Matrix Group Integrals, Surfaces, and Mapping Class Groups II: $\mathrm{O}(n)$ and $\mathrm{Sp}(n)$

Michael Magee and Doron Puder*

December 6, 2022

Abstract

Let w be a word in the free group on r generators. The expected value of the trace of the word in r independent Haar elements of $\mathrm{O}(n)$ gives a function $\mathcal{T}r_w^{\mathrm{O}}(n)$ of n. We show that $\mathcal{T}r_w^{\mathrm{O}}(n)$ has a convergent Laurent expansion at $n=\infty$ involving maps on surfaces and L^2-Euler characteristics of mapping class groups associated to these maps. This can be compared to known, by now classical, results for the GUE and GOE ensembles, and is similar to previous results concerning $\mathrm{U}(n)$, yet with some surprising twists.

A priori to our result, $\mathcal{T}r_w^{\mathrm{O}}(n)$ does not change if w is replaced with $\alpha(w)$ where α is an automorphism of the free group. One main feature of the Laurent expansion we obtain is that its coefficients respect this symmetry under $\mathrm{Aut}(\mathbf{F}_r)$.

As corollaries of our main theorem, we obtain a quantitative estimate on the rate of decay of $\mathcal{T}r_w^{\mathrm{O}}(n)$ as $n\to\infty$, we generalize a formula of Frobenius and Schur, and we obtain a universality result on random orthogonal matrices sampled according to words in free groups, generalizing a theorem of Diaconis and Shahshahani.

Our results are obtained more generally for a tuple of words $w_1,\ldots,w_\ell$, leading to functions $\mathcal{T}r_{w_1,\ldots,w_\ell}^{\mathrm{O}}$. We also obtain all the analogous results for the compact symplectic groups $\mathrm{Sp}(n)$ through a rather mysterious duality formula.

Contents

*D. P. was supported by the Israel Science Foundation (grant No. 1071/16).

1 Introduction

Let $\mathbf{F}_r$ be the free group on r generators with a fixed basis $B = \{x_1, \ldots, x_r\}$. For any word $w \in \mathbf{F}_r$ and group H there is a *word map*

$$w : H^r \to H$$

defined by substitutions, for example, if $r = 2$ and $w = x_1^2 x_2^{-2}$ then $w(h_1, h_2) = h_1^2 h_2^{-2}$. In this paper we consider the case that H is a compact orthogonal or symplectic group. For $n \in \mathbf{Z}_{\geq 1}$, the orthogonal group $\mathrm{O}(n) = \mathrm{U}(n, \mathbb{R})$ is the group of $n \times n$ real matrices A such that $A^T A = AA^T = \mathrm{Id}_n$. For $n \in \mathbf{Z}_{\geq 1}$ the compact symplectic group $\mathrm{Sp}(n) = \mathrm{U}(n, \mathbb{H})$ is the group of $n \times n$ quaternion matrices A such that $A^* A = AA^* = \mathrm{Id}_n$ (here A^* is the adjoint matrix $\overline{A^T}$). However, we use here the isomorphic description of $\mathrm{Sp}(n)$ as $2n \times 2n$ complex matrices which are both unitary and complex-symplectic in the sense that they preserve a skew-symmetric form, namely, we use the definition,

$$\mathrm{Sp}(n) \stackrel{\text{def}}{=} \left\{ A \in M_{2n}(\mathbb{C}) \,\middle|\, A^* A = AA^* = \mathrm{Id}_{2n} \text{ and } A^T J A = J \right\}, \tag{1.1}$$

where J is the matrix

$$J \stackrel{\text{def}}{=} \begin{pmatrix} 0_n & \mathrm{Id}_n \\ -\mathrm{Id}_n & 0_n \end{pmatrix}. \tag{1.2}$$

The isomorphism $U(n, \mathbb{H}) \cong \mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{C})$ is given by

$$A + Bj \mapsto \begin{pmatrix} A & -B \\ \overline{B} & \overline{A} \end{pmatrix} \tag{1.3}$$

where $A, B \in M_n(\mathbb{C})$ (see [Hal15, §1.2.8] for more details). Notice, in particular, that all matrices in $\mathrm{Sp}(n)$, as defined in (1.1), have real trace.

We fix words $w_1, \ldots, w_\ell \in \mathbf{F}_r$. Our main object of study is the following type of matrix integral

$$\mathcal{T}r^G_{w_1,\ldots,w_\ell}(n) \stackrel{\text{def}}{=} \int_{G(n)^r} \mathrm{tr}(w_1(g_1, \ldots, g_r)) \cdots \mathrm{tr}(w_\ell(g_1, \ldots, g_r)) d\mu_n(g_1) \ldots d\mu_n(g_r)$$

where $G(n)$ is either the orthogonal group $\mathrm{O}(n)$ or compact symplectic group $\mathrm{Sp}(n)$ defined in (1.1), and μ_n is the Haar probability measure on the respective group. We write $G = \mathrm{O}$ or Sp respectively to refer to the type of group being studied.

These values, the expected value of the trace and product of traces, are natural objects of study. First, the measure induced by a word w on $\mathrm{O}(n)$ or $\mathrm{Sp}(n)$ (see Remark 1.16 for the precise definition) is completely determined by these values when $w_1, \ldots, w_\ell$ are various (positive) powers of w. Second, these values are parallel to well-studied quantities in various models of random matrices and in free probability, and are related to questions on representation varieties and more. See the introduction of [MP19] for more details.

Our paper is centered around writing the functions $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ as Laurent series in the variable $(n+1-\alpha_G)^{-1}$ with positive radii of convergence[1], where $\alpha_{\mathrm{O}} = 2$ and $\alpha_{\mathrm{Sp}} = \frac{1}{2}$ are called 'Jack parameters' in the literature [Nov17]. In §1.3 we discuss the history of these parameters and their necessary role in the current paper.

In fact, the functions $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ are not only meromorphic at ∞, but actually given by rational functions of n, with integer coefficients, when n is sufficiently large. This is a reasonably straightforward consequence of the Weingarten Calculus and is proved in Corollaries 3.5 and A.6 below. See Table 1 for some examples.

The ability to find a Laurent series at ∞ for $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ using diagrammatic expansions will not be surprising to experts in Free Probability Theory and in particular, those who have worked with the Weingarten Calculus. *However, this is not the main point of this paper.* Rather, the main point is what we now explain. The integrals $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ have a priori symmetries under $\mathrm{Aut}(\mathbf{F}_r)$, the automorphism group of $\mathbf{F}_r$, according to the following lemma.

Lemma 1.1. *If α is an automorphism of $\mathbf{F}_r$, then $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n) = \mathcal{T}r^{G}_{\alpha(w_1),\ldots,\alpha(w_\ell)}(n)$.*

This is proved in the case $\ell = 1$ in [MP16, §2.2] and the case of general ℓ can be proved using the same argument. In particular, if $w \in \mathbf{F}_r$ then $\mathcal{T}r^{G}_{w}(n)$ is a function of n that only depends on w up to automorphism; in other words, it reflects algebraic properties of w, not just combinatorial properties. So a priori, any Laurent series expansion of $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ will involve $\mathrm{Aut}(\mathbf{F}_r)$-invariants of $w_1,\ldots,w_\ell$ and this brings us to the true goal of the paper: *to find Laurent series expansions at ∞ for $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ in terms of $\mathrm{Aut}(\mathbf{F}_r)$-invariants of $w_1,\ldots,w_\ell$.* Indeed, our main result, Theorem 1.5, as well as its corollaries in Section 1.2, are all given in terms of $\mathrm{Aut}(\mathbf{F}_r)$-invariants of the words.

In fact, the expressions for $G = \mathrm{O}$ or Sp are closely related; we will prove

Theorem 1.2. *There is[2] $N = N(w_1,\ldots,w_\ell) \geq 0$ such that when $2n \geq N$ we have*

$$\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n) = (-1)^{\ell}\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(-2n).$$

The quantity $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(-2n)$ is interpreted using the rational function form of $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}$ for $2n \geq N$. In particular, this identity relates the Laurent series at ∞ of $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n)$ and $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$.

This type of duality between O and Sp has been observed previously in various contexts, some of which we discuss in §1.3.

1.1 The main theorem

Understanding the functions $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ involves studying maps from surfaces to the wedge of r circles, denoted by $\bigvee^r S^1$, as was also the case in [MP19], where the corresponding theorems were proved for the functions $\mathcal{T}r^{U}_{w_1,\ldots,w_\ell}(n)$ that arise from compact unitary groups $\mathrm{U}(n) = \mathrm{U}(n,\mathbb{C})$. One central difference that appears here is that for unitary groups, only orientable surfaces featured in the description of $\mathcal{T}r^{U}_{w_1,\ldots,w_\ell}(n)$, whereas the description of $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$ for $G = \mathrm{O}$ or Sp also involves non-orientable surfaces as well.

We call the point in $\bigvee^r S^1$ at which the circles are wedged together o. We fix an orientation of each circle that gives us an identification $\pi_1(\bigvee^r S^1, o) \cong \mathbf{F}_r$ and we identify the generator $x \in B$ with the loop that traverses the circle S^1_x corresponding to x according to its given orientation.

[1] Of course, since $(n+1-\alpha_G)^{-1}$ and n^{-1} are related by $z \mapsto \frac{z}{(z+1-\alpha_G)}$, which fixes 0 and is a local biholomorphism there, Laurent series in $(n+1-\alpha_G)^{-1}$ give rise to Laurent series in n^{-1}.

[2] We define N precisely in (3.2).

ℓ	$w_1,\ldots,w_\ell$	$\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$	$\chi_{\max}$	Admissible maps with $\chi(\Sigma)=\chi_{\max}$
1	x^2y^2	$\frac{1}{n}$	-1	one $(P_{2,1},f)$ with $\mathrm{MCG}(f)=\{1\}$
	x^4y^4	$\frac{1}{n}$	-1	one $(P_{2,1},f)$ with $\mathrm{MCG}(f)=\{1\}$
	$[x,y]^2$	$\frac{n^3+n^2-2n-4}{n(n+2)(n-1)}$	0	one $(P_{1,1},f)$ with $\mathrm{MCG}(f)=\{1\}$
	$xy^3x^{-1}y^{-1}$	0	-2	one $(P_{3,1},f)$ with $\mathrm{MCG}(f)\cong\mathbf{Z}$
	$xy^4x^{-1}y^{-2}$	$\frac{1}{n}$	-1	one $(P_{2,1},f)$ with $\mathrm{MCG}(f)=\{1\}$
	$xyx^2yx^3y^2$	$\frac{3n+2}{n(n+2)(n-1)}$	-2	three $(P_{3,1},f)$ with $\mathrm{MCG}(f)=\{1\}$
2	w,w for $w=x^2y^2$	$\frac{n^3+n^2+2n+4}{n(n+2)(n-1)}$	0	one annulus with $\mathrm{MCG}(f)=\{1\}$
	w,w for $w=x^2y$	1	0	one annulus with $\mathrm{MCG}(f)=\{1\}$
3	w,w,w for $w=x^2y^2$	$\frac{3(n^4+3n^3-2n^2+6n+16)}{(n-2)(n-1)n(n+2)(n+4)}$	-1	three (annulus $\sqcup P_{2,1}$) with $\mathrm{MCG}(f)=\{1\}$

Table 1: Some examples of the rational expression for $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$. All these examples contain words in $\mathbf{F}_2$ with generators $\{x,y\}$. The notation $[x,y]$ is for the commutator $xyx^{-1}y^{-1}$. We let $\chi_{\max}=\chi_{\max}(w_1,\ldots,w_\ell)$ denote the maximal Euler characteristic of a surface in $\mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$ – see Page 16. See Definitions 1.3 and 1.4 for the notions of admissible maps and $\mathrm{MCG}(f)$ appearing in the right column. In that column, $P_{g,b}$ denotes the non-orientable surface of genus g with b boundary components (so $\chi(P_{g,b})=2-g-b$). We give more details in Section 6.3 and Table 2.

Definition 1.3 (Admissible maps)**.** Given $w_1,\ldots,w_\ell\in\mathbf{F}_r$, consider pairs (Σ,f) where

- Σ is a compact surface, not necessarily connected, with ℓ ordered and oriented boundary components $\delta_1,\ldots,\delta_\ell$ and a marked point v_j on each δ_j. Note the orientation of the jth boundary component specifies a generator $[\delta_j]$ of $\pi_1(\delta_j,v_j)$. We also require that Σ has no closed components[3].

- $f:\Sigma\to\bigvee^r S^1$ is a continuous map such that $f(v_j)=o$ and $f_*([\delta_j])=w_j\in\pi_1(\bigvee^r S^1,o)\cong\mathbf{F}_r$ for all $1\le j\le\ell$.

We call such a pair an *admissible map* (for $w_1,\ldots,w_\ell$). Consider two admissible maps (Σ,f) and (Σ',f') with boundary components $\delta_1,\ldots,\delta_\ell$ and $\delta'_1,\ldots,\delta'_\ell$ and marked points $\{v_j\}$ and $\{v'_j\}$ respectively. We say (Σ,f) and (Σ',f') are *equivalent* and write $(\Sigma,f)\approx(\Sigma',f')$ if

[3] So every connected component of a surface Σ in an admissible map is either the orientable $\Sigma_{g,b}$ of genus $g\ge0$ and with $b\ge1$ boundary components, or the non-orientable $P_{g,b}$ of genus $g\ge1$ and with $b\ge1$ boundary components. Recall that the Euler characteristics of these surfaces are $\chi(\Sigma_{g,b})=2-2g-b$ and $\chi(P_{g,b})=2-g-b$.

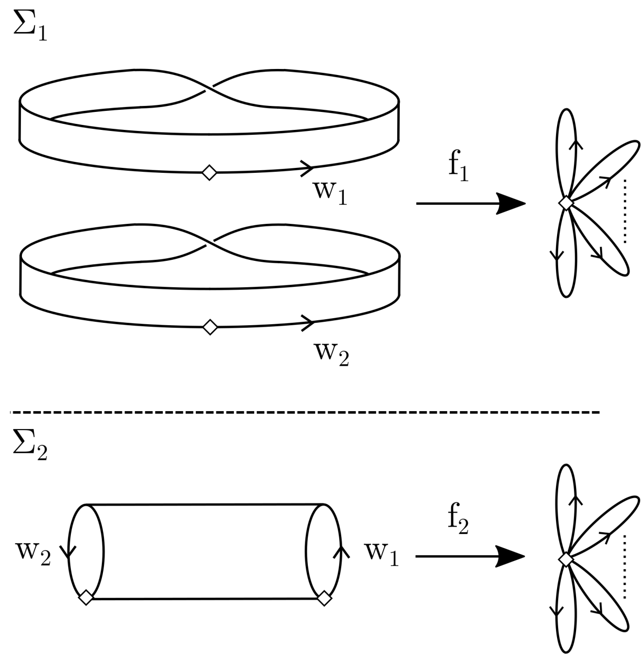

Figure 1.1: Schematic illustration of two admissible maps (Σ_1, f_1) and (Σ_2, f_2) when $\ell = 2$ and there are two words w_1, w_2. Base points v_j are marked by diamonds, and an oriented boundary component labeled by w_j maps to w_j under $(f_i)_*$, for the relevant $i = 1, 2$.

- there exists a homeomorphism $F : \Sigma \to \Sigma'$ such that $F(v_j) = v'_j$ and $F_*([\delta_j]) = [\delta'_j]$ for all $1 \leq j \leq \ell$. Here F_* are the maps induced on the relevant fundamental groups. This condition says that F preserves the orientation of each boundary component. And,
- there is a homotopy between $f' \circ F$ and f on Σ. This homotopy is relative to the points $\{v_j\}$.

The set $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is defined to be the resulting collection of equivalence classes $[(\Sigma, f)]$ of admissible maps for $w_1, \ldots, w_\ell$. $\mathsf{Surfaces}^*$

Figure 1.1 illustrates the concept of admissible maps. Definition 1.3 modifies the definition of $\mathsf{Surfaces}(w_1, \ldots, w_\ell)$ given in [MP19, Def. 1.3], by dropping the requirement that Σ is orientable, and any compatibility between the orientations of boundary components of Σ. Hence $\mathsf{Surfaces}(w_1, \ldots, w_\ell) \subset \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$. The set $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ will index the summation in our Laurent series expansion of $\mathcal{T}r^{O}_{w_1,\ldots,w_\ell}(n)$, but to understand the contribution of a given $[(\Sigma, f)]$ we must take into account the internal symmetries of this pair.

If Σ is a compact surface, the mapping class group of Σ, denoted $\mathrm{MCG}(\Sigma)$, is the collection of homeomorphisms of Σ that fix the boundary pointwise, modulo homeomorphisms that are isotopic to the identity through homeomorphisms of this type. Note that for Σ fixed, $\mathrm{MCG}(\Sigma)$ acts on the collection of homotopy classes $[f]$ of f such that $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$. To take into account the function f in our definition of symmetries, we make the following definition.

Definition 1.4. For $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$, we define $\mathrm{MCG}(f)$ to be the stabilizer of $[f]$ in $\mathrm{MCG}(\Sigma)$. This is a well-defined subgroup of $\mathrm{MCG}(\Sigma)$ up to conjugation, and so a well-defined group up to isomorphism. $\mathrm{MCG}(f)$

A certain integer invariant of $\mathrm{MCG}(f)$, where $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$, appears in our formula for $\mathcal{T}r^{G}_{w_1,\ldots,w_\ell}(n)$. The invariant that appears is the L^2*-Euler characteristic,* denoted by $\chi^{(2)}(\mathrm{MCG}(f))$, and defined precisely in §5.1. This is defined for a class of groups that we will prove in

§5.2 contains all $\mathrm{MCG}(f)$ where $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$. Moreover, for certain $[(\Sigma, f)]$, including all those contributing to the 'leading and second-leading order' terms of the Laurent series expansion of $\mathcal{T}r^G_{w_1,\ldots,w_\ell}(n)$, we will prove in §5.3 that $\chi^{(2)}(\mathrm{MCG}(f))$ coincides with a much tamer invariant: the Euler characteristic of a finite CW-complex that is an Eilenberg-Maclane space of type $K(\mathrm{MCG}(f), 1)$. We make this point precise in Theorem 5.13. Some examples are given in Table 1.

Another simpler type of Euler characteristic also appears in our formula. If $[(\Sigma, f)]$ is in $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ we write $\chi(\Sigma)$ for the usual topological Euler characteristic of Σ. Clearly this does not depend on the representative chosen for $[(\Sigma, f)]$.

We can now state our main theorem.

Theorem 1.5. *There is*[4] $M = M(w_1, \ldots, w_\ell) \geq 0$ *such that for* $n > M$, $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ *is given by the following absolutely convergent Laurent series in* $(n-1)^{-1}$*:*

$$\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n) = \sum_{[(\Sigma,f)] \in \mathsf{Surfaces}^*(w_1,\ldots,w_\ell)} (n-1)^{\chi(\Sigma)} \chi^{(2)}(\mathrm{MCG}(f)). \tag{1.4}$$

This also gives a Laurent series expansion in $(n + \frac{1}{2})^{-1}$ *for* $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n)$ *in view of Theorem 1.2.*

Remark 1.6. We have $\chi(\Sigma) \leq \ell$ for $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ (see Lemma 2.4 and its preceding paragraph). Hence the powers of $n-1$ that appear in (1.4) are bounded above.

Remark 1.7. In the course of the proof of Theorem 1.5 we prove that for each fixed χ_0, there are only finitely many $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ such that $\chi(\Sigma) = \chi_0$ and $\chi^{(2)}(\mathrm{MCG}(f)) \neq 0$. We also prove that each $\chi^{(2)}(\mathrm{MCG}(f)) \in \mathbf{Z}$, so the Laurent series has integer coefficients.

Remark 1.8. Although Theorem 1.5 holds if some of the $w_i = 1$, it simplifies the paper to assume that all $w_i \neq 1$, which we do from now on.

Remark 1.9. As claimed above, the statement of Theorem 1.5 indeed gives an expansion of $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ in terms of $\mathrm{Aut}(\mathbf{F}_r)$-invariants of the words. Indeed, if $\alpha \in \mathrm{Aut}(\mathbf{F}_r)$, there is a corresponding map $g_\alpha \colon (\bigvee^r S^1, o) \to (\bigvee^r S^1, o)$ such that $(g_\alpha)_* = \alpha$, and then every $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ corresponds to $[(\Sigma, g_\alpha \circ f)] \in \mathsf{Surfaces}^*(\alpha(w_1), \ldots, \alpha(w_\ell))$, satisfying $\mathrm{MCG}(f) \cong \mathrm{MCG}(g_a \circ f)$. For a more detailed argument, see [Bro22, Lem. 4.4].

Theorem 1.5 can be viewed as analogous to the genus expansions of GOE matrix integrals in terms of Euler characteristics of mapping class groups obtained by Goulden-Jackson [GJ97] and Goulden, Harer, and Jackson [GHJ01]. These results extended previous results of Harer and Zagier [HZ86] on the GUE ensemble.

The connection between $\mathrm{O}(n)$ matrix integrals and not-necessarily-oriented surfaces was previously pointed out by Mingo and Popa [MP13] and Redelmeier [Red15].

1.2 Consequences of the main theorem

Theorem 1.5 has several important corollaries. Firstly, for a fixed word $w \in \mathbf{F}_r$ the rate of decay of $\mathcal{T}r^G_w(n)$ as $n \to \infty$ can be bounded in terms of the *square length of* w and the *commutator length of* w.

Definition 1.10. The *square length of* $w \in \mathbf{F}_r$, denoted $\mathrm{sql}(w)$, is the minimum number s such that w can be written as the product of s squares, or ∞ if it is not possible to write w as the product of squares. The *commutator length of* w, denoted $\mathrm{cl}(w)$, is the minimum number g such that w can be written as the product of g commutators, or ∞ if w cannot be written as the product of commutators.

One has the elementary inequality

$$\mathrm{sql}(w) \leq 2\mathrm{cl}(w) + 1. \tag{1.5}$$

[4] See (3.9) for the precise definition of M.

This follows from the identities

$$[a,b]=(ab)^2(b^{-1}a^{-1}b)^2(b^{-1})^2,\quad a^2[b,c]=(a^2ba^{-1})^2(ab^{-1}a^{-1}ca^{-1})^2(ac^{-1})^2$$

for any $a,b,c\in\mathbf{F}_r$. In particular, if $\mathrm{sql}(w)=\infty$ then $\mathrm{cl}(w)=\infty$.

Corollary 1.11. *For $G=\mathrm{O}$ or Sp, we have*

$$\mathcal{T}r_w^G(n)=O\left(n^{1-\min(\mathrm{sql}(w),2\mathrm{cl}(w))}\right)$$

as $n\to\infty$. We interpret the right hand side as 0 when $\mathrm{sql}(w)=\infty$.

Remark 1.12. The inequality (1.5) implies that $\mathcal{T}r_w^G(n)=O(n^{1-\mathrm{sql}(w)})$ unless (1.5) is an equality. In that case, $\mathrm{sql}(w)=2\mathrm{cl}(w)+1$ and $\mathcal{T}r_w^G(n)=O(n^{2-\mathrm{sql}(w)})$. An analogous result holds for $G=\mathrm{U}$ the unitary group: $\mathcal{T}r_w^{\mathrm{U}}(n)=O\left(n^{1-2\mathrm{cl}(w)}\right)$ [MP19, Cor. 1.8]. Yet another analogous result, albeit with quite a different flavor, holds in the case of the symmetric group. Let $\mathcal{T}r_w^{\mathrm{Sym}}(n)$ denote the trace of the $(n-1)$-dimensional irreducible "standard" representation of $\mathrm{Sym}(n)$. Then $\mathcal{T}r_w^{\mathrm{Sym}}(n)=C\cdot n^{1-\pi(w)}+O\left(n^{-\pi(w)}\right)$, where $\pi(w)$ is the smallest rank of a subgroup of $\mathbf{F}_r$ containing w as a non-free-generator and $C=C(w)$ is some positive integer [PP15, Thm 1.8]. See also [MP21, Thm. 1.11] for similar results in the case of generalized symmetric groups.

Remark 1.13. Although we have stated Corollary 1.11 as a corollary of Theorem 1.5, the proof is significantly simpler and does not require any discussion of L^2-invariants. We explain the proof of Corollary 1.11 in §3.5 below.

Remark 1.14. The consequence of Corollary 1.11 that $\mathcal{T}r_w^G(n)=0$ when w cannot be written as the product of squares can also be proved directly from the definition of $\mathcal{T}r_w^G(n)$ – see Lemmas 2.6 and 2.3.

Remark 1.15. When $w=x_1^2x_2^2\dots x_s^2\in\mathbf{F}_r$, $r\geq s$, a result of Frobenius and Schur [FS06] (for $s=1$) and a straightforward generalization (see [MP21, §2] or [PS14, Prop. 3.1(3)]) give

$$\mathcal{T}r_w^{\mathrm{O}}(n)=\frac{1}{n^{s-1}},\quad \mathcal{T}r_w^{\mathrm{Sp}}(n)=\frac{(-1)^s}{(2n)^{s-1}}. \tag{1.6}$$

In fact, analogs of these formulas hold for any compact group G. Thus Theorem 1.5 and Corollary 1.11 can be viewed as a generalization of these formulas to arbitrary words in $\mathbf{F}_r$. However, this generalization is not as simple as it might appear on the surface. For example, when $w=x_1^2x_2^2$, combining Theorem 1.5 with (1.6) gives for large n

$$n^{-1}=\sum_{[(\Sigma,f)]\in\mathsf{Surfaces}^*(x_1^2x_2^2)}(n-1)^{\chi(\Sigma)}\chi^{(2)}(\mathrm{MCG}(f)).$$

Since n^{-1} does not agree with any Laurent polynomial of $n-1$ for large n, this implies that there are infinitely many different χ_0 such that there is $[(\Sigma,f)]$ in $\mathsf{Surfaces}^*(x_1^2x_2^2)$ with $\chi(\Sigma)=\chi_0$ and $\chi^{(2)}(\mathrm{MCG}(f))\neq 0$. We analyze these $[(\Sigma,f)]$ in Example 6.2.

Remark 1.16. In another direction, in [MP21] we show using in part Corollary 1.11 that the word $x_1^2\dots x_s^2$ is uniquely determined, up to automorphisms, by the 'word measures' induced by the word on compact groups, that is, the pushforwards of Haar measures on G^r under the word map w.

Corollary 1.17. *If all $w_1,\dots,w_\ell$ are not equal to 1, the limit $\lim_{n\to\infty}\mathcal{T}r^{\mathrm{O}}_{w_1,\dots,w_\ell}(n)$ exists, and is an integer that counts the number of pairs $[(\Sigma,f)]\in\mathsf{Surfaces}^*(w_1,\dots,w_\ell)$ such that all the connected components of Σ are annuli or Möbius bands.*

In algebraic terms, this integer is the weighted number of ways to partition $w_1, \ldots, w_\ell$ into singletons and pairs, such that the word in every singleton is a square, and every pair $\{w, w'\}$ has the property that w' is conjugate to either w or w^{-1}. The weight given to such a partition is

$$\prod_{\{w,w'\}} \max\{d \geq 1 \ : \exists u \text{ s.t. } w = u^d\}$$

where $\{w, w'\}$ run over the pairs of the partition.

We prove Corollary 1.17 in §6.1. It implies, in turn, the following corollary:

Corollary 1.18. *Suppose that $w \neq 1$ and $w = u^d$ with $d \geq 1$ such that $u \in \mathbf{F}_r$ is not a proper power of any other element of $\mathbf{F}_r$. Then for all $\ell \geq 1$ and $j_1, \ldots, j_\ell \in \mathbf{Z}$, the limit*

$$\lim_{n\to\infty} \mathcal{T}r^{\mathrm{O}}_{w^{j_1},\ldots,w^{j_\ell}}(n)$$

exists and only depends on d and the j_k, not on u. Moreover, this collection of limits determines d. The same result holds with O replaced by Sp.

The phenomenon observed in Corollary 1.18 is also known to be present for other families of groups including unitary groups $U(n)$ [MŚS07, Rădu06], symmetric groups S_n [Nic94] (complemented in [LP10, HP22]), and $\mathrm{GL}_n(\mathbb{F}_q)$ [EWPS21]. See also [DP21, Thm. 1.3] for a similar phenomenon for Ginibre ensembles. We will prove Corollary 1.18 in §6.1.

In the case $r = 1$, so that there is only one letter $x = x_1$, and $w_k = x^{j_k}$ for $k = 1, \ldots, \ell$, we have

$$\mathcal{T}r^{\mathrm{O}}_{x^{j_1},\ldots,x^{j_\ell}}(n) = \int_{\mathrm{O}(n)} \mathrm{tr}\left(g^{j_1}\right) \cdots \mathrm{tr}\left(g^{j_\ell}\right) d\mu_n(g).$$

Diaconis and Shahshahani [DS94, Thm. 4] prove that in this case, $\mathcal{T}r^{\mathrm{O}}_{x^{j_1},\ldots,x^{j_\ell}}(n)$ is *exactly* the integer described in Corollary 1.17 for any n sufficiently large, and give a closed formula for this integer. Moreover in *(ibid.)* Diaconis and Shahshahani use this fact to prove that for $j \in \mathbb{N}$, the collection of random variables $\mathrm{tr}(g), \mathrm{tr}\left(g^2\right) \ldots, \mathrm{tr}\left(g^j\right)$, where g is chosen according to Haar measure on $\mathrm{O}(n)$, converge in probability as $n \to \infty$ to independent normal variables with different centers and variances. Using the method of moments, Corollary 1.18 implies that one has the same result if x_1 is replaced by any non-trivial word w that is not a proper power. More precisely, we have the following result.

Corollary 1.19 (Universality for traces of non-powers)**.** *Let $w \in \mathbf{F}_r$, $w \neq 1$, and w not a proper power of another element in $\mathbf{F}_r$. For fixed $\ell \geq 1$, consider the real-valued random variables $T_n\left(w^j\right) \stackrel{\mathrm{def}}{=} \mathrm{tr}\left(w\left(g_1, \ldots, g_r\right)^j\right)$ on the probability space $\left(\mathrm{O}(n)^r, \mu_n^r\right)$ for $j = 1, \ldots, \ell$. The $T_n\left(w^j\right)$ converge in probability as $n \to \infty$:*

$$\left(T_n(w), T_n\left(w^2\right), \ldots, T_n\left(w^\ell\right)\right) \xrightarrow[\text{probability}]{n\to\infty} \left(Z_1, Z_2, \ldots, Z_\ell\right),$$

where $Z_1, \ldots, Z_\ell$ are independent real normal random variables, such that when j is odd, Z_j has mean 0 and variance j, and when j is even, Z_j has mean 1 and variance $j + 1$.

Some related results were previously obtained by Mingo and Popa in [MP13] where they obtain 'real second order freeness' of independent Haar elements of $\mathrm{O}(n)$. This concept does not seem to imply the explicit statement of Corollary 1.19.

1.3 Duality between Sp and O, and the parameters α_G

Duality between Sp **and** O

The formula that appears in Theorem 1.2 was suggested by Deligne [Del16] in a private communication to the second named author of this paper, without the explicit calculation of the sign $(-1)^\ell$. Deligne's reasoning is that one has the identities of 'supergroups'

$$\mathrm{O}(-2n) = \mathrm{O}(0|2n) = \mathrm{Sp}(n).$$

We do not know how to make this into a rigorous concise proof of Theorem 1.2 at the moment. The proof we give in the Appendix relies on a technical combinatorial comparison of the terms arising in $\mathcal{T}r^{\mathrm{O}}_{w_1,\dots,w_\ell}$ and $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\dots,w_\ell}$ from the Weingarten Calculus.

Formulas similar to Theorem 1.2 were observed by Mkrtchyan [MKR81] in the early 1980s in the setting of O vs Sp gauge theory. The duality also shows up as a duality between the GOE and GSE ensembles [MW03]. The introduction to *(ibid.)* also contains an overview of what was known about the O-Sp duality at that time. The $\mathrm{O}(-2n) = \mathrm{Sp}(n)$ formula has more recently been interpreted in a different way, in terms of Casimir operators, by Mkrtchyan and Veselov in [MV11].

The parameters α_G

In this section we mention other occurrences of the parameters α_G in random matrix theory. We have observed in [MP19] and the current paper that it is most natural to expand $\mathcal{T}r^{G}_{w_1,\dots,w_\ell}(n)$ as a Laurent series in $(n+1-\alpha_G)^{-1}$, where $\alpha_G = 2, 1, \frac{1}{2}$ for $G = \mathrm{O}, \mathrm{U}, \mathrm{Sp}$ respectively. It is more common in the literature that the parameter

$$\beta_G \stackrel{\text{def}}{=} \frac{2}{\alpha_G}$$

appears. Hence $\beta_G = 1, 2, 4$ for $G = \mathrm{O}, \mathrm{U}, \mathrm{Sp}$ respectively. One sees that $1, 2, 4$ correspond to the dimensions of the real, complex, and quaternion numbers as real vector spaces, and this is surely the fundamental source of the parameter β_G.

One historical role that these parameters play in random matrix theory is that they give unified expressions for the joint eigenvalue distributions of Dyson's Orthogonal, Unitary, and Symplectic Circular Ensembles[5] (COE/CUE/CSE) introduced by Dyson in [Dys62]. Results of Weyl [Wey39] (for the CUE) and Dyson [Dys62, Thm. 8] (for the COE and CSE) say that the joint eigenvalue density of a matrix in one of these ensembles is given by

$$C_\beta \prod_{k\neq l} |\exp(i\theta_k) - \exp(i\theta_l)|^\beta$$

where $\exp(i\theta_k)$ are the eigenvalues of the random matrix, C_β is a normalizing constant, and $\beta = 1$ for COE, $\beta = 2$ for CUE, and $\beta = 4$ for CSE.

More recently, and in a context more closely related to the current paper, the parameters α_G appear in the work of Novaes [Nov17] who calculates Laurent series in $(n+1-\alpha_G)^{-1}$ for Weingarten functions on $G(n)$ with $G = \mathrm{U}, \mathrm{O}, \mathrm{Sp}$. The origin of the shifted parameter in (*loc. cit.*) is its use of the following result of Forrester [For06, eq. 3.10]: if B is sampled from $\mathrm{U}(n), \mathrm{O}(n)$, or $\mathrm{Sp}(n)$ according to Haar measure, then the density function of the top left $m_1 \times m_2$ submatrix A of g is given by a determinant involving A raised to an exponent that is an explicit function of α_G. This is very different to the appearance of α_G in the current paper.

Indeed, the origin of α_G in the current paper is *topological* and based on the following observations:

[5] The COE of dimension n is the space of symmetric unitary matrices. This can be identified with $\mathrm{U}(n)/\mathrm{O}(n)$ and as such, has a natural probability measure coming from Haar measure on $\mathrm{U}(n)$. The CUE of dimension n is $\mathrm{U}(n)$ with its Haar measure. The CSE of dimension $2n$ is the space of self-dual unitary matrices, that can be identified with $\mathrm{U}(2n)/\mathrm{Sp}(n)$ and hence given the probability measure coming from Haar measure on $\mathrm{U}(2n)$.

- There are exactly two types of connected surfaces with boundary that have trivial mapping class group: a disc, and a Möbius band (cf. Lemma 5.6 and Proposition 5.8).
- As a result, when we compute the terms $\chi^{(2)}(\mathrm{MCG}(f))$ that appear in Theorem 1.5, we need to enumerate surfaces that are formed by gluing together discs **and** Möbius bands.
- On the other hand, using the Weingarten calculus to expand $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ leads to a formula that involves enumerating surfaces that are formed only by gluing discs (i.e., given as CW-complexes). This formula is given in Proposition 3.6.

Hence, one must at some stage pass from a formula involving surfaces given as CW-complexes to a formula involving surfaces formed by gluing together discs and Möbius bands. This is somewhat surprisingly accomplished simply by replacing the parameter n by $n-1$ in the case $G=\mathrm{O}$, and is given by Proposition 3.13.

We suspect that all these appearances of the parameter α_G (or β_G) in different results in random matrix theory are all related, but we do not know how to directly explain this relation.

1.4 Notation and paper organization

For $n\in\mathbf{N}$, we use the notation $[n]$ for the set $\{1,2,\ldots,n\}$. If f is a map between topological spaces with base points, then f_* is the induced map between the fundamental groups of the spaces. We write $\emptyset$ for the empty set. If $w\in\mathbf{F}_r$ we write $|w|$ for the word length of w in reduced form. We write log for the natural logarithm (base e).

The paper is organized as follows. Section 2 describes how one can construct admissible maps in $\mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$ from sets of matchings of the letters of $w_1,\ldots,w_\ell$, and gives some basic definitions and facts about general elements of $\mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$. In Section 3 we discuss the Weingarten calculus, give a combinatorial formula for $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ (Theorem 3.4), derive two different Laurent series expansions of $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ (Propositions 3.6 and 3.13), and reduce our main theorem, Theorem 1.5, to a theorem about a single admissible map in $\mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$ (Theorem 3.16). Section 4 introduces the complex of transverse maps associated with some $[(\Sigma,f)]\in\mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$ and proves it is contractible, following closely with the analogous result in [MP19]. In Section 5 we finish the proof of Theorems 3.16 and 1.5, and in Section 6 we prove Corollaries 1.17, 1.18 and 1.19, and discuss some concrete examples. Finally, Appendix A gives a combinatorial proof of Theorem 1.2.

Acknowledgements

We thank Pierre Deligne for the formulation of Theorem 1.2. We also thank Marcel Novaes and Sasha Veselov for helpful discussions related to this work, and the anonymous referee for several suggestions that improved the presentation of this work.

2 Maps on surfaces

In the rest of this paper, we view $w_1,\ldots,w_\ell\in\mathbf{F}_r$ as fixed. For a given word $w_j\in\mathbf{F}_r$ we may write

$$w_j = x_{i_1^j}^{\varepsilon_1^j} x_{i_2^j}^{\varepsilon_2^j}\cdots x_{i_{|w_j|}^j}^{\varepsilon_{|w_j|}^j},\quad \varepsilon_u^j\in\{\pm1\}, i_u^j\in[r], \tag{2.1}$$

where if $i_u^i = i_{u+1}^j$, then $\varepsilon_u^j=\varepsilon_{u+1}^j$. In other words, we write each w_j in reduced form. Recall that B is the basis $x_1,\ldots,x_r$. We define the *total unsigned exponent* of a generator $x_t\in B$ in $w_1,\ldots,w_\ell$ to be total unsigned exponent

$$\sum_{j=1}^{\ell}\#\{1\le u\le |w_j| \,:\, i_u^j = t\}.$$

2.1 Construction of maps on surfaces from matchings

In this section we assume that the total unsigned exponent of x in $w_1, \ldots, w_\ell$ is even for each $x \in B$, and write $2L_x$ for this quantity. L_x

2.1.1 Matchings and permutations

A *matching* of the set $[2k] = \{1, \ldots, 2k\}$ is a partition of $[2k]$ into pairs. The collection of matchings of $[2k]$ is denoted by M_k. We use two ways to identify a matching in M_k with a permutation in S_{2k}. In M_k the first we identify a matching m with a permutation whose cycle decomposition consists of disjoint transpositions given by the matched pairs of m. We call the resulting permutation π_m. π_m

In the second, given a matching $m \in M_k$ we canonically view m as an ordered list of ordered pairs $((m_{(1)}, m_{(2)}), (m_{(3)}, m_{(4)}), \ldots, (m_{(2k-1)}, m_{(2k)}))$ with

$$m_{(1)} < m_{(3)} < \ldots < m_{(2k-1)}, \quad m_{(2r-1)} < m_{(2r)},\ r \in [k]. \tag{2.2}$$

As such, we have an embedding $M_k \to S_{2k}$, $m \mapsto \sigma_m$ by sending the matching m to the permutation $\sigma_m : i \mapsto m_{(i)}$. σ_m

Following Collins and Śniady [CŚ06] we introduce a metric on M_k as follows. For a permutation σ, write $|\sigma|$ for the minimum number of transpositions that σ can be written as a product of. For matchings $m, m' \in M_k$, we define ρ

$$\rho(m, m') = \frac{|\pi_m \pi_{m'}|}{2}. \tag{2.3}$$

Since both permutations π_m and $\pi_{m'}$ have the same sign, $\rho(m, m') \in \mathbf{Z}_{\geq 0}$.

2.1.2 Markings of $\bigvee^r S^1$

We now describe certain markings of $\bigvee^r S^1$ that will be used in our construction of maps on surfaces. For any given tuple of positive integers $\{\kappa_x\}_{x \in B}$ we will mark additional points on the circles of $\bigvee^r S^1$ as follows. On the circle corresponding to the each generator $x \in B$ we mark $\kappa_x + 1$ distinct ordered points $(x, 0), \ldots, (x, \kappa_x)$, placed consecutively along the circle according to its orientation, and disjoint from o. This is illustrated in the bottom part of Figure 2.1 for the case $B = \{x, y\}$ and $\kappa_x = \kappa_y = 1$.

2.1.3 Construction of a map on a circle from a combinatorial word

Recall our ongoing assumptions that all words are $\neq 1$. For each word $w \neq 1$ we construct an oriented marked circle $C(w)$. Write

$$w = x_{j_1}^{\varepsilon_1} x_{j_2}^{\varepsilon_2} \ldots x_{j_{|w|}}^{\varepsilon_{|w|}} \tag{2.4}$$

in reduced form (as in (2.1)). Begin with $|w|$ disjoint copies of $[0, 1]$, and denote the uth copy by $[0, 1]_u$. Give each interval the orientation from 0 to 1. On each interval choose arbitrarily a map

$$\gamma_u : [0, 1]_u \to \bigvee\nolimits^r S^1$$

such that $\gamma_u(0) = \gamma_u(1) = o$, $\gamma_u : (0, 1)_u \to S^1_{x_{j_u}} - \{o\}$ is a diffeomorphism, and the loop in $\bigvee^r S^1$ parameterized by γ_u, based at o, corresponds to $x_{j_u}^{\varepsilon_u} \in \mathbf{F}_r$ at the level of the fundamental group. Now cyclically concatenate all the intervals and maps together to obtain a circle $C(w)$ and a map $\gamma_w : C(w) \to \bigvee^r S^1$.

Let v_w be the initial point $0_1 \in [0, 1]_1$ of this circle $C(w)$. The map γ_w has the property that $\gamma_w(v_w) = o$ and $(\gamma_w)_*$ maps a generator of $\pi_1(C(w), v_w)$ to $w \in \mathbf{F}_r$. We give $C(w)$ the orientation such that the order of the intervals read, beginning at v_w and following the orientation, matches the left to right order of (2.4). As such, the intervals of $C(w)$ are in one-to-one correspondence with the letters of w.

To clarify and summarize, by definition $w \in \mathbf{F}_r = \pi_1(o, \bigvee^r S^1)$ corresponds to a homotopy class of a loop based at o. What we have done here is pick a particular representative, γ_w, of this homotopy class such that the sequence of circles traversed by the loop is prescribed by w through (2.4), and each traversal of a circle is monotone.

2.1.4 Construction of a map on a surface from words and matchings

In this section we will describe a construction that takes in the following

Input.

- A tuple $\kappa = \{\kappa_x\}_{x \in B}$ of non-negative integers.
- A collection $\mathbf{m} = \{(m_{x,0}, \ldots, m_{x,\kappa_x})\}_{x \in B}$ that contains for each generator $x \in B$, an ordered $(\kappa_x + 1)$-tuple $(m_{x,0}, \ldots, m_{x,\kappa_x})$ of elements in M_{L_x}, where $2L_x$ is the unsigned exponent of x in $w_1, \ldots, w_\ell$. We denote by $\mathrm{MATCH}^\kappa = \mathrm{MATCH}^\kappa(w_1, \ldots, w_\ell)$ the collection of all possible such $\mathbf{m}$, for fixed κ. We write $\mathrm{MATCH}^* = \cup_{\kappa \in \mathbf{Z}^B_{\geq 0}} \mathrm{MATCH}^\kappa$. If $\mathbf{m} \in \mathrm{MATCH}^\kappa$ we will say $\kappa(\mathbf{m}) = \kappa$.
  Warning: our notation MATCH in this paper is not exactly the same as in [MP19][a].

[a] In [MP19] the matchings are more restrictive: it was assumed that the matchings $m_{x,i}$ were subordinate to a partition of $[2L_x]$ into two blocks, corresponding to the occurrences of x^{+1} and x^{-1} in $w_1, \ldots, w_\ell$, whereas here our matchings are arbitrary matchings of $[2L_x]$.

The output of our construction is the following

MATCH$^\kappa$
MATCH*

Output.

- A surface $\Sigma_\mathbf{m}$ with ℓ oriented boundary components $C(w_1), \ldots, C(w_\ell)$, with a given CW-complex structure, and a marked point v_j on each boundary component $C(w_j)$.
- A continuous function $f_\mathbf{m} : \Sigma_\mathbf{m} \to \bigvee^r S^1$ such that $f_\mathbf{m}|_{C(w_j)}$ is the function γ_{w_j} constructed in §2.1.3. In particular, $f_\mathbf{m}(v_j) = o$ for each $1 \leq i \leq \ell$ and $(f_\mathbf{m})_*$ maps the generator of $\pi_1(C(w_j), v_j)$ specified by the orientation of $C(w_j)$ to w_j.

Note that the resulting $(\Sigma_\mathbf{m}, f_\mathbf{m})$ is an admissible map for $w_1, \ldots, w_\ell$, as in Definition 1.3. As will become clearer in the sequel (mostly §4), different collections $\mathbf{m}$ of matchings may result in the same equivalence class of admissible pairs $(\Sigma_\mathbf{m}, f_\mathbf{m})$.

The construction is essentially the same as in [MP19, §2.2, §2.5], except in the current case there are less restrictions on the matchings present, so the construction can lead to non-orientable surfaces. The construction of a surface from a single matching per basis element (so $\kappa_x = 0$ for all $x \in B$) already appears in [Cul81]. We give the details now.

I. The one-skeleton. We first perform the construction of §2.1.3 for each word w_j. This gives us a collection of oriented based circles $(C(w_j), v_{w_j})$ that are respectively subdivided into $|w_j|$ intervals, and maps $\gamma_{w_j} : C(w_j) \to \bigvee^r S^1$. These oriented circles will form the oriented boundary

$$\delta\Sigma_\mathbf{m} \stackrel{\mathrm{def}}{=} \bigcup_{j=1}^{\ell} C(w_j)$$

of $\Sigma_\mathbf{m}$ and the map

$$\gamma \stackrel{\mathrm{def}}{=} \cup_{j=1}^{\ell} \gamma_{w_j} : \delta\Sigma_\mathbf{m} \to \bigvee^r S^1$$

will be the restriction of $f_\mathbf{m}$ to the boundary of $\Sigma_\mathbf{m}$. We write $v_j = v_{w_j}$ in the sequel for the marked points on the $C(w_j)$.

For each $x \in B$, the preimage

$$\gamma^{-1}(S^1_x - \{o\})$$

is a disjoint union of $2L_x$ open sub-intervals of $\delta\Sigma_{\mathbf{m}}$. This collection of intervals is denoted by $\mathcal{I}_x = \mathcal{I}_x(w_1, \ldots, w_\ell)$. We identify $\mathcal{I}_x$ with $[2L_x]$ in some arbitrary but fixed way. Moreover, each $\mathcal{I}_x$ element of $\mathcal{I}_x$ corresponds to an appearance of x or x^{-1} in a unique w_j.

Next we add to the one-skeleton of $\Sigma_{\mathbf{m}}$ by gluing some arcs by their endpoints to $\delta\Sigma_{\mathbf{m}}$. We will call these arcs *matching arcs*. For each $x \in B$ we mark points $(x, 0), \ldots, (x, \kappa_x)$ on S^1_x as in §2.1.2. matching arcs Since γ maps each interval $I \in \mathcal{I}_x$ monotonically to S^1_x, there are uniquely specified distinct points $p_I(0), \ldots, p_I(\kappa_x)$ in I such that $\gamma(p_I(k)) = (x, k)$ for $0 \leq k \leq \kappa_x$. $p_I(k)$

Now, for every pair of intervals I and J that are matched by $m_{x,k}$, glue a matching arc to $\delta\Sigma_{\mathbf{m}}$ with endpoints at $p_I(k)$ and $p_J(k)$. Carrying out this process for all generators x, now every $p_I(k)$ point in $\delta\Sigma_{\mathbf{m}}$ is the endpoint of a unique arc. Call the resulting one dimensional CW-complex $\Sigma^{(1)}_{\mathbf{m}}$. It consists of $\delta\Sigma_{\mathbf{m}}$ together with the matching-arcs described in the current paragraph. More precisely, we consider $\Sigma^{(1)}_{\mathbf{m}}$ to have the following 0 and 1-cells:

- The 0-cells (vertices) are just the points $p_I(k)$. Since there are $2L_x$ intervals in $\mathcal{I}_x$ labeled by x and $\kappa_x + 1$ points $p_I(k)$ in each such interval, there are $2\sum_{x \in B} L_x(\kappa_x + 1)$ vertices in the 0-skeleton.
- The 1-cells (edges) are of two different types. Firstly, there is an edge in $\delta\Sigma_{\mathbf{m}}$ for each connected component of $\delta\Sigma_{\mathbf{m}} - \{p_I(k)\}$. Hence there are $2\sum_{x \in B} L_x(\kappa_x + 1)$ such edges. Secondly, there is an edge for each matching-arc. Since the matching arcs give a matching of the points $p_I(k)$, there are $\sum_{x \in B} L_x(\kappa_x + 1)$ such edges. Hence in total there are $3\sum_{x \in B} L_x(\kappa_x + 1)$ edges.

Additionally, each component $C(w_j)$ of $\delta\Sigma_{\mathbf{m}}$ contains a marked point v_j. These are not part of the CW-complex structure (nor is any point of $\gamma^{-1}(o)$). We define $f^{(1)}_{\mathbf{m}}$ on $\Sigma^{(1)}_{\mathbf{m}}$ by $f^{(1)}_{\mathbf{m}}|_{\delta\Sigma_{\mathbf{m}}} = \gamma$ and by requiring that $f^{(1)}_{\mathbf{m}}$ is constant on all matching arcs. This completely specifies $f^{(1)}_{\mathbf{m}}$, since the endpoints of matching arcs are in $\delta\Sigma_{\mathbf{m}}$, so the value of $f^{(1)}_{\mathbf{m}}$ on matching arcs is specified by γ, and by construction of the matching arcs, the two endpoints of any two matching arcs have the same value under γ.

II. The two-skeleton. Next we complete the construction of $\Sigma_{\mathbf{m}}$ by gluing in discs that will be the 2-cells of the CW-complex. We glue in different types of discs as follows.

Type-(x, k) discs. For fixed x, if $0 \leq k < \kappa_x$, let $R_{x,k}$ be the open interval in $S^1_x - \{(x, i)\} - \{o\}$ that abuts the points (x, k) and $(x, k+1)$. Let $\overline{R_{x,k}}$ be the closure of this component. The preimage $(f^{(1)}_{\mathbf{m}})^{-1}(\overline{R_{x,k}})$ is a collection of disjoint cycles in $\Sigma^{(1)}_{\mathbf{m}}$ and for each of these cycles we glue the boundary of a disc simply along the cycle. These discs are called *type-(x, k) discs.*

Type-o discs. Let R_o be the star-like connected component of $\bigvee^r S^1 - \{(x, i)\}$ that contains the point o and abuts points of the form $(x, 0)$ and (x, κ_x) (for all x). Let $\overline{R_o}$ be the closure of this component. The preimage $(f^{(1)}_{\mathbf{m}})^{-1}(\overline{R_o})$ is a one-dimensional subcomplex in $\Sigma^{(1)}_{\mathbf{m}}$. We consider simple cycles c in $(f^{(1)}_{\mathbf{m}})^{-1}(\overline{R_o})$ with the property that $f^{(1)}_{\mathbf{m}}|_c$ <u>never traverses a point</u> $(x, 0)$ or (x, κ_x). Namely, when going along c, whenever $f^{(1)}_{\mathbf{m}}(c)$ reaches some $(x, 0)$ or (x, κ_x), it then stays at this point for a while (as c itself travels along a matching-arc) and then leaves in a backtracking move. For such a cycle we glue the boundary of a disc simply along the cycle. We call the disc a *type-o* disc.

The resulting (oriented) surface obtained by gluing in these discs is $\Sigma_{\mathbf{m}}$. Its CW-complex structure is the CW-complex we described for $\Sigma^{(1)}_{\mathbf{m}}$ together with the glued discs as two-cells.

Note that each disc D of $\Sigma_{\mathbf{m}}$ has its boundary mapped to a nullhomotopic curve in $\bigvee^r S^1$; for type-(x, k) discs this is because the boundary is mapped into an interval, and for type-o discs this is due to our condition that $f^{(1)}_{\mathbf{m}}$ never traverses a point (x, k) when restricted to the boundary of the disc. Hence we can extend $f^{(1)}_{\mathbf{m}}$ to a continuous function from the disc to $\overline{R_o}$ or $\overline{R_{x,k}}$ respectively, such that only the points in the preimage of $\{(x, k)\}$ are in $\Sigma^{(1)}_{\mathbf{m}}$. In other words, the extended function maps the interior of the disc to either R_o or some $R_{x,k}$. We pick such an extension for each disc and this defines

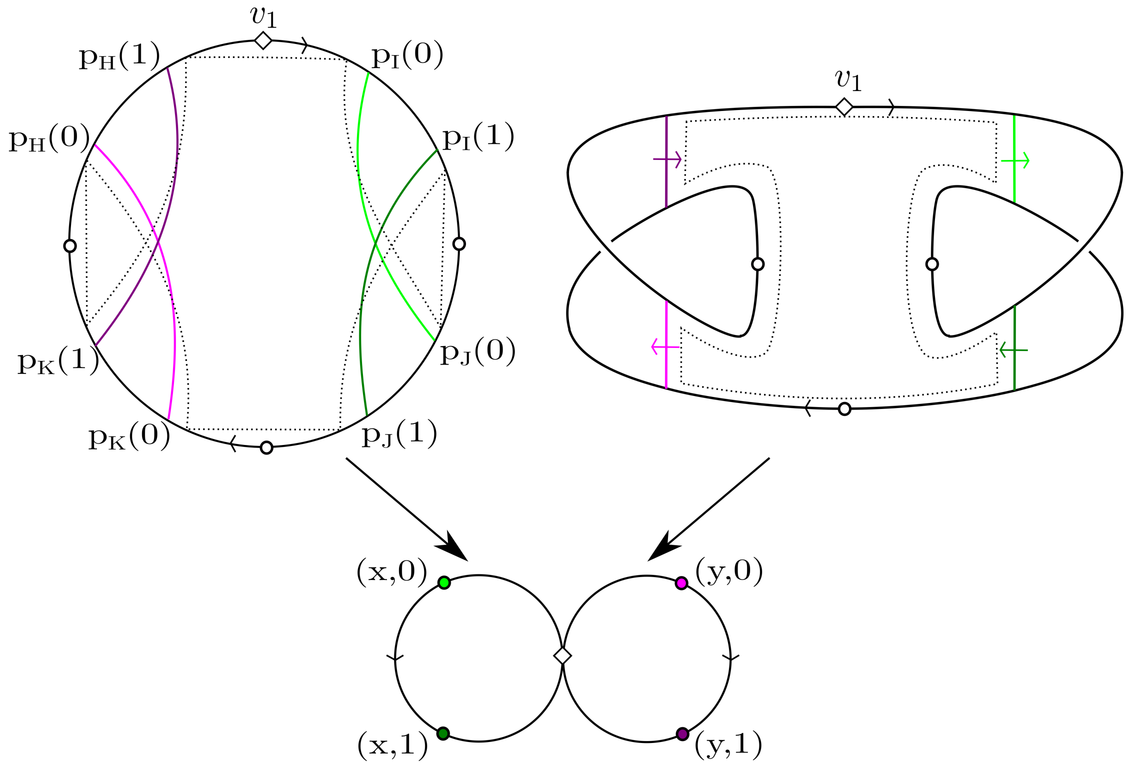

Figure 2.1: Constructing a pair $(\Sigma_{\mathbf{m}}, f_{\mathbf{m}})$ in the case that $r = 2$, $\ell = 1$, and $w_1 = x^2y^2$. We write x, y for the generators of $\mathbf{F}_2$. Here $\kappa_x = \kappa_y = 1$, so there are two matchings per generator. Also, $L_x = L_y = 1$ so $|M_x| = |M_y| = 1$ and the matchings are dictated. The top left picture shows the matching arcs attached to $C(w_1)$. The top right picture shows the constructed $\Sigma_{\mathbf{m}}$. The dotted line in both top pictures follows along the matching arcs to show how the single type-o disc should be glued to $\Sigma_{\mathbf{m}}^{(1)}$. We have called the boundary intervals I, J, K, H. The colors of the points marked in $\vee^2 S^1$ in the bottom picture match with the matching arcs in their preimage under $f_{\mathbf{m}}$. The arrows in the top right picture mark the normal direction on the matching arcs coming from the fixed orientation on the petals of the bouquet.

$f_{\mathbf{m}}$ on all of $\Sigma_{\mathbf{m}}$. Note that the extension at every disk is unique up to homotopy, and therefore $f_{\mathbf{m}}$ in general is well-defined up to homotopy. *This concludes the construction of the pair* $(\Sigma_{\mathbf{m}}, f_{\mathbf{m}})$*.* This construction is illustrates in Figure 2.1.

2.1.5 The Euler characteristic of the constructed surface $\Sigma_{\mathbf{m}}$

In this section we calculate the Euler characteristic of $\Sigma_{\mathbf{m}}$. Recall the definition of ρ from (2.3) and that $\mathbf{m} \in \mathrm{MATCH}^{\kappa}(w_1, \ldots, w_\ell)$ is a collection of matchings.

Lemma 2.1. *We have*

$$\chi(\Sigma_{\mathbf{m}}) = -\sum_{x \in B} L_x + \#\{\textit{type-o discs of } \Sigma_{\mathbf{m}}\} - \sum_{x \in B, 0 \le k < \kappa_x} \rho(m_{x,k}, m_{x,k+1}). \tag{2.5}$$

Proof. We have

$$\begin{aligned}\chi(\Sigma_{\mathbf{m}}) &= V - E + F \\ &= 2\sum_{x\in B} L_x(\kappa_x+1) - 3\sum_{x\in B} L_x(\kappa_x+1) + \#\{\text{type-}o \text{ discs of } \Sigma_{\mathbf{m}}\} + \\ &\quad \sum_{x\in B, 0\le k<\kappa_x} \#\{\text{type-}(x,k) \text{ discs of } \Sigma_{\mathbf{m}}\} \\ &= -1\sum_{x\in B} L_x(\kappa_x+1) + \#\{\text{type-}o \text{ discs of } \Sigma_{\mathbf{m}}\} + \sum_{x\in B, 0\le k<\kappa_x} \#\{\text{type-}(x,k) \text{ discs of } \Sigma_{\mathbf{m}}\}. \quad (2.6)\end{aligned}$$

Recall from §2.1.1 that we can associate to $m \in M_k$ a permutation π_m in S_{2k} all of whose cycles have length 2. As one follows along the boundary of some type-(x,k) disc, namely, along one of the disjoint cycles in $\left(f_{\mathbf{m}}^{(1)}\right)^{-1}\left(\overline{R_{x,k}}\right)$, the matching arcs alternate between arcs mapped to (x,k) and arcs mapped to $(x,k+1)$, and so the boundary corresponds to a cycle of $\pi_{m_{x,k}}\pi_{m_{x,k+1}}$. In the other orientation of the boundary of the same disk we get a second cycle of $\pi_{m_{x,k}}\pi_{m_{x,k+1}}$, so every type-$(x,k)$ disc corresponds to exactly two cycles of $\pi_{m_{x,k}}\pi_{m_{x,k+1}}$. On the other hand, every cycle of $\pi_{m_{x,k}}\pi_{m_{x,k+1}}$ corresponds a to unique type-(x,k) disc. For example, if $L_x = 3$ and, through the identification of $\mathcal{I}_x$ with $[2L_x]$, $\pi_{m_{x,k}} = (12)\,(34)\,(56)$ and $\pi_{m_{x,k+1}} = (13)\,(24)\,(56)$, then $\pi_{m_{x,k}}\pi_{m_{x,k+1}} = (14)\,(23)\,(5)\,(6)$, and $(\Sigma_{\mathbf{m}}, f_{\mathbf{m}})$ contains two type (x,k)-discs: one corresponding to the cycles (14) and (23) of $\pi_{m_{x,k}}\pi_{m_{x,k+1}}$, and one corresponding to the cycles (5) and (6).

Since for any $\sigma \in S_{2L_x}$ we have $|\sigma| = 2L_x - \#\{\text{cycles of } \sigma\}$, we have

$$\begin{aligned}\rho(m_{x,k}, m_{x,k+1}) = \frac{\left|\pi_{m_{x,k}}\pi_{m_{x,k+1}}\right|}{2} &= L_x - \frac{\#\{\text{cycles of } \pi_{m_{x,k}}\pi_{m_{x,k+1}}\}}{2} \\ &= L_x - \#\{\text{type-}(x,k) \text{ discs of } \Sigma_{\mathbf{m}}\}.\end{aligned}$$

Therefore $\#\{\text{type-}(x,k) \text{ discs of } \Sigma_{\mathbf{m}}\} = L_x - \rho(m_{x,k}, m_{x,k+1})$ and from (2.6) we get

$$\begin{aligned}\chi(\Sigma_{\mathbf{m}}) &= -1\sum_{x\in B} L_x(\kappa_x+1) + \#\{\text{type-}o \text{ discs of } \Sigma_{\mathbf{m}}\} + \sum_{x\in B, 0\le k<\kappa_x} L_x - \rho(m_{x,k}, m_{x,k+1}) \\ &= -\sum_{x\in B} L_x + \#\{\text{type-}o \text{ discs of } \Sigma_{\mathbf{m}}\} - \sum_{x\in B, 0\le k<\kappa_x} \rho(m_{x,k}, m_{x,k+1}). \quad (2.7)\end{aligned}$$

□

2.2 The anatomy of $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$

The passage between the algebraic and topological view on $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ stems from the following result of Culler [Cul81, 1.1]:

Lemma 2.2 (Culler)**.** *An element $w \in \mathbf{F}_r$ is a product of g commutators (resp. s squares) if and only if there exists an admissible map (Σ, f) for w with Σ a genus-g orientable surface (resp. a connected sum of s copies of $\mathbf{R}P^2$) with a disc removed.*

We first describe when $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is empty.

Lemma 2.3. *The following are equivalent:*

1. $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ *is non-empty.*
2. *The unsigned exponent of each $x \in B$ in $w_1, \ldots, w_\ell$ is even.*
3. *$w_1 w_2 \cdots w_\ell$ can be written as a product of squares in $\mathbf{F}_r$.*

Proof. This is very close to standard facts, in particular the results appearing in Culler [Cul81], but we give a proof here for completeness.

We first prove that the first statement implies the second. Consider the map

$$h : \pi_1(\bigvee^r S^1, o) \to H_1(\bigvee^r S^1, \mathbf{Z}/2\mathbf{Z})$$

which induces a map $H : \mathbf{F}_r \to (\mathbf{Z}/2\mathbf{Z})^{|B|}$. Then H maps $x_i \in B$ to the corresponding standard generator e_i of $(\mathbf{Z}/2\mathbf{Z})^{|B|}$. If $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ and $\delta_1, \ldots, \delta_\ell$ are the boundary components of Σ then we have

$$0 = \sum_{i=1}^{\ell} [\delta_i] \in H_1(\Sigma, \mathbf{Z}/2\mathbf{Z}).$$

Applying the map induced by f on homology to this equation implies $\sum_{i=1}^{\ell} H(w_i) = 0$, which means the unsigned exponent of each $x \in B$ in $w_1, \ldots, w_\ell$ is even.

Conversely, if the unsigned exponent of each $x \in B$ in $w_1, \ldots, w_\ell$ is even, then the construction of §2.1 shows that $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is non-empty. This proves that the second statement implies the first.

The second statement is easily seen to be equivalent to the statement that the unsigned exponent of each $x \in B$ in $w_1 w_2 \ldots w_\ell$ is even. By what we have proved, this is equivalent to $\mathsf{Surfaces}^*(w_1 w_2 \ldots w_\ell) \neq \emptyset$. Then by Lemma 2.2, this is equivalent to $w_1 w_2 \ldots w_\ell$ being either the product of commutators or the product of squares. But since any commutator is a product of squares (see §1.2), this proves that the second statement is equivalent to the third statement. □

For given $w_1, \ldots, w_\ell$, let $\chi_{\max}$

$$\chi_{\max}(w_1, \ldots, w_\ell) \stackrel{\text{def}}{=} \max\{\chi(\Sigma) \,:\, [(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)\},$$

or $-\infty$ if $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is empty. Note that if $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is non-empty the maximum clearly exists. Indeed, any connected component of Σ with $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ must contain a boundary component, so Σ has at most ℓ components. Moreover, any connected component of Σ has Euler characteristic at most 1, so $\chi(\Sigma) \leq \ell$.

Lemma 2.4. *Assume that all $w_j \neq 1$. Then $\chi_{\max}(w_1, \ldots, w_\ell) \leq 0$. Moreover, an admissible map $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ satisfies $\chi(\Sigma) = 0$ if and only if all the connected components of Σ are annuli or Möbius bands.*

Proof. As mentioned above, if $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ then writing $\Sigma = \cup_{k=1}^m \Sigma_k$ as a union of connected components, we have $m \leq \ell$, each Σ_k has at least one boundary component, and $\chi(\Sigma_k) \leq 1$. However, $\chi(\Sigma_k)$ is equal to 1 only if Σ_k is a disc, which can only happen if the w_j that marks the boundary component of the disc is $w_j = 1$. This is because the boundary component of the disc is nullhomotopic in the disc. So given all $w_j \neq 1$, we have $\chi(\Sigma_k) \leq 0$. Given $\chi(\Sigma) = 0$, this implies $\chi(\Sigma_k) = 0$ for $1 \leq k \leq m$. Finally, by the classification of surfaces, we have $\chi(\Sigma_k) = 0$ only if Σ_k is an annulus or a Möbius band. Conversely, if all the connected components of Σ are annuli or Möbius bands then obviously $\chi(\Sigma) = 0$. □

Lemma 2.5. *If $w \in \mathbf{F}_r$ then $\chi_{\max}(w) = 1 - \min(\mathrm{sql}(w), 2\mathrm{cl}(w))$.*

Proof. Since any $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w)$ has Σ connected, Culler's Lemma 2.2 is easily seen to imply the result, since the Euler characteristic of a genus g orientable surface with one boundary component is $1 - 2g$ and the Euler characteristic of a connected sum of s copies of $\mathbf{R}P^2$ with a disc removed is $1 - s$. □

2.3 Proof of Theorem 1.5 when $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is empty

Lemma 2.6. *If the total unsigned exponent of some $x \in B$ in $w_1, \ldots, w_\ell$ is not even, then $\mathcal{T}r^G_{w_1,\ldots,w_\ell}(n) = 0$ for any $n \geq 1$ and for $G = \mathrm{O}, \mathrm{Sp}$.*

Proof. Suppose, for example, that $x = x_1$. Let u be the (odd) total unsigned exponent of x_1 in $w_1, \ldots, w_\ell$. Note that minus the identity, $-I_n$, is in the center of $\mathrm{O}(n)$. By the translation-invariance of Haar measure, we have

$$\begin{aligned}\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n) &= \int_{\mathrm{O}(n)^r} \mathrm{tr}(w_1(g_1, \ldots, g_r)) \ldots \mathrm{tr}(w_\ell(g_1, \ldots, g_r)) d\mu_n(g_1) \ldots d\mu_n(g_r) \\ &= \int_{\mathrm{O}(n)^r} \mathrm{tr}(w_1(-I_n g_1, \ldots, g_r)) \ldots \mathrm{tr}(w_\ell(-I_n g_1, \ldots, g_r)) d\mu_n(g_1) \ldots d\mu_n(g_r) \\ &= (-1)^u \int_{\mathrm{O}(n)^r} \mathrm{tr}(w_1(g_1, \ldots, g_r)) \ldots \mathrm{tr}(w_\ell(g_1, \ldots, g_r)) d\mu_n(g_1) \ldots d\mu_n(g_r) \\ &= -\int_{\mathrm{O}(n)^r} \mathrm{tr}(w_1(g_1, \ldots, g_r)) \ldots \mathrm{tr}(w_\ell(g_1, \ldots, g_r)) d\mu_n(g_1) \ldots d\mu_n(g_r) \\ &= -\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n).\end{aligned}$$

The proof for $G = \mathrm{Sp}$ is the same, using that $-I_{2n}$ is in the center of $\mathrm{Sp}(n)$. □

Proof of Theorem 1.5 when $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ *is empty.* If $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is empty, then by Lemma 2.3, the total unsigned exponent of some $x \in B$ in $w_1, \ldots, w_\ell$ is not even. Therefore by Lemma 2.6, $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n) = 0$ for any $n \geq 1$. □

2.4 Incompressible and almost-incompressible maps

Certain types of admissible maps (Σ, f) will pay a special role later in the paper (see §5.3). These are the *incompressible* and *almost-incompressible* maps.

Definition 2.7. Let (Σ, f) be an admissible map. We say that (Σ, f) is *compressible* if there is a non-nullhomotopic simple closed curve $c \subset \Sigma$ such that $f(c)$ is nullhomotopic in $\bigvee^r S^1$. We call c a *compressing curve.* We say (Σ, f) is *incompressible* if it is not compressible. If every compressing curve is non-generic in the sense of Definition 5.7 below, namely, if every compressing curve is either one-sided[6] or bounds a Möbius band, we say that (Σ, f) is *almost-incompressible.*

It is clear that if (Σ, f) is (almost-) incompressible and $(\Sigma', f') \approx (\Sigma, f)$ then (Σ', f') is also (almost-) incompressible. So there is a well-defined notion of an element $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ being (almost-) incompressible.

Lemma 2.8. *Let $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$. If $\chi(\Sigma) = \chi_{\max}(w_1, \ldots, w_\ell)$ then $[(\Sigma, f)]$ is incompressible, and if $\chi(\Sigma) = \chi_{\max}(w_1, \ldots, w_\ell) - 1$ then $[(\Sigma, f)]$ is almost-incompressible.*

Proof. Denote $\chi_{\max} \stackrel{\text{def}}{=} \chi_{\max}(w_1, \ldots, w_\ell)$ and assume that $\chi(\Sigma) \geq \chi_{\max} - 1$. If (Σ, f) is compressible, let $c \subset \Sigma$ be a compressing curve. We can cut Σ along c to obtain a new surface Σ_{cut} with either $b = 1$ or $b = 2$ new boundary components. Then we can cap discs on any new boundary components of Σ_{cut} to obtain a new surface Σ'. Moreover, since $f(c)$ was nullhomotopic, it is possible to extend f across these discs to obtain a new pair (Σ', f') with $\chi(\Sigma') = \chi(\Sigma) + b$.

If $b = 1$, then (Σ', f') is admissible, since we must have cut along a one-sided curve and hence did not create any new connected components. Since $\chi(\Sigma') = \chi(\Sigma) + 1$, we must have $\chi(\Sigma) = \chi_{\max} - 1$.

[6] A curve is *one-sided* if a thickening of the curve is a Möbius band, or equivalently, cutting along the curve results in a surface with only one new boundary component.

If $b = 2$ and every connected component of Σ' has boundary, then the pair (Σ', f') is admissible and $\chi(\Sigma') = \chi(\Sigma) + 2$, in contradiction to $\chi(\Sigma) \geq \chi_{\max} - 1$. However, in Σ' we might have created at most one closed component S, in which case the pair (Σ', f') is not admissible. In this case we can simply delete S to obtain an admissible map (Σ'', f''). Since c was non-nullhomotopic, S is not a sphere, hence $\chi(S) \leq 1$ and $\chi(\Sigma'') = \chi(\Sigma') - \chi(S) \geq \chi(\Sigma') - 1 = \chi(\Sigma) + 1$ and therefore $\chi(\Sigma) \leq \chi_{\max} - 1$. Furthermore, if c was generic, S is also not $\mathbf{R}P^2$, and then $\chi(S) \leq 0$ and $\chi(\Sigma) \leq \chi_{\max} - 2$, a contradiction. □

3 A combinatorial Laurent series for $\mathcal{T}r^G_{w_1,\ldots,w_\ell}(n)$

Since we have proved Theorem 1.5 when $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell) = \emptyset$, we assume the contrary for the rest of the paper, and hence in view of Lemma 2.3 we assume that the total unsigned exponent of each x in $w_1, \ldots, w_\ell$ is even. As in Section 2.1, we write $2L_x$ for the total unsigned exponent of x in $w_1, \ldots, w_\ell$.

3.1 The Weingarten calculus

Denote by $\langle \bullet, \bullet \rangle = \langle \bullet, \bullet \rangle_{\mathrm{O}}$ the standard inner form on $\mathbf{R}^n$, and by $\langle \bullet, \bullet \rangle_{\mathrm{Sp}}$ the symplectic form on $\mathbf{C}^{2n}$ given by

$$\langle v, w \rangle_{\mathrm{Sp}} \stackrel{\text{def}}{=} v^T J w$$

(recall (1.2)). We let $\{e_i\}$ be the standard basis of $\mathbf{R}^n$ or $\mathbf{C}^{2n}$. The *Weingarten calculus* allows one to interpret integrals of products of matrix coefficients in $G(n)$, for $G = \mathrm{O}, \mathrm{Sp}$, in terms of matchings. The source of the Weingarten calculus is Brauer-Schur-Weyl duality between G and an appropriate Brauer algebra. e_i

Let $\mathbf{Q}(n)$ denote the ring of rational functions of n with rational coefficients. The following Theorem was proved for $G = \mathrm{O}$ by Collins and Śniady in [CŚ06, Cor. 3.4]. The corresponding theorem for $G = \mathrm{Sp}$ had its proof outlined by Collins and Śniady in [CŚ06, Thm. 4.1 and following discussion]. The theorem was subsequently precisely stated and proved by Collins and Stolz in [CS08, Prop. 3.2]; see also Matsumoto [Mat13, Thm. 2.4]. Recall that M_k marks the set of matchings on $[2k]$.

Theorem 3.1. *For $G = \mathrm{O}, \mathrm{Sp}$, there are unique, computable, functions*

$$\mathrm{Wg}^G_k : M_k \times M_k \to \mathbf{Q}(n)$$

with the following properties. For $m_1, m_2 \in M_k$ and $n \in \mathbf{Z}_{\geq 1}$, let us write $\mathrm{Wg}^G_k(m_1, m_2; n)$ for the evaluation of the rational function $\mathrm{Wg}^G_k(m_1, m_2)$ at n. Assume that $n \geq k$ if $G = \mathrm{O}$ and $2n \geq k$ if $G = \mathrm{Sp}$. We have

$$\int_{G(n)} g_{i_1 j_1} \ldots g_{i_{2k} j_{2k}} d\mu_n(g) = \sum_{m_1, m_2 \in M_k} \delta^G_{\mathbf{i}, m_1} \delta^G_{\mathbf{j}, m_2} \mathrm{Wg}^G_k(m_1, m_2; n).$$

Here $\mathbf{i} = (i_1, \ldots, i_{2k})$, $\mathbf{j} = (j_1, \ldots, j_{2k})$, and

$$\delta^G_{\mathbf{i}, m} \stackrel{\text{def}}{=} \prod_{r=1}^{k} \langle e_{i_{m_{(2r-1)}}}, e_{i_{m_{(2r)}}} \rangle_G$$

where $m_{(j)}$ are as in (2.2). The integral of the product of an odd number of matrix coefficients is 0.

The function Wg^G_k is called the *Weingarten function of G*. Although here we take Theorem 3.1 as the definition of the Weingarten functions, it is certainly worth pointing out that the papers [CŚ06, Mat13] contain explicit formulas for the Weingarten functions. It follows from these formulas that for any $m_1, m_2 \in M_k$, the rational function $\mathrm{Wg}^G_k(m_1, m_2)$ can be computed by an (explicit) algorithm in a finite number of steps.

The following lemma of Matsumoto [Mat13, §2.3.2] shows that the Weingarten functions of O and Sp are related in a simple way.

Lemma 3.2. *For $n \in \mathbf{Z}_{\geq 0}$ with $2n \geq k$ and $m_1, m_2 \in M_k$*

$$\mathrm{Wg}_k^{\mathrm{Sp}}(m_1, m_2; n) = (-1)^k \mathrm{sign}(\sigma_{m_1}\sigma_{m_2}^{-1})\mathrm{Wg}_k^{\mathrm{O}}(m_1, m_2; -2n).$$

The following theorem of Collins and Śniady gives the full Laurent expansion for $\mathrm{Wg}_k^{\mathrm{O}}(m_1, m_2)$ at $n = \infty$ and estimates the order of vanishing. In view of Lemma 3.2, one has analogous results for $G = \mathrm{Sp}$.

Theorem 3.3. *Fix $m_1, m_2 \in M_k$.*

1. *[CŚ06, Lem. 3.12] We have*

$$\mathrm{Wg}_k^{\mathrm{O}}(m_1, m_2; n) = n^{-k} \sum_{\ell \geq 0} \sum_{\substack{m_1 = m'_0, m'_1, \ldots, m'_\ell = m_2 \in M_k \\ m'_i \neq m'_{i+1}}} (-1)^\ell \, n^{-\rho(m'_0, m'_1) - \ldots - \rho(m'_{\ell-1}, m'_\ell)}.$$

   The sum is absolutely convergent for $n \geq k$.

2. *[CŚ06, Thm. 3.13]*

$$\mathrm{Wg}_k^{\mathrm{O}}(m_1, m_2; n) = O_{n \to \infty}(n^{-k-\rho(m_1, m_2)}).$$

3.2 A rational function form of $\mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}$

We wish to find a rational function of n which agrees with the integral

$$\mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}(n) = \int_{\mathrm{O}(n)^r} \mathrm{tr}(w_1(g_1, \ldots, g_r)) \ldots \mathrm{tr}(w_\ell(g_1, \ldots, g_r)) d\mu_n(g_1) \ldots d\mu_n(g_r) \tag{3.1}$$

for sufficiently large n, depending only on $w_1, \ldots, w_\ell$. We denote by $\kappa \equiv 1$ the tuple that assigns 1 to each $x \in B$. We now define N

$$N = N(w_1, \ldots, w_\ell) \stackrel{\mathrm{def}}{=} \max\{L_x \, : \, x \in B\}. \tag{3.2}$$

Theorem 3.4. *For $n \geq N$, we have*

$$\mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}(n) = \sum_{\mathbf{m} \in \mathrm{MATCH}^{\kappa \equiv 1}} n^{\#\{\textit{type-o discs of } \Sigma_{\mathbf{m}}\}} \prod_{x \in B} \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0}, m_{x,1}; n). \tag{3.3}$$

Proof. Assume $n \geq N$. We assume each $w_j = x_{i_1^j}^{\varepsilon_1^j} x_{i_2^j}^{\varepsilon_2^j} \ldots x_{i^j_{|w_j|}}^{\varepsilon^j_{|w_j|}}$ is written as a reduced word, as in (2.1), and aim to evaluate (3.1). Hence we have

$$\mathrm{tr}(w_j(g_1, \ldots, g_r)) = \sum_{q_1^j, \ldots, q^j_{|w_j|}} \left(g_{i_1^j}^{\varepsilon_1^j}\right)_{q_1^j q_2^j} \left(g_{i_2^j}^{\varepsilon_2^j}\right)_{q_2^j q_3^j} \cdots \left(g_{i^j_{|w_j|}}^{\varepsilon^j_{|w_j|}}\right)_{q^j_{|w_j|} q_1^j}. \tag{3.4}$$

It will be helpful to think about the indices appearing in the above expression in the following alternative way. Recall from §2.1.3 that we constructed a collection of marked circles $\bigcup_{j=1}^\ell C(w_j)$ and a map $\gamma : \bigcup_{j=1}^\ell C(w_j) \to \bigvee^r S^1$. As in §2.1.2, mark two points $(x, 0)$ and $(x, 1)$ on the circle in $\bigvee^r S^1$ corresponding to $x \in B$ for each such x. As in §2.1.4, mark points $p_I(0)$ and $p_I(1)$ on $\bigcup_{j=1}^\ell C(w_j)$. Recall the collection of intervals $\mathcal{I}_x = \mathcal{I}_x(w_1, \ldots, w_\ell)$, and that we have identified $\mathcal{I}_x$ with $[2L_x]$ for each $x \in B$. Let $\mathcal{I} = \cup_{x \in B} \mathcal{I}_x$. For $x \in B$ we will let $\mathcal{I}_x^\pm$ be the collection of intervals that correspond to occurrences of $x^\pm$ in $w_1, \ldots, w_\ell$. We denote $\mathcal{I}^\pm = \cup_{x \in B} \mathcal{I}_x^\pm$. $\mathcal{I}, \mathcal{I}_x^\pm, \mathcal{I}^\pm$

For each x and $k \in \{0,1\}$ we accordingly identify the points $\{p_I(k) : I \in \mathcal{I}_x\}$ with the set $[2L_x]$. This allows us to think of the choices of $q_1^j, \ldots, q_{|w_j|}^j$ over the various w_j as an assignment

$$\mathbf{a} : \{p_I(k) \;:\; k = 0, 1\,\} \to [n]$$

with the property that two immediately adjacent marked points p, q in $\bigcup_{j=1}^{\ell} C(w_j)$ that are not of the form $\{p,q\} = \{p_I(0), p_I(1)\}$ (i.e., internal to some interval) must have $\mathbf{a}(p) = \mathbf{a}(q)$. Write $\mathcal{A} =$ $\mathcal{A}$
$\mathcal{A}(w_1, \ldots, w_\ell)$ for the collection of all such assignments $\mathbf{a}$.

To each interval $I \in I_x$ we attach the group element $g(I) = g_i$ where $x = x_i \in B$. Using that $g^{-1} = g^T$, we can rewrite the product over k of the expressions in (3.4) as

$$\prod_{j=1}^{\ell} \operatorname{tr}(w_j(g_1, \ldots, g_r)) = \sum_{\mathbf{a} \in \mathcal{A}} \prod_{x \in B} \prod_{I \in \mathcal{I}_x} g(I)_{\mathbf{a}(p_I(0))\mathbf{a}(p_I(1))}. \tag{3.5}$$

For $\mathbf{a} \in \mathcal{A}$ and $\mathbf{m} = \{(m_{x,0}, m_{x,1})\}_{x \in B} \in \mathrm{MATCH}^{\kappa \equiv 1}$ a collection of matchings, we say $\mathbf{a} \vdash \mathbf{m}$ if $\mathbf{a} \vdash \mathbf{m}$
whenever $m_{x,i}$ matches $p_I(i)$ and $p_J(i)$, these points are assigned the same index by $\mathbf{a}$. In this case we also say that $m_{x,0}$ and $m_{x,1}$ respect $\mathbf{a}$.

If $x = x_i$, a direct consequence of Theorem 3.1 is

$$\int_{g \in \mathrm{O}(n)} \prod_{I \in \mathcal{I}_x} g_{\mathbf{a}(p_I(0))\mathbf{a}(p_I(1))} d\mu_n(g_i) = \sum_{\substack{\text{matchings } m_{x,0} \text{ of } \{p_I(0) : I \in \mathcal{I}_x\} \text{ that respect } \mathbf{a} \\ \text{matchings } m_{x,1} \text{ of } \{p_I(1) : I \in \mathcal{I}_x\} \text{ that respect } \mathbf{a}}} \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0}, m_{x,1}; n).$$

Hence

$$\int_{\mathrm{O}(n)^r} \prod_{x \in B} \prod_{I \in \mathcal{I}_x} g(I)_{\mathbf{a}(p_I(0))\mathbf{a}(p_I(1))} d\mu_n(g_1) \ldots d\mu_n(g_r) = \sum_{\mathbf{m} \in \mathrm{MATCH}^{\kappa \equiv 1} \,:\, \mathbf{a} \vdash \mathbf{m}} \prod_{x \in B} \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0}, m_{x,1}; n). \tag{3.6}$$

Reordering summation and integration, we obtain

$$\begin{aligned} \mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}(n) &= \sum_{\mathbf{a} \in \mathcal{A}} \sum_{\mathbf{m} \in \mathrm{MATCH}^{\kappa \equiv 1} \,:\, \mathbf{a} \vdash \mathbf{m}} \prod_{x \in B} \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0}, m_{x,1}; n). \\ &= \sum_{\mathbf{m} \in \mathrm{MATCH}^{\kappa \equiv 1}} \#\{\mathbf{a} \in \mathcal{A} \,:\, \mathbf{a} \vdash \mathbf{m}\} \prod_{x \in B} \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0}, m_{x,1}; n). \end{aligned} \tag{3.7}$$

For fixed $\mathbf{m}$, the condition $\mathbf{a} \vdash \mathbf{m}$ holds if and only if for every type-o disc of $\Sigma_{\mathbf{m}}$, $\mathbf{a}$ is constant on the set of $p_I(k)$ that meet the boundary of that disc. Hence

$$\#\{\mathbf{a} \in \mathcal{A} \,:\, \mathbf{a} \vdash \mathbf{m}\} = n^{\#\{\text{type-}o \text{ discs of } \Sigma_{\mathbf{m}}\}}.$$

This proves the theorem. □

Theorem 3.4 has the following easy corollary that was stated in the Introduction.

Corollary 3.5. *There is a computable rational function* $\overline{\mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}} \in \mathbf{Q}(n)$ *such that for* $n \geq N$, $\mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}(n)$ *is given by evaluating* $\overline{\mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}}$ *at* n.

Proof. The formula given in Theorem 3.4 expresses $\mathcal{T}r^{\mathrm{O}}_{w_1, \ldots, w_\ell}$ as a finite sum of computable rational functions since $\mathrm{MATCH}^{\kappa \equiv 1}$ is finite. □

3.3 First Laurent series expansion at infinity

Due to Theorem 1.2, we only need to discuss $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ throughout the rest of the paper. The full Laurent series expansion of $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ at $n=\infty$ involves elements of $\mathrm{MATCH}^*(w_1,\ldots,w_\ell)$ with extra restrictions. Recall that an element of $\mathrm{MATCH}^* = \mathrm{MATCH}^*(w_1,\ldots,w_\ell)$ is, for some $\kappa = \{\kappa_x\}_{x\in B} \in \mathbf{Z}^B_{\geq 0}$, a collection $\mathbf{m} = \{(m_{x,0},\ldots,m_{x,\kappa_x})\}_{x\in B}$ of tuples of matchings, where $m_{x,i}$ is a matching of $[2L_x]$. For any κ we write $\overline{\mathrm{MATCH}}^\kappa = \overline{\mathrm{MATCH}}^\kappa(w_1,\ldots,w_\ell)$ for the elements $\mathbf{m}$ of MATCH^κ with the additional constraint that $m_{x,i} \neq m_{x,i+1}$ for each $0 \leq i < \kappa_x$, and similarly define $\overline{\mathrm{MATCH}}^*$. We will also write $|\kappa| = \sum_{x\in B}\kappa_x$. $\overline{\mathrm{MATCH}}^\kappa$ $\overline{\mathrm{MATCH}}^*$ $|\kappa|$

Proposition 3.6. *For $n \geq N$, $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ is given by the following absolutely convergent series:*

$$\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n) = \sum_{\mathbf{m}\in\overline{\mathrm{MATCH}}^*} (-1)^{|\kappa(\mathbf{m})|} n^{\chi(\Sigma_\mathbf{m})}. \tag{3.8}$$

Proof. Putting the power series expansion in n^{-1} for the orthogonal Weingarten function given in Theorem 3.3 into Theorem 3.4 gives

$$\begin{aligned}
\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n) &= \sum_{\mathbf{m}\in\mathrm{MATCH}^{\kappa\equiv 1}} n^{\#\{o\text{-discs of }\Sigma_\mathbf{m}\}} \prod_{x\in B} \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0},m_{x,1};n) \\
&= \sum_{\mathbf{m}=\{(m_{x,0},m_{x,1})\}_{x\in B}} n^{\#\{o\text{-discs of }\Sigma_\mathbf{m}\}} \prod_{x\in B} n^{-L_x} \cdot \\
&\quad \cdot \left\{ \sum_{\kappa_x\geq 0} \sum_{\substack{m_{x,0}=m'_{x,0},\ldots,m'_{x,\kappa_x}=m_{x,1}\in M_{L_x} \\ m'_{x,i}\neq m'_{x,i+1}}} (-1)^{\kappa_x} n^{-\rho\left(m'_{x,0},m'_{x,1}\right)-\ldots-\rho\left(m'_{x,\kappa_x-1},m'_{x,\kappa_x}\right)} \right\} \\
&= n^{-\sum_{x\in B}L_x} \sum_{(\kappa_x)_{x\in B}\in\mathbf{Z}^B_{\geq 0}} (-1)^{\sum_{x\in B}\kappa_x} \cdot \\
&\quad \cdot \left\{ \sum_{\mathbf{m}'\in\overline{\mathrm{MATCH}}^\kappa} n^{\#\{o\text{-discs of }\Sigma_{\mathbf{m}'}\}\}} \prod_{x\in B} n^{-\rho\left(m'_{x,0},m'_{x,1}\right)-\ldots-\rho\left(m'_{x,\kappa_x-1},m'_{x,\kappa_x}\right)} \right\}
\end{aligned}$$

Here we used the fact that the number of type-o discs of $\Sigma_{\mathbf{m}'}$ are the same as those of $\Sigma_\mathbf{m}$, since they only depend on the outer matchings $(m'_{x,0},m'_{x,\kappa_x})_{x\in B}$, that are the same as in $\mathbf{m}$. Also note that each of the expressions for Wg^{O} are absolutely convergent when $n \geq N$, and we only used finitely many such expressions, corresponding to the finitely many choices of $\mathbf{m} \in \mathrm{MATCH}^{\kappa\equiv 1}(w_1,\ldots,w_\ell)$. Using Lemma 2.1, the above can be rewritten as

$$\sum_{\mathbf{m}'\in\overline{\mathrm{MATCH}}^*} (-1)^{|\kappa(\mathbf{m}')|} n^{\chi(\Sigma_{\mathbf{m}'})}.$$

□

3.4 Signed matchings and a new Laurent series expansion

In the section we modify our previous definitions to get a new combinatorial Laurent expansion for $\mathcal{T}r^{O}_{w_1,\ldots,w_\ell}(n)$ in the shifted parameter $(n-1)^{-1}$, or equivalently, an expansion for $\mathcal{T}r^{O}_{w_1,\ldots,w_\ell}(n+1)$ in the parameter n^{-1}. The reason for doing this is that we want to add into our Laurent expansion additional surfaces that are constructed from discs and Möbius bands. This has no analog in [MP19]. The resulting marked surfaces are the ones that are not stabilized by any non-trivial element of the

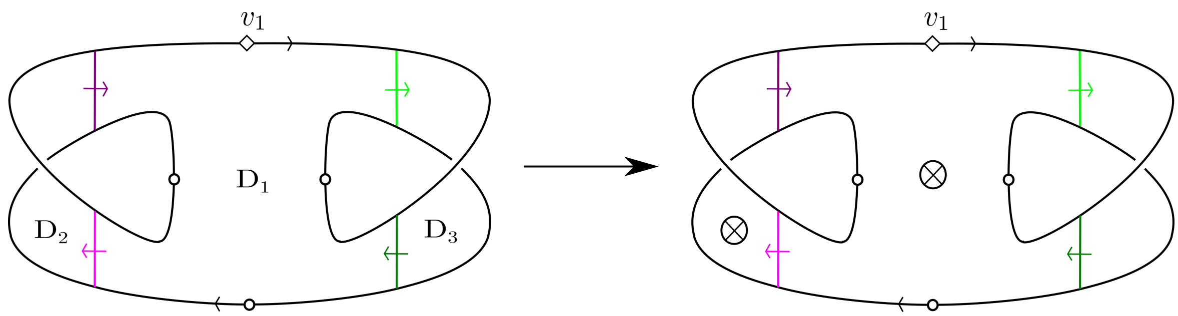

Figure 3.1: In this example $\mathbf{m}$ is the same matching data as from Figure 2.1. On the left is $\Sigma_{\mathbf{m}}$ with 2-cells D_1, D_2, D_3 labeled. Here, ε is given by $\varepsilon(D_1) = \varepsilon(D_2) = -1$ and $\varepsilon(D_3) = 1$. The resulting $\Sigma_{\mathbf{m},\varepsilon}$ is drawn on the right, where we draw a $\otimes$ on the surface to mean an $\mathbf{R}P^2$ has been connected summed there.

mapping class group of the surface; see Lemma 5.9 for the precise statement. The introduction of these extra surfaces is essential in allowing us to give clean expressions for the coefficients as in Theorem 1.5. We formalize this as follows.

Definition 3.7. Let $\mathrm{SMATCH}^* = \mathrm{SMATCH}^*(w_1, \ldots, w_\ell)$ be the collection of pairs $(\mathbf{m}, \varepsilon)$ where SMATCH*

- $\mathbf{m} \in \mathrm{MATCH}^*(w_1, \ldots, w_\ell)$,
- ε is a function from the 2-cells of $\Sigma_{\mathbf{m}}$ to $\{-1, 1\}$,
- if $m_{x,i} = m_{x,i+1}$ then at least one type-(x, i) disc of $\Sigma_{\mathbf{m}}$ must be assigned -1 by ε.

Let $\kappa(\mathbf{m}, \varepsilon) \stackrel{\text{def}}{=} \kappa(\mathbf{m})$. We call a pair $(\mathbf{m}, \varepsilon)$ a *signed matching.*

Definition 3.8. Given $(\mathbf{m}, \varepsilon) \in \mathrm{SMATCH}^*$, we construct a new pair $(\Sigma_{\mathbf{m},\varepsilon}, f_{\mathbf{m},\varepsilon})$ where $\Sigma_{\mathbf{m},\varepsilon}$ is a surface and $f_{\mathbf{m},\varepsilon} : \Sigma_{\mathbf{m},\varepsilon} \to \bigvee^r S^1$ as follows:

- Let $\Sigma_{\mathbf{m},\varepsilon}$ be the surface obtained by connected summing a real projective plane $\mathbf{R}P^2$ onto each 2-cell of $\Sigma_{\mathbf{m}}$ that is assigned -1 by ε.
- On the neighborhood of each disc that was cut out to perform a connected sum, homotope $f_{\mathbf{m}}$ to be a constant other than $\{o\} \cup \{(x, k)\}$, while maintaining the property that the only points in the preimage of $\{(x, k)\}$ are in $\Sigma_{\mathbf{m}}^{(1)}$. This is possible because $f_{\mathbf{m}}$ maps each open 2-cell of $\Sigma_{\mathbf{m}}$ to a contractible piece of $\bigvee^r S^1$. Now extend the function by the constant to the added $\mathbf{R}P^2$. Performing this homotopy then extension for each $\mathbf{R}P^2$ added to $\Sigma_{\mathbf{m}}$ yields $f_{\mathbf{m},\varepsilon}$. See Figure 3.1.

The resulting pair $(\Sigma_{\mathbf{m},\varepsilon}, f_{\mathbf{m},\varepsilon})$ is an admissible map in the sense of Definition 1.3.

Remark 3.9. Note that $\Sigma_{\mathbf{m},\varepsilon}$ is no longer a CW-complex; rather it is a CW-complex with some 2-cells replaced by Möbius bands.

The following Lemma is an obvious consequence of the fact that $\chi(\mathbf{R}P^2) = 1$.

Lemma 3.10. *The Euler characteristic of* $\Sigma_{\mathbf{m},\varepsilon}$ *is*

$$\chi(\Sigma_{\mathbf{m},\varepsilon}) = \chi(\Sigma_{\mathbf{m}}) - |\varepsilon^{-1}(\{-1\})|.$$

Definition 3.11. There is a map $\mathsf{forget} : \mathrm{SMATCH}^* \to \overline{\mathrm{MATCH}}^*$ as follows. Given $(\mathbf{m}, \varepsilon)$, let $\mathsf{forget}(\mathbf{m}, \varepsilon) \in \overline{\mathrm{MATCH}}^*$ be the set of matching data obtained by repeatedly replacing pairs of the form $m_{x,i} = m_{x,i+1}$ by $m_{x,i}$, and re-indexing, until there are no consecutive duplicate matchings. Note that because of this removal of duplicates, forget <u>does not</u> respect the $\mathbf{Z}^B$ gradings of SMATCH^* and $\overline{\mathrm{MATCH}}^*$ by κ.

We now define M

$$M = M(w_1, \dots, w_\ell) = \frac{\max\{L_x \,:\, x \in B\}}{\log 2} = \frac{N}{\log 2} \tag{3.9}$$

where log denotes the natural logarithm. This M is the quantity that appears in Theorem 1.5. Note that $M > N$. The reason for this choice of parameter will be explained in the proof of the next lemma.

Lemma 3.12. *For $n > M$, for any $\mathbf{m} \in \overline{\mathrm{MATCH}}^*$ we have*

$$(-1)^{|\kappa(\mathbf{m})|}(n+1)^{\chi(\Sigma_{\mathbf{m}})} = \sum_{(\mathbf{m}', \varepsilon) : \mathsf{forget}(\mathbf{m}', \varepsilon) = \mathbf{m}} (-1)^{|\kappa(\mathbf{m}', \varepsilon)|} n^{\chi(\Sigma_{\mathbf{m}', \varepsilon})}, \tag{3.10}$$

where the right hand side is absolutely convergent.

Proof. Write $\mathbf{m} = \{(m_{x,0}, \dots, m_{x,\kappa_x})\}_{x \in B}$. To obtain $(\mathbf{m}', \varepsilon)$ as in the right hand side of (3.10) from $\mathbf{m}$ we make the following choices, that we split into two types.

A. For each $x \in B$ and $0 \le k \le \kappa_x$ we choose $d(x,k) \in \mathbf{Z}_{\ge 0}$ and replace $m_{x,k}$ with $d(x,k) + 1$ repeats of $m_{x,k}$. Let $\mathbf{m}' = \{(m'_{x,0}, \dots, m'_{x,\kappa'_x})\}_{x \in B}$ be the resulting new tuples of matchings. Let $J_{x,k}$ be the collection of k' such that $m'_{x,k'} = m'_{x,k'+1}$ and $m_{x,k'}$ was formed by duplicating $m_{x,k}$. Hence $|J_{x,k}| = d(x,k)$. For each $k' \in J_{x,k}$ we furthermore have to choose $\varepsilon_{x,k'}$ on the type-(x,k') discs of $\Sigma_{\mathbf{m}'}$, such that $\varepsilon_{x,k'}$ assigns -1 to at least one of these discs. Note that all of these type-(x,k') discs are rectangles, and there are L_x of them.

B. Independently of the above, we choose some ε_0 on the 2-cells of $\Sigma_{\mathbf{m}}$, since these correspond bijectively to the two cells of $\Sigma_{\mathbf{m}'}$ that were not created by the previous step.

Consider the generating function

$$G(n) = \sum_{(\mathbf{m}', \varepsilon) : \mathsf{forget}(\mathbf{m}', \varepsilon) = \mathbf{m}} (-1)^{|\kappa(\mathbf{m}', \varepsilon)| - |\kappa(\mathbf{m})|} n^{\chi(\Sigma_{\mathbf{m}', \varepsilon}) - \chi(\Sigma_{\mathbf{m}})}. \tag{3.11}$$

Whatever choices we make in the two steps above, they affect both $\chi(\Sigma_{\mathbf{m}', \varepsilon}) - \chi(\Sigma_{\mathbf{m}})$ and $|\kappa(\mathbf{m}', \varepsilon)| - |\kappa(\mathbf{m})|$ independently of one another. Therefore the generating function G splits as a product over $m_{x,k}$ (type **A** above) and the discs of $\Sigma_{\mathbf{m}}$ (type **B** above).

We explain the contribution from the choices of type **A.** Since the effect of the choice made for a given $m_{x,k}$ is to contribute $d(x,k)$ to $|\kappa(\mathbf{m}', \varepsilon)| - |\kappa(\mathbf{m})|$, and for each $k' \in J_{x,k}$ the contribution of $\varepsilon_{x,k'}$ to $\chi(\Sigma_{\mathbf{m}', \varepsilon}) - \chi(\Sigma_{\mathbf{m}})$ is $-|\varepsilon^{-1}_{x,k'}(\{-1\})|$, the multiplicative contribution from a fixed $m_{x,k}$ to $G(n)$ is

$$\begin{aligned}
\sum_{d(x,k) \ge 0} (-1)^{d(x,k)} \prod_{k' \in J_{x,k}} \sum_{\varepsilon_{x,k'} \not\equiv 1} n^{-|\varepsilon^{-1}_{x,k'}(\{-1\})|} &= \sum_{d(x,k) \ge 0} (-1)^{d(x,k)} \prod_{k' \in J_{x,k}} \left((1 + n^{-1})^{L_x} - 1\right) \\
&= \sum_{d(x,k) \ge 0} (-1)^{d(x,k)} \left((1 + n^{-1})^{L_x} - 1\right)^{d(x,k)} \\
&= \frac{1}{1 + (1 + n^{-1})^{L_x} - 1} = \frac{1}{(1 + n^{-1})^{L_x}}.
\end{aligned}$$

All the sums are absolutely convergent when $\left|(1 + n^{-1})^{L_x} - 1\right| < 1$ that holds when $n > M$ (this is the reason for the choice of M). Multiplying all these contributions together over all $m_{x,k}$ the total multiplicative contribution is

$$\prod_{x \in B} \prod_{0 \le k \le \kappa_x} \frac{1}{(1 + n^{-1})^{L_x}} = \frac{1}{(1 + n^{-1})^{\sum_{x \in B} (\kappa_x + 1) L_x}}. \tag{3.12}$$

Now we explain the contribution from choices of type **B.** The choices made contribute 0 to $|\kappa(\mathbf{m}',\varepsilon)|-|\kappa(\mathbf{m})|$ and $-|\varepsilon_0^{-1}(\{-1\})|$ to $\chi(\Sigma_{\mathbf{m}',\varepsilon})-\chi(\Sigma_{\mathbf{m}})$. Hence the multiplicative contribution of these choices to $G(n)$ is more simply

$$\sum_{\varepsilon_0} n^{-|\varepsilon_0^{-1}(\{1\})|}=(1+n^{-1})^{\#\{\text{discs of }\Sigma_{\mathbf{m}}\}}. \tag{3.13}$$

Multiplying (3.12) and (3.13) together and using (2.6) we obtain

$$\begin{aligned} G(n) &= (1+n^{-1})^{\#\{\text{discs of }\Sigma_{\mathbf{m}}\}-\sum_{x\in B}(\kappa_x+1)L_x} \\ &= (1+n^{-1})^{\chi(\Sigma_{\mathbf{m}})}=\left(\frac{n+1}{n}\right)^{\chi(\Sigma_{\mathbf{m}})}. \end{aligned} \tag{3.14}$$

Equating (3.11) and (3.14) and rearranging gives the result. □

Proposition 3.13. *For $n>M$, $\mathcal{T}r^{\mathrm{O}}_{w_1,\dots,w_\ell}(n+1)$ is given by the following absolutely convergent series:*

$$\mathcal{T}r^{\mathrm{O}}_{w_1,\dots,w_\ell}(n+1)=\sum_{(\mathbf{m},\varepsilon)\in\mathrm{SMATCH}^*}(-1)^{|\kappa(\mathbf{m},\varepsilon)|}n^{\chi(\Sigma_{\mathbf{m},\varepsilon})}. \tag{3.15}$$

Proof. Assume $n>M$. Since $M>N$, we have by Proposition 3.6 and Lemma 3.12

$$\begin{aligned} \mathcal{T}r^{\mathrm{O}}_{w_1,\dots,w_\ell}(n+1) &= \sum_{\mathbf{m}\in\overline{\mathrm{MATCH}}^*}(-1)^{|\kappa(\mathbf{m})|}(n+1)^{\chi(\Sigma_{\mathbf{m}})} \\ &= \sum_{\mathbf{m}\in\overline{\mathrm{MATCH}}^*}\ \sum_{(\mathbf{m}',\varepsilon):\mathsf{forget}(\mathbf{m}',\varepsilon)=\mathbf{m}}(-1)^{|\kappa(\mathbf{m}',\varepsilon)|}n^{\chi(\Sigma_{\mathbf{m}',\varepsilon})}, \end{aligned}$$

where the right hand side is absolutely convergent. This clearly gives (3.15). □

Corollary 3.14. *For fixed $w_1,\dots,w_\ell$ and fixed $\chi_0\in\mathbf{Z}$, there are only finitely many elements $(\mathbf{m},\varepsilon)$ of $\mathrm{SMATCH}^*(w_1,\dots,w_\ell)$ with $\chi(\Sigma_{\mathbf{m},\varepsilon})=\chi_0$.*

Proof. This could be proven by a direct combinatorial argument similarly to [MP19, Claim 2.10]. It is also a direct consequence of the sum (3.15) in Proposition 3.13 being absolutely convergent for $n>M$. Indeed, if there were infinitely many $\mathbf{m}\in\mathrm{SMATCH}^*$ with $\chi(\Sigma_{\mathbf{m}})=\chi_0$ then there would be infinitely many summands in (3.15) with absolute value n^{χ_0}. □

Since every $(\mathbf{m},\varepsilon)\in\mathrm{SMATCH}^*$ gives rise to an admissible map $(\Sigma_{\mathbf{m},\varepsilon},f_{\mathbf{m},\varepsilon})$, it makes sense to partition elements of SMATCH^* according to the equivalence class of $(\Sigma_{\mathbf{m},\varepsilon},f_{\mathbf{m},\varepsilon})$ in $\mathsf{Surfaces}^*(w_1,\dots,w_\ell)$. Thus, given $[(\Sigma,f)]\in\mathsf{Surfaces}^*(w_1,\dots,w_\ell)$ we define $\mathrm{SMATCH}^*(\Sigma,f)$

$$\begin{aligned} \mathrm{SMATCH}^*(\Sigma,f) &\stackrel{\text{def}}{=} \mathrm{SMATCH}^*(w_1,\dots,w_\ell;\Sigma,f)\stackrel{\text{def}}{=} \\ &\stackrel{\text{def}}{=} \{(\mathbf{m},\varepsilon)\in\mathrm{SMATCH}^*(w_1,\dots,w_\ell)\ :\ (\Sigma_{\mathbf{m},\varepsilon},f_{\mathbf{m},\varepsilon})\approx(\Sigma,f)\}. \end{aligned}$$

Then we can rewrite Proposition 3.13 as,

Corollary 3.15. *For $n>M$,*

$$\mathcal{T}r^{\mathrm{O}}_{w_1,\dots,w_\ell}(n+1)=\sum_{[(\Sigma,f)]\in\mathsf{Surfaces}^*(w_1,\dots,w_\ell)}n^{\chi(\Sigma)}\sum_{(\mathbf{m},\varepsilon)\in\mathrm{SMATCH}^*(\Sigma,f)}(-1)^{|\kappa(\mathbf{m},\varepsilon)|}.$$

Corollary 3.15 reduces the proof of our main theorem (Theorem 1.5), when all unsigned exponents of $x\in B$ in $w_1,\dots,w_\ell$ are even, to the following.

Theorem 3.16. *For $[(\Sigma,f)]\in\mathsf{Surfaces}^*(w_1,\dots,w_\ell)$, the L^2-Euler characteristic $\chi^{(2)}(\mathrm{MCG}(f))$ is well-defined, and given by*

$$\chi^{(2)}(\mathrm{MCG}(f))=\sum_{(\mathbf{m},\varepsilon)\in\mathrm{SMATCH}^*(\Sigma,f)}(-1)^{|\kappa(\mathbf{m},\varepsilon)|}. \tag{3.16}$$

The proof of Theorem 3.16 will be the subject of §4 and §5.

3.5 Proof of Corollary 1.11

Proof of Corollary 1.11. By Lemma 2.5, if $1-\min(\mathrm{sql}(w), 2\mathrm{cl}(w)) = -\infty$, then $\mathsf{Surfaces}^*(w)$ is empty, and Lemmas 2.3 and 2.6 tell us $\mathcal{T}r_w^{\mathrm{O}}(n) = 0$ for all n, which gives the result.

If $1-\min(\mathrm{sql}(w), 2\mathrm{cl}(w))$ is finite, then by Proposition 3.6, (3.8) is a convergent Laurent series in n^{-1} with positive radius of convergence, and the order of the zero at ∞ is at least

$$-\max_{\mathbf{m}\in\mathrm{MATCH}^*(w)} \chi(\Sigma_{\mathbf{m}}).$$

Note that we only have a bound here, since there could be cancellations between the coefficients of n^χ for any given χ. On the other hand, every $\mathbf{m} \in \mathrm{MATCH}^*$ gives an admissible map $(\Sigma_{\mathbf{m}}, f_{\mathbf{m}})$ so $\max_{\mathbf{m}\in\mathrm{MATCH}^*(w)} \chi(\Sigma_{\mathbf{m}}) \leq \chi_{\max}(w)$. This implies

$$\mathcal{T}r_w^{\mathrm{O}}(n) = O\left(n^{\chi_{\max}(w)}\right)$$

as $n \to \infty$. Finally, Lemma 2.5 tells us we can replace this by $O(n^{1-\min(\mathrm{sql}(w),2\mathrm{cl}(w))})$ as stated in Corollary 1.11. ☐

Remark 3.17. In fact, every admissible map $(\Sigma, f) \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ of maximal Euler characteristic (so $\chi(\Sigma) = \chi_{\max}(w_1, \ldots, w_\ell)$) can be constructed from a suitable $\mathbf{m} \in \mathrm{MATCH}^*(w_1, \ldots, w_\ell)$, so we have

$$\max_{\mathbf{m}\in\mathrm{MATCH}^*(w_1,\ldots,w_\ell)} \chi(\Sigma_{\mathbf{m}}) = \chi_{\max}(w_1, \ldots, w_\ell).$$

Indeed, this is a simple generalization of [Cul81, Thm. 1.5]. The leading exponent of $\mathcal{T}r^{O}_{w_1,\ldots,w_\ell}(n)$ is strictly smaller than $\chi_{\max}(w_1, \ldots, w_\ell)$ only if the coefficients $\chi^{(2)}(\mathrm{MCG}(f))$ of the maps $[\Sigma, f] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ with $\chi(\Sigma) = \chi_{\max}(w_1, \ldots, w_\ell)$ sum up to zero.

4 The transverse map complex

In this section we follow [MP19, §3] closely. Our goal here is to extend the results of (*loc. cit.*) to surfaces that might be non-orientable. Since our aim is to prove Theorem 3.16, we now fix an admissible map (Σ, f) for $w_1, \ldots, w_\ell$. Recall the notation from Definition 1.3. It will be useful to mark an additional set of points V_o on the boundary of Σ with the following properties:

- V_o contains the original marked points $\{v_j\} \subset \delta\Sigma$.
- $V_o \subset f^{-1}(\{o\})$, and V_o is finite.
- $|\delta_j \cap V_o| = |w_j|$ and so $\delta_j - V_o$ consists of $|w_j|$ intervals. Ordering these intervals according to the orientation of δ_j, beginning at v_j, the uth interval, directed according to δ_j, maps under f to a loop in $\bigvee^r S^1$ that corresponds to $x_{j_u^k}^{\varepsilon_u^k} \in \pi_1(\bigvee^r S^1, o) \cong \mathbf{F}_r$.

We fix this choice of V_o henceforth.

4.1 Transverse maps on possibly non-orientable surfaces

We use the terminology *arc* to refer to an embedding of a closed interval in a compact surface such that the endpoints of the arc are in the boundary of the surface and these are the only points of the arc in the boundary. We use the terminology *curve* to refer to an embedding of a circle in a surface, disjoint from the boundary of the surface. Note that our notion of curve is what is usually referred to as a simple closed curve.

Definition 4.1. Let Σ be a compact surface. A continuous function $f : \Sigma \to \bigvee^r S^1$ is said to be *transverse* to a point $p \in \bigvee^r S^1 - \{o\}$ if $f^{-1}(\{p\})$ is a disjoint union of arcs and curves, and every arc or curve in the preimage has a tubular neighborhood that is cut into two halves by the arc or curve, and the two halves map under f to two different (local) sides of p in $\bigvee^r S^1$.

Note that this definition prevents one-sided curves in $f^{-1}(\{p\})$ (see Footnote 6 for the definition of a one-sided curve).

Definition 4.2. A *transverse map* on Σ is a tuple $\kappa = \{\kappa_x\}_{x \in B}$, a choice for each $x \in B$ of $\kappa_x + 1$ distinct *transversion points* $(x, 0), \dots, (x, \kappa_x)$ in S^1_x, ordered according to the orientation of S^1_x, and a continuous function $g : \Sigma \to \bigvee^r S^1$ that is transverse to all the points $\{(x, j) : x \in B, 0 \leq j \leq \kappa_x\}$, such that $g^{-1}\{o\} \cap \delta\Sigma = V_o$.

We say that a transverse map g *realizes* (Σ, f) if g is homotopic to f relative to V_o.

Let $J_{x,j}$ be the connected component of $\bigvee^r S^1 - \{(x, j) : x \in B, 0 \leq j \leq \kappa_x\}$ that is bounded by the points (x, j) and $(x, j+1)$. Let U_o be the connected component that contains o. We call a connected component of $g^{-1}(U_o)$ an o-zone of g and a connected component of $g^{-1}(J_{x,j})$ an (x, j)-zone of g, or if we do not care about x and j, simply an x-zone. We say that a transverse map is *filling* if all its zones are topological discs. We say the map is *almost-filling* if all its zones are discs or Möbius bands.

Two transverse maps g_1 and g_2 on Σ are said to be *isotopic* if they are homotopic through transverse maps with the same parameters κ. In this homotopy, the points (x, j) are allowed to vary continuously in $S^1_x - \{o\}$.

We refer to transverse maps realizing (Σ, f) simply as *transverse maps.* As in [MP19, §3], we think of isotopy classes of transverse maps as isotopy classes of colored arcs and curves with assigned normal direction: the (x, j)-colored arcs and curves are the components of $g^{-1}\{(x, j)\}$ and the normal direction to the curve is given by the order in which the two local sides of the arc and curve map to two local sides of (x, j) in S^1_x, with the order coming from the fixed orientation of S^1_x.

The following definition is the same as in [MP19, §3].

Definition 4.3 (Loose and strict transverse maps)**.** We say a transverse map g is *loose* if it satisfies

Restriction 1 There are no o-zones or z-zones containing no element of V_o with the property that all the bounding arcs and curves of the zone are pointing inwards, or all pointing outwards, and all the bounding arcs and curves have the same color. Note this rules out the possibility that there is a zone that is bounded by one curve, e.g. a disc or a Möbius band.

Restriction 2 Any segment of the boundary of Σ that is bounded by two same colored endpoints of arcs, that are both directed inwards or both outwards, must contain an element of V_o and hence be part of an o-zone.

The transverse map g is called *strict* if it also satisfies

Restriction 3 For every $x \in B$ and $0 \leq j < \kappa_x$ there must be an (x, j)-zone that is neither a rectangle (bounded by two arcs and two boundary segments) nor an annulus bounded by two curves.

Remark 4.4. Note that **Restriction 2** implies that if g is a transverse map with parameters $\{\kappa_x\}_{x \in B}$, any connected component sub-interval of $\delta\Sigma - V_o$ contains for some $x \in B$ exactly $\kappa_x + 1$ points that for some order of the sub-interval map to $(x, 0), \dots, (x, \kappa_x)$ respectively.

Example 4.5. Consider the pairs $(\Sigma_{\mathbf{m}}, f_{\mathbf{m}})$ constructed in §2.1.4. Each of these are admissible maps. Each $f_{\mathbf{m}}$ is a loose transverse map on $\Sigma_{\mathbf{m}}$, and it is, furthermore, strict, if and only if $\mathbf{m} \in \overline{\mathrm{MATCH}}^*$. In this case, V_o are the endpoints of the intervals used in §2.1.3. The zones of $f_{\mathbf{m}}$ are the 2-cells of $\Sigma_{\mathbf{m}}$, hence $f_{\mathbf{m}}$ is filling.

Example 4.6. Consider now the pairs $(\Sigma_{\mathbf{m},\varepsilon}, f_{\mathbf{m},\varepsilon})$ constructed in Definition 3.8. Each of these are admissible maps, and $f_{\mathbf{m},\varepsilon}$ is a strict transverse map on $\Sigma_{\mathbf{m},\varepsilon}$. Now, the zones of $f_{\mathbf{m},\varepsilon}$ may be either discs or Möbius bands, depending on ε. In this case, $f_{\mathbf{m},\varepsilon}$ is almost-filling.

4.2 Polysimplicial complexes of transverse maps

Definition 4.7. The *poset of transverse maps realizing* (Σ, f), denoted $(\mathcal{T}, \preceq)$, has underlying set $\mathcal{T} = \mathcal{T}(\Sigma, f)$ of isotopy classes $[g]$ of *strict* transverse maps g realizing (Σ, f). The partial order $\preceq$ is defined by $[g_2] \preceq [g_1]$ if g_2 is obtained from g_1 by forgetting transversion points. (After we forget transversion points we re-index the remaining $(x, i_0), \ldots, (x, i_r) \mapsto (x, 0), \ldots, (x, r)$.)

As in [MP19, §3] we have the following lemmas.

Lemma 4.8. *If g is a strict transverse map realizing (Σ, f) and g' is a transverse map obtained from g by forgetting points of transversion then g' is also a strict transverse map realizing (Σ, f).*

Proof. Same as [MP19, Lem. 3.7]. □

Lemma 4.9. *$\mathcal{T} = \mathcal{T}(\Sigma, f)$ is not empty.*

Proof. Same as [MP19, Lem. 3.8]. □

A polysimplex is a subset of $\mathbf{R}^k$ of the form $\Delta_{k_1} \times \Delta_{k_2} \times \ldots \times \Delta_{k_r}$ where $\sum_{j=1}^{r} k_j = k$ and the Δ_{k_j} are standard simplices in $\mathbf{R}^{k_j}$. The polysimplex $\Delta_{k_1} \times \Delta_{k_2} \times \ldots \times \Delta_{k_r}$ has dimension k. A *polysimplicial complex* is the natural generalization of a simplicial complex that allows cells to be polysimplices.

Definition 4.10 (Complex of transverse maps)**.** The *complex of transverse maps realizing* (Σ, f) is the polysimplicial complex with a polysimplex $\mathrm{poly}([g]) \cong \prod_{x \in B} \Delta_{\kappa_x}$ for each element $[g]$ of $\mathcal{T}(\Sigma, f)$ with associated parameters $\{\kappa_x\}_{x \in B}$. The faces of $\mathrm{poly}([g])$ are $\mathrm{poly}([g'])$ where $[g'] \preceq [g]$. The resulting polysimplicial complex is denoted $|\mathcal{T}|_{\mathrm{poly}} = |\mathcal{T}(\Sigma, f)|_{\mathrm{poly}}$. It can be naturally identified with a closed subset of Euclidean space and is given the subspace topology.

Remark 4.11. Lemma 4.8 implies that the face relations of $|\mathcal{T}|_{\mathrm{poly}}$ make sense: the property of being a strict transverse map is preserved under passing to sub-faces, and it is obvious that if g_1 is a transverse map realizing (Σ, f) and g_2 is obtained from g_1 by forgetting transversion points, then g_1 and g_2 have the same underlying map and hence g_2 realizes (Σ, f).

Also note that **Restriction 3** implies that any minimal element $[g]$ of $\mathcal{T}$ corresponds to exactly one vertex of any given polysimplex of $|\mathcal{T}|_{\mathrm{poly}}$ containing $[g]$, so $|\mathcal{T}|_{\mathrm{poly}}$ is really a polysimplicial complex.

The poset $(\mathcal{T}, \preceq)$ also gives rise to a simplicial complex called the *order complex* and denoted by $|\mathcal{T}|$. The $k-$simplices of $|\mathcal{T}|$ are chains

$$[g_0] \precneqq [g_1] \precneqq \cdots \precneqq [g_k]$$

in $(\mathcal{T}, \preceq)$, and passing to sub-faces corresponds to deleting elements from chains.

Fact 4.12. *[MP19, Claim 3.10] $|\mathcal{T}|$ is the barycentric subdivision of $|\mathcal{T}|_{\mathrm{poly}}$. In particular, $|\mathcal{T}|$ and $|\mathcal{T}|_{\mathrm{poly}}$ are homeomorphic.*

The proof of this fact has nothing to do with the issue of whether Σ is orientable, so carries over to the current situation.

Lemma 4.13. *The complex $|\mathcal{T}|_{\mathrm{poly}}$ is finite dimensional with* $\dim(|\mathcal{T}|_{\mathrm{poly}}) \leq \frac{\ell}{2} - \chi(\Sigma)$.

Proof. The proof is along the same lines as the proof of [MP19, Lem. 3.12]. The point of the proof is that given a transverse map g

$$\chi(\Sigma) = \sum_{\Sigma'} \left(\chi(\Sigma') - \frac{1}{2} \#\{\text{arcs of } g \text{ in the boundary of } \Sigma'\} \right) \tag{4.1}$$

where the sum is over zones of g and an arc is counted twice for Σ' if it meets Σ' on both sides.

First we note that because of **Restriction 1** $\chi(\Sigma') - \frac{1}{2}\#\{\text{arcs of } g \text{ in the boundary of } \Sigma'\}$ is positive only when Σ' is a disc that meets exactly one arc, and on one side. This is still true after dropping the assumption that Σ is orientable, using the classification of surfaces. Each such zone must be an o-zone containing a point v_j, and this can only happen if w_j is not cyclically reduced. Thus each of these zones contributes $1/2$ to (4.1) hence the contribution of such zones to $\chi(\Sigma)$ is at most $\ell/2$.

As in [MP19], the zones Σ' that contribute 0 to $\chi(\Sigma)$ include annuli bounded by two curves and rectangles. As Σ is not necessarily orientable, there is now the extra possibility of a Möbius band bounded by a curve. However, this is forbidden by **Restriction 1.**

Every x-zone Σ' not considered thus far contributes at most -1 to $\chi(\Sigma)$. Indeed, $\chi(\Sigma') - \frac{1}{2}\#\{\text{arcs of } g \text{ in the boundary of } \Sigma'\}$ is an integer, since every x-zone meets an even number of arcs, and we have classified the zones that contribute ≥ 0. Moreover, for each $x \in B$ and $0 \leq k < \kappa_x$, there is an (x,k)-zone contributing at most -1 to (4.1) by **Restriction 3**. Hence $\dim([g]) = \sum_{x\in B} \kappa_x \leq \ell/2 - \chi(\Sigma)$. □

The main goal of this §4 is to record the following theorem.

Theorem 4.14. *The polysimplicial complex* $|\mathcal{T}|_{\text{poly}}$ *is contractible.*

The motivation for this theorem is that $|\mathcal{T}|_{\text{poly}}$ carries an action of $\text{MCG}(f)$ that will allow us to calculate $\chi^{(2)}(\text{MCG}(f))$ in terms of the orbits of $\text{MCG}(f)$ on $\mathcal{T}$. On the other hand, these orbits can be related to the terms in (3.16) (see Lemma 5.11).

The proof of Theorem 4.14 is the same as the proof of [MP19, Thm. 3.14]. However, there is one minor point that needs adjusting. Here we refer to terminology of [MP19] to explain the adjustment for the sake of completeness. In the classification of maximal null-arc systems on pages 384-385 of *(ibid.)*, it is argued that any component of the complement of a maximal system Ω of null-arcs that has one boundary component consisting of a closed null-arc, contains a pair of pants disjoint from the curves of g, where g is an auxiliary transverse map with $\kappa_x = 0$ for all x and such that the arcs and curves of g are disjoint from Ω. This should be replaced by the following analysis. Let Σ' be a component of the complement of Ω that is bounded by a single closed null-arc. If Σ' is orientable then it contains a pair of pants disjoint from the curves of g, and this contradicts the maximality of Ω as in [MP19, pg. 385]. If Σ' is not orientable, then Σ' contains a simple closed curve γ that bounds a Möbius band, both of which are disjoint from the arcs and curves of g. On the Möbius band consider the waist curve. We can add a new null-arc that is disjoint from the old ones and essentially crosses the waist curve of the Möbius band and is hence not homotopic to any null-arc in Ω. This contradicts the maximality of Ω.

5 The action of $\text{MCG}(f)$ on the transverse map complex.

5.1 L^2-invariants

For a discrete group G, the L^2-Euler characteristic $\chi^{(2)}(G)$ is defined as follows. First of all, we make the following definition.

Definition 5.1. We say that X is a G-CW-complex if X is a CW-complex, with a cellular action of G, such that if $g \in G$ preserves an open cell of X, then g acts as the identity on that cell.

For a discrete group G, the *group von Neumann algebra* $\mathcal{N}(G)$ is the algebra of G-equivariant bounded operators on $\ell^2(G)$. Let X be a G-CW-complex, and let $C_*(X)$ be the singular chain complex of X. Since $C_*(X)$ is a complex of left $\mathbf{Z}[G]$-modules, we can form the chain complex

$$\ldots \to \mathcal{N}(G) \otimes_{\mathbf{Z}[G]} C_{p+1}(X) \xrightarrow{d_{p+1}} \mathcal{N}(G) \otimes_{\mathbf{Z}[G]} C_p(X) \xrightarrow{d_p} \mathcal{N}(G) \otimes_{\mathbf{Z}[G]} C_{p-1}(X) \to \ldots.$$

This is a complex of left Hilbert $\mathcal{N}(G)$-modules, following [Lüc02, Def. 1.15], and the boundary maps are bounded G-equivariant operators between Hilbert spaces. We define

$$H_p^{(2)}(X;G) \stackrel{\text{def}}{=} \frac{\ker(d_p)}{\text{closure}(\text{image}(d_{p+1}))}.$$

Each of these homology groups is also a Hilbert $\mathcal{N}(G)$-module and hence has an associated *von Neumann dimension [Lüc02, Def. 6.20]*

$$b_p^{(2)}(X;G) \stackrel{\text{def}}{=} \dim_{\mathcal{N}(G)} H_p^{(2)}(X;G) \in [0,\infty].$$

Following [Lüc02, Def. 6.79], let

$$\chi^{(2)}(X;G) \stackrel{\text{def}}{=} \sum_{p\in\mathbf{Z}_{\geq 0}} (-1)^p b_p^{(2)}(X,G)$$

if the sum is absolutely convergent. Note this assumes at the very least that all the $b_p^{(2)}(X;G)$ are finite. If EG is a contractible G-CW-complex with a free action of G, and the sum defining $\chi^{(2)}(EG;G)$ is absolutely convergent, then we define

$$b_p^{(2)}(G) \stackrel{\text{def}}{=} b_p^{(2)}(EG;G), \quad \chi^{(2)}(G) \stackrel{\text{def}}{=} \chi^{(2)}(EG;G).$$

These quantities do not depend on EG, so give invariants of G (when they are defined). The reason for this is that $b_p^{(2)}(EG;G)$ is invariant under G-equivariant homotopy equivalence of EG [Lüc02, Thm. 6.54], and EG always exists and is unique up to such homotopy equivalences [tD72, tD87].

In this paper we will calculate $\chi^{(2)}(\text{MCG}(f))$, for $[(\Sigma,f)] \in \mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$, by other means, which make use of the following definition.

Definition 5.2. Let $\mathcal{B}_\infty$ be the class of discrete groups such that all $b_p^{(2)}(G)$ are defined and equal to 0 for all $p \in \mathbf{Z}_{\geq 0}$.

If G is a discrete group and X a G-CW-complex, and σ is a cell of X, then we define the *isotropy group* G_σ to be the stabilizer of σ in G. We use the convention that $\frac{1}{|G_\sigma|} = 0$ if G_σ is infinite. We will use the following theorem to calculate $\chi^{(2)}(\text{MCG}(f))$.

Theorem 5.3. *Let G be a discrete group, and X be a G-CW-complex with the following properties*

- *X is acyclic.*
- *All the isotropy groups of G_σ are either infinite and in the class $\mathcal{B}_\infty$, or finite.*
- *We have*

$$\sum_{[\sigma]\in G\backslash X} \frac{1}{|G_\sigma|} < \infty. \tag{5.1}$$

Then $\chi^{(2)}(G)$ is well-defined and given by

$$\chi^{(2)}(G) = \sum_{[\sigma]\in G\backslash X} (-1)^{\dim(\sigma)} \frac{1}{|G_\sigma|}.$$

Theorem 5.3 is a synthesis of results in Lück [Lüc02, Thm. 6.80(1), Ex. 6.20].

As in [MP19, §4], we use the following theorem, essentially due to Cheeger and Gromov (cf. [CG86, Cor. 0.6]), as a source of groups lying in $\mathcal{B}_\infty$. The precise statement we need can be deduced from [Lüc02, Thm. 7.2, items (1) and (2)]. Recall that a discrete group is called *amenable* if it has a finitely additive left invariant probability measure.

Theorem 5.4 (Cheeger-Gromov)**.** *If G is a discrete group containing a normal infinite amenable subgroup then $G \in \mathcal{B}_\infty$.*

5.2 Proof of Theorem 3.16

When we apply Theorem 5.3 to prove Theorem 3.16, we will take $X = |\mathcal{T}|_{\mathrm{poly}}$. Recall that $\mathcal{T} = \mathcal{T}(\Sigma, f)$ for $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$. Let $\Gamma = \mathrm{MCG}(f)$. We now prepare the necessary inputs for Theorem 5.3 in the following lemmas.

Lemma 5.5. *The action of* Γ *on* $\mathcal{T}$ *makes* $|\mathcal{T}|_{\mathrm{poly}}$ *into a* Γ*-CW-complex.*

Proof. The same as the proof of [MP19, Lem. 4.5]. □

Analogously to [MP19, §4], we let $(\mathcal{T}_\infty, \preceq)$ denote the subposet of $\mathcal{T}$ consisting of isotopy classes of transverse maps that are not almost-filling. Recall a transverse map is almost-filling if its zones are discs or Möbius bands. This definition marks an essential departure from [MP19], where the presence of Möbius bands is not possible due to the surfaces under consideration being orientable. In fact, this difference is what is responsible for the shift by the Jack parameter in our main theorem (Theorem 1.5). The reason discs and Möbius bands are singled out here is because these are precisely the type of zones that have trivial mapping class group: $\mathcal{T}_\infty$

Lemma 5.6. *The mapping class group*[7] *of a disc or a Möbius band is trivial.*

Proof. The statement for a disc is the Alexander Lemma [FM12, Lem. 2.1]. The statement for a Möbius band can be found in Epstein [Eps66, Thm. 3.4]. □

Definition 5.7. We say that a two-sided simple closed curve in Σ is *generic* if it is not homotopic to a boundary component, and does not bound either a disc or a Möbius band.

We also need the following proposition that appears in [Stu06, Prop. 4.4].

Proposition 5.8. *If* Σ *is any surface with boundary, and* $c_1, \ldots, c_r$ *are a collection of disjoint, pairwise non-isotopic, generic two-sided simple closed curves in* Σ*, then the Dehn twists in* $c_1, \ldots, c_r$ *generate a subgroup of* $\mathrm{MCG}(\Sigma)$ *that is isomorphic to* $\mathbf{Z}^r$*.*

Lemma 5.9. *The isotropy groups* $\Gamma_{[g]}$ *of the action of* Γ *on* $\mathcal{T}$ *can be classified as follows*

- $\Gamma_{[g]} = \{\mathrm{id}\}$ *if* $[g] \in \mathcal{T} - \mathcal{T}_\infty$,
- $\Gamma_{[g]}$ *is infinite and in the class* $\mathcal{B}_\infty$ *if* $[g] \in \mathcal{T}_\infty$.

Proof. The proof of the first statement (when $[g]$ is almost-filling) is similar to the proof of [MP19, Lem. 4.7], but incorporating Lemma 5.6 instead of simply the Alexander Lemma.

The proof of the statement given when $[g] \in \mathcal{T}_\infty$ is similar to the proof of [MP19, Lem. 4.8]. Given $[g] \in \mathcal{T}_\infty$, we create a list of simple closed curves $c_1, \ldots, c_k$ as follows. For every boundary component of any zone of g, add to the list the simple closed curve that follows close to the boundary component inside the zone. After doing so, remove repeats of isotopic curves (e.g. if a zone of g is bounded by a simple closed curve, then in the previous step isotopic curves were created on both sides). Also remove any curves that are not generic.

Note that by construction the c_i are pairwise non-isotopic, disjoint, two-sided, and generic. We should check that the collection of c_i is not empty. Indeed, since $[g] \in \mathcal{T}_\infty$, some zone Z of g is not a Möbius band or a disc. Hence Z must have a boundary component that gave rise to a c_i that is generic.

Any Dehn twist D_{c_i} in one of the c_i is in $\Gamma_{[g]}$, since c_i is disjoint from the arcs and curves of g, and $[g]$ is determined by these. Since mapping classes in $\Gamma_{[g]}$ have representatives that respect the zones of g, elements of $\Gamma_{[g]}$ permute the isotopy classes of the c_i. Hence the group generated by the D_{c_i} (we choose one Dehn twist for each c_i) is a normal subgroup of $\Gamma_{[g]}$. By Proposition 5.8, this subgroup is isomorphic to $\mathbf{Z}^d$ with $d \geq 1$. Since $\mathbf{Z}^d$ is amenable by a result of von Neumann [von29], we deduce from Theorem 5.4 that $\Gamma_{[g]}$ is in $\mathcal{B}_\infty$. □

[7] Mapping classes fix the boundary pointwise.

Recall the sets of signed matchings $\mathrm{SMATCH}^* = \mathrm{SMATCH}^*(w_1, \ldots, w_\ell)$ from §3.4. Let $[g] \in \mathcal{T} - \mathcal{T}_\infty$. We will now describe how $[g]$ naturally defines an element $(\mathbf{m}([g]), \varepsilon([g]))$ of SMATCH^*.

The points V_o cut $\delta\Sigma$ into intervals, which by design, are naturally identified with the sub-intervals of $\cup_{j=1}^{\ell} C(w_j)$ that were used in their construction. Let $\kappa = \{\kappa_x\}_{x \in B}$ be the parameters of $[g]$. Consider, for each $x \in B$ and $0 \le k \le \kappa_x$, the collection $A(x,k)$ of arcs of g that are in the preimage of (x,k). These arcs naturally give a matching $m_{x,k}(g)$ of $\mathcal{I}_x(w_1, \ldots, w_\ell)$. This matching does not change under isotopy of g. Hence reading off all the matchings as x and k vary, we obtain a tuple of matchings

$$\mathbf{m}([g]) = \{(m_{x,0}(g), \ldots, m_{x,\kappa_x}(g))\}_{x \in B} \in \mathrm{MATCH}^*$$

with $\kappa(\mathbf{m}([g])) = \kappa$. Note that **Restriction 3,** together with $[g] \in \mathcal{T} - \mathcal{T}_\infty$ implies that if $m_{x,i} = m_{x,i+1}$ then at least one of the (x,i)-zones of $[g]$ is a Möbius band.

By **Restriction 1**, together with $[g] \in \mathcal{T} - \mathcal{T}_\infty$, we have that $g^{-1}(\{(x,i)\})$ contains no curves but only arcs. By construction, the matching arcs of $\Sigma_{\mathbf{m}}$ corresponding to $m_{x,i}$ are in one-to-one correspondence with the connected components of $g^{-1}(\{(x,i)\})$, and any zone of g corresponds to a 2-cell of $\Sigma_{\mathbf{m}}$ by matching up the matching arcs on the boundary of the 2-cell. However, a zone of g that is a Möbius band may correspond to a 2-cell of $\Sigma_{\mathbf{m}}$ that is a disc. To record this discrepancy, we define $\varepsilon = \varepsilon([g])$ to assign 1 to each 2-cell of $\Sigma_{\mathbf{m}}$ that corresponds to a zone of g that is a disc, and define ε to assign -1 to any 2-cell of $\Sigma_{\mathbf{m}}$ that corresponds to a zone of g that is a Möbius band.

Thus we have defined a map

$$(\mathbf{m}, \varepsilon) : \mathcal{T} - \mathcal{T}_\infty \to \mathrm{SMATCH}^*.$$

Lemma 5.10. *Let* $[(\Sigma_i, f_i)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ *and* $[g_i] \in \mathcal{T}(\Sigma_i, f_i)$ *for* $i = 1, 2$. *Then* $(\mathbf{m}([g_1]), \varepsilon([g_1])) = (\mathbf{m}([g_2]), \varepsilon([g_2]))$ *if and only if there is a homeomorphism*

$$\phi : \Sigma_1 \to \Sigma_2$$

that respects all markings of the boundaries of the two surfaces and such that

$$[g_2 \circ \phi] = [g_1]$$

(as isotopy classes of transverse maps).

Proof. First suppose that $(\mathbf{m}([g_1]), \varepsilon([g_1])) = (\mathbf{m}([g_2]), \varepsilon([g_2]))$. By design of the map $(\mathbf{m}, \varepsilon)$, there are homeomorphisms

$$\phi_i : \Sigma_{\mathbf{m}([g_i]), \varepsilon([g_i])} \to \Sigma_i$$

such that for $i = 1, 2$, $f_{\mathbf{m}([g_i]), \varepsilon([g_i])} \circ \phi_i^{-1}$ is a transverse map on Σ_i that is isotopic to g_i, and the ϕ_i respect the boundary markings of the surfaces. The map $\phi = \phi_2 \circ \phi_1^{-1}$ satisfies $[g_2 \circ \phi] = [g_1]$.

On the other hand, if ϕ is as in the statement of the lemma, then it is not hard to see that $(\mathbf{m}([g_1]), \varepsilon([g_1])) = (\mathbf{m}([g_2]), \varepsilon([g_2]))$. □

Lemma 5.11. *The map* $(\mathbf{m}, \varepsilon)$ *has the following properties:*

1. $(\mathbf{m}, \varepsilon)$ *is invariant under* Γ, *that is, for* $\gamma \in \Gamma = \mathrm{MCG}(f)$ *and* $[g] \in \mathcal{T} - \mathcal{T}_\infty$, $(\mathbf{m}(\gamma[g]), \varepsilon(\gamma[g])) = (\mathbf{m}([g]), \varepsilon([g]))$.
2. *The image of* $(\mathbf{m}, \varepsilon)$ *is* $\mathrm{SMATCH}^*(\Sigma, f)$.
3. $(\mathbf{m}, \varepsilon)$ *descends to a bijection* $\widehat{(\mathbf{m}, \varepsilon)} : \Gamma \backslash (\mathcal{T} - \mathcal{T}_\infty) \to \mathrm{SMATCH}^*(\Sigma, f)$ *that respects the* $\mathbf{Z}_{\ge 0}^{B}$ *gradings of the two sets given by the two incarnations of* κ *(on transverse maps and signed matchings).*

Proof. Part 1. This is a special case of one of the implications of Lemma 5.10.

Part 2. This follows from the fact that given any $(\mathbf{m}, \varepsilon) \in \text{SMATCH}^*(\Sigma, f)$, by definition $(\Sigma_{\mathbf{m},\varepsilon}, f_{\mathbf{m},\varepsilon}) \approx (\Sigma, f)$. So there is a homeomorphism $h : \Sigma_{\mathbf{m},\varepsilon} \to \Sigma$ such that $f_{\mathbf{m},\varepsilon} \circ h^{-1}$ is homotopic to f relative to V_o. On the other hand, $f_{\mathbf{m},\varepsilon} \circ h^{-1}$ is a strict transverse map realizing (Σ, f), all of whose zones are discs or Möbius bands by construction. Hence $[f_{\mathbf{m},\varepsilon} \circ h^{-1}] \in \mathcal{T}(\Sigma, f) - \mathcal{T}_\infty(\Sigma, f)$ with $\mathbf{m}([f_{m,\varepsilon} \circ h^{-1}]) = \mathbf{m}$ and $\varepsilon(f_{\mathbf{m},\varepsilon} \circ h^{-1}) = \varepsilon$ (by Part 1).

Part 3. The fact that $(\mathbf{m}, \varepsilon)$ descends to a surjective map $\widehat{(\mathbf{m}, \varepsilon)} : \Gamma \backslash (\mathcal{T} - \mathcal{T}_\infty) \to \text{SMATCH}^*(\Sigma, f)$ follows from Parts 1 and 2. We need to prove $\widehat{(\mathbf{m}, \varepsilon)}$ is injective, in other words, if $(\mathbf{m}([g_1]), \varepsilon([g_1])) = (\mathbf{m}([g_2]), \varepsilon([g_2]))$ then there is some $\gamma \in \text{MCG}(f)$ such that $\gamma([g_1]) = [g_2]$. The needed γ is furnished by Lemma 5.10 (taking $\Sigma_1 = \Sigma_2 = \Sigma$). □

Corollary 5.12. *$\Gamma \backslash (\mathcal{T} - \mathcal{T}_\infty)$ is finite.*

Proof. By Lemma 5.11, Part 3, $\Gamma \backslash (\mathcal{T} - \mathcal{T}_\infty)$ has the same cardinality as $\text{SMATCH}^*(\Sigma, f)$, which is finite by Corollary 3.14, applied with $\chi_0 = \chi(\Sigma)$. □

Proof of Theorem 3.16. Combining Theorems 4.14 and 5.3 (with $X = |\mathcal{T}|_{\text{poly}}$ and $G = \Gamma$) together with Lemmas 5.5 and 5.9 and Corollary 5.12 shows that $\chi^{(2)}(\Gamma)$ is well-defined and given by

$$\chi^{(2)}(\Gamma) = \sum_{[[g]] \in \Gamma \backslash (\mathcal{T} - \mathcal{T}_\infty)} (-1)^{|\kappa([g])|}.$$

Finally, Lemma 5.11, Part 3, shows that the above sum can be replaced by

$$\chi^{(2)}(\Gamma) = \sum_{(\mathbf{m},\varepsilon) \in \text{SMATCH}^*(\Sigma, f)} (-1)^{|\kappa(\mathbf{m},\varepsilon)|}$$

that gives the formula stated in Theorem 3.16. □

5.3 The L^2-Euler characteristic is the usual one for almost-incompressible maps

Recall Definition 2.7 of incompressible and almost-incompressible maps.

Theorem 5.13. *If $[(\Sigma, f)]$ is an almost-incompressible element of $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ then there exists a finite CW-complex $X(f)$ such that $X(f)$ is an Eilenberg-Maclane space of type $K(\text{MCG}(f), 1)$ and*

$$\chi^{(2)}(\text{MCG}(f)) = \chi(X(f))$$

where the right hand side is the usual topological Euler characteristic.

Remark 5.14. Note that by Lemma 2.8, Theorem 5.13 applies to all $[(\Sigma, f)]$ with $\chi(\Sigma) \geq \chi_{\max}(w_1, \ldots, w_\ell) - 1$.

Remark 5.15. Since such $X(f)$ are unique up to weak homotopy equivalence, $\chi(X(f))$ is an invariant of $\text{MCG}(f)$ usually simply denoted by $\chi(\text{MCG}(f))$.

The proof of Theorem 5.13 relies on the following lemma.

Lemma 5.16. *Let $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ be almost-incompressible and $\mathcal{T} = \mathcal{T}(\Sigma, f)$. Then all $[g] \in \mathcal{T}$ are almost-filling, namely, $\mathcal{T}_\infty(\Sigma, f)$ is empty.*

Proof. Assume that $[g] \in \mathcal{T}$ is not almost-filling. Then there must be a generic simple closed curve c in Σ that is disjoint from the arcs and curves of g. Indeed, if g contains any curves then we can take c to be parallel to one of these curves, and c would then be generic by **Restriction 1**. Otherwise, if there is a zone of g that is not a topological disc nor a Möbius band, then we can take c to be any generic simple closed curve in this zone. Since c lives in only one zone of g, $g(c)$ is confined to a contractible region of $\bigvee^r S^1$, hence is nullhomotopic. Hence $f(c)$ is also nullhomotopic, since by assumption g is homotopic to f. This contradicts our assumption, hence all $[g] \in \mathcal{T}(\Sigma, f)$ are almost-filling. □

Proof of Theorem 5.13. Let $X(f) = \mathrm{MCG}(f)\backslash|\mathcal{T}|_{\mathrm{poly}}$. This is a finite CW-complex by Corollary 5.12 and Lemma 5.16. Since by Lemma 5.9 the action of $\mathrm{MCG}(f)$ on $|\mathcal{T}|_{\mathrm{poly}}$ is free and by Theorem 4.14 $|\mathcal{T}|_{\mathrm{poly}}$ is contractible, standard arguments show that $X(f)$ is connected, $\pi_1(X(f)) \cong \mathrm{MCG}(f)$, and the higher homotopy groups $\pi_k(X(f)) = 0$ for $k \geq 2$. This is the statement that $X(f)$ is an Eilenberg-Maclane space of type $K(\mathrm{MCG}(f), 1)$.

For the equality of Euler characteristics, we may use Theorem 5.3 to obtain

$$\chi^{(2)}(\mathrm{MCG}(f)) = \sum_{[\sigma]\in \mathrm{MCG}(f)\backslash\mathcal{T}} (-1)^{\dim(\sigma)} = \sum_{\tau \text{ a cell of } X(f)} (-1)^{\dim(\tau)} = \chi(X(f)).$$

Note that the use of Theorem 5.3 is valid since the sum in (5.1) is finite by finiteness of $X(f)$. □

Corollary 5.17. *For fixed $w_1, \ldots, w_\ell$, there are only finitely many almost-incompressible elements in* $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$.

Proof. Suppose $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is almost-incompressible. By Lemmas 4.9 and 5.16, $\mathcal{T}(\Sigma, f)$ is non-empty, and all its elements are almost-filling. By Lemma 4.8, there is $[g] \in \mathcal{T}(\Sigma, f)$ with $\kappa([g]) = (0, \ldots, 0)$. Recalling the maps $(\mathbf{m}, \varepsilon)$ from §5.2 and relying on Lemma 5.10, the pair $(\mathbf{m}([g]), \varepsilon([g]))$ can only be obtained under the map $(\mathbf{m}, \varepsilon)$ on $\mathcal{T}(\Sigma, f)$ for one particular $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$. Therefore, the cardinality of the almost-incompressible maps in $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ is at most the cardinality of the set

$$\{(\mathbf{m}, \varepsilon) \in \mathrm{SMATCH}^*(w_1, \ldots, w_\ell) \,|\, \kappa(\mathbf{m}) = (0, 0, \ldots, 0)\}.$$

The latter set is clearly finite, since its elements are obtained by choosing a finite number of matchings $\mathbf{m}$ of a finite set and then a finite number of possible maps ε from the discs of $\Sigma_{\mathbf{m}}$ to $\{\pm 1\}$. □

Remark 5.18. The counting yielding the upper bound in the proof of the last corollary is much redundant. First, if $(\mathbf{m}, \varepsilon) \in \mathrm{SMATCH}^*(w_1, \ldots, w_\ell)$ and $(\Sigma_{\mathbf{m},\varepsilon}, f_{\mathbf{m},\varepsilon})$ is almost-filling, then ε assigns $+1$ to each 2-cell in $\Sigma_{\mathbf{m}}$ except for, possibly, at most one 2-cell in every connected component of $\Sigma_{\mathbf{m}}$. Indeed, if there were two zones in the same connected component of $\Sigma_{\mathbf{m},\varepsilon}$ which are Möbius bands, then a simple closed curve tracing the boundary of one of these zones, continuing to the second zone, tracing its boundary and going back to the first zone along a parallel path (thus creating a kind of a barbell-shape) would be a generic, compressing curve. Second, it is not hard to see that moving a single Möbius band from one disc of $\Sigma_{\mathbf{m}}$ to another disc in the same connected component, does not alter the element in $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$. Therefore, a given matching $\mathbf{m}$ corresponds to at most $2^{\#\text{ connected components of }\Sigma_{\mathbf{m}}}$ almost-incompressible maps in $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$.

6 Remaining proofs and some examples

6.1 Proof of Corollaries 1.17, 1.18 and 1.19

We begin with the following lemma.

Lemma 6.1. *If all $w_j \neq 1$, $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$, and all the connected components of Σ are annuli or Möbius bands, then $\mathrm{MCG}(f)$ is trivial and, thus, $\chi^{(2)}(\mathrm{MCG}(f)) = 1$.*

Proof. Note that $\mathrm{MCG}(f)$ is the product of the mapping class groups of f restricted to the various connected components of Σ, so it is sufficient to prove this when $\ell = 2$ and Σ is an annulus or $\ell = 1$ and Σ is a Möbius band. In the latter case, the whole mapping class group is trivial as stated in Lemma 5.6 (due to [Eps66, Thm. 3.4]), and so, in particular, $\mathrm{MCG}(f) = \{1\}$.

Finally, suppose that Σ is an annulus. The mapping class group of Σ is isomorphic to $\mathbf{Z}$ and generated by a Dehn twist D in a curve parallel to the boundary. Consider a directed arc β connecting

v_1 to v_2 (v_i is the marked point on the boundary component δ_i). Then $f(\beta)$ is a loop in $\bigvee^r S^1$ based at o, and we write $f_*(\beta) \in \pi_1(\bigvee^r S^1, o) = \mathbf{F}_r$ for the class of this loop. For every $n \in \mathbf{Z}$, $D^n(\beta)$ is also an arc in A with the same endpoints. If one boundary component of Σ is labeled w_1, then $f_*(D^n(\beta)) = w_1^{\pm n} f_*(\beta) \neq f_*(\beta)$ for all $0 \neq n \in \mathbf{Z}$ since $w_1 \neq 1$. Hence $D^n \notin \mathrm{MCG}(f)$ and so $\mathrm{MCG}(f) = \{1\}$. □

Lemma 6.2. *Let w_1 and w_2 be two words in $\mathbf{F}_r$, both $\neq 1$. Let d be the maximal integer such that $w_1 = u^d$ with $u \in \mathbf{F}_r$. The number of elements $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, w_2)$ such that Σ is an annulus is*

$$\begin{cases} d & \text{if } w_1 \text{ is conjugate to either } w_2 \text{ or } w_2^{-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Assume that $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, w_2)$ is an annulus. Then, $f : \Sigma \to \bigvee^r S^1$ defines a free homotopy between $f(\delta_1)$ and $f(\delta_2)$. Since free homotopy classes of oriented curves in $\bigvee^r S^1$ correspond to conjugacy classes in $\mathbf{F}_r$, this shows that w_1 must be conjugate to either w_2 or w_2^{-1}, depending on whether the orientations of the two boundary components of Σ agree or not.

Since no non-identity element of $\mathbf{F}_r$ is conjugate to its inverse, w_1 cannot be conjugate to both w_2 and w_2^{-1}. Without loss of generality, we assume from now on that w_1 is conjugate to w_2. Let β be a directed arc as in the proof of Lemma 6.1, connecting v_1 to v_2. Denote $b = f_*(\beta) \in \mathbf{F}_r$, and note that $w_1 = b w_2 b^{-1}$. Also note that as β cuts Σ into a disc, the map f is completely determined, up to homotopy, by $f_*(\beta)$. As in the proof of Lemma 6.1,

$$\{([f] \circ [\rho])_* (\beta) \,|\, [\rho] \in \mathrm{MCG}(\Sigma)\} = \{w_1^n f_*(\beta)\}_{n \in \mathbf{Z}}.$$

But as the centralizer of w_1 in $\mathbf{F}_r$ is $\langle u \rangle$, with $u \in \mathbf{F}_r$ the d-th root of w_1 as in the statement of the lemma, we have $\{c \in \mathbf{F}_r \,|\, w_1 = c w_2 c^{-1}\} = \{u^n b \,|\, n \in \mathbf{Z}\}$. Thus, there are exactly d distinct orbits of possible values of $f_*(\beta)$ under the action of $\mathrm{MCG}(\Sigma)$, and therefore exactly d classes of annuli in $\mathsf{Surfaces}^*(w_1, w_2)$. □

Lemma 6.3. *Let $w \neq 1$, $w \in \mathbf{F}_r$. The number of elements $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w)$ such that Σ is a Möbius band is 0 if w is not a square and 1 if w is a square in $\mathbf{F}_r$.*

Proof. If w is not a square in $\mathbf{F}_r$ then there are no $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w)$ with Σ a Möbius band by Lemma 2.2. So suppose that w is a square in $\mathbf{F}_r$. Then by Lemma 2.2, there is at least one $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w)$ with Σ a Möbius band. Let $[(\Sigma, f)]$ be of this form. Then up to homotopy, there is a unique arc $\alpha(\Sigma)$ in Σ joining v_1 to itself and not separating Σ [Eps66, Proof of Thm. 3.4]. We have $f^*(a)^2 = w$, which uniquely specifies $f_*(\alpha)$. Let u be the unique solution of $u^2 = w$ in $\mathbf{F}_r$.

Let (M_1, f_1) and (M_2, f_2) be admissible maps for w with the M_i Möbius bands. For $i = 1, 2$ let α_i be an embedded directed arc from v_1 to itself in M_i that does not separate M_i. Since by the previous paragraph $(f_1)_*(\alpha_1) = (f_2)_*(\alpha_2) = u$, the homeomorphism h from M_1 to M_2 that preserves the markings on boundaries and maps α_1 to α_2, has the property that $f_2 \circ h$ is homotopic to f_1 and this shows $[(M_1, f_1)] = [(M_2, f_2)]$. *Hence there is exactly one element $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w)$ with Σ a Möbius band.* □

Proof of Corollary 1.17. We assume all $w_j \neq 1$ and examine the expansion given in Theorem 1.5. The limit $\lim_{n\to\infty} \mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ exists, since there are no $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ with $\chi(\Sigma) > 0$ by Lemma 2.4. Moreover, Lemma 2.4 gives that

$$\mathsf{Surfaces}^*_0(w_1, \ldots, w_\ell) \stackrel{\text{def}}{=} \{[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_1, \ldots, w_\ell) \,:\, \chi(\Sigma) = 0\}$$

consists precisely of surfaces all the connected components of which are annuli or Möbius bands. Thus

$$\lim_{n\to\infty} \mathcal{T}r^{O}_{w_1,\ldots,w_\ell}(n) = \sum_{[(\Sigma,f)]\in\mathsf{Surfaces}^*_0(w_1,\ldots,w_\ell)} \chi^{(2)}(\mathrm{MCG}(f)) = \sum_{[(\Sigma,f)]\in\mathsf{Surfaces}^*_0(w_1,\ldots,w_\ell)} 1$$
$$= \left|\mathsf{Surfaces}^*_0(w_1,\ldots,w_\ell)\right|,$$

the second equality following from Lemma 6.1. This proves the first statement of the corollary.

Every admissible pair $[(\Sigma,f)] \in \mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$ induces a partition on $\{w_1,\ldots,w_\ell\}$ where every block consists of the words associated with one connected component of Σ. The algebraic characterization given in the statement of Corollary 1.17 follows from the geometric part using these partitions and Lemmas 6.2 and 6.3. $\square$

Proof of Corollary 1.18. Assume $w = u^d$, where $u \neq 1$ is not a proper power. The first statement of this corollary follows readily from the algebraic characterization of $\lim_{n\to\infty}\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$ in Corollary 1.17: the valid partitions of the words $w^{j_1},\ldots,w^{j_\ell}$ depend only on $j_1,\ldots,j_\ell$, and the weight of every partition depends only on d and not on u. The collection of limits determines d using, for example, the following two equalities:

$$\lim_{n\to\infty}\mathcal{T}r^{\mathrm{O}}_{w}(n) = \begin{cases}1 & \text{if } d \text{ is even,}\\ 0 & \text{if } d \text{ is odd,}\end{cases} \quad\text{and}\quad \lim_{n\to\infty}\mathcal{T}r^{\mathrm{O}}_{w,w}(n) = \begin{cases}d+1 & \text{if } d \text{ is even,}\\ d & \text{if } d \text{ is odd.}\end{cases}$$

The analogous result for $\mathrm{Sp}(n)$ now follows from Theorem 1.2. $\square$

Proof of Corollary 1.19. Corollary 1.18 shows that the joint moments of $T_n(w),\ldots,T_n(w^\ell)$ converge to the same values as the joint moments of $T_n(x),\ldots,T_n(x^\ell)$ for some $x\in B$, as long as $w\neq 1$ and is not a proper power. By the method of moments one can now deduce the corollary, using that Diaconis and Shahshahani have shown in [DS94, §3] that these limits of moments are precisely those of the multivariate normal distribution described in the statement. $\square$

6.2 An example: non-orientable surface words

Fix $s \geq 1$. Let $w_s = x_1^2\cdots x_s^2 \in \mathbf{F}_s$. Here we analyze $\mathsf{Surfaces}^*(w_s)$. Note that $\chi_{\max}(w_s) = 1 - \mathrm{sql}(w_s) = 1-s$, so there is no admissible map $(\Sigma,f)\in\mathsf{Surfaces}^*(w_s)$ with $\chi(\Sigma) > 1-s$.

Claim 6.4. *Let $t\in\mathbf{Z}_{\geq 0}$ be a non-negative integer. Then there is exactly one $[(\Sigma_t,f_t)]\in\mathsf{Surfaces}^*(w_s)$ with $\chi(\Sigma_t) = 1-s-t$ that can be realized by an almost-filling strict transverse map. In particular, there is at most one $[(\Sigma_t,f_t)]\in\mathsf{Surfaces}^*(w_s)$ with $\chi^{(2)}(\mathrm{MCG}(f_t))\neq 0$ and $\chi(\Sigma_t) = 1-s-t$. In fact,*

$$\chi^{(2)}(\mathrm{MCG}(f_t)) = \begin{cases}\delta_{t,0} & \text{if } s=1,\\ (-1)^t\binom{t+s-2}{s-2} & \text{if } s\geq 2,\end{cases} \tag{6.1}$$

where $\delta_{t,0}$ is the Kronecker delta.

Note that for all $t\geq 0$, Σ_t is the non-orientable surface of genus $s+t$ with one boundary component (namely, the connected sum of $s+t$ copies of $\mathbf{R}P^2$, with a disc removed). When $t=0$, $\mathrm{MCG}(f_0)$ is trivial and $\chi^{(2)}(\mathrm{MCG}(f_0)) = \chi(\mathrm{MCG}(f_0)) = 1$. When $t=1$, a simple analysis gives that $\mathrm{MCG}(f_1)\cong\pi_1(\Sigma_0)\cong\mathbf{F}_s$, in which case $\chi^{(2)}(\mathrm{MCG}(f_1)) = \chi(\mathrm{MCG}(f_1)) = \chi(\mathbf{F}_s) = 1-s$. This agrees with the $t=1$ case in (6.1). It intrigues us to wonder whether $\mathrm{MCG}(f_t)$ is related to some well-known group when $t\geq 2$.

Proof of Claim 6.4. If $[(\Sigma,f)]\in\mathsf{Surfaces}^*(w_s)$ satisfies $\chi^{(2)}(\mathrm{MCG}(f))\neq 0$, then it must be realized by some almost-filling strict transverse map, by Theorem 3.16 and Lemma 5.11. So the second statement of the claim follows from the first one.

Now fix $t \geq 0$ and assume that $[(\Sigma, f)] \in \mathsf{Surfaces}^*(w_s)$ satisfies $\chi(\Sigma) = 1 - s - t$ and is realized by some almost-filling strict transverse map g. As g is almost-filling, it has only arcs and no curves. Note that each letter x_i appears exactly twice in $w_s = x_1^2 \cdots x_s^2$, so there is only one possible matching for every letter x_i, matching the two occurrences of x_i in w_s. Therefore, there is a single (x_i, j)-zone for every $1 \leq i \leq s$ and $0 \leq j < \kappa_{x_i}([g])$ which must be a Möbius band. It is easy to check that there is one o-zone in g, which may be a disc or a Möbius band. A simple Euler characteristic calculation shows that exactly t of the zones of g are Möbius bands. In particular, $|\kappa([g])| \in \{t-1, t\}$.

Now we modify $[g]$ by forgetting all points of transversion (x, i) with $i \geq 1$. Let $[h] \in \mathcal{T}(\Sigma, f)$ be the resulting transverse map. This $[h]$ has exactly one zone, it is an o-zone, and by the previous paragraph, this zone was obtained by gluing t Möbius bands together, and is, thus, necessarily the non-orientable surface of genus t with one boundary component. Because the matchings in $[h]$ are dictated and so is the topological type of its sole zone, any $[h']$ obtained in the same way from some other $[(\Sigma', f')]$ with the same properties, would be equivalent to $[h]$. Namely, we could find $[h'] \in \mathcal{T}(\Sigma', f')$ with $[h'] = [h \circ \phi]$ for $\phi : \Sigma' \to \Sigma$ a homeomorphism respecting boundary markings. The same ϕ shows $(\Sigma', f') \approx (\Sigma, f)$. Hence there is exactly one $[(\Sigma_t, f_t)] \in \mathsf{Surfaces}^*(w_s)$ *with* $\chi(\Sigma_t) = 1 - s - t$ *which can be realized by an almost-filling strict transverse map.*

It is left to prove the equality (6.1). We prove it in two different ways. First, as mentioned above, any almost-filling $[g] \in \mathcal{T}(\Sigma_t, f_t)$ satisfies $|\kappa([g])| \in \{t-1, t\}$, and the same analysis shows that there is exactly one $\mathrm{MCG}(f)$-orbit of almost-filling strict transverse maps in $\mathcal{T}(\Sigma_t, f_t)$ for every valid choice of κ. There are $\binom{t+s-1}{t}$ possible $\kappa \in (\mathbf{Z}_{\geq 0})^s$ with $|\kappa| = t$, each contributing $(-1)^t$ to (3.16), and $\binom{t-1+s-1}{t-1}$ possible $\kappa \in (\mathbf{Z}_{\geq 0})^s$ with $|\kappa| = t-1$, each contributing $(-1)^{t-1}$ to (3.16). The total sum is precisely the one specified in (6.1).

The second proof uses the Frobenius-Schur type formula (1.6), by which for $s = 1$

$$\mathcal{T}r^{\mathrm{O}}_{x^2}(n+1) = 1,$$

and for $s \geq 2$,

$$\mathcal{T}r^{\mathrm{O}}_{w_s}(n+1) = \frac{1}{(n+1)^{s-1}} = \frac{1}{n^{s-1}}\left(1 - \frac{1}{n} + \frac{1}{n^2} - \dots\right)^{s-1} = \frac{1}{n^{s-1}} \sum_{t=0}^{\infty} (-1)^t \binom{t+s-2}{s-2} \frac{1}{n^t}.$$

Combining these two expressions with Theorem 1.5, we see that $\chi^{(2)}(\mathrm{MCG}(f_t))$ is given by (6.1). ☐

6.3 More examples

We give here some more details about the examples from Table 1. We elaborate on the exact contributions, in the language of Theorem 1.5, to the two leading terms with exponents $\chi_{\max}$ and $\chi_{\max} - 1$. The data is summarized in Table 2. The analysis of these examples was carried out with the help of a SageMath script, and using various observations and considerations. We do not describe the analysis here as we do not see it as crucial – we only aim to give a sense of how our main theorem plays out in concrete examples.

The fourth column of Table 2 specifies the rational expressions for $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n+1)$ (unlike the expression in Table 1 which gave the expression for $\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n)$), as well as the coefficients of $n^{\chi_{\max}}$ and of $n^{\chi_{\max}-1}$ in the Laurent expansion. The fifth column is the same as the fifth one in Table 1, while the sixth column lists the equivalence classes of maps $[(\Sigma, f)]$ in $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$ with $\chi(\Sigma) = \chi_{\max} - 1$. Note that by Lemma 2.8, when $\chi(\Sigma) = \chi_{\max}$ the maps are incompressible, and when $\chi(\Sigma) = \chi_{\max} - 1$, the maps are always almost-incompressible and sometimes even incompressible (in the table we point out specifically the cases where the stronger condition holds).

Moreover, by Corollary 5.17, there are finitely many such equivalence classes in $\mathsf{Surfaces}^*(w_1, \ldots, w_\ell)$, so we can indeed list them all. By Theorem 5.13, in all these cases, we get

ℓ	$w_1,\ldots,w_\ell$	$\chi_{\max}$	$\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(n+1)$ and two leading terms	Admissible maps with $\chi(\Sigma)=\chi_{\max}$	Admissible maps with $\chi(\Sigma)=\chi_{\max}-1$
1	x^2y^2	-1	$\frac{1}{n+1}=\frac{1}{n}-\frac{1}{n^2}+\ldots$	one $P_{2,1}$ w. $\mathrm{MCG}(f)=\{1\}$	one $P_{3,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{F}_2$
	x^4y^4	-1	$\frac{1}{n+1}=\frac{1}{n}-\frac{1}{n^2}+\ldots$	one $P_{2,1}$ w. $\mathrm{MCG}(f)=\{1\}$	two incompr. $P_{3,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{Z}$; one $P_{3,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{F}_2$
	$[x,y]^2$	0	$\frac{n^3+4n^2+3n-4}{(n+1)(n+3)n}=1+\frac{0}{n}+\ldots$	one $P_{1,1}$ w. $\mathrm{MCG}(f)=\{1\}$	one $P_{2,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{Z}$
	$xy^3x^{-1}y^{-1}$	-2	$0=\frac{0}{n^2}+\frac{0}{n^3}+\ldots$	one $P_{3,1}$ w. $\mathrm{MCG}(f)\cong\mathbb{Z}$	one $P_{4,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{Z}\times\mathbf{F}_3$
	$xy^4x^{-1}y^{-2}$	-1	$\frac{1}{n+1}=\frac{1}{n}-\frac{1}{n^2}+\ldots$	one $P_{2,1}$ w. $\mathrm{MCG}(f)=\{1\}$	three incompr. $P_{3,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{Z}$; one $P_{3,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{F}_2$
	$xyx^2yx^3y^2$	-2	$\frac{3n+5}{(n+1)(n+3)n}=\frac{3}{n^2}-\frac{7}{n^3}+\ldots$	three $P_{3,1}$ w. $\mathrm{MCG}(f)=\{1\}$	one $P_{4,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{F}_3$; one $P_{4,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{F}_6$
2	w,w for $w=x^2y^2$	0	$\frac{n^3+4n^2+7n+8}{(n+1)(n+3)n}=1+\frac{0}{n}+\ldots$	one A w. $\mathrm{MCG}(f)=\{1\}$	one $P_{1,2}$ w. $\mathrm{MCG}(f)\cong\mathbf{Z}$
	w,w for $w=x^2y$	0	$1=1+\frac{0}{n}+\ldots$	one A w. $\mathrm{MCG}(f)=\{1\}$	one $P_{1,2}$ w. $\mathrm{MCG}(f)\cong\mathbf{Z}$
3	w,w,w for $w=x^2y^2$	-1	$\frac{3(n^4+7n^3+13n^2+15n+24)}{(n-1)n(n+1)(n+3)(n+5)}=\frac{3}{n}-\frac{3}{n^2}+\ldots$	three $A\sqcup P_{2,1}$ w. $\mathrm{MCG}(f)=\{1\}$	three $A\sqcup P_{3,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{F}_2$; three $P_{1,2}\sqcup P_{2,1}$ w. $\mathrm{MCG}(f)\cong\mathbf{Z}$

Table 2: This table gives more details about the examples from Table 1. Here A denotes an annulus, and as in Table 1, $P_{g,b}$ denotes the non-orientable surface of genus g with b boundary components (so $\chi(P_{g,b})=2-g-b$).

concrete, finite CW-complexes of type $K(\mathrm{MCG}(f),1)$ for these maps, which means we can understand the groups pretty well. Indeed, we were able to compute the exact isomorphism type of the groups $\mathrm{MCG}(f)$ in all cases mentioned in the table. The fact all groups but one are free is probably only due to the fact that the words in these examples are rather short, which means the complexes associated with them tend to have low dimensions.

A Proof of Theorem 1.2: relationship between O and Sp

Here we prove Theorem 1.2. Throughout this appendix, fix n with $2n \geq N(w_1,\ldots,w_\ell)$, the latter defined in (3.2). For $i \in [2n]$ denote

$$\hat{i} \stackrel{\text{def}}{=} \begin{cases} i+n & \text{if } 1 \leq i \leq n, \\ i-n & \text{if } n+1 \leq i \leq 2n, \end{cases}$$

and

$$\xi(i) \stackrel{\text{def}}{=} \operatorname{sign}\left(n+\frac{1}{2}-i\right) = \begin{cases} 1 & \text{if } 1 \leq i \leq n, \\ -1 & \text{if } n+1 \leq i \leq 2n. \end{cases}$$

Recall that we think of $\mathrm{Sp}(n)$ as a subgroup of $\mathrm{GL}_{2n}(\mathbf{C})$, and that the matrix J was defined in (1.2). The following lemma follows easily from (1.3) and the fact that $A^{-1} = J^T A^T J$ for $A \in \mathrm{Sp}(n)$.

Lemma A.1. *If $A \in \mathrm{Sp}(n)$ and $i,j \in [2n]$, then*

$$\left(A^{-1}\right)_{i,j} = \xi(i)\,\xi(j)\,A_{\hat{j},\hat{i}}. \tag{A.1}$$

Our first goal is to obtain an analog of Theorem 3.4 for $\mathrm{Sp}(n)$, namely, to obtain a formula for $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n)$ as a finite sum over systems of matchings, only with an additional sign associated with every such system; see Proposition A.5 for the precise statement.

We recall some of the notation we use here. Let $2L = 2\sum_{x\in B} L_x = \sum_{j=1}^{\ell} |w_j|$ denote the total number of letters in $w_1,\ldots,w_\ell$. The jth boundary component of every surface in $\mathsf{Surfaces}^*(w_1,\ldots,w_\ell)$ is subdivided to $|w_j|$ intervals corresponding to the letters of w_j, and we denoted by $\mathcal{I},\mathcal{I}^+,\mathcal{I}^-$ the sets of all $2L$ intervals, the subset of intervals corresponding to positive letters and its complement, respectively. Likewise, we denote by $\mathcal{I}_x,\mathcal{I}_x^+,\mathcal{I}_x^-$ the analogous sets of intervals corresponding to the instances of $x \in B$. We again identify $\mathcal{I}_x$ with the set $[2L_x]$, for each $x \in B$, in the same way as in Section 2.1.4. Similarly to the notation from Section 2.1.4, we denote by $\mathcal{A} = \mathcal{A}(w_1,\ldots,w_\ell)$ the set of index assignments

$$\mathbf{a}\colon \{p_I(k) \mid I \in \mathcal{I}, k \in \{0,1\}\} \to [2n],$$

where for every two immediately adjacent marked points p,q in $\cup_{j=1}^{\ell} C(w_j)$ that belong to different intervals in $\mathcal{I}$ we have $\mathbf{a}(p) = \mathbf{a}(q)$. (Note the range here is $[2n]$ and not $[n]$ as in Section 2.1.4). Given $\mathbf{a} \in \mathcal{A}$, let $\hat{\mathbf{a}}$ be the assignment obtained after applying (A.1), namely,

$$\hat{\mathbf{a}}(p_I(i)) = \begin{cases} \mathbf{a}(p_I(i)) & \text{if } I \in \mathcal{I}^+ \\ \widehat{\mathbf{a}(p_I(i))} & \text{if } I \in \mathcal{I}^-. \end{cases}$$

As we shall use Theorem 3.1 for evaluating $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n)$, we need the following expression which gathers the total sign contribution for a given system of matchings $\mathbf{m} = \{(m_{x,0},m_{x,1})\}_{x\in B} \in \mathrm{MATCH}^{\kappa\equiv 1}$ and an assignment $\mathbf{a}$. Recall the notation $\delta^{\mathrm{Sp}}_{\mathbf{i},m}$ from Theorem 3.1. We let

$$\begin{aligned}\Delta(\mathbf{a},\mathbf{m}) &\stackrel{\text{def}}{=} \left[\prod_{I\in\mathcal{I}^-} \xi(\mathbf{a}(p_I(0)))\,\xi(\mathbf{a}(p_I(1)))\right] \cdot \prod_{\substack{x\in B\\ k=0,1}} \delta^{\mathrm{Sp}}_{\hat{\mathbf{a}}|_{\{p_I(k)\,|\,I\in\mathcal{I}_x\}},m_{x,k}} \\ &= \left[\prod_{I\in\mathcal{I}^-} \xi(\mathbf{a}(p_I(0)))\,\xi(\mathbf{a}(p_I(1)))\right] \cdot \prod_{\substack{x\in B\\ k=0,1}} \prod_{\substack{(p_I(k),p_J(k))\\ \text{matched by } m_{x,k}}} \left\langle e_{\hat{\mathbf{a}}(p_I(k))}, e_{\hat{\mathbf{a}}(p_J(k))}\right\rangle_{\mathrm{Sp}},\end{aligned}$$

where in the innermost product, each matched pair appears once and is given its predetermined order. Note that for $i,j \in [2n]$, we have

$$\langle e_i, e_j\rangle_{\mathrm{Sp}} = e_i^T J e_j = \delta_{\hat{i},j}\,\xi(i), \tag{A.2}$$

where here δ is the Kronecker delta. Also notice that $\Delta(\mathbf{a},\mathbf{m}) \in \{-1,0,1\}$. We say $\mathbf{a} \vdash^* \mathbf{m}$ if $\Delta(\mathbf{a},\mathbf{m}) \neq 0$. Therefore,

$$\Delta(\mathbf{a},\mathbf{m}) = \mathbf{1}_{\mathbf{a}\vdash^*\mathbf{m}} \cdot \left[\prod_{I\in\mathcal{I}^-} \xi\left(\mathbf{a}(p_I(0))\right)\xi\left(\mathbf{a}(p_I(1))\right)\right] \cdot \prod_{\substack{x\in B\\ k=0,1}} \prod_{\substack{(p_I(k),p_J(k))\\ \text{matched by } m_{x,k}}} \xi\left(\hat{\mathbf{a}}\left(p_I(k)\right)\right). \tag{A.3}$$

Definition A.2. Let $\mathbf{m} \in \mathrm{MATCH}^{\kappa\equiv 1}$. Call a matching arc of $\mathbf{m}$ *orientable* if it pairs an interval in $\mathcal{I}^{\pm}$ with an interval in $\mathcal{I}^{\mp}$, and *non-orientable* otherwise. Let m be one of the matchings in $\mathbf{m}$. In every pair $\left(m_{(2t-1)}, m_{(2t)}\right)$ we think of the corresponding matching arc in $\Sigma_{\mathbf{m}}$ as directed from its *origin* – the interval corresponding to $m_{(2t-1)}$, to its *terminus* – the interval associated with $m_{(2t)}$. Let D be a type-o disc of $\Sigma_{\mathbf{m}}$. Every interval in $\mathcal{I}$ that meets ∂D has an orientation coming from the given orientation of $\partial\Sigma_{\mathbf{m}}$. We say that two intervals that meet ∂D are *co-oriented* (relative to D) if their orientation induces the same orientation on ∂D, and *counter-oriented* otherwise. Note that a matching arc is orientable if and only if it matches two co-oriented intervals meeting δD.

In the computation of $\Delta(\mathbf{a},\mathbf{m})$, we attribute every sign that appears in (A.3) to one of the type-o discs of $\Sigma_{\mathbf{m}}$. Indeed, every matching arc is at the boundary of exactly one type-o disc, and every $p_I(k)$ also belongs to exactly one type-o disc.

Lemma A.3 (Computation of $\Delta(\mathbf{a},\mathbf{m})$)**.** *Assume $\mathbf{a},\mathbf{a}_1,\mathbf{a}_2 \in \mathcal{A}$ and $\mathbf{m} \in \mathrm{MATCH}^{\kappa\equiv 1}$, and let D be a type-o disc in $\Sigma_{\mathbf{m}}$.*

1. *The number of non-orientable matching arcs along ∂D is even.*
2. *If $\mathbf{a} \vdash^* \mathbf{m}$, the total sign contribution of D to $\Delta(\mathbf{a},\mathbf{m})$ is the product of:*
   (i) the sign of the index[8] *given by $\mathbf{a}$ at the origin of every matching arc with origin in $\mathcal{I}^+$,*
   (ii) the sign of the index given by $\mathbf{a}$ at the terminus of every matching arc with terminus in $\mathcal{I}^-$, and
   (iii) (-1) for every matching arc with origin in $\mathcal{I}^-$.
3. *If $\mathbf{a} \vdash^* \mathbf{m}$ and $p_I(k), p_J(k)$ are matched by any $m_{x,k}$ then $\mathbf{a}(p_I(k)) \equiv \mathbf{a}(p_J(k)) \bmod n$, and moreover, $\mathbf{a}(p_I(k)) = \mathbf{a}(p_J(k))$ if and only if $m_{x,k}$ corresponds to an orientable matching arc.*
4. *For fixed $\mathbf{m}$, the number of $\mathbf{a}$ with $\mathbf{a} \vdash^* \mathbf{m}$ is $(2n)^{\#\{\text{type-}o\text{ discs of }\Sigma_{\mathbf{m}}\}}$.*
5. *If $\mathbf{a}_1,\mathbf{a}_2 \vdash^* \mathbf{m}$, then $\Delta(\mathbf{a}_1,\mathbf{m}) = \Delta(\mathbf{a}_2,\mathbf{m})$.*

The final statement allows us to define:

Definition A.4. For every $\mathbf{m} \in \mathrm{MATCH}^{\kappa\equiv 1}$ we let $\Delta(\mathbf{m}) \stackrel{\text{def}}{=} \Delta(\mathbf{a},\mathbf{m})$, defined by any $\mathbf{a} \vdash^* \mathbf{m}$.

Proof of Lemma A.3. We prove part by part.

Part 1. The first point is due to the fact that the boundary components of Σ have built-in orientation, and along the boundary of D, the orientation of intervals meeting $\partial\Sigma$ is preserved when going along an orientable matching arc, and flipped along a non-orientable matching are. But ∂D is a loop, so the number of orientation flips must be even.

Part 2. This follows from (A.3) by checking case by case over all possibilities.

[8] Here and elsewhere in this appendix, the "sign of an index" i is $\xi(i)$.

Part 3. We have $\mathbf{a} \vdash^* \mathbf{m}$ if and only if for all ordered matched pairs $p_I(k), p_J(k)$ of any $m_{x,k}$

$$\mathbf{a}\left(p_I(k)\right) = \begin{cases} \mathbf{a}\left(\widehat{p_J(k)}\right) & \text{if } I \text{ and } J \text{ are in the same set } \mathcal{I}^{\pm}, \\ \mathbf{a}\left(p_J(k)\right) & \text{if } I \text{ and } J \text{ are in } \mathcal{I}^{\pm} \text{ and } \mathcal{I}^{\mp}, \text{ respectively.} \end{cases} \tag{A.4}$$

This means that when $\mathbf{a} \in \mathcal{A}$ and $\mathbf{a} \vdash^* \mathbf{m}$, there is a constraint on the values of $\mathbf{a}$ at every pair of points that are adjacent on the boundary of some type-o disc of $\Sigma_{\mathbf{m}}$. This is similar to the situation for the orthogonal group, but the constraints are more complicated now. The constraint implies that the values of $\mathbf{a}$ on the points $p_I(k)$ in the boundary of a fixed type-o disc D of $\Sigma_{\mathbf{m}}$ are determined by the value at any fixed point p_D on the boundary of that disc. The values of $\mathbf{a}$ are constant along segments of ∂D, except for segments that are matching arcs joining intervals in the same set $I_x^{\pm}$, across which the value of $\mathbf{a}$ jumps by $n \bmod 2n$. These are the non-orientable matching arcs defined in Definition A.2.

Part 4. It now follows from Parts 1 and 3 that if for each type-o disc D of $\Sigma_{\mathbf{m}}$, we choose $\mathbf{a}(p)$ for some p in ∂D, then there exists a unique $\mathbf{a} \in \mathcal{A}$ with these prescribed values and such that $\mathbf{a} \vdash^* \mathbf{m}$. Hence, for any $\mathbf{m}$, there are $(2n)^{\#\{\text{type-}o \text{ discs of } \Sigma_{\mathbf{m}}\}}$ elements of $\mathcal{A}$ with $\mathbf{a} \vdash^* \mathbf{m}$.

Part 5. We need to show that for $\mathbf{m}$ fixed, all the $\Delta(\mathbf{a}, \mathbf{m})$ have the same sign. Indeed, we collect the contribution to the sign of every o-disc D separately, and show it does not depend on the particular assignment of indices along ∂D. There are two options for the signs of these indices, where one is a complete negation of the other. Recall that the sign of $\Delta(\mathbf{a}, \mathbf{m})$ splits up into three types of contributions according to Part 2. The contribution from (iii) clearly does not depend on $\mathbf{a}$. Now consider the (i)- and (ii)-type contributions.

- If α is an orientable matching arc, its (i)- and (ii)-type contributions to $\Delta(\mathbf{a}_i, \mathbf{m})$ are always 1 in total. This is surely the case if α is directed from $I \in I_x^-$ to $J \in I_x^+$. But it is also the case when α is directed the other way round, as the signs of both indices at its endpoints are identical. Hence the contributions of type (i) and type (ii) of orientable matching arcs to either $\Delta(\mathbf{a}_1, \mathbf{m})$ or $\Delta(\mathbf{a}_2, \mathbf{m})$ is equal to 1.
- Note from the discussion in the proof of Part 3, that $\mathbf{a}_1$ and $\mathbf{a}_2$ are related by a sequence of the following type of *flip-moves*: choose a type-o disc D of $\Sigma_{\mathbf{m}}$, and modify $\mathbf{a}_1$ by adding n to $\mathbf{a}(p_I(k))$ modulo $2n$, for every $p_I(k)$ that meets ∂D. Now for any given non-orientable matching arc α, its (i)- and (ii)-type contribution is the sign of one of the endpoints. Hence the effect of a flip-move on $\mathbf{a}_1$ at a disc D is to change the type (i) and (ii) contributions to $\Delta(\mathbf{a}_1, \mathbf{m})$ by $(-1)^{\#\text{non-orientable matching arcs of } \mathbf{m} \text{ meeting } D}$. On the other hand, by Part 1, the total number of non-orientable matching arcs of $\mathbf{m}$ meeting D is even.

This concludes the proof of Lemma A.3. □

We can now prove the analog of Theorem 3.4 for $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n)$.

Proposition A.5. *For* $2n \geq N$

$$\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n) = \sum_{\mathbf{m} \in \mathrm{MATCH}^{\kappa \equiv 1}} (2n)^{\#\{\textit{type-}o \textit{ discs of } \Sigma_{\mathbf{m}}\}} \Delta(\mathbf{m}) \prod_{x \in B} \mathrm{Wg}^{\mathrm{Sp}}_{L_x}(m_{x,0}, m_{x,1}; n),$$

with $\Delta(\mathbf{m}) \in \{1, -1\}$ *as defined in Definition A.4.*

Proof. Let $g(I)$ be as in §3.2. Assume $2n \geq N$. By the same arguments that led to (3.7), incorporating (A.1) and using Theorem 3.1, we have

$$\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n) = \sum_{\mathbf{a}\in\mathcal{A}(w_1,\ldots,w_\ell)} \sum_{\mathbf{m}\in\mathrm{MATCH}^{\kappa\equiv 1}} \Delta(\mathbf{a},\mathbf{m}) \prod_{x\in B} \mathrm{Wg}^{\mathrm{Sp}}_{L_x}(m_{x,0}, m_{x,1}; n).$$

This formula was the original motivation for introducing $\Delta(\mathbf{a},\mathbf{m})$*.* Now using Lemma A.3, Parts 4 and 5, and interchanging the sums over $\mathbf{a}$ and $\mathbf{m}$ gives

$$\begin{aligned}\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n) &= \sum_{\mathbf{m}\in\mathrm{MATCH}^{\kappa\equiv 1}} \left(\prod_{x\in B} \mathrm{Wg}^{\mathrm{Sp}}_{L_x}(m_{x,0}, m_{x,1}; n)\right)\left(\sum_{\mathbf{a}\in\mathcal{A}(w)} \Delta(\mathbf{a},\mathbf{m})\right)\\ &= \sum_{\mathbf{m}\in\mathrm{MATCH}^{\kappa\equiv 1}} (2n)^{\#\{\text{type-}o\text{ discs of }\Sigma_{\mathbf{m}}\}}\, \Delta(\mathbf{m}) \prod_{x\in B} \mathrm{Wg}^{\mathrm{Sp}}_{L_x}(m_{x,0}, m_{x,1}; n)\end{aligned}$$

as required. □

We can now prove the main result of this subsection and show that $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n) = (-1)^\ell \cdot \mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}(-2n)$ for large n.

Proof of Theorem 1.2. It follows from Proposition A.5 and Lemma 3.2 that

$$\begin{aligned}\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n) &= \sum_{\mathbf{m}\in\mathrm{MATCH}^{\kappa\equiv 1}} (2n)^{\#\{o\text{-discs of }\Sigma_{\mathbf{m}}\}}\, \Delta(\mathbf{m}) \prod_{x\in B} \mathrm{Wg}^{\mathrm{Sp}}_{L_x}(m_{x,0}, m_{x,1}; n)\\ &= \sum_{\mathbf{m}\in\mathrm{MATCH}^{\kappa\equiv 1}} (2n)^{\#\{o\text{-discs of }\Sigma_{\mathbf{m}}\}}\\ &\qquad \cdot \prod_{x\in B} (-1)^{L_x} \cdot \mathrm{sign}\left(\sigma^{-1}_{m_{x,0}} \cdot \sigma_{m_{x,1}}\right) \cdot \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0}, m_{x,1}; -2n) \cdot \Delta(\mathbf{m})\\ &= \sum_{\mathbf{m}\in\mathrm{MATCH}^{\kappa\equiv 1}} (-2n)^{\#\{o\text{-discs of }\Sigma_{\mathbf{m}}\}} \prod_{x\in B} \mathrm{Wg}^{\mathrm{O}}_{L_x}(m_{x,0}, m_{x,1}; -2n)\, \Xi(w_1,\ldots,w_\ell;\mathbf{m}) \quad \text{(A.5)}\end{aligned}$$

where

$$\Xi(w_1,\ldots,w_\ell;\mathbf{m}) \stackrel{\text{def}}{=} (-1)^{\#\{\text{type-}o\text{ discs of }\Sigma_{\mathbf{m}}\}} \cdot (-1)^L \cdot \Delta(\mathbf{m}) \cdot \prod_{x\in B} \mathrm{sign}\left(\sigma^{-1}_{m_{x,0}}\sigma_{m_{x,1}}\right).$$

We now show that $\Xi(w_1,\ldots,w_\ell;\mathbf{m})$ is independent of $\mathbf{m}$ and equal to $(-1)^\ell$. This will complete the proof by combining (A.5) with Theorem 3.4.

Our strategy for proving that $\Xi(w_1,\ldots,w_\ell;\mathbf{m}) \equiv (-1)^\ell$ consists of three parts:

1. The fact there are $r = |B|$ different types of letters in $w_1,\ldots,w_\ell$ can be ignored, and all letters may be considered as identical.
2. If $\mathcal{I} = \mathcal{I}^+$, there is one particular set of matchings $\mathbf{m}$ for which $\Xi(w_1,\ldots,w_\ell;\mathbf{m}) = (-1)^\ell$.
3. The value of $\Xi(w_1,\ldots,w_\ell;\mathbf{m})$ does not change if we make small local changes: (a) flipping the direction of one matching arc, (b) exchanging the termini of two matching arcs, or (c) flipping the orientation of one of the letters in the word from positive to negative or vice versa.

During this proof we consider the matchings $m_{x,i}$ of $\mathbf{m}$ as matchings of the letters of the words $w_1,\ldots,w_\ell$, this is possible since the letters are in one-to-one correspondence with the intervals $\mathcal{I}$.

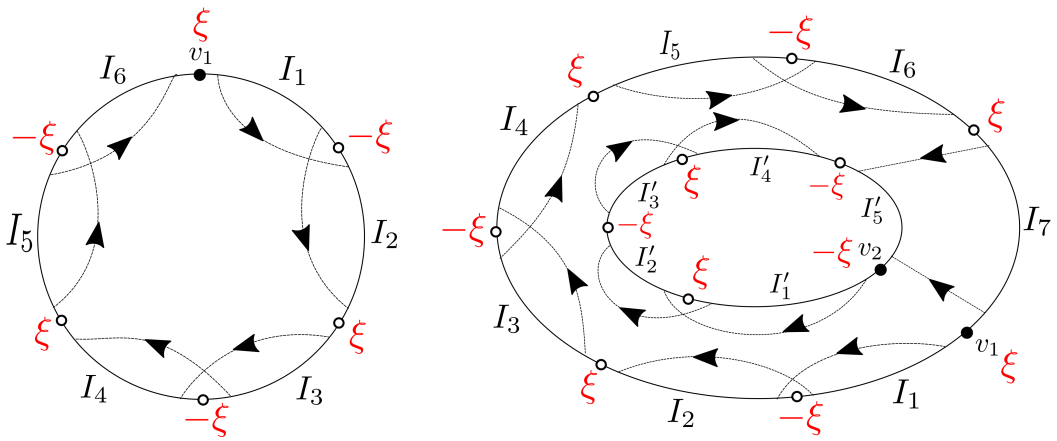

Figure A.1: On the left hand side, there is one word of even length (6 in this example) with all letters positive, and $m_0 = m_1$ match $I_1 \to I_2$, $I_3 \to I_4$ and $I_5 \to I_6$. An easy computation gives that $\Xi(w; m_0, m_1) = -1$ in this case. On the right hand side, there are two words of odd length each (7 and 5 in this example) with all letters positive. Here, $m_0 = m_1$ match $I_1 \to I_2$, $I_3 \to I_4$, $I_5 \to I_6$, $I_1' \to I_2'$, $I_3' \to I_4'$ and $I_7 \to I_5'$. An easy analysis gives that $\Xi(w_1, w_2; m_0, m_1) = 1$ in this case.

Part I: Consider all letters as identical First, recall the definition from §2.1.1 of the permutation $\sigma_m \in S_{2k}$ associated with the matching m belonging to $\mathbf{m}$, and note that the order of the pairs in m does not affect the sign of σ_m, nor $\Delta(\mathbf{m})$, so we ignore it here. (In contrast, the order within each pair does affect these quantities.) As a result, we can treat all matchings $\{m_{x,0}\}_{x\in B}$ as a single matching $m_0 \in M_L$ of the whole collection of intervals $\mathcal{I}$, where we keep track of the order within each pair, namely, of which endpoint is the origin and which the terminus of every matching arc. Similarly, we replace $\{m_{x,1}\}_{x\in B}$ with a single matching $m_1 \in M_L$ of $\mathcal{I}$. The corresponding permutations σ_{m_0} and σ_{m_1} lie in S_{2L}. From every pair of matchings $m_0, m_1 \in M_L$, we can construct a corresponding surface Σ_{m_0,m_1} as in §2.1. We define $\Delta(m_0, m_1)$ accordingly. It is thus enough to show that for every $m_0, m_1 \in M_L$,

$$\Xi(w_1, \ldots, w_\ell; m_0, m_1) \stackrel{\text{def}}{=} (-1)^{\#\{\text{type-}o\text{ discs of }\Sigma_{m_0,m_1}\}} \cdot (-1)^L \cdot \Delta(m_0, m_1) \cdot \operatorname{sign}\left(\sigma_{m_0}^{-1}\sigma_{m_1}\right) = (-1)^\ell. \quad \text{(A.6)}$$

Part II: Particular matchings m_0, m_1 Next, we show that the equality (A.6) holds for a particular pair $m_0, m_1 \in M_L$, when all letters in $w_1, \ldots, w_\ell$ are positive, namely, when $\mathcal{I} = \mathcal{I}^+$. The pair will satisfy $m_0 = m_1$, and so $\operatorname{sign}\left(\sigma_{m_0}^{-1}\sigma_{m_1}\right) = 1$. We partition the words $w_1, \ldots, w_\ell$ into singletons of even-length words and pairs of odd-length words. The matchings m_0 and m_1 will only pair letters of words in the same block of this partition. It is enough to prove (A.6) for every connected component of Σ_{m_0,m_1} separately.

First consider the case of a single, even-length word w (which, by abuse of notation, has length $2L$). Let each of m_0, m_1 pair the first interval to the second, the third to the fourth, and so forth. It is easy to check that in this case, there is exactly one type-o disc, with $2L$ non-orientable matching arcs at its boundary, all directed, say, clockwise. In every compatible assignment of indices $\mathbf{a} \vdash^* (m_0, m_1)$, the sign ξ flips along every matching arc, and as all letters are positive, exactly half of the matching arcs contribute (-1) (see Lemma A.3, Part 2), so $\Delta(m_0, m_1) = (-1)^L$ in this case. Hence the left hand side of (A.6) is $(-1)^1 \cdot (-1)^L \cdot (-1)^L \cdot 1 = (-1)$, which is the desired outcome as $\ell = 1$. See the left hand side of Figure A.1.

Second, consider the case of a pair of odd-length words w_1, w_2, of total length $2L$. Let each of m_0, m_1 pair the first interval of each word with the second one, the third with the fourth and so on, and pair the last interval of $C(w_1)$ with the last interval of $C(w_2)$. Again, it is easy to verify there is a single type-o disc in Σ_{m_0,m_1}, with $2L$ non-orientable matching arcs at its boundary. At the boundary of

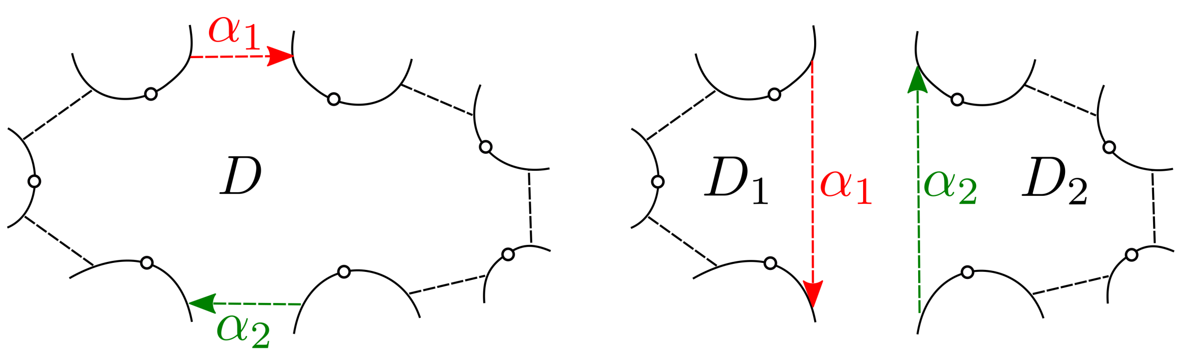

Figure A.2: The left part depicts a type-o disc D with two co-directed matching arcs α_1 and α_2 at its boundary. Switching their termini results in splitting D into two separate type-o discs: D_1 and D_2, as in the right hand side. This move flips both sign $\left(\sigma_{m_0}^{-1}\sigma_{m_1}\right)$ and $(-1)^{\#\{\text{type-}o\text{ discs of }\Sigma_{m_0,m_1}\}}$, but leaves $\Delta(m_0,m_1)$, and therefore also $\Xi(w_1,\ldots,w_\ell;m_0,m_1)$, unchanged.

the type-o disc there are $|w_1|$ successive o-points of $C(w_1)$, and then $|w_2|$ o-points of $C(w_2)$, where the matching arcs separating these two sequences are the two matchings arcs connecting the last interval of w_1 with the last interval of w_2. In every compatible assignment $\mathbf{a}\vdash^*(m_0,m_1)$, the signs ξ alternate, and so every pair of matching arcs connecting the same two intervals of the same word contributes (-1) to $\Delta(m_0,m_1)$. However, both matching arcs connecting the last intervals have the same sign at their origins, and so their contribution is 1. This shows that $\Delta(m_0,m_1)=(-1)^{L-1}$ in this case. Hence the left hand side of (A.6) is $(-1)^1\cdot(-1)^L\cdot(-1)^{L-1}\cdot 1=1$, which is the desired outcome as $\ell=2$. See the right hand side of Figure A.1.

Part III: Ξ is invariant under local modifications Finally, we show that the three local modifications we specified above do not alter the value of $\Xi(w_1,\ldots,w_\ell;m_0,m_1)$. As applying suitable steps of all three types leads from the instance described in part II of this proof to any given pair of matchings m_0,m_1 and to any orientation of the $2L$ letters (positive/negative), this will complete the proof. Note that none of these changes affect the total number of letters, L, so we ought to show that they do not alter the product $(-1)^{\#\{\text{type-}o\text{ discs of }\Sigma_{m_0,m_1}\}}\cdot\Delta(m_0,m_1)\cdot\text{sign}\left(\sigma_{m_0}^{-1}\sigma_{m_1}\right)$.

We begin with flipping the direction of one matching arc. Obviously, this does not change the number of type-o discs. It does change the sign of one of σ_{m_0} or σ_{m_1}, and therefore the sign of $\sigma_{m_0}^{-1}\sigma_{m_1}$, but it also changes the contribution of this matching arc to $\Delta(m_0,m_1)$: this follows from a simple case-by-case analysis of whether the origin of the matching arc is in $\mathcal{I}^+$ or in $\mathcal{I}^-$, and likewise the terminus of the arc. The analysis is based on Lemma A.3 and (A.4).

Next, consider a switch between the termini of two matching arcs α_1 and α_2 of, say, m_0. This switch changes the sign of σ_{m_0} and therefore of $\sigma_{m_0}^{-1}\sigma_{m_1}$. We distinguish between three cases and show that in each one of them, there is one more sign change that cancels with the change in sign$(\sigma_{m_0}^{-1}\sigma_{m_1})$:

- Assume that α_1 and α_2 both belong to the same type-o disc D and are directed along the same orientation of ∂D. Then switching the termini splits D into two discs, and so the sign of $(-1)^{\#\{\text{type-}o\text{ discs of }\Sigma_{m_0,m_1}\}}$ flips. Any compatible assignment $\mathbf{a}$ before the switch remains compatible after it, and the combined sign contribution of the two arcs (as in Lemma A.3, Part 2) remains unchanged. See Figure A.2

- Assume that α_1 and α_2 both belong to the same type-o disc D and are directed along different orientations of ∂D. Of the two components of $\partial D\backslash(\alpha_1\cup\alpha_2)$, one, denoted C_o, has the origins of α_1 and α_2 as endpoints, and the other, denoted C_t, has the two termini as endpoints. Switching the termini corresponds to reflecting C_t – see Figure A.3. By the definition of compatible assignments,

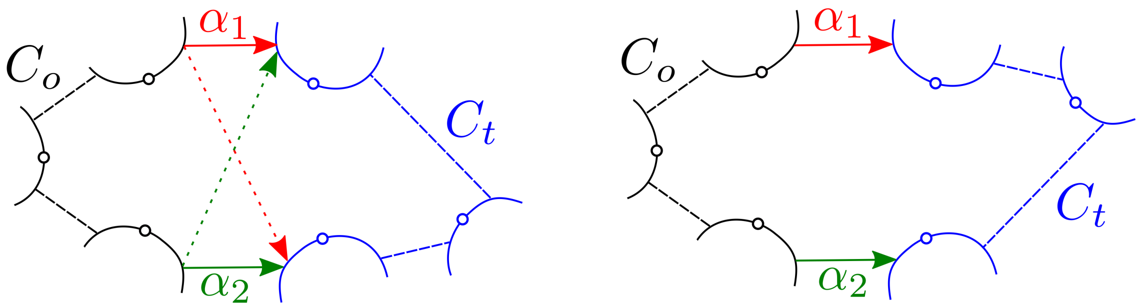

Figure A.3: The left part depicts a type-o disc D with two counter-directed matching arcs α_1 and α_2 at its boundary. The two connected components of $D\backslash(\alpha_1 \cup \alpha_2)$ are denoted C_o and C_t. Switching the termini of α_1 and α_2 results in reflecting C_t, as in the right hand side. This move flips both sign $\left(\sigma_{m_0}^{-1}\sigma_{m_1}\right)$ and $\Delta\left(m_0, m_1\right)$, but leaves $(-1)^{\#\{\text{type-}o \text{ discs of } \Sigma_{m_0,m_1}\}}$, and therefore also $\Xi\left(w_1, \ldots, w_\ell; m_0, m_1\right)$, unchanged.

every piece of $\partial D \cap \partial\Sigma$ is assigned a well-defined index in $[2n]$, and by Lemma A.3, Part 3, two different pieces of $\partial D \cap \partial\Sigma$ are assigned the same index if and only if the corresponding orientations induced by $\partial\Sigma$ induce, in turn, the same orientation on ∂D. This means that if we preserve the assignment along C_o, the signs along C_t must be flipped. The number of type-o discs is preserved. In the terminology of Lemma A.3, type-(iii) contributions to $\Delta\left(m_0, m_1\right)$ do not change. The sign contributions of the matching arcs along C_o do not change. Also, the contribution of orientable arcs along C_t does not change, nor does the type-(i) contribution of α_1 and α_2. However, $\Delta\left(m_0, m_1\right)$ does flip. To see this, denote by ∂_1, ∂_2 the two connected components of $\partial D \cap \partial\Sigma$ which contain (as endpoints) the two termini $t\left(\alpha_1\right)$ and $t\left(\alpha_2\right)$, respectively (∂_1 and ∂_2 may be equal). Let i_1 and i_2 be the indices corresponding to ∂_1 and ∂_2 in some compatible assignment (before the flip of α_1 and α_2).

If the orientation of $\partial\Sigma$ along ∂_1 and ∂_2 induces the same orientation on ∂D, then $i_1 = i_2$ and of the two intervals at the termini of α_1 and α_2, one is in $\mathcal{I}^+$ and the other in $\mathcal{I}^-$. Thus the total type-(ii) contribution of α_1 and α_2 flips. As C_t contains an even number of non-orientable matching arcs in this case, the total sign contribution of the non-orientable arcs along C_t is preserved (as in the proof of Lemma A.3, Part 5).
If the orientation of $\partial\Sigma$ along ∂_1 and ∂_2 induces different orientations on ∂D, then $i_2 = \widehat{i_1}$ and the two letters at the termini are both positive or both negative. In this case, the total type-(ii) contribution of α_1 and α_2 is unchanged, the total type-(i) and type-(ii) contribution of every orientable arc alongs C_t is unchanged, but the same contribution of every non-orientable arc along C_t is flipped, and the total number of non-orientable arcs along C_t is odd (by our assumption about ∂_1 and ∂_2).

- The third and last case is the one where α_1 and α_2 belong to different type-o discs. Switching their termini then leads to merging the two discs into one. In the united type-o disc, the two arcs are "co-oriented", so this case is the reverse of the first one, and $\Delta\left(m_0, m_1\right)$ remains unchanged.

The final small change we consider is that of flipping some letter from being positive to negative, namely, of flipping an interval in some $C\left(w_j\right)$ from $\mathcal{I}^\pm$ to $\mathcal{I}^\mp$. Here, sign $\left(\sigma_{m_0}^{-1}\sigma_{m_1}\right)$ is unchanged. By the first local modification in this part of the proof, we may assume without loss of generality that this letter is at the termini of two matching arcs, α_1 and α_2. A similar argument as in the previous paragraph would show that:

- Assume that α_1 and α_2 belong to the same type-o disc D with the same orientation. The flip of the letter then splits D into two type-o discs. Denote by ∂_1 and ∂_2 the pieces of $\partial D \cap \partial\Sigma$ at the termini of α_1 and α_2. They must be counter-oriented. We may preserve the same assignment of indices as before the flip of the letter, but then the type-(ii) contribution of both arcs flips when the letter is flipped. No other change in sign contributions occurs.

- Assume that α_1 and α_2 belong to the same type-o disc D with opposite orientations. The flip of the letter preserves the number of type-o discs and corresponds to reflecting C_t. Here ∂_1 and ∂_2 are co-oriented and the signs along C_t must be flipped. There is no change to $\Delta(m_0, m_1)$: the total type-(ii) contribution of α_1 and α_2 is 1 before and after the flip, and the number of non-orientable arcs along C_t is even.

- If α_1 and α_2 belong to different type-o disc, the flip is the reverse of the first case.

This completes the proof of Theorem 1.2. □

We now have the analog of Corollary 3.5 for $G = \mathrm{Sp}$:

Corollary A.6. *There is a rational function* $\overline{\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}} \in \mathbf{Q}(n)$ *such that for* $2n \geq \max\{L_x : x \in B\}$, $\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}(n)$ *is given by evaluating* $\overline{\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}}$ *at* n.

Proof. Theorem 1.2 shows that we can obtain $\overline{\mathcal{T}r^{\mathrm{Sp}}_{w_1,\ldots,w_\ell}}$ by switching n with $-2n$ in $\overline{\mathcal{T}r^{\mathrm{O}}_{w_1,\ldots,w_\ell}}$ (the rational function from Corollary 3.5) and multiplying by $(-1)^\ell$. □

References

[Bro22] Y. Brodsky. Word measures on unitary groups: Improved bounds for small representations. preprint arxiv:2208.11957, 2022. 6

[CG86] J. Cheeger and M. Gromov. L_2-cohomology and group cohomology. *Topology*, 25(2):189–215, 1986. 29

[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. 11, 18, 19

[CS08] B. Collins and M. Stolz. Borel theorems for random matrices from the classical compact symmetric spaces. *Ann. Probab.*, 36(3):876–895, 2008. 18

[Cul81] M. Culler. Using surfaces to solve equations in free groups. *Topology*, 20(2):133–145, 1981. 12, 15, 16, 25

[Del16] P. Deligne. Private communication, 2016. 9

[DP21] G. Dubach and Y. Peled. On words of non-Hermitian random matrices. *Ann. Probab.*, 49(4):1886–1916, 2021. 8

[DS94] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31(A):49–62, 1994. 8, 35

[Dys62] F. J. Dyson. Statistical theory of the energy levels of complex systems. I. *J. Math. Phys.*, 3(1):140–156, 1962. 9

[Eps66] D. B. A. Epstein. Curves on 2-manifolds and isotopies. *Acta Math.*, 115:83–107, 1966. 30, 33, 34

[EWPS21] D. Ernst-West, D. Puder, and M. Seidel. Word measures on $GL_N(q)$ and free group algebras. preprint arxiv:2110.11099, 2021. 8

[FM12] B. Farb and D. Margalit. *A primer on mapping class groups*, volume 49 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2012. 30

[For06] P. J. Forrester. Quantum conductance problems and the Jacobi ensemble. *J. Phys. A*, 39(22):6861–6870, 2006. 9

[FS06] G. Frobenius and I. Schur. Über die reellen darstellungen der endlichen gruppen. *Sitzungsberichte Akademie der Wissenschaften zu Berlin*, pages 186–208, 1906. 7

[GHJ01] I. P. Goulden, J. L. Harer, and D. M. Jackson. A geometric parametrization for the virtual Euler characteristics of the moduli spaces of real and complex algebraic curves. *Trans. Amer. Math. Soc.*, 353(11):4405–4427, 2001. 6

[GJ97] I. P. Goulden and D. M. Jackson. Maps in locally orientable surfaces and integrals over real symmetric matrices. *Canad. J. Math.*, 49(5):865–882, 1997. 6

[Hal15] B. Hall. *Lie groups, Lie algebras, and representations.* Springer, 2015. 2

[HP22] L. Hanany and D. Puder. Word measures on symmetric groups. *Int. Math. Res. Not. IMRN*, 2022. appeared online, doi.org/10.1093/imrn/rnac084. 8

[HZ86] J. L. Harer and D. Zagier. The Euler characteristic of the moduli space of curves. *Invent. Math.*, 85(3):457–485, 1986. 6

[LP10] N. Linial and D. Puder. Word maps and spectra of random graph lifts. *Random Structures and Algorithms*, 37(1):100–135, 2010. 8

[Lüc02] W. Lück. *L^2-invariants: theory and applications to geometry and K-theory*, volume 44 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics.* Springer-Verlag, Berlin, 2002. 29

[Mat13] S. Matsumoto. Weingarten calculus for matrix ensembles associated with compact symmetric spaces. *Random Matrices Theory Appl.*, 2(2):26 pages, 2013. 18

[MKR81] The equivalence of $\mathrm{Sp}(2n)$ and $\mathrm{SO}(-2n)$ gauge theories. *Phys. Lett. B*, 105(2):174–176, 1981. 9

[MP13] J. A. Mingo and M. Popa. Real second order freeness and Haar orthogonal matrices. *J. Math. Phys.*, 54(5):35 pages, 2013. 6, 8

[MP16] M. Magee and D. Puder. Word measures on unitary groups. arXiv preprint 1509.07374 v2, 2016. 3

[MP19] M. Magee and D. Puder. Matrix group integrals, surfaces, and mapping class groups I: $U(n)$. *Invent. Math.*, 218(2):341–411, 2019. 3, 5, 7, 9, 10, 12, 21, 24, 25, 26, 27, 28, 29, 30

[MP21] M. Magee and D. Puder. Surface words are determined by word measures on groups. *Israel J. Math.*, 241(2):749–774, 2021. 7

[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. 8

[MV11] R. L. Mkrtchyan and A. P. Veselov. On duality and negative dimensions in the theory of Lie groups and symmetric spaces. *J. Math. Phys.*, 52(8):10 pages, 2011. 9

[MW03] M. Mulase and A. Waldron. Duality of orthogonal and symplectic matrix integrals and quaternionic Feynman graphs. *Comm. Math. Phys.*, 240(3):553–586, 2003. 9

[Nic94] A. Nica. On the number of cycles of given length of a free word in several random permutations. *Random Structures Algorithms*, 5(5):703–730, 1994. 8

[Nov17] M. Novaes. Expansion of polynomial lie group integrals in terms of certain maps on surfaces, and factorizations of permutations. *J. Phys. A Math. Theor.*, 50(7):22 pages, 2017. 3, 9

[PP15] D. Puder and O. Parzanchevski. Measure preserving words are primitive. *J. Amer. Math. Soc.*, 28(1):63–97, 2015. 7

[PS14] O. Parzanchevski and G. Schul. On the Fourier expansion of word maps. *Bull. Lond. Math. Soc.*, 46(1):91–102, 2014. 7

[Răd06] F. Rădulescu. Combinatorial aspects of Connes's embedding conjecture and asymptotic distribution of traces of products of unitaries. In *Proc. Oper. Alg. Conf., Bucharest*. Theta Foundation, 2006. 8

[Red15] C. Redelmeier. Topological expansion for Haar-distributed orthogonal matrices and second-order freeness of orthogonally invariant ensembles. preprint, arXiv:1511.01087, 2015. 6

[Stu06] M. Stukow. Dehn twists on nonorientable surfaces. *Fund. Math.*, 189(2):117–147, 2006. 30

[tD72] T. tom Dieck. Orbittypen und äquivariante Homologie. I. *Arch. Math. (Basel)*, 23:307–317, 1972. 29

[tD87] T. tom Dieck. *Transformation groups*, volume 8 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1987. 29

[von29] J. von Neumann. Zur allgemeinen Theorie des Maßes. *Fundam. Math.*, 13:73–116, 1929. 30

[Wey39] H. Weyl. *The Classical Groups. Their Invariants and Representations*. Princeton University Press, Princeton, N.J., 1939. 9

Michael Magee, Department of Mathematical Sciences, Durham University, Lower Mountjoy, DH1 3LE Durham, United Kingdom
`michael.r.magee@durham.ac.uk`

Doron Puder, School of Mathematical Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel
`doronpuder@gmail.com`