THE MASTER FIELD AND FREE BROWNIAN MOTIONS

NATHAN XIONG AND PU YU

ABSTRACT. The master field on the plane is the large N limit of the Wilson loop functionals from the two-dimensional Yang–Mills holonomy process. In this paper, we redefine the master field purely through free Brownian motions, so that its definition is independent from finite N Yang–Mills theory. From this aspect, we prove that the master field does not depend on the lasso basis chosen on a graph. We also give a new, elementary proof for the Makeenko–Migdal equations, which allow us to efficiently calculate the master field of any loop via a system of differential equations. While previous work in this field is mostly differential geometric in nature, our proofs all use combinatorial techniques, heavily utilizing the moment-cumulant relation from free probability.

1. INTRODUCTION

1.1. **Background.** Quantum gauge theories are used to describe the fundamental interactions between elementary particles, with the particular interaction characterized by the structure group. In his 1974 seminal paper, 't Hooft [tH74] investigated quantum gauge theories with structure group $\mathrm{U}(N)$ when $N \to \infty$. He noticed that the $1/N$ expansion simplifies considerably in the large N limit. Following this discovery, many physicists began to study the large N limit of gauge theories (see [DK93, Dou95, KK80, Pol80, Kaz81]), and the concept of the *master field*, a deterministic large N limit for gauge theories, arose.

The idea of a mathematically rigorous master field was first introduced by Singer [Sin95] on the Euclidean plane. The first formal definitions of the master field on the plane were later given by Anshelevitch and Sengupta [AS12] as well as by Lévy [Lév17]. As conjectured by Singer [Sin95], the master field is naturally described using the framework of free probability.

The large N limit of two-dimensional Yang–Mills theory also has connections to string theory (see [GT93a, GT93b, GM95]) but not through the master field.

1.2. **Defining the master field on the plane.** Although rigorously making mathematical sense of Yang–Mills theory remains a major open problem, in two-dimensional spacetime, thanks to a series of works [Dri89, Fin91, Wit91, Sen97, Lév03, Lév10], we now have a relatively clear understanding through stochastic calculus. In this section, we briefly recall the basics of 2D Yang–Mills theory, which leads to the *master field* on the plane. See [Lév20, Section 1] for more details.

1.2.1. *The Yang–Mills measure.* A Yang–Mills theory is specified by a compact surface Σ, a connected compact Lie group G, called the *structure group*, and a principal G-bundle $\pi\colon P \to \Sigma$. Further, we require that Σ has a volume form and that the Lie algebra $\mathfrak{g}$ of G is endowed with an invariant scalar product. For this paper, we will take $\Sigma = \mathbb{R}^2$, which is not technically a compact surface, but we can think of it as the limit of a sequence of disks with increasing radius. We also take $G = \mathrm{U}(N)$, so that $\mathfrak{g} = \mathfrak{u}(N)$ with scalar product

$$\langle X, Y\rangle = N\mathrm{Tr}(X^*Y).$$

The mathematical approach to the Yang–Mills measure constructs a probability measure on the image of $\mathcal{A}(P)$, the space of connections on P, under the *holonomy* mapping. Let $\mathsf{L}_o(\Sigma)$ denote the set of rectifiable loops on Σ based at some origin o. Any connection ω induces a holonomy h, which is a *multiplicative* map from $\mathsf{L}_o(\Sigma)$ to G, in the sense that

- For any $l \in \mathsf{L}_o(\Sigma)$, if l^{-1} denotes the same loop traversed backwards, then $h(l^{-1}) = h(l)^{-1}$.
- For any $l_1, l_2 \in \mathsf{L}_o(\Sigma)$, if $l_1 l_2$ denotes the concatenated loop, then $h(l_1 l_2) = h(l_2)h(l_1)$.

In general, we consider the space $\mathcal{M}(\mathsf{L}_o(\Sigma), G)$ of multiplicative maps from $\mathsf{L}_o(\Sigma)$ to G. Then, recalling that $\mathrm{Aut}(P)$, the group of bundle automorphisms, is the gauge group, it turns out that the following holonomy mapping is injective

$$\mathrm{hol}\colon \mathcal{A}(P)/\mathrm{Aut}(P) \to \mathcal{M}(\mathsf{L}_o(\Sigma), G)/G,$$

where G acts on $\mathcal{M}(\mathsf{L}_o(\Sigma), G)$ by conjugation. Thus, we can equivalently define the Yang–Mills measure on $\mathcal{M}(\mathsf{L}_o(\Sigma), G)$ and no information will be lost. In doing so, the Yang–Mills measure becomes a collection $(H_l)_{l\in \mathsf{L}_o(\Sigma)}$ of G-valued random variables indexed by loops in $\mathsf{L}_o(\Sigma)$.

1.2.2. *Yang–Mills theory on planar graph.* The traditional approach to constructing a Yang–Mills theory first defines a lattice Yang–Mills theory and then takes the limit to a continuous setting. The two-dimensional lattice Yang–Mills theory is defined on the *configuration space* of a graph. Let $\mathsf{P}(\mathbb{R}^2)$ be the set of Lipschitz continuous maps $c\colon [0,1] \to \mathbb{R}^2$ up to reparametrization and $\mathbb{G} = (\mathbb{V}, \mathbb{E}, \mathbb{F})$ be a *planar graph* whose edge set $\mathbb{E} \subset \mathsf{P}(\mathbb{R}^2)$. Then exactly one of the faces in $\mathbb{F}$ is unbounded, and we denote this face by F_∞. Let $\mathbb{F}^b = \mathbb{F} \setminus \{F_\infty\}$ be the set of bounded faces. For any $F \in \mathbb{F}^b$, write $|F|$ for the area of F. Also, let ∂F be the loop formed by starting from an arbitrary vertex adjacent to F and going clockwise once along the boundary of F. This loop is not well defined because it does not have a base point, yet it turns out that this does not matter by the conjugate invariance of the heat kernel (see below). Let $\mathsf{P}(\mathbb{G})$ be the set of paths on $\mathbb{G}$ formed by concatenating edges of $\mathbb{G}$, and let $\mathsf{L}_o(\mathbb{G})$ be the subset of $\mathsf{P}(\mathbb{G})$ consisting of loops on $\mathbb{G}$ based at o.

The invariant scalar product on $\mathfrak{g}$ determines a Laplace-Beltrami operator Δ. Then, the *heat kernel* on G is the unique function $p : \mathbb{R}_+^* \times G \to \mathbb{R}_+^*$ satisfying the heat equation $(\partial_t - \frac{1}{2}\Delta)p = 0$ such that $p_t(g)\,dg$ converges weakly to the Dirac measure at the identity of G as $t \to 0$, where dg is the Haar measure on G. By definition (see [Itô50]), $p_t(g)\,dg$ gives the distribution of the Brownian motion on G at time t.

A useful fact about the p_t is that they are conjugate invariant. That is, for all $x, y \in G$, we have $p_t(x) = p_t(yxy^{-1})$. This invariance means that the quantity $p_t(h(\partial F))$ is well defined for all $h \in \mathcal{M}(\mathsf{P}(\mathbb{G}), G)$. So, we can define the Yang–Mills measure on $\mathcal{M}(\mathsf{P}(\mathbb{G}), G)$ restricted to $\mathbb{G}$ by the following *Driver–Sengupta formula*,

$$\mu_{\mathsf{YM}}^{\mathbb{G}}(dh) = \prod_{F\in\mathbb{F}^b} p_{|F|}(h(\partial F))\,dh \tag{1}$$

where dh is the normalized Haar measure on $\mathcal{M}(\mathsf{P}(\mathbb{G}), G)$.

1.2.3. *The Yang–Mills holonomy process.* The most useful property of the discrete Yang–Mills measure is that it is invariant under refinement [Lév03, Section 1.6]. That is, if $\mathbb{G}_1$ and $\mathbb{G}_2$ are graphs such that $\mathbb{G}_2$ is *finer* than $\mathbb{G}_1$ (every path on $\mathbb{G}_1$ is also a path on $\mathbb{G}_2$), then under the natural restriction mapping $\mathcal{M}(\mathsf{P}(\mathbb{G}_2), G) \to \mathcal{M}(\mathsf{P}(\mathbb{G}_1), G)$, the pushforward of the measure $\mu_{\mathsf{YM}}^{\mathbb{G}_2}$ is the measure $\mu_{\mathsf{YM}}^{\mathbb{G}_1}$. This result, along with the following continuity condition, is enough for us to just take the limit of these graphs. If c_1 and c_2 are paths, define the *length metric* to be

$$d(c_1, c_2) = |\ell(c_1) - \ell(c_2)| + \inf_{\varphi_1,\varphi_2} \sup_{t\in[0,1]} \{|c_1(\varphi_1(t)) - c_2(\varphi_2(t))|\},$$

where $\ell(c)$ is the length of path c and the infimum is taken over all reparametrizations φ_1, φ_2 of c_1, c_2 by $[0, 1]$. Under the length metric, $\mathsf{P}(\Sigma)$ becomes a metric space. Now, we can define the G-valued random variables $(H_l)_{l\in\mathsf{L}_o(\mathbb{R}^2)}$ hinted to in Section 1.2.1.

Proposition 1 (The Yang–Mills holonomy process)**.** *The Yang–Mills holonomy process is the unique collection $(H_l)_{l\in\mathsf{L}_o(\mathbb{R}^2)}$ of G-valued random variables satisfying the following conditions.*

(a) For any $l_1, l_2 \in \mathsf{L}_o(\mathbb{R}^2)$, the equalities $H_{l_1^{-1}} = H_{l_1}^{-1}$ and $H_{l_1 l_2} = H_{l_2} H_{l_1}$ hold almost surely.

(b) The random variables are stochastically continuous in the space of loops, i.e., if $(l_n)_{n\geq 0}$ is a sequence of loops converging to l, then the sequence $(H_{l_n})_{n\geq 0}$ converges in probability to H_l.

(c) The finite-dimensional distributions are described by (1).

1.2.4. *The master field.* After constructing the Yang–Mills holonomy process with structure group $\mathrm{U}(N)$, $\mathrm{SO}(N)$, or $\mathrm{Sp}(N)$, the master field is traditionally defined to be the large N limit of the Wilson loop functionals. Remarkably, this limit does not depend on the choice of the structure group.

Proposition 2 (See also Chapter 6 of [Lév17])**.** *For each $N \geq 1$, let $(H_{N,l})_{l\in\mathsf{L}_o(\mathbb{R}^2)}$ be the Yang–Mills holonomy process on $\mathbb{R}^2$ with structure group $\mathrm{U}(N)$, $\mathrm{SO}(N)$ or $\mathrm{Sp}(N)$. Then for every $l \in \mathsf{L}_o(\mathbb{R}^2)$, the convergence*

$$\frac{1}{N}\mathrm{Tr}(H_{N,l}) \xrightarrow{N\to\infty} \Phi(l)$$

holds in probability toward a deterministic limit called the master field.

1.3. **Defining the master field through free Brownian motions.**

1.3.1. *Free probability and free Brownian motions.* A good introductory survey on free probability is given by Mingo and Speicher [MS17]. Our main tool from free probability will be the *moment-cumulant relation.*

Proposition 3 (Moment-cumulant relation; see also Section 3.2 of [Spe14])**.** *Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space. The* cumulants $\kappa_n \colon \mathcal{A}^n \to \mathbb{C}$ *are multilinear functionals obeying the relation*

$$\varphi(a_1 \cdots a_n) = \sum_{\pi\in NC(n)} \kappa_\pi(a_1, \ldots, a_n)$$

for all $a_1, \ldots, a_n \in \mathcal{A}$. The set $NC(n)$ consists of the noncrossing partitions on $\{1, \ldots, n\}$, and κ_π represents the product of the cumulants specified by the blocks in π.

Cumulants have the very useful property that mixed cumulants of free random variables vanish.

Proposition 4 (See also Section 3.3 of [Spe14]). *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. Then $x, y \in \mathcal{A}$ are free if and only if $\kappa_n(a_1, \dots, a_n) = 0$ whenever $n \geq 2$, all a_i are either x or y, and $a_i \neq a_j$ for some i, j.*

Finally, we define free Brownian motions. Biane [Bia97] proved that the free Brownian motion is actually the large N limit of the Brownian motion on $\mathrm{U}(N)$ (as defined in [Itô50]), where the convergence refers to convergence in non-commutative distribution, i.e., all moments converge.

Theorem 5 ([Bia97]). *The free (multiplicative) Brownian motion is a collection $(u_t)_{t\geq 0}$ of unitary random variables in a non-commutative probability space $(\mathcal{A}, \varphi)$ satisfying the following three properties.*

- *For all $0 \leq s < t$, the random variable $u_t u_s^*$ has the same distribution as u_{t-s}.*
- *For all $0 \leq t_1 < \cdots < t_n$, the random variables $u_{t_1}, u_{t_2} u_{t_1}^*, \dots, u_{t_n} u_{t_{n-1}}^*$ are mutually free.*
- *For all $n \in \mathbb{Z}$, we have*

$$\varphi(u_t^n) = e^{-\frac{nt}{2}} \sum_{k=0}^{n-1} \frac{(-t)^k}{k!} n^{k-1} \binom{n}{k+1}.$$

1.3.2. *The group of loops in a graph.* Say that two loops in $\mathsf{L}_o(\mathbb{G})$ are equivalent if one can be obtained from the other through a finite sequence of insertions and deletions of sub-loops of the form ee^{-1}, where e is an edge of $\mathbb{G}$. After we quotient out $\mathsf{L}_o(\mathbb{G})$ by this equivalence relation, we obtain $\mathsf{RL}_o(\mathbb{G})$, the space of *reduced loops* on $\mathbb{G}$. Lévy [Lév17] proved that $\mathsf{RL}_o(\mathbb{G})$ is a free group with rank equal to the number of bounded faces in $\mathbb{G}$. One such basis for $\mathsf{RL}_o(\mathbb{G})$ is the *lasso basis*, whose definition we recall below.

We follow the definition given by Lévy [Lév17] for the dual graph $\mathbb{G}^* = (\mathbb{V}^*, \mathbb{E}^*, s, t)$ of a planar graph $\mathbb{G}$. Pick a spanning tree T on $\mathbb{G}$. Let T^* be the dual tree defined on $\mathbb{G}^*$. For two vertices v_1, v_2, let $[v_1, v_2]_{\mathsf{T}}$ be the shortest path along T from v_1 to v_2. Then for each bounded face $F \in \mathbb{F}^b$, consider the first dual edge $e^* \in \mathbb{E}^*$ along the path $[F, F_\infty]_{\mathsf{T}^*}$. Let $e \in \mathbb{E}$ be the edge corresponding to e^*, and let $\underline{e}$ denote the starting vertex of e. Finally, let $\partial_e F$ be the loop formed by tracing the boundary of F once, starting with e. Define the *lasso* for F to be the reduced loop

$$\lambda_F = [o, \underline{e}]_{\mathsf{T}} \partial_e F [\underline{e}, o]_{\mathsf{T}}.$$

The lasso basis is then the set $\{\lambda_F : F \in \mathbb{F}^b\}$.

1.3.3. *The master field on the plane.* In this paper, we define the master field on the plane purely through free Brownian motions, such that the definition is independent from finite N Yang–Mills theory.

Proposition 6 (The master field; see also Chapter 0 of [Lév17]). *The master field is a collection $(h_l)_{l\in\mathsf{L}_o(\mathbb{R}^2)}$ of random variables in a non-commutative probability space $(\mathcal{A}, \tau)$, where $\mathcal{A}$ is involutive and τ is tracial. The distribution of the $(h_l)_{l\in\mathsf{L}_o(\mathbb{R}^2)}$ is uniquely determined by the following three properties.*

(1) The equalities $h_{l_1^{-1}} = h_{l_1}^* = h_{l_1}^{-1}$ and $h_{l_1 l_2} = h_{l_2} h_{l_1}$ hold for any two loops $l_1, l_2 \in \mathsf{L}_o(\mathbb{R}^2)$.

(2) The $(h_l)_{l\in\mathsf{L}_o(\mathbb{R}^2)}$ are continuous in the space of loops, i.e., if $(l_n)_{n\geq 0}$ is a sequence of loops converging to a loop l, then the sequence $(h_{l_n})_{n\geq 0}$ converges in distribution to h_l.

(3) *Let $\mathbb{G}$ be a planar graph in $\mathbb{R}^2$ containing the vertex o. For any lasso basis $\{\lambda_F : F \in \mathbb{F}^b\}$ on $\mathbb{G}$, the random variables $(h_{\lambda_F})_{F\in\mathbb{F}^b}$ are mutually free, and for each F, the distribution of h_{λ_F} is given by a free Brownian motion stopped at time $|F|$, where $|F|$ denotes the Euclidean area of F.*

Following Lévy [Lév17], we think of the master field as the function $\Phi\colon \mathsf{L}_o(\mathbb{R}^2) \to \mathbb{C}$ with

$$\Phi(l) = \tau(h_l).$$

To calculate $\tau(h_l)$ for an arbitrary loop in $\mathsf{L}_o(\mathbb{R}^2)$, take any graph on which l lives and pick some lasso basis, and then write l as a product of lassos λ_F. The first property of the master field lets us rewrite h_l as a product of h_{λ_F} terms, each of which is understood explicitly by the third property.

1.4. **The Makeenko–Migdal equations.** The Makeenko–Migdal equations give an efficient way to actually compute the master field $\Phi(l)$ for some loop l. The initial proof of these equations given by Makeenko and Migdal [MM79] relies on integration by parts with respect to an ill-defined integral. Rigorous proofs were later given by Dahlqvist [Dah16], Lévy [Lév17], and Driver, Hall, and Kemp [DHK17]. All of these proofs first show an analogue of the Makeenko–Migdal equations in the finite N Yang–Mills theory and then take $N \to \infty$.

The main idea of the equations is to treat $\Phi(l)$ as a function of the areas of the bounded faces delimited by l. The theorem then gives us a system of first-order differential equations which we can solve to find $\Phi(l)$.

Theorem 7 (Makeenko–Migdal [MM79])**.** *Let l be a loop in $\mathsf{L}_o(\mathbb{R}^2)$. Fix a point of self-intersection of l which has exactly two ingoing strands and two outgoing strands. Let l_1 and l_2 be the two loops formed by swapping which outgoing strand connects to each ingoing strand. Label the four surrounding faces F_1, F_2, F_3, F_4 cyclically so that F_1 is between two outgoing strands (see Figure 1).*

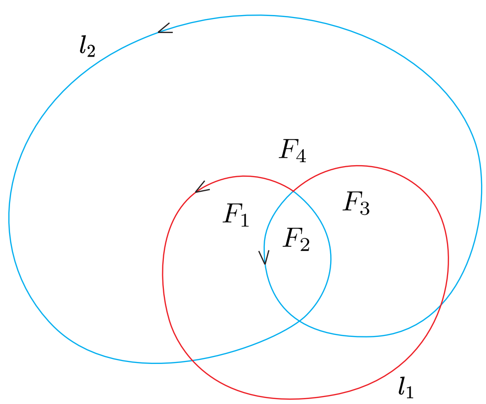

FIGURE 1. Setup for the Makeenko–Migdal equations

Then, $\Phi(l)$ satisfies the Makeenko–Migdal equation

$$\left(\frac{d}{d|F_1|} - \frac{d}{d|F_2|} + \frac{d}{d|F_3|} - \frac{d}{d|F_4|}\right)\Phi(l) = \Phi(l_1)\Phi(l_2). \tag{2}$$

If any of $F_1, \ldots, F_4$ is the unbounded face, we just replace its corresponding term on the left by 0.

1.5. **Main Results.** In this paper, we treat the master field as an object in its own right by defining it purely through free Brownian motions as in Section 1.3, so that its definition is independent from finite N Yang–Mills theory. Under this new definition, we prove in Section 2 that the master field on a graph does not depend on the lasso basis chosen for that graph.

Theorem 8. *Define the master field on the plane as in Proposition 6. Let l be a loop in $\mathsf{L}_o(\mathbb{R}^2)$ and $\mathbb{G}$ be a graph induced by l. The value of $\Phi(l)$ does not depend on the lasso basis chosen on $\mathbb{G}$.*

We also present in Section 3 a new proof for the Makeenko–Migdal equations on the master field. While previous proofs of these equations first prove a finite N analogue using tools from differential geometry and then take $N \to \infty$, we directly attack the $N = \infty$ case. As a result, our proof is more elementary and combinatorial in nature.

2. The master field does not depend on the lasso basis

2.1. **Setup and definitions.** Let l be a loop in $\mathsf{L}_o(\mathbb{R}^2)$ and $\mathbb{G} = (\mathbb{V}, \mathbb{E}, \mathbb{F})$ be a graph induced by l. We assume that $\mathbb{G}$ has finitely many faces. If $\mathbb{G}$ has infinitely many faces, then since l is rectifiable, for any $\varepsilon > 0$, only finitely many faces have area greater than ε. We can then use the continuity of Φ to evaluate $\Phi(l)$ by passing to a limit.

Let T and $\widetilde{\mathsf{T}}$ be two distinct spanning trees on $\mathbb{G}$, and suppose that they define lasso bases λ and $\widetilde{\lambda}$ respectively. For any spanning tree T and a fixed orientation $\mathbb{E}^+ \subset \mathbb{E}$, define the *beta basis* to be the set $\{\beta_e : e \in \mathbb{E}^+ \setminus \mathsf{T}\}$ consisting of the reduced loops

$$\beta_e = [o, \underline{e}]_{\mathsf{T}} e [\overline{e}, o]_{\mathsf{T}},$$

where $\underline{e}$ (resp. $\overline{e}$) denotes the starting (resp. ending) vertex of e. It is easy to see that the beta basis is always a basis for $\mathsf{RL}_o(\mathbb{G})$. Furthermore, if $\psi(e_1, \ldots, e_n)$ is a word in the edges $e_1, \ldots, e_n$ which forms a loop, then the loops $\psi(e_1, \ldots, e_n)$ and $\psi(\beta_{e_1}, \ldots, \beta_{e_n})$ are equal in $\mathsf{RL}_o(\mathbb{G})$. That is, we can always freely replace e by β_e. Let β and $\widetilde{\beta}$ be the beta bases under T and $\widetilde{\mathsf{T}}$ respectively.

Finally, let ψ_1 and ψ_2 be the words in the lasso bases λ and $\widetilde{\lambda}$ respectively corresponding to the loop l. Our goal is to prove $\Phi(\psi_1) = \Phi(\psi_2)$. Here and throughout this section, we treat Φ not as a function of loops but as the state of a non-commutative probability space generated by the mutually free Brownian motions specified from a lasso basis on $\mathbb{G}$.

2.2. **Proof of Theorem 8.** We use strong induction on the number of bounded faces in $\mathbb{G}$. The base case is $|\mathbb{F}^b| = 1$. In this case, the graph $\mathbb{G}$ must be a cycle (it could have extra edges sticking out from the cycle, but these do not affect anything because they do not border a bounded face). By the symmetry of the cycle graph, any lasso basis will yield the same value for $\Phi(l)$.

Henceforth, assume that Theorem 8 holds for all graphs with less than N bounded faces. We prove that all graphs $\mathbb{G}$ with N bounded faces must satisfy the conclusion of Theorem 8. There are two cases to consider depending on the structure of T and $\widetilde{\mathsf{T}}$.

2.2.1. *Case 1.* There exists some edge $e \in \mathbb{E}^+$ which borders the unbounded face and $e \notin \mathsf{T}, \widetilde{\mathsf{T}}$ (see Figure 2).

Recall that ψ_1 and ψ_2 are words in the lasso bases λ and $\widetilde{\lambda}$ respectively for the loop l. By assumption, the face A is adjacent to the unbounded face in the dual graph through the dual edge e^*. Then according to Lévy's [Lév17] explicit change of basis formula between the lasso and beta

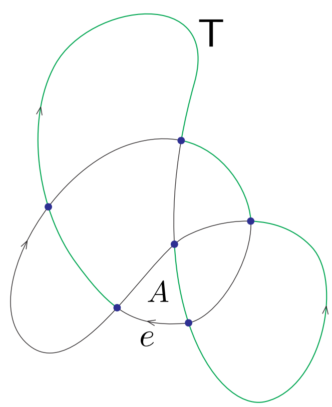

FIGURE 2. The edge e belongs to neither T nor $\widetilde{\mathsf{T}}$.

bases, the λ_A (and $\widetilde{\lambda}_A$) terms appear in ψ_1 (and ψ_2) only when we traverse e. In particular, if we compare ψ_1 and ψ_2, the λ_A and $\widetilde{\lambda}_A$ terms occur in the "same" positions because the positions of e in the loop l depend only on l itself. Expand $\Phi(\psi_1)$ into a sum of cumulants using the moment-cumulant relation. By Proposition 4, all nonzero cumulants will only involve one lasso (and its inverse). We condition our sum on blocks of λ_A and λ_A^{-1} terms which include some fixed λ_A (or λ_A^{-1}).

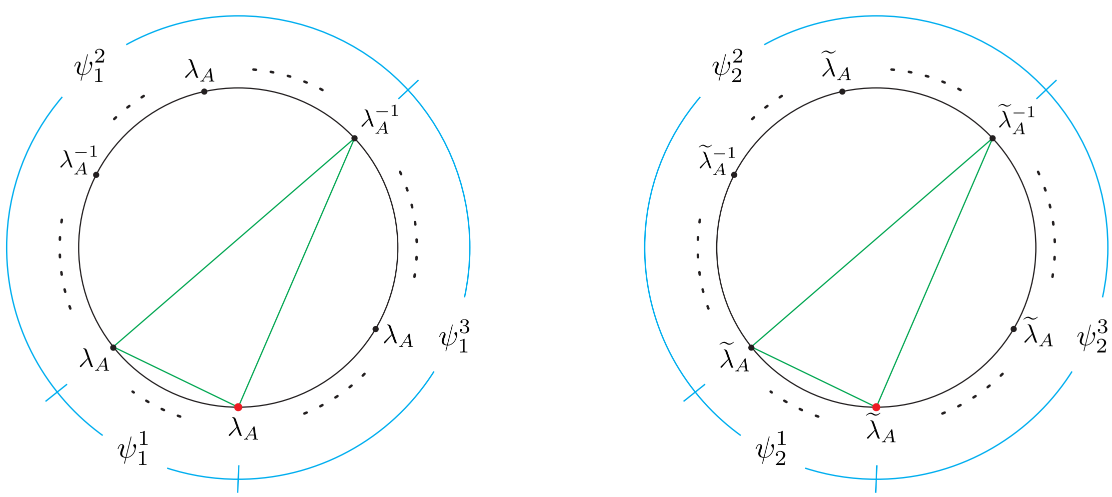

FIGURE 3. Every block of λ_A and λ_A^{-1} terms divides the word ψ_1 into subwords $\psi_1^1, \psi_1^2, \ldots$. The word ψ_2 follows a similar division into subwords $\psi_2^1, \psi_2^2, \ldots$.

This block of λ_A and λ_A^{-1} terms divides the word ψ_1 into distinct subwords, such that no other block of the noncrossing partition can contain terms from different subwords without breaking the "noncrossing" condition (see Figure 3). So when we sum over all noncrossing partitions containing this particular block, we are actually summing over all noncrossing partitions on the different subwords of ψ_1. If we then apply the moment-cumulant relation in reverse on each subword, the cumulants for each subword combine into the moment for that subword. That is, if the block B splits ψ_1 into the subwords $\psi_1^1, \ldots, \psi_1^{|B|}$, we have

$$\Phi(\psi_1) = \sum_{B \in S_1(A)} \kappa_{|B|}(h_{\lambda_A}^{\varepsilon_1}, \ldots, h_{\lambda_A}^{\varepsilon_{|B|}}) \prod_{i=1}^{|B|} \Phi(\psi_1^i),$$

where $S_1(A)$ consists of the blocks of λ_A and λ_A^{-1} terms which include some fixed λ_A (or λ_A^{-1}) and $\varepsilon_1, \ldots, \varepsilon_{|B|} \in \{1, *\}$ are chosen appropriately. Using the same idea and analogous notation for the lasso basis $\widetilde{\lambda}$,

$$\Phi(\psi_2) = \sum_{B \in S_2(A)} \kappa_{|B|}(h_{\widetilde{\lambda}_A}^{\varepsilon_1}, \ldots, h_{\widetilde{\lambda}_A}^{\varepsilon_{|B|}}) \prod_{i=1}^{|B|} \Phi(\psi_2^i),$$

where $S_2(A)$ consists of the blocks of $\widetilde{\lambda}_A$ and $\widetilde{\lambda}_A^{-1}$ terms which include some fixed $\widetilde{\lambda}_A$ (or $\widetilde{\lambda}_A^{-1}$). Because the positions of the λ_A and λ_A^{-1} in ψ_1 are the same as the positions of the $\widetilde{\lambda}_A$ and $\widetilde{\lambda}_A^{-1}$ in ψ_2, we only need to show that for each block B, the summands above are equal. Clearly, the cumulants are equal, so we just need to prove that

$$\prod_{i=1}^{|B|} \Phi(\psi_1^i) = \prod_{i=1}^{|B|} \Phi(\psi_2^i).$$

Because ψ_1 and ψ_2 have the same combinatorial structure as loops, for each i, the subwords ψ_1^i and ψ_2^i describe the same subloop. If ψ_1^i (and ψ_2^i) does not include λ_A (and $\widetilde{\lambda}_A$), then this subloop does not actually pass through e. So, we can imagine deleting e from $\mathbb{G}$ to obtain a graph with less bounded faces than $\mathbb{G}$. We can still define this subloop validly on this graph, so we can just apply the inductive hypothesis to obtain $\Phi(\psi_1^i) = \Phi(\psi_2^i)$.

If ψ_1^i (and ψ_2^i) does include λ_A (and $\widetilde{\lambda}_A$), then the subloop cannot necessarily be defined on a graph with less bounded faces than $\mathbb{G}$. Instead, we repeat the above process with $\psi_1 = \psi_1^i$ and $\psi_2 = \psi_2^i$. We expand $\Phi(\psi_1^i)$ using the moment-cumulant relation conditioned on blocks of λ_A and λ_A^{-1} terms which include some fixed λ_A (or λ_A^{-1}). The analogous expansion is done for $\Phi(\psi_2^i)$, and we continue this downward process until eventually none of the moments in the expansion of $\Phi(\psi_1^i)$ contain λ_A and λ_A^{-1} terms. At this point, the edge e is not in any of the subloops, so every subloop can be defined on a graph with less bounded faces than $\mathbb{G}$. By the inductive hypothesis applied to each subloop, the moments for each subloop must be equal. Since the corresponding cumulants are also always equal, we can sum everything to get $\Phi(\psi_1^i) = \Phi(\psi_2^i)$.

Having proved $\Phi(\psi_1^i) = \Phi(\psi_2^i)$ for all i, we get $\Phi(\psi_1) = \Phi(\psi_2)$, as desired.

2.2.2. *Case 2.* For every edge $e \in \mathbb{E}^+$ bordering the unbounded face, either $e \in \mathsf{T}$ or $e \in \widetilde{\mathsf{T}}$.

Call a bounded face F an *exterior face* if some edge of F also borders the unbounded face. We first consider the case when $\mathbb{G}$ has exactly one exterior face (see Figure 4). There is exactly one edge $e \in \mathbb{E}^+$ bordering the unbounded face which does not belong to T. Let A be the exterior face of $\mathbb{G}$.

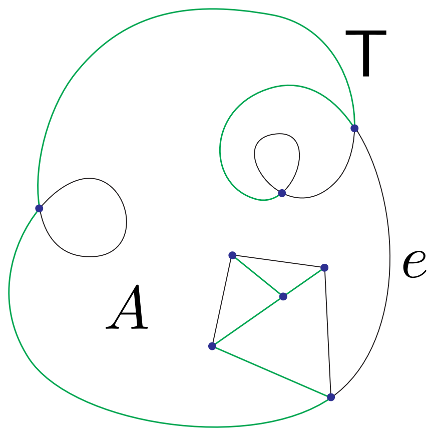

FIGURE 4. A graph with one exterior face

Let $\overline{\mathsf{T}}$ be the spanning tree T, except that the edge e, which borders the exterior face A and the unbounded face, is shifted one edge left to $\overline{e}$ (see Figure 5). Note that throughout this section, we use $\overline{e}$ to refer to this edge, not the ending vertex of e. It suffices to prove that the value of $\Phi(l)$ is equal under T and $\overline{\mathsf{T}}$. In particular, this would imply that we can continue "rotating" this special edge e until it matches the corresponding edge $\widetilde{e} \notin \widetilde{\mathsf{T}}$ on the boundary of $\mathbb{G}$. Then we can finish by applying the argument in Section 2.2.1.

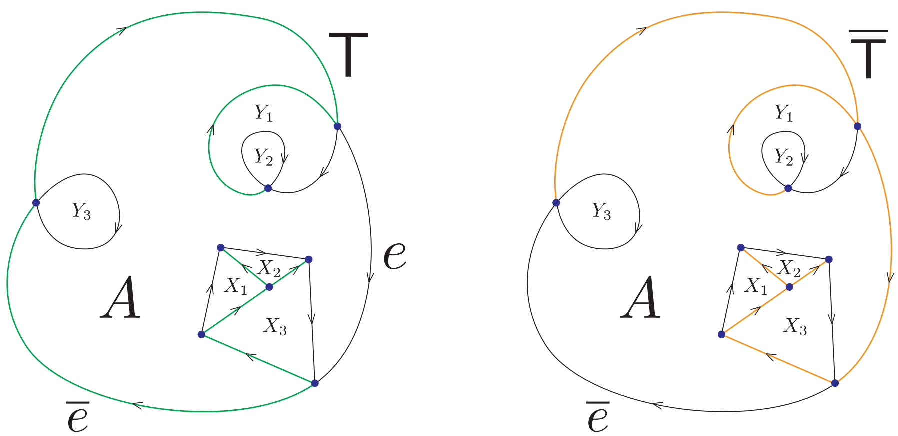

FIGURE 5. "Rotate" the edge e in T once to the left to obtain the spanning tree $\overline{\mathsf{T}}$.

Now, assume that the common vertex of e and $\overline{e}$ has degree greater than 2, i.e., there is some loop starting at this common vertex which lies within the boundary of $\mathbb{G}$. Further assume without loss of generality that the orientation of e and $\overline{e}$ follows that of Figure 5. Let $X_1, \ldots, X_n$ denote the faces in the loop between e and $\overline{e}$, and let $Y_1, \ldots, Y_m$ denote the other faces in $\mathbb{G}$ (excluding A). It is always possible to order the $X_1, \ldots, X_n$ and the $Y_1, \ldots, Y_m$ and to orient the edges such that

$$\beta_e = \lambda_A \lambda_{Y_1} \cdots \lambda_{Y_m} \lambda_{X_1} \cdots \lambda_{X_n}.$$

For convenience, define the words $\alpha = AY_1 \cdots Y_m$ and $\sigma = \alpha X_1 \cdots X_n$. Let λ_α and λ_σ be the corresponding words in the lasso basis, i.e.., $\lambda_\alpha = \lambda_A \lambda_{Y_1} \cdots \lambda_{Y_m}$ and $\lambda_\sigma = \beta_e$. Through a series of omitted lasso basis calculations, we can obtain the following change of basis formula.

Proposition 9. *Let λ and $\overline{\lambda}$ be the lasso bases under T and $\overline{\mathsf{T}}$ respectively.*

- *For all $1 \leq i \leq m$, we have $\overline{\lambda}_{Y_i} = \lambda_{Y_i}$.*
- *For all $1 \leq i \leq n$, we have $\overline{\lambda}_{X_i} = \lambda_\sigma \lambda_{X_i} \lambda_\sigma^{-1}$.*
- *We have $\overline{\lambda}_A = \lambda_\sigma \lambda_\alpha^{-1} \lambda_A \lambda_\alpha \lambda_\sigma^{-1}$.*

To prove $\Phi(\psi_1) = \Phi(\psi_2)$, we prove the following more general statement.

Proposition 10. *Let $\psi(X_1, \ldots, X_n, A)$ denote any word in the alphabet $X_1, \ldots, X_n, A$, interpreted as a loop under the lasso basis λ. Define $\psi(\overline{X}_1, \ldots, \overline{X}_n, \overline{A}) = \psi(\sigma X_1 \sigma^{-1}, \ldots, \sigma X_n \sigma^{-1}, \sigma\alpha^{-1} A \alpha \sigma^{-1})$. Then,*

$$\Phi(\psi(X_1, \ldots, X_n, A)) = \Phi(\psi(\overline{X}_1, \ldots, \overline{X}_n, \overline{A})).$$

Indeed, it is obvious that proving

$$\Phi(\psi(X_1, \ldots, X_n, Y_1, \ldots, Y_m, A)) = \Phi(\psi(\overline{X}_1, \ldots, \overline{X}_n, \overline{Y}_1, \ldots, \overline{Y}_m, \overline{A}))$$

would complete the proof for all $\mathbb{G}$ with one exterior face. However, since $\overline{Y}_i = Y_i$ for all $1 \leq i \leq m$, it suffices to just prove the equality in Proposition 10. To prove this, we first need a useful lemma on the "conjugate invariance" property of Φ.

Lemma 11. *Let $A_1, \ldots, A_k$ be an alphabet of k letters such that the k random variables associated with $A_1, \ldots, A_k$ are mutually free. If W is some word in this alphabet not containing A_1 or A_1^{-1}, then*

$$\Phi(\psi(A_1, A_2, \ldots, A_k)) = \Phi(\psi(WA_1W^{-1}, A_2, \ldots, A_k)).$$

Proof. The proof is very similar to the argument in Section 2.2.1. We apply the moment-cumulant relation to expand $\Phi(\psi(A_1, \ldots, A_k))$ into a sum of cumulants. Since the k random variables are mutually free, all the nonzero cumulants will only involve one random variable (and its inverse). We condition our sum on blocks of A_1 and A_1^{-1} terms which include some fixed A_1 (or A_1^{-1}). Such a block subdivides the word $\psi(A_1, \ldots, A_k)$ into distinct subwords, such that no other block can contain terms from different subwords. We use the same strategy to expand $\Phi(\psi(WA_1W^{-1}, A_2, \ldots, A_k))$ into cumulants. This sum runs over the same blocks because W does not contain A_1 or A_1^{-1}.

Like in Section 2.2.1, it suffices to prove that the Φ values of any two corresponding subwords are equal. Because $\psi(WA_1W^{-1}, A_2, \ldots, A_k)$ is obtained by replacing every A_1 in $\psi(A_1, \ldots, A_k)$ by WA_1W^{-1}, every subword in the former word starts with W (or W^{-1}) and ends with W^{-1} (or W). Hence, these two ends cancel out when we apply Φ onto this subword. However, the resulting subword is not necessarily equal to the corresponding subword in $\psi(A_1, \ldots, A_k)$ because there could still be W and W^{-1} terms in the middle of the subword. Again, we use the same idea as in 2.2.1 by applying the moment-cumulant relation onto the subword. Eventually, none of the moments in the expansion will contain A_1 or A_1^{-1} terms, and hence W and W^{-1} terms. At this point, the equality becomes trivial. ∎

Obviously, we can extend Lemma 11 to replacing any A_i with WA_iW^{-1} for W not containing A_i or A_i^{-1}. We will use this lemma extensively throughout our proof of Proposition 10. The idea is to slowly transform $\Phi(\psi(\overline{X}_1,\ldots,\overline{X}_n,\overline{A}))$ into $\Phi(\psi(X_1,\ldots,X_n,A))$ through a series of replacements of the form $X\mapsto WXW^{-1}$, where W does not contain X or X^{-1}. By Lemma 11, the Φ value remains constant throughout these transformations.

Proof of Proposition 10. First, note that

$$\begin{aligned}\sigma X_n\sigma^{-1} &= \alpha X_1\cdots X_nX_nX_n^{-1}\cdots X_1^{-1}\alpha^{-1}\\ &= \alpha X_1\cdots X_{n-1}X_nX_{n-1}^{-1}\cdots X_1^{-1}\alpha^{-1}.\end{aligned}$$

Define $\sigma_n=\alpha X_1\cdots X_{n-1}$, so that $\overline{X}_n=\sigma X_n\sigma^{-1}=\sigma_nX_n\sigma_n^{-1}$. Note that the word σ_n does not contain X_n. For all X_n in $\psi(\overline{X}_1,\ldots,\overline{X}_n,\overline{A})$, transform $X_n\mapsto\sigma_n^{-1}X_n\sigma_n$. All $\overline{X}_n$ are hence transformed $\overline{X}_n=\sigma_nX_n\sigma_n^{-1}\mapsto\sigma_n\sigma_n^{-1}X_n\sigma_n\sigma_n^{-1}=X_n$. For $i<n$, the $\overline{X}_i$ are transformed

$$\begin{aligned}\sigma X_i\sigma^{-1} &\mapsto \alpha X_1\cdots X_{n-1}\sigma_n^{-1}X_n\sigma_nX_i\sigma_n^{-1}X_n^{-1}\sigma_nX_{n-1}^{-1}\cdots X_1^{-1}\alpha^{-1}\\ &= X_n\sigma_nX_i\sigma_n^{-1}X_n^{-1}.\end{aligned}$$

For now, we will ignore A and just say that it was transformed $\overline{A}\mapsto\overline{A}_n$. Then by Lemma 11,

$$\Phi(\psi(\overline{X}_1,\ldots,\overline{X}_n,\overline{A}))=\Phi(\psi(X_n\sigma_nX_1\sigma_n^{-1}X_n^{-1},\ldots,X_n\sigma_nX_{n-1}\sigma_n^{-1}X_n^{-1},X_n,\overline{A}_n)).$$

For the next step, note that

$$\begin{aligned}X_n\sigma_nX_{n-1}\sigma_n^{-1}X_n^{-1} &= X_n\alpha X_1\cdots X_{n-1}X_{n-1}X_{n-1}^{-1}\cdots X_1^{-1}\alpha^{-1}X_n^{-1}\\ &= X_n\alpha X_1\cdots X_{n-2}X_{n-1}X_{n-2}^{-1}\cdots X_1^{-1}\alpha^{-1}X_n^{-1}.\end{aligned}$$

Define $\sigma_{n-1}=X_n\alpha X_1\cdots X_{n-2}$, so that $X_n\sigma_nX_{n-1}\sigma_n^{-1}X_n^{-1}=\sigma_{n-1}X_{n-1}\sigma_{n-1}^{-1}$. Note that the word σ_{n-1} does not contain X_{n-1}. For all X_{n-1} in $\psi(X_n\sigma_nX_1\sigma_n^{-1}X_n^{-1},\ldots,X_n\sigma_nX_{n-1}\sigma_n^{-1}X_n^{-1},X_n,\overline{A}_n)$, transform $X_{n-1}\mapsto\sigma_{n-1}^{-1}X_{n-1}\sigma_{n-1}$. All $X_n\sigma_nX_{n-1}\sigma_n^{-1}X_n^{-1}$ are transformed $X_n\sigma_nX_{n-1}\sigma_n^{-1}X_n^{-1}=\sigma_{n-1}X_{n-1}\sigma_{n-1}^{-1}\mapsto X_{n-1}$. For $i<n-1$, the $X_n\sigma_nX_i\sigma_n^{-1}X_n^{-1}$ are transformed

$$\begin{aligned}X_n\sigma_nX_i\sigma_n^{-1}X_n^{-1} &\mapsto X_n\alpha X_1\cdots X_{n-2}\sigma_{n-1}^{-1}X_{n-1}\sigma_{n-1}X_i\sigma_{n-1}^{-1}X_{n-1}^{-1}\sigma_{n-1}X_{n-2}^{-1}\cdots X_1^{-1}\alpha^{-1}X_n^{-1}\\ &= X_{n-1}\sigma_{n-1}X_i\sigma_{n-1}^{-1}X_{n-1}^{-1}.\end{aligned}$$

Again, suppose that $\overline{A}_n$ was transformed $\overline{A}_n\mapsto\overline{A}_{n-1}$. By Lemma 11,

$$\Phi(\psi(\overline{X}_1,\ldots,\overline{X}_n,\overline{A}))=\Phi(\psi(X_{n-1}\sigma_{n-1}X_1\sigma_{n-1}^{-1}X_{n-1}^{-1},\ldots,X_{n-1},X_n,\overline{A}_{n-1})).$$

For the third step, note that

$$\begin{aligned}X_{n-1}\sigma_{n-1}X_{n-2}\sigma_{n-1}^{-1}X_{n-1}^{-1} &= X_{n-1}X_n\alpha X_1\cdots X_{n-2}X_{n-2}X_{n-2}^{-1}\cdots X_1^{-1}\alpha^{-1}X_n^{-1}X_{n-1}^{-1}\\ &= X_{n-1}X_n\alpha X_1\cdots X_{n-3}X_{n-2}X_{n-3}^{-1}\cdots X_1^{-1}\alpha^{-1}X_n^{-1}X_{n-1}^{-1}.\end{aligned}$$

Define $\sigma_{n-2}=X_{n-1}X_n\alpha X_1\cdots X_{n-3}$, and transform $X_{n-2}\mapsto\sigma_{n-2}^{-1}X_{n-2}\sigma_{n-2}$. Repeat this process all the way to X_1 and σ_1, where in general we define $\sigma_i=X_{i+1}\cdots X_n\alpha X_1\cdots X_{i-1}$. After applying these n transformations in order, Lemma 11 gives

$$\Phi(\psi(\overline{X}_1,\ldots,\overline{X}_n,\overline{A}))=\Phi(\psi(X_1,\ldots,X_n,\overline{A}_1)).$$

Finally, we just need to examine the $\overline{A}$ term. Recall that
$$\overline{A} = \alpha X_1 \cdots X_n \alpha^{-1} A \alpha X_n^{-1} \cdots X_1^{-1} \alpha^{-1}.$$
After the transformation $X_n \mapsto \sigma_n^{-1} X_n \sigma_n$, the term $\overline{A}$ is transformed into
$$\overline{A}_n = X_n \alpha X_1 \cdots X_{n-1} \alpha^{-1} A \alpha X_{n-1}^{-1} \cdots X_1^{-1} \alpha^{-1} X_n^{-1}.$$
After the transformation $X_{n-1} \mapsto \sigma_{n-1}^{-1} X_{n-1} \sigma_{n-1}$, the term $\overline{A}_n$ is transformed into
$$\overline{A}_{n-1} = X_{n-1} X_n \alpha X_1 \cdots X_{n-2} \alpha^{-1} A \alpha X_{n-2}^{-1} \cdots X_1^{-1} \alpha^{-1} X_n^{-1} X_{n-1}^{-1}.$$
Continuing this pattern, after n transformations, we have
$$\begin{aligned} \overline{A}_1 &= X_1 \cdots X_n \alpha \alpha^{-1} A \alpha \alpha^{-1} X_n^{-1} \cdots X_1^{-1} \\ &= X_1 \cdots X_n A X_n^{-1} \cdots X_1^{-1}. \end{aligned}$$
Apply the final transformation $A \mapsto X_n^{-1} \cdots X_1^{-1} A X_1 \cdots X_n$, and Lemma 11 gives
$$\Phi(\psi(\overline{X}_1, \ldots, \overline{X}_n, \overline{A})) = \Phi(\psi(X_1, \ldots, X_n, A)),$$
as desired. □

Now, consider the case when $\mathbb{G}$ has more than one exterior face. Since T is a spanning tree, there exists some edge $e \in \mathbb{E}^+$ on the boundary with $e \notin \mathsf{T}$. Let A be the bounded face bordered by e. Similarly, there exists some edge $\widetilde{e} \in \mathbb{E}^+$ on the boundary with $\widetilde{e} \notin \widetilde{\mathsf{T}}$. Let $\widetilde{A}$ be the bounded face bordered by $\widetilde{e}$.

If $A = \widetilde{A}$, the situation is almost identical to the case when $\mathbb{G}$ has one exterior face (see Figure 6).

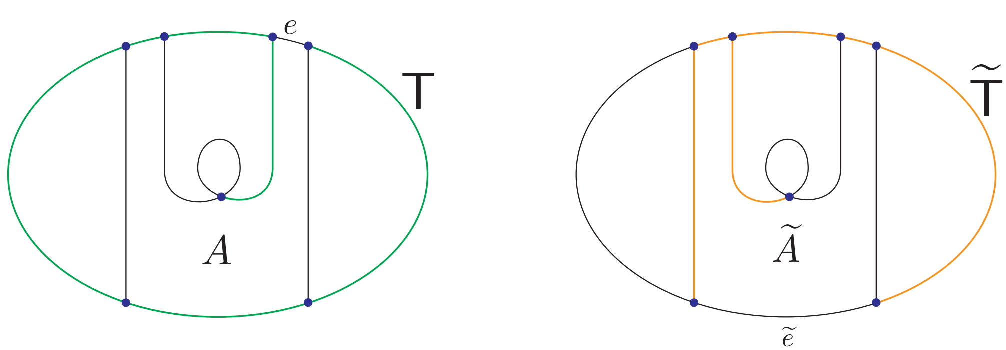

FIGURE 6. In this case, there is some face $A = \widetilde{A}$ which has both an edge $e \notin \mathsf{T}$ on the boundary and an edge $\widetilde{e} \notin \widetilde{\mathsf{T}}$ on the boundary.

We want to apply the same "rotation" above to rotate e to $\widetilde{e}$ so that we can reduce the problem to 2.2.1. The only difference here is that the edges of A bordering the unbounded face do not necessarily form a cycle which we can naturally rotate along. However, it is not hard to see that we can slightly generalize our proof above to allow us to skip over edges of A which do not border the unbounded face.

Otherwise, assume $A \neq \widetilde{A}$ (see Figure 7). This is possible because $\mathbb{G}$ has more than one exterior face.

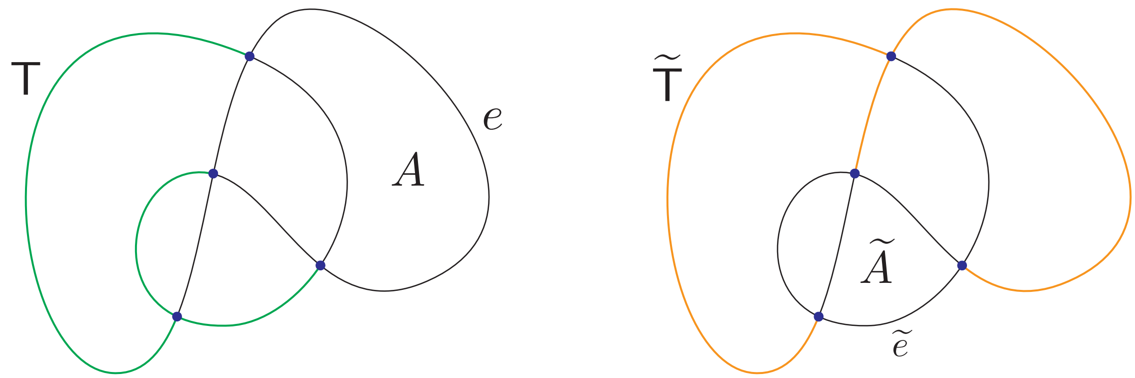

FIGURE 7. In this case, the faces A and $\widetilde{A}$ are distinct.

Define a new spanning tree $\mathcal{T}$ as follows. If $\widetilde{e} \notin \mathsf{T}$, let $\mathcal{T} = \mathsf{T}$. Otherwise, let $e_{\widetilde{A}}$ be the edge whose dual edge $e^*_{\widetilde{A}}$ is the first edge traversed on the path $[\widetilde{A}, F_\infty]_{\mathsf{T}^*}$. By definition, $e_{\widetilde{A}} \notin \mathsf{T}$. Let $\mathcal{T}$ be the tree T with $\widetilde{e}$ replaced by $e_{\widetilde{A}}$. It is easy to check that $\mathcal{T}$ is indeed a spanning tree (see Figure 8).

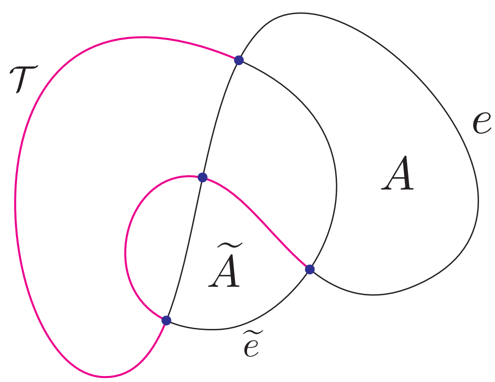

FIGURE 8. Construct a spanning tree $\mathcal{T}$ which does not include e or $\widetilde{e}$.

Now, since T and $\mathcal{T}$ both do not include the edge e, from Section 2.2.1 we know that the $\Phi(l)$ values under these two spanning trees must agree. Similarly, since $\widetilde{\mathsf{T}}$ and $\mathcal{T}$ both do not include the edge $\widetilde{e}$, the $\Phi(l)$ values under these two spanning trees must agree. Putting these two equalities together, we get that the $\Phi(l)$ values under T and $\widetilde{\mathsf{T}}$ must agree, i.e., $\Phi(\psi_1) = \Phi(\psi_2)$, hence completing the proof.

3. PROOF FOR THE MAKEENKO–MIGDAL EQUATIONS

3.1. **Proof of Theorem 7.** Let l be a loop in $\mathsf{L}_o(\mathbb{R}^2)$, and define l_1 and l_2 like in the statement of Theorem 7. The loop l naturally induces a planar graph $\mathbb{G} = (\mathbb{V}, \mathbb{E}, \mathbb{F})$ along with an orientation $\mathbb{E}^+ \subset \mathbb{E}$. Fix a point of self-intersection with exactly two ingoing edges and two outgoing edges.

For each sufficiently small $\varepsilon > 0$, we will define a loop l'_ε by slightly modifying $\mathbb{G}$ and l. This family of loops will be defined such that as $\varepsilon \to 0$, the loops l'_ε will converge to l. Henceforth, we fix ε and omit the ε in the subscript of l'_ε. Also note that in this section, we assume that Φ is a well-defined function on $\mathsf{L}_o(\mathbb{R}^2)$.

To modify $\mathbb{G}$, add one vertex on each of the four edges connected to the point of self-intersection, and join the four new vertices cyclically to create four new faces. If ε is small enough, it is always possible to do this such that these four faces all have area ε. Without loss of generality, orient the four new edges in counter-clockwise direction. For later convenience, we will still use $\mathbb{G} = (\mathbb{V}, \mathbb{E}, \mathbb{F})$ to refer to this modified graph. If we apply this modification to the graph in Figure 1, we obtain the graph in Figure 9.

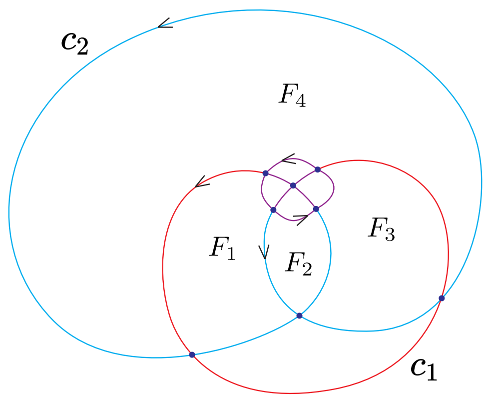

FIGURE 9. Build four new faces, each with area ε, around the point of self-intersection.

Zoom in and label the four new faces and edges as shown in Figure 10.

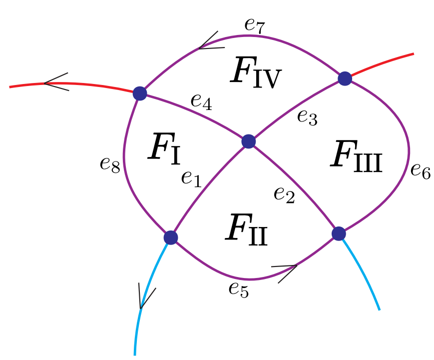

FIGURE 10. Label the four new faces $F_{\mathrm{I}}, F_{\mathrm{II}}, F_{\mathrm{III}}, F_{\mathrm{IV}}$ and the eight new edges $e_1, \ldots, e_8$.

Note that the highlighted red and blue paths in Figure 9 are no longer loops, and we call them c_1 and c_2 respectively instead. The loop l can be written as $l = e_1 c_1 e_2 e_4 c_2 e_3$. We define the modified loop l' to be

$$l' = e_2^{-1} e_5^{-1} c_1 e_2 e_4 c_2 e_7 e_4^{-1}.$$

The key idea of our proof is to focus on the value

$$\lim_{\varepsilon \to 0} \frac{\Phi(l') - \Phi(l)}{\varepsilon}.$$

We will prove that this expression is equal to both sides of (2). Note that we will only actually take the right limit $\varepsilon \to 0^+$. We could easily modify the definition of l' to include negative ε. However, because $(\Phi(l') - \Phi(l))/\varepsilon$ is a smooth function in ε (it is polynomial in ε and $e^{-\varepsilon/2}$) over some domain of the form $(-\delta, \delta) \setminus \{0\}$, it suffices to only consider $\varepsilon > 0$.

3.1.1. *Left hand side of* (2). Imagine erasing the four edges e_1, e_3, e_6, e_8, thus deleting the four faces $F_{\mathrm{I}}, F_{\mathrm{II}}, F_{\mathrm{III}}, F_{\mathrm{IV}}$ (see Figure 11).

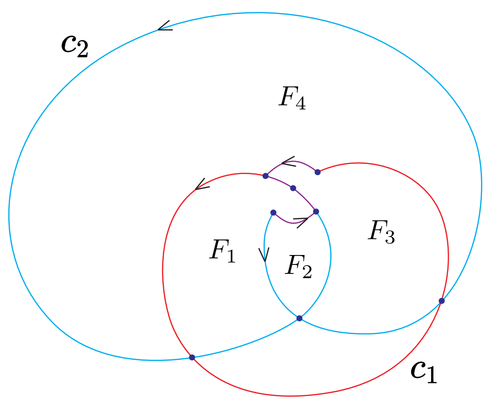

FIGURE 11. The edges e_1, e_3, e_6, e_8 have been erased from $\mathbb{G}$.

Our definition of l' still stands because it does not involve any of e_1, e_3, e_6, e_8. Furthermore, the value of $\Phi(l')$ remains unchanged because the loop itself did not change. Note that this new graph is almost the same as the original $\mathbb{G}$ (before the modification), except that the areas of F_1 and F_3 have both increased by ε, while the areas of F_2 and F_4 have both decreased by ε. So by treating $\Phi(l)$ as a function of the areas of the bounded faces delimited by l, basic properties of derivatives imply the equality

$$\lim_{\varepsilon \to 0} \frac{\Phi(l') - \Phi(l)}{\varepsilon} = \left(\frac{d}{d|F_1|} - \frac{d}{d|F_2|} + \frac{d}{d|F_3|} - \frac{d}{d|F_4|} \right) \Phi(l).$$

3.1.2. *Right hand side of* (2). First, choose a spanning tree T on $\mathbb{G}$ which includes the edges e_2, e_5, e_7, e_8 (see Figure 12 for an example).

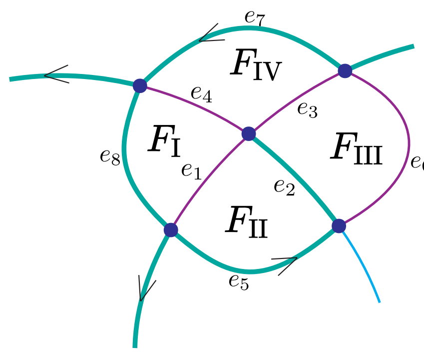

FIGURE 12. The spanning tree (highlighted in green) includes the edges e_2, e_5, e_7, e_8.

Let $\{\lambda_F : F \in \mathbb{F}^b\}$ and $\{\beta_e : e \in \mathbb{E}^+ \setminus \mathsf{T}\}$ be the lasso and beta bases respectively. For convenience, abbreviate $\lambda_i = \lambda_{F_i}$ and $\beta_i = \beta_{e_i}$. We have

$$\begin{aligned}\beta_1 &= \lambda_{\mathrm{II}}\\ \beta_3 &= \lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}\\ \beta_4 &= \lambda_{\mathrm{I}}\lambda_{\mathrm{II}}\\ \beta_6 &= \lambda_{\mathrm{III}}\lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}.\end{aligned}$$

For later convenience, define the loops

$$\gamma_i = [o, \underline{c_i}]_{\mathsf{T}} c_i [\overline{c_i}, o]_{\mathsf{T}}$$

for $i = 1, 2$, where $\underline{c_i}$ (resp. $\overline{c_i}$) is the starting (resp. ending) vertex of the path c_i. Because we are working in $\mathsf{RL}_o(\mathbb{G})$, we can always replace c_i by γ_i and the loop will remain unchanged. Hence, we can write $l' = e_2^{-1}e_5^{-1}\gamma_1 e_2 e_4 \gamma_2 e_7 e_4^{-1}$ and $l = e_1\gamma_1 e_2 e_4 \gamma_2 e_3$.

Our next replacement involves the beta basis. Recall that for any loop l written as a concatenation of edges, we can replace each edge e in the concatenation by β_e and the loop will remain unchanged in $\mathsf{RL}_o(\mathbb{G})$. Applying this replacement to the loop l', we get

$$l' = \gamma_1\beta_4\gamma_2\beta_4^{-1},$$

where we understand γ_1 and γ_2 to be words in the beta basis. We can then change from the beta basis into the lasso basis to get

$$l' = \gamma_1\lambda_{\mathrm{I}}\lambda_{\mathrm{II}}\gamma_2\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1},$$

where we now understand γ_1 and γ_2 to be words in the lasso basis. We can do the same process with the loop l to get

$$l = \beta_1\gamma_1\beta_4\gamma_2\beta_3 = \lambda_{\mathrm{II}}\gamma_1\lambda_{\mathrm{I}}\lambda_{\mathrm{II}}\gamma_2\lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}.$$

Now that we have explicitly expanded both l' and l in terms of the lasso basis, we can compute the value of $\Phi(l') - \Phi(l)$. First, we will show that the terms $\lambda_{\mathrm{I}}, \lambda_{\mathrm{II}}, \lambda_{\mathrm{III}}, \lambda_{\mathrm{IV}}$ and their inverses can only appear in γ_1 and γ_2 through a fixed pattern.

Lemma 12. *Treat γ_1 and γ_2 as words in the lasso basis. Then the terms $\lambda_{\mathrm{I}}, \lambda_{\mathrm{II}}, \lambda_{\mathrm{III}}, \lambda_{\mathrm{IV}}$ and their inverses can only appear in γ_1 and γ_2 through the word $\lambda_{\mathrm{III}}\lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}$ or its inverse.*

Proof. It suffices to prove that for each edge $e \in \mathbb{E}^+ \setminus \mathsf{T}$, the expansion of β_e in terms of the lasso basis only contains the terms $\lambda_{\mathrm{I}}, \lambda_{\mathrm{II}}, \lambda_{\mathrm{III}}, \lambda_{\mathrm{IV}}$ and their inverses through the word $\lambda_{\mathrm{III}}\lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}$ or its inverse. We already know that the statement is true for e_1, e_3, e_4, e_6 from our calculations above. And since $\beta_6 = \lambda_{\mathrm{III}}\lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}$, Lévy's [Lév17] explicit formula for the change of basis between the beta and lasso bases quickly implies the result for the other edges of $\mathbb{G}$. ☐

We also need a short lemma on the cumulants of free Brownian motions.

Lemma 13. *Let u_t be a free Brownian motion at time t. For all $n \geq 3$ and $\varepsilon_1, \ldots, \varepsilon_n \in \{1, *\}$, the constant and linear terms in the expansion of $\kappa_n(u_t^{\varepsilon_1}, \ldots, u_t^{\varepsilon_n})$ as a power series in t are both zero.*

Proof. First, if all the $\varepsilon_1, \ldots, \varepsilon_n$ are equal to 1 or $*$, it is well known that

$$\kappa_n(u_t, \ldots, u_t) = \kappa_n(u_t^*, \ldots, u_t^*) = e^{-nt/2}\frac{(-n)^{n-1}}{n!} \cdot t^{n-1}.$$

Otherwise, induct downwards using the recursive formula

$$\kappa_n(u_t^{\varepsilon_1}, \ldots, u_t^{\varepsilon_n}) = -\sum_{m=1}^{n-1} \kappa_m(u_t^{\varepsilon_1}, \ldots, u_t^{\varepsilon_m}) \cdot \kappa_{n-m}(u_t^{\varepsilon_{m+1}}, \ldots, u_t^{\varepsilon_n})$$

proven by Demni, Guay-Paquet, and Nica [DGPN15]. Note that to use this formula, we require that $\varepsilon_1 = 1$ and $\varepsilon_n = *$, but this can always be achieved by taking a suitable cyclic permutation of $(\varepsilon_1, \ldots, \varepsilon_n)$. The cases $n = 1, 2$ can easily be calculated explicitly. ☐

Now, we are ready to attack $\Phi(l')$ and $\Phi(l)$. We only need the $O(\varepsilon)$ term in $\Phi(l') - \Phi(l)$. Recall our expansions for l and l' in the lasso basis:

$$\begin{aligned} l &= \lambda_{\mathrm{II}}\gamma_1\lambda_{\mathrm{I}}\lambda_{\mathrm{II}}\gamma_2\lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}, \\ l' &= \gamma_1\lambda_{\mathrm{I}}\lambda_{\mathrm{II}}\gamma_2\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}. \end{aligned}$$

The two expressions only differ in that l contains an extra λ_{II} at the start and an extra λ_{IV} after γ_2. For clarity, these two terms are marked in red in the diagrams below. We will focus on these two terms when we expand $\Phi(l)$ using the moment-cumulant relation. By Proposition 4, all the nonzero cumulants will only involve one lasso (and its inverse). There are three cases to consider for this noncrossing partition.

Case 1: The terms λ_{II} and λ_{IV} are each in a block with size one. Then, we can naturally project the rest of the blocks in the noncrossing partition onto the blocks of a noncrossing partition in the expansion of $\Phi(l')$. Hence, if we sum over all such noncrossing partitions, this case contributes $e^{-\varepsilon}\Phi(l') = \Phi(l') - \varepsilon\Phi(l') + O(\varepsilon^2)$ to the value of $\Phi(l)$.

Case 2: Exactly one of the terms λ_{II} and λ_{IV} is in a block with size greater than one. If this block happened to have size greater than two, then Lemma 13 would imply that the contribution from this cumulant would be $O(\varepsilon^2)$. Otherwise, assume that this block has size two. We take subcases depending on where the other term in this block is located. The first subcase is shown in Figure 13.

In this subcase, we can assume that the terms $\lambda_{\mathrm{I}}^{-1}$ (marked in magenta below) and λ_{IV} are both in blocks of size one, as otherwise the contribution from this cumulant would be $O(\varepsilon^2)$. After making these assumptions, we can naturally project the rest of the blocks in the noncrossing partition onto

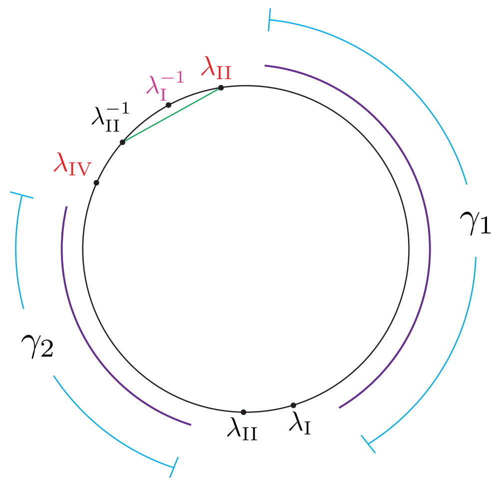

FIGURE 13. In the noncrossing partition, the terms λ_{II} (marked red) and $\lambda_{\mathrm{II}}^{-1}$ form a block.

the blocks of a noncrossing partition for the expansion of $\Phi(l')$ where the terms $\lambda_{\mathrm{I}}^{-1}$ and $\lambda_{\mathrm{II}}^{-1}$ are both in blocks with size one. If this subset of noncrossing partitions for $\Phi(l')$ (the noncrossing partitions for which $\lambda_{\mathrm{I}}^{-1}$ and $\lambda_{\mathrm{II}}^{-1}$ are both in blocks with size one) contributes $e^{-\varepsilon}f(\varepsilon)$ to the value of $\Phi(l')$, where $f(\varepsilon)$ is some power series in ε, then this subcase contributes $e^{-\varepsilon}(1-e^{-\varepsilon})f(\varepsilon)$ to the value of $\Phi(l)$. Since $\Phi(l') = e^{-\varepsilon}f(\varepsilon) + O(\varepsilon)$, we deduce that this subcase contributes $\varepsilon\Phi(l') + O(\varepsilon^2)$ to the value of $\Phi(l)$.

The second subcase is shown in Figure 14.

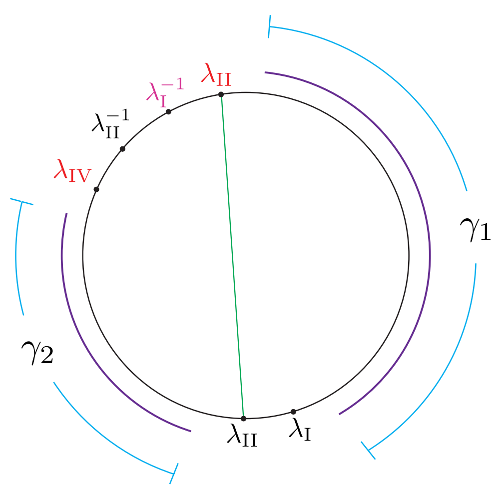

FIGURE 14. In the noncrossing partition, the terms λ_{II} (marked red) and λ_{II} form a block.

Similar to the previous subcase, we can assume that the other four terms not in γ_1 and γ_2 are all in blocks with size one, as otherwise the contribution of the cumulant would be $O(\varepsilon^2)$. Note that the block containing the terms λ_{II} (marked red) and λ_{II} essentially separates γ_1 from γ_2. This implies that no block of the noncrossing partition can contain terms from both γ_1 and γ_2. In particular, when we sum over all noncrossing partitions containing the green block, we are actually summing over all noncrossing partitions on γ_1 and then on γ_2. Hence, this subcase contributes $-\varepsilon e^{-3\varepsilon}\Phi(\gamma_1)\Phi(\gamma_2) = -\varepsilon\Phi(\gamma_1)\Phi(\gamma_2) + O(\varepsilon^2)$ to the value of $\Phi(l)$.

For the third subcase, we can assume that one of the terms λ_{II} and λ_{IV} is connected to a term in either γ_1 or γ_2. Assume without loss of generality that the block includes λ_{II}. As expected, we can assume that there are no other blocks with size greater than one that contain one of the terms $\lambda_{\mathrm{I}}, \lambda_{\mathrm{II}}, \lambda_{\mathrm{III}}, \lambda_{\mathrm{IV}}$ or their inverses. If such a block existed, the contribution of this cumulant would be $O(\varepsilon^2)$. Then by Lemma 12, the other term in the block containing λ_{II} belongs to the word $\lambda_{\mathrm{III}}\lambda_{\mathrm{IV}}\lambda_{\mathrm{II}}^{-1}\lambda_{\mathrm{I}}^{-1}$ or its inverse. We can then consider the noncrossing partition which instead has a block containing λ_{IV} (marked red) and the λ_{IV} or $\lambda_{\mathrm{IV}}^{-1}$ in that particular word, with everything else unchanged (see Figure 15).

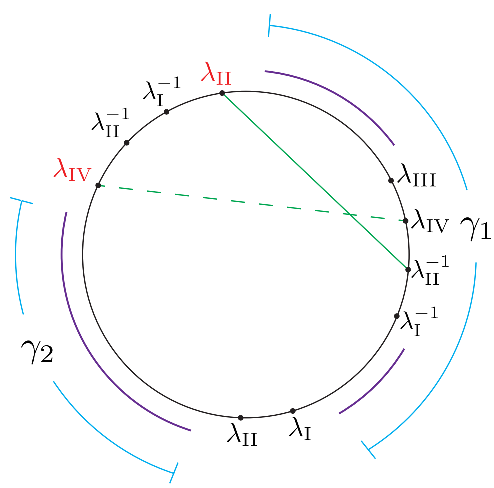

FIGURE 15. We can pair up noncrossing partitions when exactly one of λ_{II} (marked red) and λ_{IV} (marked red) is in a block with a term in γ_1 or γ_2.

This allows us to pair up the noncrossing partitions within this subcase. However, the contribution from each pair of cumulants is always $-\varepsilon e^{-\varepsilon}g(\varepsilon) + (1 - e^{-\varepsilon})g(\varepsilon)$ for some power series $g(\varepsilon)$. In particular, this contribution is always $O(\varepsilon^2)$. Thus, this subcase only contributes $O(\varepsilon^2)$ to the value of $\Phi(l)$.

Overall, this case contributes $\varepsilon\Phi(l') - \varepsilon\Phi(\gamma_1)\Phi(\gamma_2) + O(\varepsilon^2)$ to the value of $\Phi(l)$.

Case 3: Both λ_{II} and λ_{IV} are in blocks with size greater than one. Clearly, this case contributes $O(\varepsilon^2)$ to the value of $\Phi(l)$.

Finally, we can sum over all three cases to get

$$\begin{aligned}\Phi(l) &= \Phi(l') - \varepsilon\Phi(l') + \varepsilon\Phi(l') - \varepsilon\Phi(\gamma_1)\Phi(\gamma_2) + O(\varepsilon^2)\\ &= \Phi(l') - \varepsilon\Phi(\gamma_1)\Phi(\gamma_2) + O(\varepsilon^2).\end{aligned}$$

To finish, we take advantage of the continuity of Φ to get

$$\lim_{\varepsilon\to 0}\frac{\Phi(l') - \Phi(l)}{\varepsilon} = \lim_{\varepsilon\to 0}\Phi(\gamma_1)\Phi(\gamma_2) = \Phi(l_1)\Phi(l_2).$$

This completes the proof.

4. Acknowledgements

We would like to thank the MIT PRIMES program for making this research possible. We are also grateful to Minjae Park, Joshua Pfeffer, and Scott Sheffield for helpful discussions.

References

[AS12] M. Anshelevich and A. N. Sengupta. Quantum free Yang–Mills on the plane. *J. Geom. Phys.*, 62(2):330–343, 2012.

[Bia97] P. Biane. Free Brownian motion, free stochastic calculus and random matrices. In *Free Probability Theory*, volume 12 of *Fields Inst. Commun.*, pages 1–19. Amer. Math. Soc., Providence, RI, 1997.

[Dah16] A. Dahlqvist. Free energies and fluctuations for the unitary Brownian motion. *Comm. Math. Phys.*, 348:395–444, 2016.

[DGPN15] N. Demni, M. Guay-Paquet, and A. Nica. Star-cumulants of free unitary Brownian motion. *Adv. Appl. Math.*, 69:1–45, 2015.

[DHK17] B. K. Driver, B. C. Hall, and T. Kemp. Three proofs of the Makeenko-Migdal equations for Yang-Mills theory on the plane. *Comm. Math. Phys.*, 351:741–774, 2017.

[DK93] M. R. Douglas and V. A. Kazakov. Large N phase transition in continuum QCD_2. *Physics Letters B*, 319(1):219–230, 1993.

[Dou95] M. R. Douglas. *Conformal Field Theory Techniques in Large N Yang-Mills Theory*, pages 119–135. Springer US, Boston, MA, 1995.

[Dri89] B. K. Driver. YM_2: continuum expectations, lattice convergence, and lassos. *Comm. Math. Phys.*, 123(4):575–616, 1989.

[Fin91] D. S. Fine. Quantum Yang–Mills on a Riemann surface. *Comm. Math. Phys.*, 140(2):321–338, 1991.

[GM95] D. J. Gross and A. Matytsin. Some properties of large-N two-dimensional Yang-Mills theory. *Nuclear Phys. B*, 437(3):541–584, 1995.

[GT93a] D. J. Gross and W. Taylor. Twists and Wilson loops in the string theory of two-dimensional QCD. *Nuclear Phys. B*, 403(1):395–449, 1993.

[GT93b] D. J. Gross and W. Taylor. Two-dimensional QCD is a string theory. *Nuclear Phys. B*, 400(1):181–208, 1993.

[Itô50] K. Itô. Brownian motions in a Lie group. *Proc. Japan Acad.*, 26(8):4–10, 1950.

[Kaz81] V. A. Kazakov. Wilson loop average for an arbitrary contour in two-dimensional $\mathrm{U}(N)$ gauge theory. *Nuclear Phys. B*, 179(2):283–292, 1981.

[KK80] V. A. Kazakov and I. K. Kostov. Non-linear strings in two-dimensional $\mathrm{U}(\infty)$ gauge theory. *Nuclear Phys. B*, 176(1):199–215, 1980.

[Lév03] T. Lévy. *Yang–Mills measure on compact surfaces*. American Mathematical Soc., 2003.

[Lév10] T. Lévy. Two-dimensional Markovian holonomy fields. *Astérisque*, (1), 2010.

[Lév17] T. Lévy. The master field on the plane. *Astérisque*, 388:x+201, 2017.

[Lév20] T. Lévy. Two-dimensional quantum Yang–Mills theory and the Makeenko–Migdal equations. In *Frontiers in Analysis and Probability*, pages 275–325. Springer, 2020.

[MM79] Y. Makeenko and A. Migdal. Exact equation for the loop average in multicolor QCD. *Phys. Lett. B*, 88(1):135–137, 1979.
[MS17] J. Mingo and R. Speicher. *Free probability and random matrices*, volume 35. Springer, 2017.
[Pol80] A. M. Polyakov. Gauge fields as rings of glue. *Nuclear Phys. B*, 164:171–188, 1980.
[Sen97] A. N. Sengupta. *Gauge theory on compact surfaces*, volume 600. American Mathematical Soc., 1997.
[Sin95] I. M. Singer. On the master field in two dimensions. In *Functional Analysis on the Eve of the 21st Century*, volume 131 of *Progr. Math.*, pages 263–281. Birkhäuser Boston, Boston, MA, 1995.
[Spe14] R. Speicher. Free probability and random matrices. *arXiv preprint arXiv:1404.3393*, 2014.
[tH74] G. 't Hooft. A planar diagram theory for strong interactions. *Nuclear Phys. B*, 72(3):461–473, 1974.
[Wit91] E. Witten. On quantum gauge theories in two dimensions. *Comm. Math. Phys.*, 141(1):153–209, 1991.

Phillips Academy Andover, Andover, MA 01810, USA
Email address: nxiong22@andover.edu

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Email address: puyu1516@mit.edu