depth in $u_{m'}(\vec{\alpha})$ for $m' < m$). First, if η is any γ-arc in $\Sigma_{\vec{\alpha}}$ corresponding to a backtracking step at distance n, it necessarily enters and exists D through two arcs in the same block of P_D and these two crossings disappear in $h_{m,n}(\vec{\alpha})$, hence this leaf is indeed pruned. Second, any piece η of the arc $\gamma_{m'}$ for some $m' \leq m$ which is allocated by two successive crossings of $\vec{\alpha}$-arcs in $\Sigma_{\vec{\alpha}}$ and which is contained in a type-o disc D of $\vec{\alpha}$ satisfies the following:

- If η enters and exists D through two arcs in the same block of P_D, then these two crossings disappear in $h_{m,n}(\vec{\alpha})$, and the corresponding subword p_ip_i (or q_iq_i) of $u_{m'}(\vec{\alpha})$ is reduced.
- If η enters and exists D through two arcs e_1 and e_2 in two different blocks (or "z_i-corridors") B_1 and B_2, respectively, of P_D, then it necessarily does not cross any other block. I.e., there cannot be two other arcs, e_3 and e_4 at the same block B_3 of P_D, $B_3 \neq B_1, B_2$, with the cyclic order of the four being e_1, e_3, e_2, e_4, because η does not intersect the γ-arcs. Thus, in minimal position, the only crossings of η with arcs in $h_{m,n}(\vec{\alpha})$ are with the arc through which it leaves B_1 and then through the arc through which it enters B_2. By definition of P_D, the first arc has the same color as e_1, and the second arc has the same color as e_2. Thus, in this case, there is no change to the part of $u_{m'}(\vec{\alpha})$ corresponding to η, when moving from $\vec{\alpha}$ to $h_{m,n}(\vec{\alpha})$.

There are also pieces of $\gamma_{m'}$ at its very beginning or very end which may be contained in type-o discs of $\vec{\alpha}$. The same argument shows there is no change in the subword of $u_{m'}(\vec{\alpha})$ read along such segments when applying $h_{m,n}(\vec{\alpha})$.

By Corollary A.4, if we want to show that $h_{m,n}$ induces a deformation retract $|h_{m,n}| : |\mathcal{P}_{m,n}| \to |\mathcal{P}_{m,n-1}|$, we have left to show that $h_{m,n}$ is order-preserving. So assume $\vec{\alpha} \preceq \vec{\beta}$, and both are in $\mathcal{P}_{m,n}$. We need to show that $h_{m,n}(\vec{\alpha}) \preceq h_{m,n}(\vec{\beta})$. This follows from two properties expressed in the following two lemmas:

Lemma 6.17. *Let $\vec{\alpha} \preceq \vec{\beta}$ in $\mathcal{P}(f)$. We divide the word $u_m(\vec{\alpha})$ to subwords $x_1, \ldots, x_t$ by grouping together successive crossings with arcs at the boundary of the same type-o disc. So $u_m(\vec{\alpha}) = x_1 * x_2 * \ldots * x_t$, with $*$ denoting concatenation and each x_j of length 2 except for, possibly, x_1 and x_t, which may be of length 1. Each x_j corresponds to a segment η_j of γ_m (allocated by the two crossings). Since the type-o discs of $\vec{\beta}$ can be thought of as being contained inside the type-o discs of $\vec{\alpha}$, we let y_j ($1 \leq j \leq t$) be the subword of $u_m(\vec{\beta})$ which corresponds to η_i and then $u_m(\vec{\beta}) = y_1 * \ldots * y_t$. We claim that for every j, the vertex in $\mathbb{T}_{2r,2}$ that the walk of $u_m(\vec{\beta})$ visits at the beginning of y_j, is the same as the vertex visited by $u_m(\vec{\alpha})$ at the beginning of x_j.*

Proof. It is enough to show that x_j and y_j are equivalent through reduction for every j. Indeed, assume that x_j corresponds to the type-o disc D of $\vec{\alpha}$ and that the partition of this disc inside the set of partitions leading from $\vec{\alpha}$ to $\vec{\beta}$ is P_D. Because P_D is non-crossing, there is a clear order on the set of blocks of P_D (or "z_i-corridors") crossed by η_j (η_j has to exit a block immediately after entering it, before entering the next block). We are done as entering and exiting a block of P_D corresponds to a pair of backtracking steps in y_j. □

Lemma 6.18. *Assume that $\vec{\alpha} \preceq \vec{\beta}$ in $\mathcal{P}_{m,n}$. Let η be a γ-arc in $\vec{\alpha}$. Assume that the $\vec{\beta}$-arcs intersected by η are $\beta_1, \beta_2, \ldots, \beta_{2\ell}$. Then they are all of the same color and represent ℓ leaves of depth n in $\vec{\beta}$.*

Note that since η begins and ends in (the boundary of) type-z_i discs of $\vec{\alpha}$, it must indeed intersect an even number of arcs of $\vec{\beta}$ (recall that the type-o discs of $\vec{\beta}$ can be assumed to be contained in type-o discs of $\vec{\alpha}$). Of course, $\ell = 0$ is possible.

Proof. Assume $\ell > 0$ (otherwise the statement is trivial). Let D be the type-o disc of $\vec{\alpha}$ in which η is embedded. Since η represents a leaf in $u_m(\vec{\alpha})$, it enters and exits D through equally-colored arcs α_1 and α_2, and assume w.l.o.g. these are p_1-arcs. Now consider the partition P_D of the arcs of D which is part of the set of partitions yielding $\vec{\beta}$ from $\vec{\alpha}$. By construction, the two arcs β_{2i} and β_{2i+1} $(1 \leq i \leq \ell - 1)$ are formed by rewiring of the $\vec{\alpha}$-arcs in the same block of P_D, and thus are of the same color.

Since η is a γ-arc, then, by definition, the piece of walk in $u_m(\vec{\alpha})$ it corresponds to moves from a vertex at distance $n-1$ from $\otimes_m$ to a vertex of distance n and back. By the previous lemma, the piece of walk represented by η in $u_m(\vec{\beta})$ also starts at the same vertex of $\mathbb{T}_{2r,2}$, at distance $n-1$ from $\otimes_m$. The arc β_1 is formed by the rewiring of the block containing α_1, and thus has also color p_1. Thus, after the intersection of η with β_1, the walk $u_m(\vec{\beta})$ is at distance n from $\otimes_m$. But, and this is the crux of this lemma, $\vec{\beta} \in \mathcal{P}_{m,n}$ so $\text{depth}(u_m(\vec{\beta})) \leq n$. So the next step of $u_m(\vec{\beta})$ must backtrack, hence β_2 is also of color p_1. We already know that β_2 and β_3 have the same color, so β_3 is also of color p_1 and represents a step to the vertex at distance n. The same argument as before now shows that β_4 must also be a p_1-arc and represents a backtracking step. Repeating these arguments proves the lemma. $\square$

We now reach the endgame. Assume that $\vec{\alpha} \preceq \vec{\beta}$ and both are in $\mathcal{P}_{m,n}$. We already know that $\vec{\alpha} \preceq h_{m,n}(\vec{\alpha})$, and that $\vec{\alpha} \preceq \vec{\beta} \preceq h_{m,n}(\vec{\beta})$ so $\vec{\alpha} \preceq h_{m,n}(\vec{\beta})$. We need to show that $h_{m,n}(\vec{\alpha}) \preceq h_{m,n}(\vec{\beta})$, namely, that the partitions in type-o discs of $\vec{\alpha}$ yielding $h_{m,n}(\vec{\beta})$ are *coarser* than those yielding $h_{m,n}(\vec{\alpha})$. To see this, it is convenient to think of these partitions at the type-o disc D as partitions of the neighboring type-z_i discs: each arc at the boundary of D separates it from some type-z_i disc[36]. The neighboring type-z_i discs in the same block are those which are merged together through new "z_i-corridors" formerly belonging to D. It is enough to show that for any γ-arc η, the two type-z_i discs of $\vec{\alpha}$ it connects are also in the same block in the partition leading from $\vec{\alpha}$ to $h_{m,n}(\vec{\beta})$. This is clearly the case by Lemma 6.18 and the fact that all depth-n leaves in $\vec{\beta}$ are pruned in $h_{m,n}(\vec{\beta})$. This completes the proof of Theorem 6.12 and thus also of Theorem **5.12** and hence of our main results, Theorems 1.10 and 1.12.

Remark 6.19. A slightly different approach for the proof of contractability would treat all guide-arcs at one shot, and define the depth of $\vec{\alpha}$ as the maximal depth of one of $u_1(\vec{\alpha}), \ldots, u_M(\vec{\alpha})$. The only subtlety is that the basepoint in $\mathbb{T}_{2r,2}$ of different $u_m(\vec{\alpha})$'s may be different, depending on the type of the disc where γ_m begins and ends. There are several ways to go around this: for example, one can prune the depth-j leaves in two steps, one for each subset of the guide-arcs. Another solution is to fix some $\vec{\alpha}_0$ which satisfies $\sigma_{\vec{\alpha}_0} = \tau_{\vec{\alpha}_0}$. It is easy to see that in this case the guide-arcs can be taken to be all inside type-o discs.

[36]More precisely, we may have to take some of the neighboring discs with multiplicity two if they have two borders with D, a p_i-border and a q_i-border (see Lemma 5.2). But the partition is colored and thus never merges these two copies together.

7 More Consequences

In this section we gather some further consequences of our analysis which are worth mentioning.

Finding all solutions and incompressible maps to the (generalized) commutator problem

Already in the late 1970's, several algorithms were found to determine the commutator length of a given word $w \in [\mathbf{F}_r, \mathbf{F}_r]$ (as mentioned on Page 11). One of these algorithms, due to Culler in [Cul81], basically follows the same argument as in Lemma 4.9 above — see Remark 4.12. By enumerating all matchings $\sigma \in \mathrm{Match}(w)$, one can find $g = \mathrm{cl}(w)$ as $\frac{1}{2}\left(1 - \max_{\sigma \in \mathrm{Match}(w)} \chi(\sigma, \sigma)\right)$, and then find representatives of every equivalence class of solutions to

$$[u_1, v_1] \cdots [u_g, v_g] = w.$$

By the same Lemma 4.9, the same algorithm extends to finding representatives for all classes in $\mathrm{Solu}(w_1, \ldots, w_\ell)$ for any $w_1, \ldots, w_\ell \in \mathbf{F}_r$, and more generally, to all incompressible $[(\Sigma, f)]$ which is admissible for $w_1, \ldots, w_\ell$.

The main additional contributions of the current paper to this problem are the following:

1. **Identifying all incompressible** (Σ, f)**.** It can be inferred from the analysis in this paper that $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$ is incompressible if and only if, roughly speaking, $\mathcal{PMP}\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$ is downward-closed. More accurately, $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$ is *compressible* if and only if there is a path (each step is between comparable elements) in the poset $\left(\mathrm{Match}(w_1, \ldots, w_\ell)^2, \preceq\right)$ from (σ, τ) to some (σ', τ') with $\chi(\sigma', \tau') > \chi(\sigma, \tau)$ and without going through elements of Euler characteristic smaller than $\chi(\sigma, \tau)$.

2. **Distinguishing equivalence classes**. The current paper yields a convenient way of distinguishing the different classes of solutions, or more generally, of incompressible maps. By Theorem 5.12, given $(\sigma, \tau) \in \mathrm{Match}(w_1, \ldots, w_\ell)$ with $\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$ incompressible, we can construct $\mathcal{PMP}\left(\Sigma_{(\sigma,\tau)}, f_{(\sigma,\tau)}\right)$ by restricting to pairs $(\sigma', \tau') \in \mathrm{Match}(w_1, \ldots, w_\ell)^2$ with $\chi(\sigma', \tau') = \chi(\sigma, \tau)$ and then taking the connected component of (σ, τ). This allows us to identify all (σ', τ') belonging to the same equivalence class of admissible incompressible (Σ, f) as (σ, τ).
   In fact, the analysis shows it is enough to follow this algorithm solely in the bottom two layers of $\mathrm{Match}(w_1, \ldots, w_\ell)^2$: namely, the pairs where $\left\|\sigma^{-1}\tau\right\|$ is 0 (so $\sigma = \tau$) or 1 (so $\sigma^{-1}\tau$ is a transposition).

A bound on the dimension of the $\mathrm{K}(G, 1)$**-complex from Theorems 1.4 and 1.12**

Recall that if (Σ, f) is admissible for $w_1, \ldots, w_\ell$ and incompressible, then $|\mathcal{PMP}(\Sigma, f)|$ is a finite $\mathrm{K}(G, 1)$-complex for $G = \mathrm{Stab}_{\mathrm{MCG}(\Sigma)}\left(\tilde{f}\right)$. We can bound the dimension of this $\mathrm{K}(\mathrm{G}, 1)$-complex in terms of $\chi(\Sigma)$.

Although we have not stressed it so far, some of the objects in this paper, such as $\mathrm{Match}(w_1, \ldots, w_\ell)$, $\mathcal{PMP}$ or $\mathcal{AP}$ depend on the particular presentation of $w_1, \ldots, w_\ell$ as in (1.1). In our analysis we assume we fix a particular presentation (e.g. the reduced one) and stick to it. We say a presentation is cyclically reduced if $x_{i_{(j+1) \bmod |w|}}^{\varepsilon_{(j+1) \bmod |w|}} \neq x_{i_j}^{-\varepsilon_j}$ for every $1 \leq j \leq |w|$. Since the objects we study depend only on the conjugacy class of the words, we can assume they are taken to be cyclically reduced.

Corollary 7.1. *Assume $w_1,\dots,w_\ell \neq 1$ and that the presentations of $w_1,\dots,w_\ell$ are cyclically reduced. If (Σ,f) is admissible for $w_1,\dots,w_\ell$ and incompressible, then the dimension of $|\mathcal{PMP}(\Sigma,f)|$ is at most $-\chi(\Sigma)$.*

Proof. It is enough to show that $\|\sigma^{-1}\tau\| \le -\chi(\Sigma)$ for every $(\sigma,\tau)\in\mathcal{PMP}(\Sigma,f)$. The rank $\|\sigma^{-1}\tau\|$ is equal to $L-\#\text{cycles}\left(\sigma^{-1}\tau\right)$ which is also equal to $\sum_c(|c|-1)$, the summation being over all cycles of $\sigma^{-1}\tau$. These cycles are in one-to-one correspondence with type-z_i discs of $\Sigma_{(\sigma,\tau)}$, and the size of a cycle is half the number of matching-edges at the boundary of the corresponding type-z_i disc. If we denote the number of matching-edges at the boundary of a disc D in $\Sigma_{(\sigma,\tau)}$ by $\deg(D)$, we obtain

$$\|\sigma^{-1}\tau\| = \sum_{D:\,\text{type}-z_i\text{ disc in }\Sigma_{(\sigma,\tau)}} \left(\frac{\deg(D)}{2}-1\right). \tag{7.1}$$

Recall that the CW-complex $\Sigma_{(\sigma,\tau)}$ has $4L$ 0-cells, $4L$ 1-cells along the boundary $\partial\Sigma$ and $2L$ 1-cells as matching-edges, so

$$\chi(\sigma,\tau)=\chi\left(\Sigma_{(\sigma,\tau)}\right)=4L-(2L+4L)+\#\{\text{discs}\}=-2L+\#\{\text{discs}\}.$$

Since every matching-edge is at the boundary of exactly two discs,

$$-\chi(\sigma,\tau)=2L-\#\{\text{discs}\}=\sum_{D:\,\text{disc}}\left(\frac{\deg(D)}{2}-1\right) \tag{7.2}$$

But when $w_1,\dots,w_\ell$ are cyclically reduced, every disc D in $\Sigma_{(\sigma,\tau)}$ has at least two matching-edges at its boundary, i.e., $\deg(D)\ge 2$. Hence the right hand side of (7.2) is an upper bound for the rank in (7.1). □

Explicit finite presentations of the stabilizers in $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ or $\mathrm{MCG}(\Sigma)$

Our analysis also yields a straight-forward algorithm to explicitly find elements in the stabilizers of solutions $\phi\in\mathrm{Hom}_w(\mathbf{F}_{2g},\mathbf{F}_r)$ or $[(\Sigma,f)]\in\mathrm{Solu}(w_1,\dots,w_\ell)$. One way to obtain this is the following. For simplicity, we restrict to the case of a single word w with $\mathrm{cl}(w)=g$ and find the stabilizer in $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$ of a solution $\phi\in\mathrm{Hom}_w(\mathbf{F}_{2g},\mathbf{F}_r)$. Let (Σ,f) be associated with the solution ϕ. Choose an arc system $\vec{\alpha}_0\in\mathcal{AP}(\Sigma,f)$ sitting above some $(\sigma_0,\tau_0)\in\mathcal{PMP}(\Sigma,f)$. Also fix generators $a_1,b_1,\dots,a_g,b_g$ to $\pi_1(\Sigma_{g,1},v_1)$ with $[a_1,b_1]\dots[a_g,b_g]=[\partial\Sigma_{g,1}]$, and for each generator write down the sequence of discs it traverses and the color of the arc it crosses at each step (a disc can be recognized after an action of $\mathrm{MCG}(\Sigma_{g,1})$ by the pieces in $\partial\Sigma_{g,1}$ it touches). Then, for any element $\theta\in\pi_1(|\mathcal{PMP}(\Sigma,f)|,(\sigma_0,\tau_0))$, lift it to $(|\mathcal{AP}(\Sigma,f)|,\vec{\alpha}_0)$ and find the corresponding element $\vec{\beta}\in\mathcal{AP}(\Sigma,f)$. For every generator a_i (or b_i), follow the same sequence of discs in $\vec{\beta}$ as it traversed in $\vec{\alpha}_0$ (this is well-defined by Lemma 5.2). This defines an element of $\pi_1(\Sigma_{g,1},v_1)$, which is exactly $\theta(a_i)$, where θ is identified with the corresponding element of the stabilizer $\mathrm{Stab}_{\mathrm{Aut}_\delta(\mathbf{F}_{2g})}(\phi)$.

As an example, let us return to the word $w=[x,y][x,z]$ and two of the elements of $\mathcal{AP}(\Sigma,f)$ drawn in Figure 6.2. Let $\vec{\alpha}_0$ be the right most element in this figure, and $\vec{\beta}$ be the left most one, both of which sit above the same element of $\mathcal{PMP}(\Sigma,f)$. These two elements are redrawn in Figure 7.1, and assume that $\theta\in\mathrm{MCG}(\Sigma_{g,1})$ maps $\vec{\alpha}_0$ to $\vec{\beta}$. Let the generators a_1,b_1,a_2,b_2 be the loops at v_1 around the four handles at the two sides of the surface, so a_1 is a clockwise loop around the top-right handle (drawn in Figure 7.1 on the right), b_1 is a counter-clockwise loop around the bottom-right handle, a_2 is clockwise around the bottom-left and b_2 is counter-clockwise around the top-left. In $\vec{\alpha}_0$, the loop corresponding to a_1 traverses the discs marked by I, II and III in the following order:

$$I \overset{p_1-\text{arc}}{\to} II \overset{q_1-\text{arc}}{\to} I \overset{p_1-\text{arc}}{\to} III \overset{q_1-\text{arc}}{\to} I \overset{q_1-\text{arc}}{\to} II \overset{p_1-\text{arc}}{\to} I.$$

Following the same pattern in $\vec{\beta}$ results in the dotted loop marked on the left side of Figure 7.1. In the generators we chose for $\pi_1(\Sigma_{g,1}, v_1)$, this new loop is $a_1a_2a_1A_2A_1$, so $\theta(a_1) = a_1a_2a_1A_2A_1$. In the same manner we can figure out how θ acts on the other generators:

$$a_1 \mapsto a_1a_2a_1A_2A_1 \quad b_1 \mapsto a_1a_2A_1A_2b_1a_1a_1A_2A_1 \quad a_2 \mapsto a_1a_2A_1 \quad b_2 \mapsto b_2a_2A_1, \tag{7.3}$$

which gives an explicit description of θ. Since in this case $|\mathcal{PMP}(\Sigma, f)|$ is a cycle with four edges, θ generates the stabilizer. The solution corresponding to the entire connected component of $\vec{\alpha}_0$ and $\vec{\beta}$ (with respect to these generators of $\pi_1(\Sigma_{g,1}, v_1)$) is $w = [x, y][x, z]$, and we deduce

$$\mathrm{Stab}_{\mathrm{Aut}_\delta(\mathbf{F}_4)}([x, y][x, z]) = \langle \theta | \rangle. \tag{7.4}$$

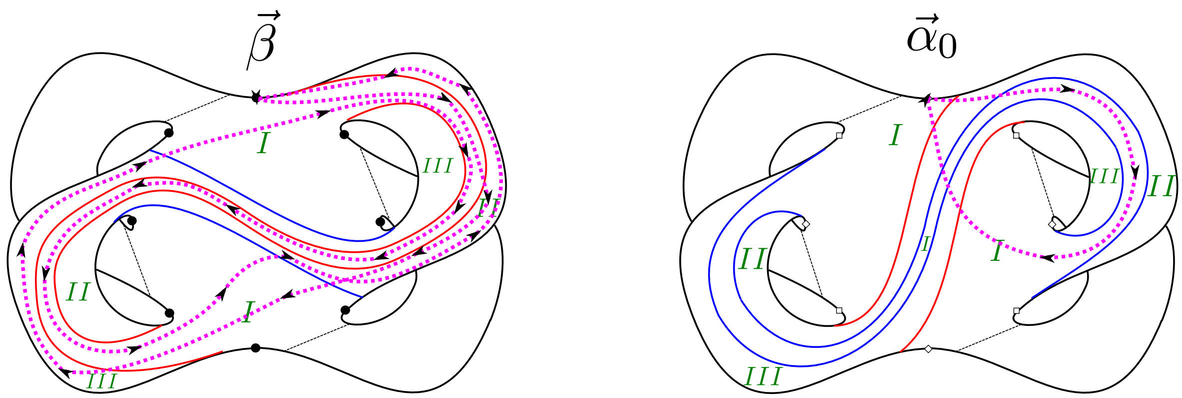

Figure 7.1: Two elements in the same connected component of $\mathcal{AP}(\Sigma, f)$ for the sole admissible (Σ, f) for $w = [x, y][x, z]$, sitting above the same element of $\mathcal{PMP}(\Sigma, f)$. The element of $\mathrm{MCG}(\Sigma_{g,1})$ mapping $\vec{\alpha}_0$ to $\vec{\beta}$ maps the generator a_1 marked in dotted pink line on the right, to the dotted pink line on the left.

We can always find an explicit presentation for the stabilizers. One method would be to find a generating set for the fundamental group of the 1-skeleton of $|\mathcal{PMP}(\Sigma, f)|$, which is free, and then add a relation for every 2-simplex. We give one more detailed presentation in Section 8.

Solvability of the word problem

Finally, let us mention another consequence of our constructions: they show that the word problem for the stabilizers is solvable. To illustrate this, use the generators we constructed in the previous paragraph. For every word in these generators, trace the lift in $|\mathcal{AP}(\Sigma, f)|$ of the corresponding loop in $|\mathcal{PMP}(\Sigma, f)|$. This word is the identity if and only if the lifted path is also closed, which can be easily checked algorithmically.

8 Examples

In this section we gather some concrete examples of the solutions of the commutator equation for a single word and their stabilizers in $\mathrm{Aut}_\delta(\mathbf{F}_{2g})$. We always denote $g = \mathrm{cl}(w)$.

- As mentioned in Remark 1.3, if $\phi \in \mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$ is injective, namely, if $\{\phi(a_1), \ldots, \phi(b_g)\}$ is a free set in $\mathbf{F}_r$, then the stabilizer of ϕ is trivial, and thus its Euler characteristic is 1. For instance,
  - If $\mathrm{cl}(w) = 1$, every solution is free.
  - The word $w = [x,y]^3$ has commutator length 2, and admits 9 equivalence classes of solutions, each of which is injective. One of them was already mentioned in Section 2.1: $[x,y]^3 = [xyX, YxyX^2][Yxy, y^2]$. The coefficient of $\frac{1}{n^3}$ in $\mathcal{T}r_{[x,y]^3}(n)$ is, therefore, 9. Each of the nine complexes $|\mathcal{PMP}(\cdot)|$ consists of a single isolated point. The full expression is $\mathcal{T}r_{[x,y]^3}(n) = \frac{9(n^2+4)}{n^5-5n^3+4n}$.
- There are also "non-injective" solutions with trivial stabilizer. For example, $w = [x,y][x^2y^2, z]$ has $\mathrm{cl}(w) = 2$ with one solution which is non-injective. Yet, $|\mathcal{PMP}(\cdot)|$ is a path composed of ten edges, and is contractible. Hence the stabilizer is trivial, and the coefficient of $\frac{1}{n^3}$ is 1. The full expression is $\frac{n^2-8}{n^5-5n^3+4n}$.
- Along the paper we mentioned the word $w = [x,y][x,z]$. We computed the only pairs of matchings poset associated with it and the corresponding simplicial complex (a cycle of length 4), showed pieces of its arc poset and also computed its stabilizer in (7.4). The Euler characteristic of this $|\mathcal{PMP}(\cdot)|$ is 0, and thus so is the coefficient of $\frac{1}{n^3}$. As we mentioned in Example 4.14, $\mathcal{T}r_{[x,y][x,z]}(n) = 0$ for $n \geq 2$ in this case.
- The leading term vanishes also for $w = [x,y][x,z][x,t]$. Here $\mathrm{cl}(w) = 3$ and there is a single equivalence class of solutions. The pairs of matchings poset $\mathcal{PMP}$ is of size 30: six of rank 0, eighteen of rank 1 and six of rank 2. Hence $|\mathcal{PMP}|$ is 2-dimensional. It consists of 30 vertices, 102 edges and 72 2-simplices, and thus $\chi(|\mathcal{PMP}|) = 0$ and the coefficient of $\frac{1}{n^5}$ is 0. In fact, here too, $\mathcal{T}r_{[x,y][x,z][x,t]}(n) \equiv 0$ (for $n \geq 3$). A closer look at $|\mathcal{PMP}|$ reveals it is homeomorphic to the cross product of S^1 with a Theta figure, so its fundamental group is isomorphic to $\mathbb{Z} \times \mathbf{F}_2$. A computation conducted as explained in Section 7 reveals that

$$\mathrm{Stab}_{\mathrm{Aut}_\delta(\mathbf{F}_6)}([x,y][x,z][x,t]) = \langle \theta_1, \theta_2, \theta_3 \mid [\theta_1,\theta_2], [\theta_1,\theta_3] \rangle,$$

where the θ_i's are given by:

	θ_1	θ_2	θ_3
$a_1 \mapsto$	$a_1^{a_1a_2a_3}$	$a_1^{a_1a_2}$	$a_1^{a_1a_3}$
$b_1 \mapsto$	$(a_2a_3a_1b_1A_1A_1A_1)^{a_1a_2a_3}$	$(a_2a_1b_1A_1A_1)^{a_1a_2}$	$(a_3a_1b_1A_1A_1)^{a_1a_3}$
$a_2 \mapsto$	$a_2^{a_1a_2a_3}$	$a_2^{a_1a_2}$	$a_2^{A_1A_3a_1a_3}$
$b_2 \mapsto$	$(a_3a_1a_2b_2A_2A_2A_2)^{a_1a_2a_3}$	$(a_1a_2b_2A_2A_2)^{a_1a_2}$	$b_2^{A_1A_3a_1a_3}$
$a_3 \mapsto$	$a_3^{a_1a_2a_3}$	a_3	$a_3^{a_1a_3}$
$b_3 \mapsto$	$(a_1a_2a_3b_3A_3A_3A_3)^{a_1a_2a_3}$	b_3	$(a_1a_3b_3A_3A_3)^{a_1a_3}$

(by u^v we mean $v^{-1}uv$, so $a_1^{a_1a_2a_3} = A_3A_2A_1a_1a_1a_2a_3 = A_3A_2a_1a_2a_3$).
- If $w = [x,y]^2$, then $\mathrm{cl}(w) = 2$ with exactly one solution. The sole $|\mathcal{PMP}|$ is 1-dimensional with 12 vertices and 16 edges. Here $\chi(|\mathcal{PMP}|) = -4$ is the leading coefficient. The stabilizer is isomorphic to $\mathbf{F}_5$. One possible generator (a primitive element of this $\mathbf{F}_5$) is given in (7.3).
- If $w = w_1w_2$ is a product of two words with disjoint letters (or more generally of two words from complementing free factors of $\mathbf{F}_r$), then $\mathcal{T}r_w(n) = \mathcal{T}r_{w_1}(n) \cdot \mathcal{T}r_{w_2}(n) \cdot n^{-1}$, the

stabilizer of a solution is the direct product of the stabilizer of the corresponding solution of w_1 and that of w_2, and the Euler characteristics of the stabilizers are multiplicative as well.

9 Some Open Problems

We mention some open problems that naturally arise from the discussion in this paper.

1. In this work we analyzed the expected trace of a random element of $\mathcal{U}(n)$, which corresponds to a natural series of (irreducible) characters ξ_n of $\mathcal{U}(n)$. As explained in Section 2.2, the more general Theorem 1.10 also gives information about other series of irreducible characters of $\mathcal{U}(n)$. A similar question was studied in [PP15] regarding the series of irreducible characters of S_n which count the number of fixed points in a permutation (minus one). It should be very interesting to realize what $\mathrm{Aut}(\mathbf{F}_r)$-invariants of words play a role in similar questions surrounding:

   - The expected trace of elements in the orthogonal group $O(n)$ or the symplectic group $Sp(n)$: as the results of Collins and Śniady [CŚ06] extend to these groups, there should be rational expressions in n as in Theorem 3.7. What is the leading term of each expression?
   - There should also be rational expressions for other series of characters of the groups S_n, $O(n)$ and $\mathrm{Sp}(n)$. What is the leading term for each series?
   - In particular, what are the $\mathrm{Aut}(\mathbf{F}_r)$-invariants of words controlling (the asymptotics of) balanced characters of $\mathcal{U}(n)$ (recall that balanced characters are those invariant under rotations - see Section 2.2).
   - What about completely different families of groups? For example, consider the action of $\mathrm{PSL}_2(q)$ on the projective line $\mathbb{P}^1(q)$. What is the expected number of fixed points in this action when $g \in \mathrm{PSL}_2(q)$ is sampled by some w-measure and q varies?
   - Is it possible to find the algebraic meaning of the other ($\mathrm{Aut}(\mathbf{F}_r)$-invariant) coefficients of the rational function $\mathcal{T}r_w(n)$?

2. In some cases, the coefficient of $\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$ we analyze in Theorem 1.10 vanishes. This is the case, for example, for $w = [x,y][x,z]$ and also for $w = [x,y][x,z][x,t]$. What is the leading coefficient in these cases? Interestingly, among the dozens of concrete examples we computed, there were a handful where the coefficient from Theorem 1.2 vanished. In all these cases the entire expression turned out to be zero, namely, $\mathcal{T}r_{w_1,\ldots,w_\ell}(n) = 0$ for any large enough n.

Acknowledgments

We would also like to thank Danny Calegari, Alexei Entin, Mark Feighn, Alex Gamburd, Peter Sarnak, Zlil Sela, Avi Wigderson and Ofer Zeitouni for valuable discussions about this work.

Appendices

A Appendix: Posets and Complexes

In this appendix we include some auxiliary general results regarding posets and complexes, which are directly used in the proofs along the paper. These results are not new.

A.1 Homotopy of poset morphisms

In our proof of contractability of the connected components of $|\mathcal{AP}(\Sigma, f)|$ in Section 6.3, we use a series of deformation retracts of simplicial complexes associated with posets (see Definition 5.10). Here, we establish a criterion which guarantees that a retract of posets $f : P_2 \to P_1$, where P_1 is a subposet of P_2, is a deformation retract of the associated simplicial complexes. This is the criterion we use in the proof of contractability.

The main ingredient in establishing this criterion deals with direct products of posets. The direct product $P \times Q$ of the posets $(P, \leq_P)$ and $(Q, \leq_Q)$ is defined on the set $P \times Q$ with partial order $(p_1, q_1) \leq_{P \times Q} (p_2, q_2)$ if and only if $p_1 \leq_P p_2$ and $q_1 \leq_Q q_2$. The following lemma is well known: see, for instance, [Wal88, Theorem 3.2].

Lemma A.1. *Let P and Q be posets. The function $\gamma\colon |P \times Q| \to |P| \times |Q|$ defined by*

$$\sum \lambda_i (p_i, q_i) \mapsto \left(\sum \lambda_i p_i, \sum \lambda_i q_i\right)$$

is an homeomorphism.

The following corollary appears in [Qui78, Section 1.3]. Recall that a map f between posets is called a poset-morphism if it is order preserving. If $f\colon P \to Q$ is a poset morphism, we let $|f|$ denote the induced map

$$|f| : |P| \to |Q|$$

defined naturally as $|f| \left(\sum \lambda_i p_i\right) = \sum \lambda_i f(p_i)$.

Corollary A.2. *Let P and Q be posets, and $f, g\colon P \to Q$ poset morphisms. If $f(p) \leq g(p)$ for every $p \in P$, then $|f|$ and $|g|$ are homotopic.*

Proof. Let $\{0 \leq 1\}$ denote the poset with two comparable elements 0 and 1. Define a map $(f, g) : P \times \{0 \leq 1\} \to Q$ by $(p, 0) \mapsto f(p)$ and $(p, 1) \mapsto g(p)$. This is clearly a poset-morphism by the assumptions, so it induces a continuous map

$$|(f, g)| : |P \times \{0 \leq 1\}| \to |Q|.$$

By Lemma A.1, there is an homeomorphism

$$|P \times \{0 \leq 1\}| \xrightarrow{\cong} |P| \times |\{0 \leq 1\}| = |P| \times [0, 1],$$

so we get that $|(f, g)|$ is a continuous map $|P| \times [0, 1] \to |Q|$. Because $|(f, g)|\Big|_{|P \times \{0\}|} \equiv |f|$ and $|(f, g)|\Big|_{|P \times \{1\}|} \equiv |g|$, the map $|(f, g)|$ is the sought after homotopy. □

Remark A.3. Note that the homotopy does not move the points where f and g agree. Namely, if $P_0 \subseteq P$ is the subposet where $f(p) = g(p)$, then $|(f, g)|(x, t) = f(x) = g(x)$ for every $x \in |P_0|$ and $t \in [0, 1]$.

Corollary A.4. *Let P be a subposet of the poset Q. Assume that $f\colon Q \to P$ satisfies the following:*

- *it is a poset morphism,*
- *it is a retract (i.e., $f\Big|_P \equiv \mathrm{id}$), and*
- *$f(q) \le q$ for every $q \in Q$, or $q \le f(q)$ for every $q \in Q$.*

Then $|f|$ is a (strong) deformation retract.

By a strong deformation retract we mean that there is a homotopy of $|f|$ with the identity on $|Q|$ which fixes the points in $|P|$ throughout the homotopy.

Proof. Simply note that the map $f\colon Q \to Q$ and the identity $\mathrm{id}\colon Q \to Q$ satisfy the conditions in Corollary A.2 hence $|f|$ is homotopic to the identity. The fact that the homotopy fixes $|P|$ pointwise follows from Remark A.3. □

A.2 Regular G-complexes

When we say that a discrete group G acts on a simplicial complex K, we mean, in particular, that the action is simplicial. Namely, we mean that G acts on the set of vertices, and the induced map on the subsets of vertices maps every simplex to a simplex. There are two natural ways to construct a quotient space for this action. One way is to construct a simplicial complex as follows: the set of vertices consists of the orbits ${}^{V(K)}/_{G}$ of vertices and whenever $(v_0, \ldots, v_r)$ is an r-simplex of K, then $([v_0], \ldots, [v_r])$ is an r-simplex of the quotient. We denote this quotient by $|{}^{K}/_{G}|$. The second way is to consider the geometric realization of K, which G clearly acts on, and take the usual quotient of an action on a topological space. We denote this quotient by ${}^{|K|}/_{G}$.

The problem is that these two quotient spaces do not coincide in general. First, if the action mixes different vertices of the same simplex, the topological quotient results in pieces which are fractions of simplices. This is the case, for example, in the case that ${}^{\mathbb{Z}}/_{2\mathbb{Z}}$ acts on a graph with a single edge by flipping the edge. Secondly, as illustrated by the action of ${}^{\mathbb{Z}}/_{2\mathbb{Z}}$ on the boundary of a square by a 180°-rotation mentioned in Remark 6.9, the orbits of the simplices in the geometric realization are not always determined by the orbits of the vertices.

These, however, can be easily remedied by adding the following assumptions:

Definition A.5. [Bre72, Definition III.1.2] A simplicial G-action on the simplicial complex K is called **regular**, if

1. If $v \in V(K)$ and $g.v$ belong to same simplex for some $g \in G$, then $g.v = v$.
2. Whenever $g_0, \ldots, g_r$ are elements of G and $(v_0, \ldots, v_r)$ and $(g_0.v_0, \ldots, g_r.v_r)$ are r-simplices of K, there is some $g \in G$ with $(g_0.v_0, \ldots, g_r.v_r) = (g.v_0, \ldots, g.v_r)$.

In other words, these additional conditions exactly guarantee that (1) the action does not "break" simplices by identifying different points of the same simplex, and that (2) the orbits of the simplices in the geometric realization can be deduced from those of the vertices.

Lemma A.6. *[Bre72, Page 117] If the action of G on the simplicial complex K is regular then*

$$|{}^{K}/_{G}| \cong {}^{|K|}/_{G}.$$

In the current paper, we are interested in G-actions on graded posets and on their corresponding simplicial complexes. Lemma A.6 translates to the following (see Definition 5.10 and the footnote on Page 38 for some of the terminology):

Corollary A.7. *Let G act on a locally-finite graded poset $(P, \leq)$ by a graded-poset action, and assume that whenever $x_0 < \ldots < x_r$ and $g_0.x_0 < \ldots < g_r.x_r$ for some $g_0, \ldots, g_r \in G$ and $x_0, \ldots, x_r \in P$, there is a $g \in G$ with $g.x_i = g_i.x_i$ for every i. Then*

$$|{}^{P}/_{G}| \cong {}^{|P|}/_{G}.$$

Proof. We only need to check that the action is regular. Item 2 of Definition A.5 holds by our extra assumption, while item 1 follows from the fact that the action preserves rank, thus guaranteeing that x and $g.x$ cannot belong to same simplex of $|P|$ unless $x = g.x$. □

Glossary

		Reference	**Remarks**
$\mathbf{F}_r$	the free group on r generators		
$x_1, \ldots, x_r$	a set of generators for $\mathbf{F}_r$		sometimes x, y, z, t used instead
$X_1, \ldots, X_r$	$X_i = x_i^{-1}$ marks the inverse		likewise, X, Y, Z, T
$\mathcal{U}(n)$	the group of $n \times n$ unitary matrices		
μ_n	the Haar measure on $\mathcal{U}(n)$		
$\mathcal{T}r_w(n)$	expected trace of $A \in \mathcal{U}(n)$ sampled according to the w-measure	(1.3)	
$\mathrm{cl}(w)$	the commutator length of w	Page 3	
$a_1, b_1, \ldots, a_g, b_g$	a set of generators for $\mathbf{F}_{2g}$		$A_1, B_1, \ldots, A_g, B_g$ mark inverses
δ_g	$[a_1, b_1] \ldots [a_g, b_g]$		
$\mathrm{Hom}_w(\mathbf{F}_{2g}, \mathbf{F}_r)$	$\{\phi \in \mathrm{Hom}(\mathbf{F}_{2g}, \mathbf{F}_r) \mid \phi(\delta_g) = w\}$		
$\mathrm{Aut}_\delta(\mathbf{F}_{2g})$	$\{\rho \in \mathrm{Aut}(\mathbf{F}_{2g}) \mid \rho(\delta_g) = \delta_g\}$		
χ	Euler characteristic of a space or a group	Page 1.1	
$\Sigma_{g,1}$	Orientable surface of genus g and one boundary component.	Section 1.2	
$\mathcal{T}r_{w_1,\ldots,w_\ell}(n)$	$\mathbb{E}\left[\prod_{i=1}^{\ell} \mathrm{tr}\left(w_i\left(U_1^{(n)}, \ldots, U_r^{(n)}\right)\right)\right]$		
$(\bigvee^r S^1, o)$	a wedge of r circles, fundamental group identified with $\mathbf{F}_r$, pointed at the wedge point o	Section 1.2, Figure 4.1	sometimes additional marked points
(Σ, f) admissible for $w_1, \ldots, w_\ell$	Σ a compact oriented surface with ℓ boundary components and $f\colon \Sigma \to \bigvee^r S^1$ maps these components to $w_1, \ldots, w_\ell$	Definition 1.5	
(Σ, f) incompressible	no essential simple closed curve mapped to nullhomotopic loop	Definition 1.11	
$\tilde{f}$	homotopy class of $f\colon \Sigma \to \bigvee^r S^1$, relative $\partial\Sigma$	Theorem 1.10	

		Reference	Remarks
$v_1, \ldots, v_\ell$	"basepoints" of Σ, one at every boundary component		
$\partial_1, \ldots, \partial_\ell$	identifications of boundary components of Σ with S^1		
f_w	a map $\left(S^1, 1\right) \to \left(\bigvee^r S^1, o\right)$ with image representing w	Page 6 and more detailed in Section 2.1	
$\mathrm{MCG}\left(\Sigma\right)$	mapping class group of Σ, consisting of mapping classes which fix $\partial\Sigma$ pointwise		
$\mathrm{chi}\left(w_1, \ldots, w_\ell\right)$		Definition 1.6	
$\left(\Sigma, f\right) \sim \left(\Sigma', f'\right)$		Page 8	
$\left[\left(\Sigma, f\right)\right]$	equivalence class of $\left(\Sigma, f\right)$	Definition 8	
$\mathrm{Solu}\left(w_1, \ldots, w_\ell\right)$	set of equivalence classes of admissible maps of maximal χ	Definition 8	
balanced set of words	words such that the total number of x_i^{+1} is the same as total number of x_i^{-1}		
$S^1\left(w\right)$	a marked circle which spells out w	Section 4 and Figure 4.1	
$o,\ p_i,\ z_i,\ q_i$	marked points on $\bigvee^r S^1$	Sections 2.1 and 4	
$p_i^{\pm},\ q_i^{\pm}$	marked points of $S^1\left(w\right)$, $\partial\Sigma$	Sections 2.1 and 4	
Wg	the Weingarten function	Definition 3.2	
$\lVert\sigma\rVert$	the norm of the permutation σ	Section 3.1	
$\mathrm{Möb}\left(\sigma\right)$	the Möbius function of σ	Proposition 3.4	
$L,\ L_i$	assuming $w_1, \ldots, w_\ell$ balanced , $2L = \sum \lvert w_i\rvert$ and L_i is the number of appearances of x_i^{+1}	Section 3	
$E^{\pm}, E_i^{\pm}$	subsets of the letter of $w_1, \ldots, w_\ell$	Section 4	
$\mathrm{Match}(w_1,\ldots,w_\ell)$	the set of bijections $E^+ \xrightarrow{\sim} E^-$ which map E_i^+ to E_i^-	Definition 4.1	
$\Sigma_{(\sigma,\tau)}$	the CW-complex associated with $\left(\sigma, \tau\right) \in \mathrm{Match}\left(w_1, \ldots, w_\ell\right)^2$	Definition 4.2	
matching-edges		Definition 4.2	
$\chi\left(\sigma, \tau\right)$	the Euler characteristic of $\Sigma_{(\sigma,\tau)}$	Definition 4.5	
type-o and type-z_i discs	types of discs in $\Sigma_{(\sigma,\tau)}$ as well as in $\Sigma_{\vec{\alpha}}$	Claims 4.4, Section 6.1	
$f_{(\sigma,\tau)}, f_{\vec{\alpha}}$	$f_{(\sigma,\tau)} \colon \Sigma_{(\sigma,\tau)} \to \bigvee^r S^1$ and $f_{\vec{\alpha}} \colon \Sigma_{\vec{\alpha}} \to \bigvee^r S^1$ are homotopy classes of maps	Definition 4.7 and Section 6.1	
$\mathcal{PMP}\left(\Sigma, f\right)$, $\lvert\mathcal{PMP}\left(\Sigma, f\right)\rvert$	the pair of matchings poset and its associated simplicial complex	Definitions 5.1 and 5.10	
$\sigma_{\vec{\alpha}}, \tau_{\vec{\alpha}}$	the pair of matchings induced by the arc system $\vec{\alpha}$	Section 6.1	

		Reference	**Remarks**
$\Sigma_{\vec{\alpha}}$	the CW-complex structure induced on Σ by the arc system $\vec{\alpha}$	Section 6.1	
$\mathcal{AP}(\Sigma, f)$, $\lvert\mathcal{AP}(\Sigma, f)\rvert$	the arc poset and its associated simplicial complex	Definitions 6.4, 5.10	
$\preceq$	partial orders defined on S_L, S_L^2, $\text{Match}(w_1, \ldots, w_\ell)^2$, $\mathcal{PMP}(\Sigma, f)$, $\mathcal{AP}(\Sigma, f)$	Sections 3.1 and 5 and Definitions 5.1 and 6.4	
graded poset		Footnote on Page 38	
x covers y			for x, y in a poset

References

[Alo86] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. Theory of computing (Singer Island, Fla., 1984).

[AV11] A. Amit and U. Vishne. Characters and solutions to equations in finite groups. *J. Algebra Appl.*, 10(4):675–686, 2011.

[BF05] M. Bestvina and M. Feighn. Counting maps from a surface to a graph. *Geom. Funct. Anal.*, 15(5):939–961, 2005.

[BIZ80] D. Bessis, C. Itzykson, and J. B. Zuber. Quantum field theory techniques in graphical enumeration. *Adv. in Appl. Math.*, 1(2):109–157, 1980.

[Bor15] C. Bordenave. A new proof of Friedman's second eigenvalue Theorem and its extension to random lifts. *ArXiv e-prints*, February 2015.

[Bre72] G. E. Bredon. *Introduction to compact transformation groups*. Elsevier, 1972.

[Bro82] K. S. Brown. *Cohomology of groups*, volume 87 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[Cal08] D. Calegari. What is... stable commutator length? *Notices Amer. Math. Soc.*, 55(9):1100–1101, 2008.

[Cal09a] D. Calegari. *scl*, volume 20 of *MSJ Memoirs*. Mathematical Society of Japan, Tokyo, 2009.

[Cal09b] D. Calegari. Stable commutator length is rational in free groups. *Journal of the American Mathematical Society*, 22(4):941–961, 2009.

[CM14] B. Collins and C. Male. The strong asymptotic freeness of Haar and deterministic matrices. *Ann. Sci. Éc. Norm. Supér. (4)*, 47(1):147–163, 2014.

[Col03] B. Collins. Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability. *International Mathematics Research Notices*, 2003(17):953–982, 2003.

[CŚ06] B. Collins and P. Śniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Comm. Math. Phys.*, 264(3):773–795, 2006.

[Cul81] M. Culler. Using surfaces to solve equations in free groups. *Topology*, 20(2):133–145, 1981.

[DS94] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *Journal of Applied Probability*, pages 49–62, 1994.

[Edm75] C. C. Edmunds. On the endomorphism problem for free groups. *Communications in Algebra*, 3(1):1–20, 1975.

[Fey48] R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20:367–387, Apr 1948.

[FM12] B. Farb and D. Margalit. *A primer on mapping class groups*, volume 49 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2012.

[FN03] W. Fenchel and J. Nielsen. *Discontinuous groups of isometries in the hyperbolic plane*, volume 29 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2003. Edited and with a preface by Asmus L. Schmidt, Biography of the authors by Bent Fuglede.

[Fri08] J. Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.*, 195(910):viii+100, 2008.

[Ful97] W. Fulton. *Young tableaux: with applications to representation theory and geometry*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, 1997.

[Gro] A. Grothendieck. Esquisse d'un programme. 1984.

[GT79] R. Z. Goldstein and E. C. Turner. Applications of topological graph theory to group theory. *Mathematische Zeitschrift*, 165(1):1–10, 1979.

[Har85] J. L. Harer. Stability of the homology of the mapping class groups of orientable surfaces. *Annals of Mathematics*, 121(2):pp. 215–249, 1985.

[Hat91] A. Hatcher. On triangulations of surfaces. *Topology Appl.*, 40(2):189–194, 1991.

[Hat02] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[Hme71] J. I. Hmelevskiĭ. Systems of equations in a free group I. *Izvestiya: Mathematics*, 5(6):1245–1276, 1971.

[HT05] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math. (2)*, 162(2):711–775, 2005.

[HZ86] J. L. Harer and D. Zagier. The Euler characteristic of the moduli space of curves. *Invent. Math.*, 85(3):457–485, 1986.

[Kon92] M. Kontsevich. Intersection theory on the moduli space of curves and the matrix Airy function. *Comm. Math. Phys.*, 147(1):1–23, 1992.

[LS77] R. C. Lyndon and P. E. Schupp. *Combinatorial group theory*. Springer-Verlag, 1977.

[LW81] R. C. Lyndon and M. J. Wicks. Commutators in free groups. *Canad. Math. Bull.*, 24(1):101–106, 1981.

[LZ04] S. K. Lando and A. K. Zvonkin. *Graphs on surfaces and their applications*, volume 141 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. With an appendix by Don B. Zagier, Low-Dimensional Topology, II.

[MŚS07] J. A. Mingo, P. Śniady, and R. Speicher. Second order freeness and fluctuations of random matrices. II. Unitary random matrices. *Adv. Math.*, 209(1):212–240, 2007.

[NS06] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.

[Pen88] R. C. Penner. Perturbative series and the moduli space of Riemann surfaces. *J. Differential Geom.*, 27(1):35–53, 1988.

[PP15] D. Puder and O. Parzanchevski. Measure preserving words are primitive. *Journal of the American Mathematical Society*, 28(1):63–97, 2015.

[Pud14] D. Puder. Primitive words, free factors and measure preservation. *Israel J. Math.*, 201(1):25–73, 2014.

[Pud15] D. Puder. Expansion of random graphs: new proofs, new results. *Inventiones Mathematicae*, 201(3):845–908, 2015.

[Qui78] D. Quillen. Homotopy properties of the poset of nontrivial p-subgroups of a group. *Adv. in Math.*, 28(2):101–128, 1978.

[Răd06] Florin Rădulescu. Combinatorial aspects of Connes's embedding conjecture and asymptotic distribution of traces of products of unitaries. In *Proceedings of the Operator Algebra Conference, Bucharest*. Theta Foundation, 2006.

[Sel01] Z. Sela. Diophantine geometry over groups. I. Makanin-Razborov diagrams. *Publ. Math. Inst. Hautes Études Sci.*, (93):31–105, 2001.

[Sha13] A. Shalev. Some results and problems in the theory of word maps. In L. Lovász, I. Ruzsa, V.T. Sós, and D. Palvolgyi, editors, *Erdős Centennial (Bolyai Society Mathematical Studies)*, pages 611–650. Springer, 2013.

[Sta12] R. P. Stanley. *Enumerative Combinatorics, Volume I.* Number 49 in Cambridge Studies in Advanced Mathematics. Cambridge university press, 2012.

[tH74] G. 't Hooft. A planar diagram theory for strong interactions. *Nuclear Physics B*, 72(3):461 – 473, 1974.

[VDN92] D. V. Voiculescu, K. J. Dykema, and A. Nica. *Free random variables*, volume 1 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 1992. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups.

[Voi85] D. Voiculescu. Symmetries of some reduced free product C^*-algebras. In *Operator algebras and their connections with topology and ergodic theory (Buşteni, 1983)*, volume 1132 of *Lecture Notes in Math.*, pages 556–588. Springer, Berlin, 1985.

[Voi91] D. Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104(1):201–220, 1991.

[Voi93] D. Voiculescu. Around quasidiagonal operators. *Integral Equations Operator Theory*, 17(1):137–149, 1993.

[Wal88] J. W. Walker. Canonical homeomorphisms of posets. *European Journal of Combinatorics*, 9(2):97–107, 1988.

[Wei78] D. Weingarten. Asymptotic behavior of group integrals in the limit of infinite rank. *Journal of Mathematical Physics*, 19(5):999–1001, 1978.

[Wit91] E. Witten. Two-dimensional gravity and intersection theory on moduli space. In *Surveys in differential geometry (Cambridge, MA, 1990)*, pages 243–310. Lehigh Univ., Bethlehem, PA, 1991.

[Xu97] F. Xu. A random matrix model from two dimensional Yang-Mills theory. *Communications in mathematical physics*, 190(2):287–307, 1997.

[ZVC80] H. Zieschang, E. Vogt, and H.-D. Coldewey. *Surfaces and planar discontinuous groups*, volume 835 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980. Translated from the German by John Stillwell.

Michael Magee,
Department of Mathematics,
Yale University,
PO Box 208283, New Haven, CT 06520 USA
michael.magee@yale.edu

Doron Puder,
School of Mathematics,
Institute for Advanced Study,
Einstein Drive, Princeton, NJ 08540 USA
doronpuder@gmail.com